

# Chapter 1. Defining Efficiency

Efficiency is one of the most important values in our culture. In the same way that societies throughout history have imposed great personal sacrifices on individuals in order to appease the gods, or to honour the memory of their ancestors, or serve the glory of the nation, in our society we have all, in one way or another, been asked to make sacrifices in the name of "efficiency." Workers are replaced by machines, social services are delivered by impersonal bureaucracies, neighbourhood shops are put out of business by giant malls and supermarkets, not because anyone seems to really like it that way, but because doing it that way is "efficient." Similarly, we have all been criticized at one time or another by others for doing things "inefficiently." Unfortunately, despite its widespread use and acceptance, the idea of efficiency is very poorly understood. Apart from the fact that the idea is seldom given a precise definition, it is also not widely understood how the centrality of this principle in our society is connected to a number of other core democratic values, such as the significance that we attach to personal freedom, pluralism, and multiculturalism. Understanding efficiency correctly, as we will see, involves understanding several key elements of cultural modernity.

## 1.1 Internal and External Criteria

Before beginning, it is important to clear up one widespread misunderstanding about the way the efficiency standard is applied to judge social states. We use the word "efficiency" to describe both social and non-social processes, e.g. we talk about efficient machines, efficient routines, etc., but also about efficient administrative procedures, efficient organizations, etc. Things are complicated by the fact that the word can mean something subtly different in these different cases. Machines, for example, are designed to do certain things for us. Increasing the efficiency of a machine means that on balance it will do what it is supposed to do better. In the abstract, a machine can be thought of as a device that takes a certain input and transforms it into output that it is of greater use to us. A more efficient machine either produces more output with the same input, or it produces the same output with diminished input, or it produces the same output in less time, etc. These all amount to the same thing. If the input is consumed, and

therefore thought of as a "cost," while the output is thought of as the "benefit," then the most efficient machine is the one that, in a given time, produces the greatest balance of benefits over costs. In this way, efficiency can be determined using what is called **cost-benefit analysis**. If one can find a way to measure the costs and the benefits on a commensurable scale, then one can subtract the costs of a given process from its benefits. Maximizing efficiency then means maximizing net benefit.

The important thing to notice about all this is that the notion of cost and benefit is relative to human evaluation. A machine is efficient if it produces something we want. If we change our minds about what we want, then a machine that was once efficient may become inefficient. Efficiency is therefore not an intrinsic or natural property of the machine. Forest fires and volcanoes are extremely 'efficient' at generating atmospheric pollution, but we do not want atmospheric pollution, and so do not normally think of them as efficient. Someday we might, because this sort of pollution diminishes the greenhouse effect, and therefore slows global warming. Similarly, internal combustion engines are far more 'efficient' at producing heat than they are torque, but we tend not to want this heat (in fact, we have to add another machine to the engine – the radiator – specifically to get rid of it). Things are different, however, if we are running the engine in order to warm up. Thus efficiency is relative to what is considered a cost and what is considered a benefit.

This means that the efficiency of a machine presupposes a certain social consensus about the value of its inputs and outputs. Thus judgments of efficiency of this type must bring in some "external" criteria to specify what the judgment is relative to. Things are different, however, in the case in individual action. We often speak of choices that people make as more or less efficient. For example, we talk about the most efficient way to get downtown, to mow the lawn, to tabulate grades, to succeed in advertising, etc. In this case, we are not evaluating the action according to some external standard, but in terms of the agent's own goals and preferences. We are seldom in a position to satisfy all of our preferences, and so choosing to pursue any particular goal will involve foregoing the opportunity to pursue some others. The things that we want, but that we choose to do without, are sacrifices. Goals that we succeed in realizing are achievements. Efficient actions are ones that produce a greater balance of achievement over sacrifice. (The view

that rational decision involves selecting the most efficient action available, in this sense, is referred to as the **maximizing conception of practical rationality**.)

The important thing to notice about this definition of efficiency is that it relies upon entirely "internal" criteria in order to judge actions. The individual's own goals supply the values that determine what counts as a cost and what counts as a benefit. So whereas a machine can be efficient only relative to the values of the people it serves, individuals' actions can be efficient if they produce more of what the individual alone wants. However, given that people can also judge each others' actions, it is also possible to judge the efficiency of an individual's action from an "external" perspective, i.e. in terms of how well their actions produce something that *others* want. And since it is often the case that not everyone wants the same thing, there is no reason that the internal criteria and the external criteria should always coincide. For instance, a worker may be very inefficient, from the standpoint of his employer, but perfectly efficient from his own point of view. If he has, like most people, a preference for leisure over labour, then he may seek to earn his wages with the least effort. Efficiency will therefore mandate that he put in just enough effort to avoid being fired. This may lead to ways of doing things that are inefficient from the employer's point of view, e.g. the worker may take the "scenic route" while making deliveries, and so on.

This leads to a certain ambiguity in our everyday use of the term efficient. When one person criticizes another for being inefficient, it can be meant as an internal evaluation: "you would be better off if you did it this way," or external: "I would be better off if you did it this way." This ambiguity has led to serious confusion in the application of the efficiency standard to the evaluation of social states. In general, if social institutions are evaluated using "external" criteria, the efficiency standard takes the form of classical **utilitarianism**. On the other hand, if "internal" criteria are used, then the efficiency standard gives rise to what is called the principle of **Pareto efficiency**.

## 1.2 Utilitarianism

There is clearly a sense in which the efficiency of social institutions, just like machines, can be evaluated using external criteria. A bureaucracy, for instance, can be evaluated by looking

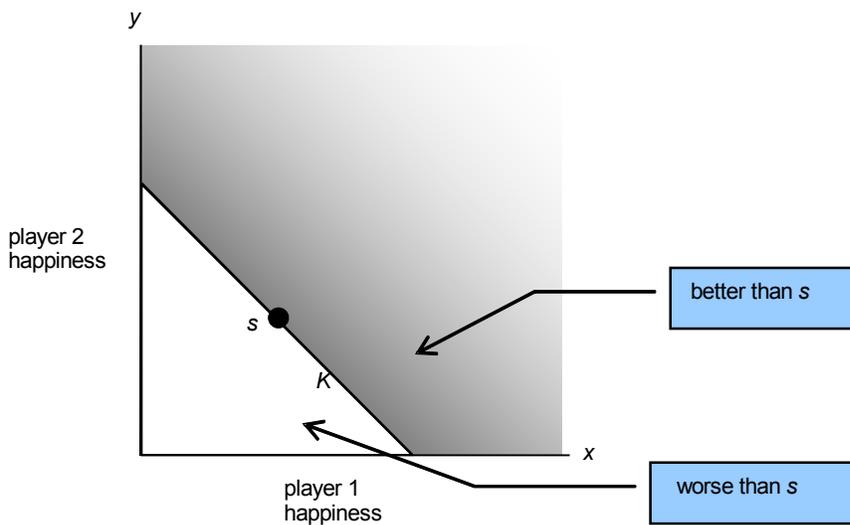
at what tasks it is supposed to perform and the resources that it consumes in performing them. This has led many theorists to think that social states in general can be evaluated in the same way. However, unlike a bureaucracy, it is unclear what purpose *society* is supposed to serve, or what function it is supposed to discharge. There have been a variety of different answers to this question, but the most common has been that the purpose of society is to advance the well-being of its members. As Rawls said, "social cooperation makes possible a better life for all." Naturally, people have different ideas about what the good life consists in. However, one characteristic everyone shares is that achieving their particular vision of the good life makes them *happy*. It can therefore be said that society, in helping each individual to achieve his or her own vision of the good life, promotes the happiness of its members. According to this view, maximizing efficiency therefore means *maximizing happiness*.

This line of thinking has been enormously influential in moral philosophy, culminating perhaps in Jeremy Bentham's 19<sup>th</sup> century "hedonic calculus." The picture that emerges with Bentham is of society as a giant mechanism that produces a certain type of generic good – pleasure – and minimizes a certain bad – pain. This good then provides an "external" standard against which different social states can be evaluated. This means that the efficiency of different social states can be judged using a very generic sort of cost-benefit analysis, simply by determining which one produces the greatest balance of pleasure over pain. Thus John Stuart Mill defined the central "utilitarian" principle as the maximization of happiness, where happiness is defined as the net amount of pleasure:

The creed which accepts as the foundation of morals "utility" or the "greatest happiness principle" holds that actions are right in proportion as they tend to promote happiness; wrong as they tend to promote unhappiness. By happiness is intended pleasure and the absence of pain; by unhappiness, pain and the privation of pleasure (Mill 1863, 10).

In order to see how this utilitarian standard works, consider Figure 1.1. If  $s$  is the status quo, then the total amount of happiness at that point will be some amount  $k$ , obtained by adding up the happiness of player 1 and the happiness of player 2. We can draw a line through this point such that every point on this line is a social state with the same total utility (this line is given by

the function  $K = x + y$ ). The utilitarian standard then suggests that any point to the north-east of *any* point on this line is preferable to the status quo. Any point on the line will be equivalent to the status quo, from a utilitarian standpoint, and any point to the south-west of any point on the line will be worse. Thus utilitarianism claims that the entire shaded region of Figure 1.1 is preferable to  $s$ , because each social state in this region is one with greater total happiness than at  $s$ .



**Figure 1.1 Utilitarianism**

This suggestion strikes most people as plausible the first time they hear it. After all, if people are happier in one state than in another, doesn't that make it better? However, closer examination of the utilitarian standard reveals that not *everyone* needs to be happier in order for a social state to be preferred. Because utilitarianism aggregates pleasure, i.e. it seeks to maximize the *sum* of happiness, it is possible for one person's pleasure to "outweigh" the pain experienced by another. That is, the greatest happiness principle does not seek to maximize the balance of pleasure over pain for each individual, but rather the balance of pleasure over pain for society as a whole. This means one person can wind up getting all the pleasure, and someone else all the pain. Specifically, any utilitarian improvement that lies to the north-west or to the south-east of

the status quo represents a net gain in happiness for "society," but this net gain comes about through a loss of happiness for one player, compensated for by a greater gain in happiness for the other.

Critics of utilitarianism were quick to point out some of the outrageous consequences that follow from this doctrine. For example, we often use very straightforward utilitarian calculations when dealing with animals. Most people consider it permissible to kill lame animals when doing so will minimize their suffering. We may kill a herd of cattle in order to prevent an infectious disease from being spread. Most people also consider it permissible to eat animals, as long as they are slaughtered humanely (i.e. in a way that minimizes pain). Naturally, it would be outrageous to treat our fellow human beings in this way. Along the same lines, it would appear to be permissible, from a utilitarian standpoint, to kill or torture one individual whenever doing so would produce enough happiness in others to outweigh the losses suffered (or more realistically, it seems permissible to discriminate against a minority, if doing so makes enough people in the majority happy).

Utilitarians have tried to avoid these consequences in various ways, but none of these responses have been able to carry general conviction. Often they have claimed that, empirically, situations in which one can truly benefit the many by sacrificing the few are unlikely to arise. This is inadequate, for a variety of reasons. Apart from the fact that such situations do arise routinely in the case of animals, and are handled largely in the way that the utilitarian calculus recommends, there is the more fundamental problem that this response does not address the *principle* of the matter. The happiness that discrimination against a minority group generates for members of the majority simply should not *count*, from the moral point of view. Thus the question of how many people benefit from the discrimination simply should not arise. The fact that one is even doing the math on these numbers suggests that a prior moral error has been committed.

Without getting further into the details of these arguments, we might just consider why these seemingly unattractive consequences turn up in the first place. The analysis developed here suggests that the problem stems from the utilitarian attempt to specify an efficiency standard for society using an "external" criterion. The argument gets off on the wrong foot the moment it takes the production of happiness as the "goal" of society. In the case of a machine, there is

someone "outside" of it whose purposes determine what counts as a cost and what counts as a benefit, and therefore whether or not it is functioning efficiently. Similarly, in the case of an individual, there can be others standing outside that individual who can impose "external" standards to determine whether he or she is functioning efficiently. In the case of society, however, there is no one outside of society who can set such a standard.

To see the problem that can arise from the absence of an external standpoint, consider the following example. In order to determine whether a social state is maximizing, there must first be some specification of what the social state is supposed to be producing. But who is to say that happiness is that goal? Many early modern Christian philosophers believed that the earth had been created by God in order to promote the happiness of "his creatures." But there have been countless other hypotheses about what God's purpose might have been, and it is unclear how we could even begin to select among these accounts. In *Hitchhiker's Guide to the Galaxy*, Douglas Adams suggests that the earth is actually a giant computer, constructed and operated by mice, designed to discover the meaning of life (Adams 1979). This hypothesis leads to a much different conception of the "purpose" of society. It suggests that social states are more or less efficient if they are assisting in the expeditious performance of this calculation, inefficient if they are not. There are two things about this suggestion that are significant. First, its ridiculousness shows that all hypotheses of this type are in principle unverifiable, and therefore of no practical use. Second, it shows that if the "purpose" is genuinely external, then there is no reason to think that it should be related to any of the purposes that members of society actually have – just as the worker's internal criteria of efficiency need not have any connection to the employer's. This reveals the underlying problem with the attempt to judge the efficiency of a social state using external criteria, and suggests that social states, unlike individuals and machines, can *only* be judged by internal criteria.

But once this is acknowledged, it changes our perspective in fundamental ways. While an external observer could have a single goal for all of society, it is not clear that society has a single internal purpose. If the efficiency of a social state is to be judged according to "internal" criteria, i.e. the goals and purposes of the agents who are affected by that social state, there is no reason to think that there will be a single standard that can be used to measure costs and benefits. It is possible to decide whether a particular agent's choices are efficient by considering the

balance that it strikes between sacrifice and achievement for that individual. But two different agents might disagree about what counts as an achievement, so that a particular choice of social state may be efficient for one of them and inefficient for the other. This suggests that if social states can only be judged by internal criteria, then it may not be possible to determine whether they are more or less efficient.

Mill saw this problem quite clearly, and tried to save utilitarianism by showing that "general happiness," which is what the greatest happiness principle seeks to maximize, is actually the goal that all persons seek, and therefore that Bentham's "external" criterion is actually an "internal" criterion shared by all agents. In doing so, he produced one of the most famously invalid arguments in the history of philosophy:

No reason can be given why the general happiness is desirable, except that each person, so far as he believes it to be attainable, desires his own happiness. This, however, being a fact, we have not only all the proof which the case admits of, but all which it is possible to require, that happiness is a good, that each person's happiness is a good to that person, and the general happiness, therefore, a good to the aggregate of persons (Mill 1872, 45).

This argument commits what is known as a compositional fallacy. From the fact that each person's own happiness is a good to that person, it simply does not follow that the happiness of all will be a good for all persons. (Consider: the fact that I want a promotion does not mean that everyone wants everyone to be promoted.) But despite the failure of this particular argument, Mill's basic idea is correct. Society cannot be judged more or less efficient using "external" criteria, it must be judged by "internal" criteria. However, the only internal criteria are those supplied by agents' own particular goals and projects. (This is not to deny that groups can have goals, it is just to say that one cannot assume that society as a whole shares a single goal.) Thus the efficiency of a social state must be determined by whether or not it is more efficient from the standpoint of the individuals who compose it. Mill simply made a mistake in determining how the social efficiency standard is to be composed from the individual standards. The first person to explicitly formulate a better standard was Vilfred Pareto (1909).

### 1.3 The Pareto Standard

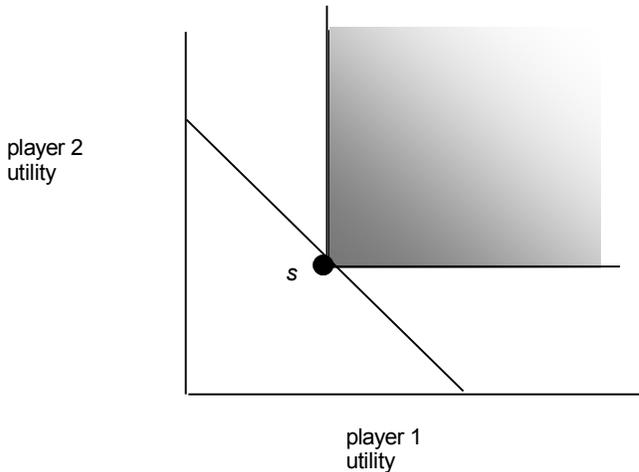
Pareto's idea is very simple: a social state is more efficient if it is more efficient for every individual in it. As we have seen, a state is more efficient for an individual if it provides her with a greater balance of achievement over sacrifice, relative to her own goals. Whereas Mill tried to aggregate all of these goals in order to calculate the "total" value of a social state, Pareto claimed that the value of a social state is determined by its capacity to satisfy each agent according to her own criteria. This idea is cashed out in the form of two definitions, which have become standard in the years following Pareto's initial discussion:

**Pareto-superiority:** Social state  $s_1$  is Pareto-superior to social state  $s_2$  if and only if at least one agent prefers  $s_1$  to  $s_2$ , and none prefer  $s_2$  to  $s_1$ .

**Pareto-optimality:** Social state  $s_1$  is Pareto-optimal if and only if there is no other social state available that is Pareto-superior to it.

The Pareto standard therefore claims that one social state is more efficient than another if and only if it is Pareto-superior to it, and that it is the most efficient if and only if it is Pareto-optimal. Intuitively, this means that society can be made more efficient if it is possible to improve the condition of at least one person without harming anyone else. The idea that we should maximize efficiency then translates into the idea that we should try to achieve Pareto-optimality – a state in which no one's condition can be improved without worsening the condition of someone else.

The Pareto-standard, like the utilitarian standard, has a very simple graphical representation. The shaded region of Figure 1.2 shows the set of points that are Pareto-superior to  $s$ . Clearly, the set of points recommended by the Pareto standard are a subset of the points recommended by the utilitarian standard. What has been eliminated are all of the points at which either player's condition would be worsened by the transition from  $s$ .



**Figure 1.2 The Pareto standard**

One thing to note about the Pareto standard is that it does not provide a complete ranking of social states. The definition of Pareto-superiority gives rise quite naturally to the opposite notion of Pareto-inferiority:

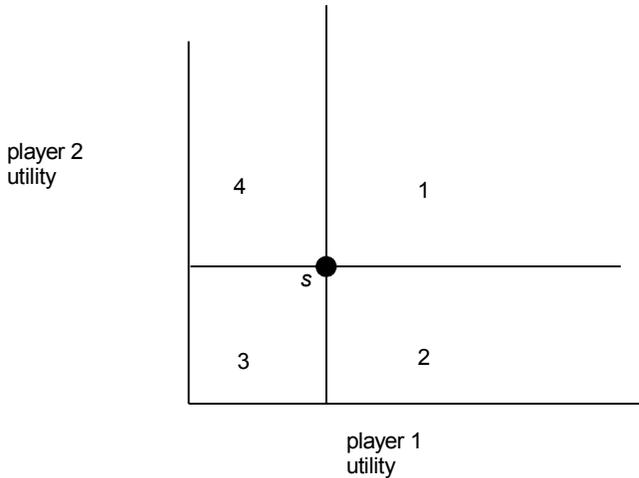
**Pareto-inferiority:** Social state  $s_1$  is Pareto-inferior to social state  $s_2$  if and only if  $s_2$  is Pareto-superior to  $s_1$ .

But these two standards leave a number of social states unranked. Any point to the NE of  $s$  is Pareto-superior to it, and points to the SW is Pareto-inferior to it, but this does not say anything about the points to the NW and the SE. The Pareto-standard simply does not rank such points. Thus some social states will be:

**Pareto-noncomparable:** Social states  $s_1$  and  $s_2$  are Pareto-noncomparable if and only if at least one agent prefers  $s_1$  to  $s_2$ , and at least one agent prefers  $s_2$  to  $s_1$ .

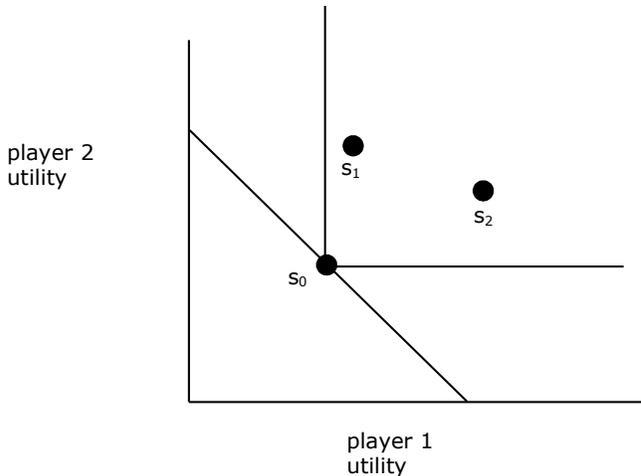
The easiest way to keep these straight is through the "numbered quadrants" representation. Treating the social state as the origin and drawing  $x$  and  $y$  axes through it produces four quadrants, shown in Figure 1.3 below. Every point in quadrant 1, including the line, is Pareto-

superior to  $s$ . Every point in quadrant 3, including the line, is Pareto-inferior to it. All points in quadrants 2 and 4 are Pareto-noncomparable to  $s$ .



**Figure 1.3 Numbered quadrants**

The existence of Pareto-noncomparable points has fairly significant consequences. For instance, if we refer to any step toward a Pareto-superior state as a *Pareto improvement*, then the Pareto standard does not rank all Pareto improvements against one another. For example, in Figure 1.4,  $s_1$  and  $s_2$  are both Pareto-superior to  $s_0$ , but they are Pareto-noncomparable to each other. (Note, however, Pareto-superiority is a transitive relation, so if  $s_1$  is Pareto-superior to  $s_2$  and  $s_2$  is Pareto-superior to  $s_3$ , then  $s_1$  is Pareto-superior to  $s_3$ .) If these two noncomparable options are the only ones available, this means that they are *both* Pareto-optimal, and so there is, from the standpoint of efficiency, no grounds for choosing between them. Thus the Pareto standard, unlike the classical utilitarian standard, does not aspire to be a general theory of justice. It is just a single principle, which must be combined with other principles in order to recommend a determinate social state.



**Figure 1.4 Pareto non-comparable points**

Thus Pareto efficiency may be "a" principle of justice, but it cannot be "the" principle of justice. It states that a social state is more efficient if it is more efficient for each individual person affected by it. For each individual, a more efficient state is one that allows her to realize more of her goals and projects (or more precisely, to achieve a greater balance of achievement over sacrifice). Thus a Pareto improvement makes it possible for at least some individuals to better achieve their goals, while not reducing anyone else's ability to do the same, regardless of what these goals are. It is simply a "no lose" transformation – not everyone may be a winner, but at least no one is a loser. This is a weak constraint, but it nevertheless has some surprisingly powerful normative consequences. Some of these will be examined below.

### 1.4 Why Pareto-efficiency?

What is it about the Pareto standard that recommends it as a way of gauging efficiency, and why should we seek to maximize efficiency, so understood? Compared to utilitarianism, it clearly offers a less controversial set of recommendations (largely because it offers *no* recommendations for the most controversial set of cases, viz. those in which the gains of some must be weighed against the losses of others). In this respect, it reveals the utilitarian principle to be in fact an amalgam of two distinct theses. The first, which may be referred to as *welfarism*, is

the claim that what *matters*, from the standpoint of justice, is human happiness, and that insofar as it is possible to promote happiness we should seek to do so. (The first quote from John Stuart Mill above, read narrowly, is an expression of this commitment to welfarism: “actions are right in proportion as they tend to promote happiness; wrong as they tend to promote unhappiness.”) The second aspect of utilitarianism is its *aggregationism*, viz. the view that, in cases where happiness for some can only be obtained by imposing suffering on others, the way to determine the best course of action is to add and subtract across individuals, then choose the action that promotes the greatest sum of happiness. If one looks back over the arguments of the previous section, it should be obvious that the aggregationism is what produces all the morally counterintuitive implications of utilitarianism, not the welfarism. Furthermore, examining Figure 1.1 again, it should be clear that utilitarianism should be understood, not only as a claim about efficiency, but also as a (morally controversial) view about how questions of fairness should be settled. As we shall see later on, it amounts to the claim that considerations of equality should be assigned absolutely no weight when it comes to determining how “just” a social arrangement is.

From this perspective, the most important conceptual gain that is achieved through the adoption of the Pareto efficiency standard is that it serves to separate out the welfarism from the aggregationism in utilitarianism, allowing us to adopt the former, while (temporarily at least) suspending our commitment to the latter. This is appealing, simply because the welfarism is what most people find attractive about utilitarianism (while the aggregationism is what generates the embarrassing counterexamples that utilitarian philosophers have sought, with varying degrees of desperation, to avoid).

The fact that the Pareto standard suspends the commitment to aggregationism has other benefits. In particular, since it does not require that we add or subtract happiness across individuals, it allows us to avoid the very thorny question of how one is supposed to *compare* happiness across individuals. In more technical terms, the Pareto standard does not require **interpersonal comparisons of utility**. There are a variety of reasons why this is appealing. In an age in which it is difficult to secure consensus over substantive questions of value, a principle of justice that does not require any judgments to be made about the value that different people attach to their projects is able to sidestep a large number of controversies. A social state is judged to be superior, according to the Pareto standard, if people *regard themselves* as better off – or at

least not worse off – under the new state. What they think about the value of each other's achievements is irrelevant. Thus the Pareto standard appears to specify conditions under which individuals may pursue their own projects, subject only to the constraint of reciprocity. Individuals are allowed to do their own thing, as long as it is not at the expense of others.

By way of contrast, it was a central characteristic of classical and medieval political theory that philosophers attempted to directly specify what should count as a good or worthwhile life. Aristotle, for instance, produced a list of the qualities of character that a person should aspire to have. Attaining the good, in his view, consisted of living a life that exhibited these specific virtues. Political communities, according to Aristotle, were established so that individuals could engage in the shared pursuit of this good. This meant that the state was thought to exist strictly to promote the realization of some particular conception of the good life. Social states were to be judged as better or worse depending upon how close they brought their members to this common ideal (a doctrine referred to as *perfectionism*).

The key weakness in this view of politics is that the kind of consensus over questions of the good life that Aristotle took for granted seems to be obtainable in only one of two ways. The first is to specify “the good” at such a high level of abstraction that pretty much anything can qualify. This works well enough theoretically, but winds up being of little value when it comes to addressing concrete political or economic questions. The second option, therefore, is to ignore and suppress dissenting views. Aristotle thought his list of virtues was definitive of the good life because he took it for granted that his own society and culture was superior to every other. This allowed him to ignore the fact that women, slaves, and “barbarians,” if asked, would probably have produced very different lists. However, as the size of states grew larger, it became increasingly difficult to ignore the enormous differences of opinion on such questions that people tend to have. The major world religions all succeeded in securing a certain level of uniformity of belief about the nature of the good among their members, but this was usually at the expense of considerable intolerance toward heretics and infidels (i.e. dissent from people inside and outside the faith community, respectively).

In Europe, this tension came to a head in the wars of religion of the 17<sup>th</sup> century. For a variety of reasons, the Christian church found itself incapable of exercising sufficient force in the repression of heresy. This generated widespread disagreement over various aspects of religious

doctrine, including important questions about the nature of the good life. Most importantly, a major disagreement erupted over whether it was best to withdraw from everyday activities and dedicate oneself to contemplation and prayer (the monastic ideal), or whether it was best to commit oneself to some worldly activity (to find one's "calling"). This dispute had enormous economic and political consequences, because a significant percentage of the land in Europe was owned by monasteries (e.g. at one point, 25 per cent of France). If the purpose of the state was to promote the good life, and the good life consisted of worldly activity, many people ("Protestants") reasoned that the government should seize this church land and integrate it into the regular economy. But many more ("Catholics") thought that the highest priority of the state should be to preserve this land for those who chose to pursue of a life of seclusion and prayer. The associated conflict raged for over a hundred years, creating untold bloodshed, but most importantly, showed no signs of ever being resolved through the exercise of force.

It is in this context that the theory of the **social contract** arose. Thomas Hobbes, who is widely regarded as the founder of this school of thought, argued that any particular conception of the good will be too unstable to serve as the basis for political order (Hobbes 1651, 39). Not only will different people inevitably disagree about the nature of the good life, he argued, but even single individuals are constantly changing their minds about what is best. The purpose of government, therefore, cannot be to promote one particular vision of the good. Hobbes suggested instead that the most important function of government is simply to enforce contracts. This provides security to all individuals, because it enables them to enter into binding agreements with one another. This allows them to pursue their particular vision of the good life without having to worry about other people undermining or blocking their efforts. The authority of the state, Hobbes claimed, comes from the fact that everyone can *agree* to be bound by its decisions. Because everyone has an interest in security, everyone would consent to the creation of an authority whose mandate it was to guarantee it. A government that was committed to a particular conception of the good, on the other hand, could not count upon unanimous consent, so long as there is anyone who did not share that view.

Hobbes is a transitional figure, because he remained partially attached to the older view that the authority of government flows from its ability to help citizens achieve a particular good. By defining "security," or "self-preservation" as this good, Hobbes found a way of guaranteeing

that the good provided by the state was one that everyone wanted. However, the idea that caught on was not that governments provide a very general good, but that the state secures the conditions under which each individual can pursue *his or her own good*. Their authority, in this case, stems from the fact that, insofar as their power is used to secure these conditions, each individual could be expected to consent to the exercise of that power, simply because it helps her to achieve her own vision of the good. This provides a basis for political authority that is neutral with respect to all of the controversial religious and philosophical questions that divide people.

Thus the authority of the state, according to the social contract view, stems from the fact that citizens voluntarily agree to be bound by its rules. They are willing to agree because the state is able to bring about a social state in which everyone is better off, or in which some are better off and none are worse off. But this is just another way of saying that governments can secure consent because they promote *efficiency* in precisely Pareto's sense. Thus the major normative intuition underlying the Pareto standard is the idea that voluntary agreements create Pareto improvements, and that, in turn, Pareto improvements can be expected to secure agreement. When people trade, or make promises, or enter into any other kind of commitment, they do so precisely because they expect to be better off by virtue of this agreement. If all parties to an agreement benefit, and everyone else is unaffected, then the overall change to society is a Pareto improvement.

Thus the normative intuition underlying the Pareto-efficiency standard is essentially *contractual*. Pareto improvements are changes that no one has any reason to reject. Making these improvements therefore means making some people better off, *under conditions that everyone can accept*. Recalling that the purpose of these normative standards is to permit cooperation, efficiency as a value permits social integration while requiring very little in the way of consensus about basic questions of value. This is one of the characteristics of the efficiency standard that makes it a characteristically modern value. Efficiency is not something that is to be considered intrinsically good, or valued for its own sake. Efficiency is valued because it is conducive to each individual's realization of his or her own conception of the good life. Pareto-optimal social arrangements are ones in which individuals are afforded the greatest opportunity for the realization of these goals, subject only to the constraint that no one has been harmed in order to provide this opportunity. Thus it provides terms under which people can cooperate despite deep

disagreement over fundamental questions of value. This explains why efficiency has become such an important moral idea in culturally and religiously plural societies.

The first theorist to make significant (implicit) use of the Pareto standard within the context of social contract theory was John Locke. Locke in fact founded his entire theory of government on the idea that it could be created through a series of “no lose” transactions (i.e. Pareto-improvements). He recognized that the social contract was mythological, in the sense that at no time did people ever get together and sign a contract in which they agree to obey the state. However, he claimed that despite the absence of an explicit contract, an *implicit contract* could be assumed. But this raised the obvious question: How is one to decide, for other people, whether they have implicitly agreed to a particular arrangement? Locke claimed that so long as they benefited from the arrangement, or were at least not harmed by it, they could be assumed to consent to it. In modern terminology, he assumes that everyone would agree to Pareto improvements, and so consent need not be explicitly procured in order to make changes. This idea plays a crucial role in Locke's theory of property. In the following passage, he claims that through their labour, people may transform resources that belong to the community into their own private property, without explicit consent. The reason they are able to do so is that initial appropriation constitutes a Pareto improvement:

God, when he gave the world in common to all mankind, commanded man also to labour, and the penury of his condition required it of him. God and his reason commanded him to subdue the earth, i.e. improve it for the benefit of life, and therein lay out something upon it that was his own, his labour. He that, in obedience to this command of God, subdued, tilled and sowed any part of it, thereby annexed to it something that was his property, which another had no title to, nor could without injury take from him. Nor was this appropriation of any parcel of land, by improving it, any prejudice to any other man, since there was still enough, and as good left; and more than the yet unprovided could use. So that, in effect, there was never the less left for others because of his enclosure for himself. For he that leaves as much as another can make use of, does as good as take nothing at all. No body could think himself injured by the drinking of another man, though he took a

good draught, who had a whole river of the same water left him to quench his thirst (Locke 1689, §§32-33).

Locke tacitly appeals to the Pareto standard twice in this passage. Property rights are conferred through labour, in Locke's view. Labour is a hardship, and so people engage in it only in order to transform nature into something that is of use to them. People are willing to invest their labour in objects because the benefit that they get from using these objects outweighs the costs associated with rendering them usable. If, however, someone other than the person who invested the labour intervenes in such a way that she receives the benefit instead, then the person who did invest the labour is harmed by this action. Because such an action does not satisfy the Pareto standard, consent cannot be assumed. Thus the object that the person invests labour in is said to be *appropriated* or “enclosed” – it can no longer be used without the *explicit* consent of the individual who gave it its value. On the other hand, Locke claims that this kind of enclosure does not harm anyone else, and so explicit consent is not required. Because people create the value that is invested in the object, as long as they leave enough raw material for others, their appropriation does not take anything away from the rest (“no body could think himself injured”). As a result, initial appropriation of unowned resources satisfies the Pareto standard.

To summarize Locke's argument: when someone invests labour in an object, then consumes it, they do no harm to others (the transformation is a Pareto improvement). If someone else consumes it, it harms the person who invested labour in it (the transformation is Pareto-noncomparable). Those implicit consent cannot be assumed, explicit consent must be obtained. An object that cannot be used by another, without explicit consent, is said to be owned. If individuals only have a right to effect Pareto improvements in the social state, then individuals have a right to acquire property through labour, and no one else has a right to use this property once it has been appropriated. Thus the institution of private property, and the emergence of the rudiments of a market economy, is a direct consequence of the use of the Pareto standard to regulate economic activity.

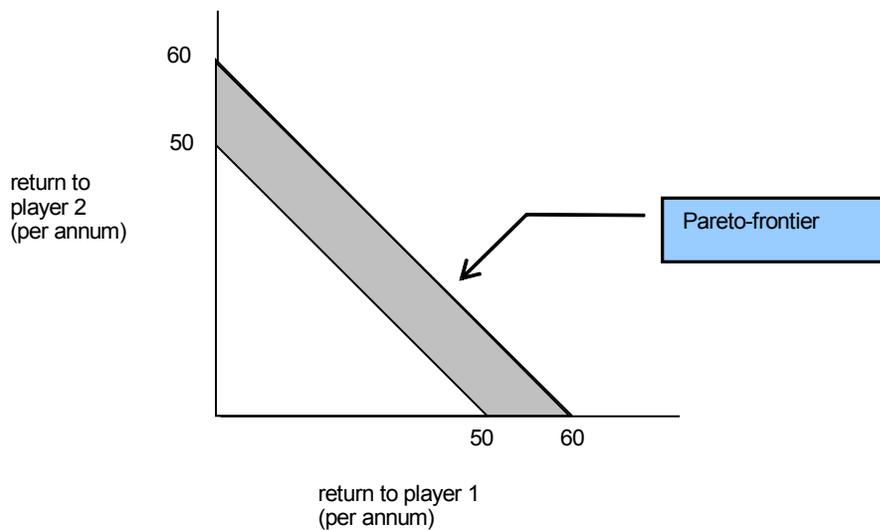
This example shows how the Pareto principle, which seems like a rather weak constraint on the selection of social states, can be used to generate surprisingly robust conclusions. Political

philosophers today are still grappling with arguments that have the same basic form as the one used by Locke over 300 years ago (the view is now known as libertarianism [see Nozick, 1973]).

### 1.5 Some objections

In the previous section, it was argued that the Pareto standard of efficiency is attractive because it provides a way of ranking social states that everyone could agree to, no matter what their particular goals and projects. It must be noted, however, that this expectation is no more than a presumption. It would be more accurate to say that people should be willing to accept Pareto improvements *unless they have some specific reason not to*. If it is clear that a particular social change would represent a Pareto improvement, then the burden of proof falls upon anyone who opposes such a change to explain what is wrong with it. There are, of course, many things that can be wrong with any particular Pareto improvement. Some of the issues that often arise include: inequality, interdependent preferences, bad preferences, paternalism and path-dependency

- *Inequality*: The most common objection to the Pareto-standard is that it ignores questions of distribution. As long as no one is losing, then any benefit conferred on any individual is legitimate, no matter how disproportionately these benefits are distributed. Issues of fairness will be discussed extensively in the second half of this book. For the moment, it is important simply to note how little the efficiency criterion says about distribution. Consider the following situation: two people have bank accounts which earn no interest on balances less than \$1000, and each of them has only \$600 in their account. If they put their money into a single account, they could earn interest at, say, five per cent. Thus depending upon how much money they put into the account they can earn between \$50 and \$60 in interest per year. The cooperative arrangements they can adopt are shown as the shaded region in figure 1.5.



**Figure 1.5 Cooperative arrangements**

Any cooperative arrangement in which they deposit between \$1000 and \$1200 in the joint account is superior to the status quo, at which they earn nothing. The Pareto standard simply recommends that they put the full \$1200 into the account. (If they put less than that amount in, then it is possible to improve at least one person's return without harming anyone, simply by increasing the amount of the deposit.) However, once the interest has been earned, the parties are free to divide it up any way they choose, and all of these distributions will be Pareto-optimal. Most importantly, the situation in which one of them gets the entire \$60, while the other gets nothing, is Pareto-optimal, because the player who gets nothing is not doing any *worse* under the cooperative arrangement than she was under the status quo, and the other person is doing better. The set of Pareto-optima is referred to as the **Pareto-frontier**, shown in Figure 1.5 as the outer boundary of the shaded region.

This is not so much a problem as it is a way of driving home the point that the efficiency standard needs to be supplemented with some conception of fairness in order to be taken seriously as a principle that can regulate cooperation. But that having been said, it should be noted that the insensitivity of the efficiency standard to distributive questions has often appealed to people precisely because it can be used as a way of excusing or justifying inequality. Part of the reason that Locke used the Pareto standard as the basis for his theory of property is that he

needed to explain how, despite the fact that God "gave the world in common to all mankind," a tiny landed aristocracy could have wound up owning all of England. Locke's view is that the condition of perfect equality was transformed into one of extreme inequality through a series of Pareto improvements. As people appropriated property, the amount of wealth increased, and while this increase accrued disproportionately to those were doing the appropriating, at no time did anyone suffer a decrease in wealth through their activities. As a result, he claims, "it is plain, that men have agreed to a disproportionate and unequal possession of the earth" (Locke 1689, §50).

- *Envy*: In order for the Pareto standard to have any applicability, it is important that agents not take an interest in each other's interests. Most obviously, if agents envy one another it may make it very difficult to make Pareto improvements. If someone else's success makes me unhappy, then it will be impossible to improve that person's condition without worsening mine.

There are ways in which these more obvious cases of envy can be handled without serious difficulty. One could, for instance, aspire to ignore envy in calculating the level of happiness associated with a particular social state. The more serious problem, however, is that people tend to evaluate their level of well-being in relative, rather than absolute terms. If you ask someone whether they feel rich or poor, they will answer by comparing themselves to their contemporaries and neighbours, not their ancestors or people in distant parts of the world. This is because our aspirations are formed, at least in part, by our expectations. Seeing a certain level of wealth around us makes it reasonable to aspire to a comparable, perhaps somewhat greater, level of wealth. This makes it easy to forget that, compared to the standard of living even a hundred years ago, our society enjoys wealth beyond the dreams of avarice. Sometimes we need to remind people of this fact, by telling them they have "problems for happy people."

The problem with these adaptive or relative preferences is that they are extremely common, and not stigmatized in the way that envy is. All the same, preferences of this type are capable of making Pareto improvements impossible. For instance, a student might quite reasonably be satisfied with her grades as long as they are above the class average. Unfortunately, in principle not every student can be above average. This means that if everyone wants to be above average, a large percentage of the class is bound to be disappointed. This makes it

impossible to effect any Pareto improvements – if one student's grades rise, bringing her above the average, it will also raise the average, usually dropping someone else below it. (One of the objections that people have to grading "on a curve" is that it imposes this structure on students, effectively guaranteeing that for every winner there will also be a loser.)

It is not hard to see how prevalent this kind of preference is. Most people will be happy with their salary if they are earning slightly more than the average of someone in their position, most people are happy with their car as long as it is nicer than the average vehicle on the road, and so on. Because no one can get a raise or buy a nicer car without affecting the relative position of everyone else, what looks like a Pareto improvement in absolute terms may not turn out to be one from the standpoint of any member of the society.

• *Adapted Preferences:* One of the greatest strengths of the Pareto standard is also one of its greatest weaknesses. It suggests that a social state is superior if each individual judges it to be superior, in terms of his capacity to achieve his own goals. It does not require that either individuals or "society" compare the relative merit of these goals, and so it can provide a basis for cooperation in contexts where there is serious disagreement over basic questions of value. However, because it does not evaluate the substantive merit of individuals' preferences, the efficiency standard can promote the realization of goals that, from some other perspective, we might regard as somewhat dubious. For instance, people may want to eat foods that are bad for them, even if they are aware of the adverse health effects. If the amount of junk food that they get can be increased without decreasing anyone else's share, then this constitutes an efficiency gain. Similarly, improvements in the distribution network for illegal drugs represent just as much of an efficiency gain as improvements in any other sector of the economy.

Our intuitions often pull in both directions on this issue. On the one hand, it is important that individuals not presume to judge each other's projects. It is far too easy to ignore people's stated preferences by treating them as brainwashed victims of some ideology, whether it be consumerism, religious fanaticism, or patriarchy. When taken too far, this kind of thinking can result in a social policy in which people are simply told what they should want, instead of being given what they do want. It also raises tricky questions about who is entitled to decide what people should want, if not the people themselves. Because of this, many social theorists are

tempted to "take preferences as given." On the other hand, it is very important to recognize that people do not, in general, form their preferences under conditions of full autonomy. People are subject to all kinds of social pressure to want one thing or another, and some of these pressures serve to reinforce unjust social arrangements.

This is a special case of a more general phenomenon, which philosophers refer to as **adaptive preference formation**. The idea is simply that the goals people set for themselves reflect their conception of what is obtainable. Some people are more realistic than others, but everyone is in the same general ballpark. For instance, the kind of features that people look for in an automobile is rather sharply constrained by the current state of automotive technology. It is therefore an adapted preference. This phenomenon has enormous political significance, because what people consider to be a reasonable expectation will be strongly influenced by their gender, class, or cultural background. People who come from disadvantaged social groups often have downwardly adapted preferences. This means that because they were raised in such a way as to expect less, they have set their sights much lower than other people. As a result, they may be happy to live under conditions that most other members of the society would find unacceptable. Thus they may not think that certain kinds of social arrangements harm them, even if most other people would.

Debates over the legitimacy of a specific pattern of preference are at the core of a number of different political debates, and cannot be settled in any formulaic way. The important point here is simply that a particular proposal to increase the efficiency of a social state can always be challenged by someone who is willing to claim that the preferences in terms of which the efficiency gain is calculated should not be held by the individuals who hold them. There are many situations in which this kind of objection can be morally compelling. But since the point of the efficiency standard is precisely to avoid making these kind of evaluations, it often seems as if we are forced to go one way or the other: either we let people make up their own minds about what is best for them, and use the efficiency standard; or we engage in substantive evaluation of their projects, in which case we may as well just evaluate the social state directly, and replace the efficiency standard with a perfectionist one. But this is too stark an opposition. A middle road would be to say that people's expressed preferences should be respected, unless there is some specific reason not to (and this reason cannot include the mere fact that their preferences are

adapted, since all preferences are adapted to some degree). This means that efficiency gains will not necessarily be determined by expressed preferences, but the burden of proof would lie on the side of anyone who wants to give people something other than what they say they want.

• *Paternalism*: As we have seen, the Pareto standard suggests conditions under which agents should be able to agree to a particular transformation of the social state. However, it obviously does not require that they must *actually* agree to any change. In fact, for theorists like Locke, this was one major advantage of the Pareto standard – it provided a rule for ascribing implicit consent to individuals. This allowed the state to proceed without securing explicit consent from those affected by its actions. Consider the following:

I say, that every man, that hath any possessions, or enjoyment of any part of the dominions of any government, doth thereby give his tacit consent, and is as far forth obliged to obedience to the laws of that government, during such enjoyment, as any one under it; whether this his possession be of land, to him and his heirs for ever, or a lodging only for a week; or whether it be barely travelling freely on the highway: and, in effect, it reaches as far as the very being of any one within the territories of that government. (Locke 1689, §119)

There is something to be said for this view, and certainly Locke's analysis here tallies with the general practice of nation-states in the world today. But bypassing consent can be a source of enormous difficulties. People may reasonably reject Pareto improvements if they are implemented in a *paternalistic* manner. To take a common example, even if you buy someone new boots, it is not a good idea to throw out their old pair without permission. Despite the fact that the net result is an improvement, people often resent having such improvements forced upon them. Similarly, if you throw a rock through someone's window, but attach a hundred-dollar bill to it in order to pay for the damage and inconvenience, that doesn't make it okay.

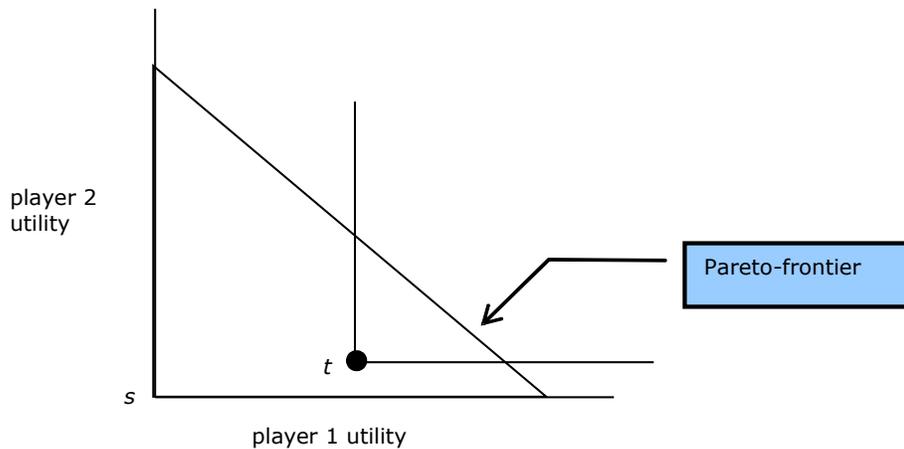
Historically, this issue is not an insignificant one. Locke's theory of property was extremely important in justifying the appropriation of Native land in British North America. While Spain and France simply asserted a "right" of conquest, the British claimed to be

establishing themselves on the American continent through a system of voluntary agreements with the indigenous peoples. Europeans could appropriate the land, they argued, without explicit consent, because they were able to use it more intensively than the Native population. As a result, their appropriation was said to generate Pareto improvements. According to Locke:

He who appropriates land to himself by his labour, does not lessen, but increases the common stock of mankind. For the provisions serving to the support of human life, produced by one acre of enclosed and cultivated land, are (to speak much within compass) ten times more than those which are yielded by an acre of land of an equal richness lying waste in common. And therefore he that enclose land, and has a greater plenty of the conveniencies of life from ten acres, than he could have from an hundred left to nature, may truly be said to give ninety acres to mankind. For his labour now supplies him with provisions out of ten acres, which were by the product of an hundred lying in common. I have here rated the improved land very low, in making its product but as ten to one, when it is much nearer an hundred to one. For I ask, whether in the wild woods and uncultivated waste of America, left to nature, without any improvement, tillage, or husbandry, a thousand acres yield the needy and wretched inhabitants as many conveniencies of life, as ten acres equally fertile land do in Devonshire, where they are well cultivated?(Locke 1689, §37)

Whatever one makes of this argument (and there is a significant element of apologetic in it), the idea that British economic activities were generating benefits for the indigenous population is what gave them the idea that they could colonize North America through contractual relations. This is what motivated the British to sign the Indian treaties that, to this day, form the legal basis of the Canadian constitutional order. The principles underlying the land appropriation, however, have not withstood the test of time, precisely because explicit consent was lacking. In this case, the idea that one can ignore other people's voice, simply because they may ultimately benefit from one's actions, proved to be a dangerous illusion.

• *Path-dependency*: One of the other significant characteristics of the Pareto-standard that may give rise to controversy is the fact that Pareto-optimal outcomes may be *path-dependent*. If the optimum is to be achieved by a series of Pareto improvements, each improvement that is taken will limit the set of optima that can be achieved through subsequent improvements. For example, consider the cooperative surplus illustrated in Figure 1.6. From the status quo  $s$ , any point on the Pareto frontier can be achieved through a series of Pareto improvements. However, if the players are given the opportunity to advance to  $t$ , player 2 might have some reservations. Despite the fact that  $t$  is a clear improvement over  $s$ , the set of Pareto-optima that are obtainable from  $t$  are considerably less favorable to player 2 than they are to player 1. In fact, by accepting  $t$ , player 2 guarantees that she will not get more than half of the cooperative surplus. She might therefore choose to reject  $t$ , if there is some possibility that another option more favorable to her will come along.



**Figure 1.6 Path dependency**

More generally, inequality generated by Pareto improvements may have a tendency to amplify over time. People may use the resources they secure under one arrangement to guarantee an even bigger "cut" for themselves in the future. They may even use these resources to benefit

themselves at the expense of others. This means that people can quite reasonably reject a Pareto improvement if they don't like the direction it is leading.

## 1.6 Implementing Efficiency

Despite these complications, the Pareto-efficiency standard still enjoys considerable intuitive plausibility. In a pure "win-win" situation, it just seems obvious that we should prefer a situation in which some people are better off to one in which they are worse off. At very least those whose condition is improved will prefer it, and we should respect their wishes. In fact, *rejecting* Pareto improvements in clear-cut cases like this seems to inflict a gratuitous harm upon those whose condition could be improved, and is therefore difficult to contemplate without legitimating some type of malice.

With this in mind, it seems that the Pareto standard should be very easy to implement. The principle merely stipulates that a social state is more efficient if it is more efficient for each member of the society. If it is assumed that individuals try to act in the most efficient manner possible, relative to their own goals and projects, it is tempting to assume that social interaction would have a tendency to generate Pareto-optimal outcomes automatically. If everyone acts efficiently, shouldn't the social state that results be Pareto-efficient? It seems to be in every agent's interest to secure a more, rather than less, efficient social state. Unfortunately, nothing could be further from the truth. In a surprisingly wide range of cases, if every individual involved in a social interaction selects the action that is most efficient for her, it will *not* produce a Pareto-optimal outcome. Considering the fact that Pareto-optima benefit everyone, this is one of the most perverse aspects of the human condition. If every individual lived on a deserted island, then having each individual act efficiently would guarantee a Pareto-optimal "social" state. Unfortunately, as soon as two of them get together, all bets are off. In order to understand why this is so, it is necessary to examine somewhat more carefully the idea of individual efficiency, and the way that individuals acting efficiently interact with one another.

**Key Words**

adaptive preference formation

cost-benefit analysis

interpersonal comparisons of utility

maximizing conception of practical rationality

Pareto efficiency

Pareto-inferiority

Pareto-noncomparable

Pareto-optimality

Pareto-superiority

Pareto frontier

social contract

utilitarianism