Utilitarianism as a Criterion for State Action

Nicolaus Tideman*
Department of Economics, Virginia Polytechnic Institute and State University
Blacksburg, VA 24061
ntideman@vt.edu
Phone: 540-231-7592; Fax: 540-231-9288

Florenz Plassmann
Department of Economics, State University of New York at Binghamton
Binghamton, NY 13902-6000
fplass@binghamton.edu
Phone: 607-777-4934; Fax: 607-777-2681

This version: March 27, 2017

Abstract
As a criterion for state action, utilitarianism faces the difficulty that it permits the reasonable expectations of some to be overridden for the benefit of others. Neither Bentham’s felicific calculus nor the Kaldor-Hicks potential compensation test can justify the coercion needed for state action. Social welfare functions, including those that incorporate Rawls’ difference principle, face the same difficulty as utilitarianism. One way of resolving this problem is to move to a constitutional framework of constrained utilitarianism, under which only policies that violate no one’s reasonable claims are considered. To ensure that no one will have a reasonable basis for objecting, there must be a Tieboutian opportunity for dissenters to form their own polities.

Key words: constrained utilitarianism, Tiebout, justified coercion, felicific calculus, potential compensation test, social welfare function, difference principle, constitutional framework
Introduction

“The greatest happiness of the greatest number” was the catch-phrase of the philosophical radicals surrounding Jeremy Bentham and James Mill, making Utilitarianism popular and helping to turn the philosophical idea of hedonism into a framework for state action. Bentham’s felicific calculus provides a straightforward blueprint for how the state would put Utilitarian principles into practice. Yet since government officials have no objective yardstick to measure the amounts of pleasure and pain that citizens will experience under different proposed policies, it is not clear, philosophically, how those officials acquire a mandate to coerce citizens to obey legislative rules.

More generally, there is the question of whether a moral theory of Utilitarianism can and should take the role of a theory that is used to justify state action. John Stuart Mill thought that this was the case, but various thinkers challenged that notion in the 20th century. We examine this debate and assess the potential for Utilitarianism to serve as a criterion for state action.

Utilitarianism as a moral imperative

In Judeo-Christian culture, Utilitarianism appeared first as a moral imperative rather than as a criterion for State action. In Leviticus 18:19 one finds, “Love your neighbor as yourself,” which is equivalent to a Utilitarian instruction. If you love your neighbor as yourself, then your neighbor’s well-being counts as much as your own in your decision about what to do, which is what Utilitarianism requires, provided that everyone who is affected by your action counts as your neighbor.

The Hebrew word that is traditionally translated as “neighbor” in this biblical verse, רֵעַ (Irakh), does not actually mean a person who lives close-by. Rather, its meaning spans “friend” to “fellow member of one’s reference group,” so that “comrade” and “compatriot” are possible translations as well. Thus, “Love your fellow citizen as yourself” might be closer to the biblical meaning of the verse.

The question of who counts as one’s neighbor was a subject of debate among Jewish scholars of the first century CE. This is revealed by the biblical account of the dialogue leading up to Jesus’ telling of The Parable of The Good Samaritan (Luke 10:25-29), in which a lawyer asks Jesus whom he ought to count as his neighbor. Jesus’ answer makes it clear that the Christian interpretation of one’s neighbor is quite expansive (Luke 10:30-37):
In reply Jesus said: “A man was going down from Jerusalem to Jericho, when he was attacked by robbers. They stripped him of his clothes, beat him and went away, leaving him half dead. A priest happened to be going down the same road, and when he saw the man, he passed by on the other side. So too, a Levite, when he came to the place and saw him, passed by on the other side. But a Samaritan, as he traveled, came where the man was; and when he saw him, he took pity on him. He went to him and bandaged his wounds, pouring on oil and wine. Then he put the man on his own donkey, brought him to an inn and took care of him. The next day he took out two denarii and gave them to the innkeeper. ‘Look after him,’ he said, ‘and when I return, I will reimburse you for any extra expense you may have.’

“Which of these three do you think was a neighbor to the man who fell into the hands of robbers?”

The expert in the law replied, “The one who had mercy on him.”

Jesus told him, “Go and do likewise.”

Thus Jesus indicates that the man’s neighbors include not only his Jewish brothers—the priest and the Levite—but also the semi-Jewish Samaritan. Jews and Samaritans had been in conflict for centuries, so the Christian answer to who counts as one’s neighbor includes anyone with whom one comes into contact, even those whom one might be inclined to dislike.

It can be debated whether the prescription to love one’s neighbor as one’s self is meant to include all people on earth. John Calvin (1555, p. 48) interpreted the parable of the Good Samaritan to mean that “the chief design is to show that the neighborhood, which lays us under obligation to mutual offices of kindness, is not confined to friends or relatives, but extends to the whole human race.” Similarly, in a nation whose founding document includes the words, “All men are created equal,” it would not be unreasonable to entertain the idea that embracing the rule of “love your neighbor as yourself” requires counting the consequences for all other people on earth as just as important as any consequences for oneself. This would be truly Utilitarian.

Utilitarianism as a moral imperative versus a standard for legislation

When it comes to putting moral imperatives into action, the question arises as to whether one ought to draw a distinction between the moral duty of each person—the subject of the original
injunction to love your neighbor as yourself—and the standard for laws that are to be implemented. Cicero (De officiis, Book I, Sections 7-14) drew a closely related distinction, between justice and beneficence, arguing that justice was strictly owed to all, while beneficence was owed to those with whom one had some close connection. However, while Cicero considered justice and beneficence as different moral duties to which a person ought to attend, we are concerned here with justice as a standard for laws that will be coercively enforced, as well as with how Utilitarianism as a criterion for this concept of justice is related to Utilitarianism as a standard for moral individual conduct. It is worth addressing these two questions together because some authors treat them as not needing to be distinguished. Jeremy Bentham (1776, preface to the fragment), for example, asserted that the principle that, “it is the greatest happiness of the greatest number that is the measure of right and wrong” was the appropriate basis for legislation as well as morals (Burns, 2005). Bentham’s writings became the foundation of classical Utilitarianism. The critical difference between morals and legislation is that, as far as morals are concerned, people who fail to abide by the proposed criterion will only be scolded. When it comes to legislation, however, those who fail to abide by the law will be subject to fines and/or imprisonment. Thus when one endorses the Utilitarian criterion as the standard for morals, one is setting forth a goal for human conduct. While all progress toward the goal is good, a lack of complete success is not an occasion for punishment. In contrast, when one endorses the Utilitarian criterion as the standard for legislation, one is setting the basis for punishment, confiscation, and coercive taxation.

This distinction between morality and justice is closely related to the distinction drawn by John Stuart Mill. Mill considered justice to be part of the general concept of morality, and he viewed sanctions as the defining characteristic that distinguishes immoral behavior from inefficient behavior. Mill (1863, p. 47) wrote:

[T]he idea of penal sanction, which is the essence of law, enters not only into the conception of injustice, but into that of any kind of wrong. We do not call anything wrong, unless we mean to imply that a person ought to be punished in some way or other for doing it; if not by law, by the opinion of his fellow-creatures; if not by opinion, by the reproaches of his own conscience. This seems the real turning point of the distinction between morality and simple expediency.
Thus people have an obligation to act morally but no obligation to act efficiently. What distinguishes obligations in justice from other moral obligations like generosity and beneficence is that unjust acts violate somebody’s rights. Morality may require us to be generous, but no person has a right to be treated with generosity and beneficence. In contrast, depriving a person of his possessions constitutes a wrong toward a specific person who is thus deprived of a right. Because having a right implies an obligation of society to defend this right, societies enforce respect for rights coercively.

Justifying coercion

Many philosophers have attempted to justify coercion by appeals to natural law. For example, Aristotle (Rhetoric 1373b2) proposed that

Universal law is the law of Nature. For there really is, as everyone to some extent divines, a natural justice and injustice that is binding on all men, even on those who have no association or covenant with each other.

The key question is how one might discover what the natural law is and what it requires people to do. Thomas Hobbes proposed that natural law could be discovered by considering people’s natural rights, leading him to identify 19 laws of nature in Leviathan. But Jeremy Bentham challenged the notions of natural law and natural rights by pointing out that political philosophers have yet to identify the underlying reasons for these laws and rights. In Theory of Legislation, Bentham (1802, Vol. 1, Ch. 13) wrote,

If “Nature” really made such and such a law, those who lay it down with so much confidence, and have modestly undertaken to interpret it, must believe that “Nature” had some reasons for making it. Would it not be safer, shorter, and more convincing, to give us those reasons at once, rather than to offer us the ipse dixit of this unknown legislator as being of itself sufficient authority?

and in Anarchical Fallacies, Bentham (1816, Art. 2) pointed out that

In proportion to the want of happiness resulting from the want of rights, a reason exists for wishing that there were such things as rights. But reasons for wishing there were such things
as rights, are not rights;—a reason for wishing that a certain right were established, is not that right—want is not supply—hunger is not bread.

That which has no existence cannot be destroyed—that which cannot be destroyed cannot require anything to preserve it from destruction. Natural rights is simple nonsense: natural and imprescriptible rights, rhetorical nonsense,—nonsense upon stilts. But this rhetorical nonsense ends in the old strain of mischievous nonsense: for immediately a list of these pretended natural rights is given, and those are so expressed as to present to view legal rights. And of these rights, whatever they are, there is not, it seems, any one of which any government can, upon any occasion whatever, abrogate the smallest particle.

Bentham (1802, Vol. 1, Ch. 11) proposed replacing the notion of natural rights with the fundamental proposition that the “sole object of government should be the greatest happiness of the community,” which the government achieves “by creating rights, which it confers upon individuals…” However, his proposal does not resolve the question of why striving for the greatest happiness of the community would justify coercing those persons who disagree with the rules that the government adopts.

John Stuart Mill (1863, p. 52) suggested that the importance of the issues to be addressed would naturally lead to Utilitarianism, offering an argument along the line given by Thomas Hobbes in *Leviathan*:

If the objector goes on to ask, why [society] ought [to defend my rights]? I can give him no other reason than general utility. If that expression does not seem to convey a sufficient feeling of the strength of the obligation, nor to account for the peculiar energy of the feeling, it is because there goes to the composition of the sentiment, not a rational only, but also an animal element, the thirst for retaliation; and this thirst derives its intensity, as well as its moral justification, from the extraordinarily important and impressive kind of utility which is concerned. The interest involved is that of security, to every one’s feelings the most vital of all interests. All other earthly benefits are needed by one person, not needed by another; and many of them can, if necessary, be cheerfully foregone, or replaced by something else; but security no human being can possibly do without; on it we depend for all our immunity from evil, and for the whole value of all and every good, beyond the passing moment; since
nothing but the gratification of the instant could be of any worth to us, if we could be
deprived of anything the next instant by whoever was momentarily stronger than ourselves.

Mill acknowledged that particular instances of coercion can be justified in other ways, but he
asserted that Utilitarianism was the only principle that would be applicable in all instances.
However, the lack of an objective yardstick that could be used to balance utility gains and losses
that accrue to different persons renders Mill’s assertion unsatisfying.

The difficulty of balancing gains and losses that accrue to different persons

The challenge to Utilitarianism came at the turn of the 20th century, with Vilfredo Pareto’s
concern (Pareto, 1971 [1906]), pp. 451–452):

Consider any position and assume that we move away from it by a very small
amount, consistent with the restrictions [on what is economically feasible]. If in so doing
the welfare of all the individuals of the collectivity is increased, it is obvious that the new
position is more advantageous to each of them; and vice versa, it is less so if the welfare
of all of the individuals is decreased. Moreover, the welfare of some of them can remain
the same, without changing these conclusions. But on the other hand, if this small
movement increases the welfare of certain individuals and decreases that of others, we
can no longer state positively that it is advantageous to the entire collectivity to carry out
that movement.

Since utility cannot be measured objectively, how can we say that a change is a good thing when
it makes some people better off and others worse off?

Nicholas Kaldor (1939) and John Hicks (1939) suggested a potential solution to Pareto’s
concern, arguing that a proposed change should be regarded as a good thing if it passes a
“potential compensation” test. That is, if a proposed change causes gains measured in dollars
that are greater than the losses measured in dollars, then an unquestionable improvement could
be achieved by combining the proposed change with transfers from those who gain from the
change to those who lose, such that everyone would be better off from the combination of the
change and the transfers. If we do not require that such transfers accompany the change, the
argument goes, then it must be because we regard the change without the transfers as better than
the change with the transfers. The change with the transfers must be a good thing because it would make everyone better off, so the change without the transfers, if we decide not to make them, must be even better than this good thing.

If the premises of this line of thinking were true, then the Kaldor-Hicks proposal would be a sound example of Utilitarian thinking. However, we lack the capacity to identify the particular persons who gain and lose as well as the magnitudes of their gains and losses. And even if we could identify those persons, we cannot count on having administrative devices by which we could extract money from those who gain to pay those who lose. Furthermore, it is possible to construct examples in which a change as well as the reversal of the change both pass the potential compensation test, leaving it unclear which situation is better (De Scitovszky, 1941). Thus a Utilitarian endorsement of the potential compensation test is unjustified.

At about the same time as Kaldor and Hicks were proposing the potential compensation test, benefit-cost analysis was being developed. This move was prompted by the efforts of the U.S. Congress to distinguish proposed public works projects that were worthwhile from those that were not. The benefit-cost test is isomorphic to the potential compensation test: a proposed public works project passes a benefit-cost test if its monetary benefits are greater than its monetary costs. The dubiousness of this criterion from a Utilitarian perspective is obvious: the proposition that a project should be adopted if and only if it passes a benefit-cost test is true as a general proposition only if the marginal utility of dollars is the same to all persons. If this is not the case, then benefit-cost analysis does not identify those programs that increase social utility, because the utility increase of a person who gains a dollar might by smaller than the utility decrease of another person who loses the dollar.

Still, the Utilitarian perspective offers a more defensible way of justifying benefit-cost analysis. Suppose that benefit-cost analysis is the only available tool for project evaluation. If there is no reason to assume that the average marginal utility of money for those who gain from such projects is less than the average marginal utility of money for those who lose, then from a Utilitarian perspective there is an expected gain from adopting any project that passes a benefit-cost test (Polinsky, 1972). If there are many projects and the gains and losses of persons are uncorrelated across projects, then the expected proportion of the population that gains increases
as the number of projects increases. This logic collapses if projects that benefit some subsets of the population, such as the rich, are more likely to be proposed.

*Justifying coercion through the maximization of social welfare*

An alternative answer to the challenge posed by Pareto came from the development of the social welfare function, a concept put forth initially by Abram Bergson (1938) and refined later by Paul Samuelson (1938). The conceptual framework of social welfare functions accepts that it is not possible to measure utility reliably. To proceed despite this limitation, the framework makes use of individual indifference curves that are measured only ordinally. The level of social welfare is determined by the combination of indifference curves attained by all members of society.

Bergson and Samuelson had somewhat different ideas about how such combinations of indifference curves would be evaluated. Bergson regarded the social welfare function as a device for communication between politicians and the economists who worked for them. The politician would specify the social welfare function, and the economist would give the politician advice about how to maximize it. From this perspective, the social welfare function is not an indicator of objective merit, but rather a mechanism for implementing a normative framework that is justified by the legitimacy of the power of politicians. However, David Braybrooke and Charles Lindblom (1963) have argued that the political decision making process is generally more complicated, in that the economists’ advice provides information that might induce a politician to refine the social welfare function. Thus it may be impossible to specify a social welfare function in the absence of information from economists regarding the cost of different alternatives, making it impossible for economists to avoid responsibility for the content of the social welfare function.

Samuelson (1948, p. 221) developed the idea of a social welfare function in greater generality. He wrote:

> Without inquiring into its origins, we take as a starting point for our discussion a function of all the economic magnitudes of system which is supposed to characterize some ethical belief—that of a benevolent despot, or a complete egoist, or “all men of good will.” A misanthrope, the state, race, or group mind, God, etc.
Thus while Bergson viewed a social welfare function as a tool for communication between politicians and economists who work for them, Samuelson assumed an ethical framework that corresponds to maximizing a function. The key characteristic that these two ideas about the source of the social welfare function have in common is that they are dictatorial, or at least single-minded. Some philosophers advocate the idea of a decision maker who does not require input from the people—for example, Plato’s *philosopher king*. Economists generally employ the idea of a “benevolent social planner” when working with social welfare functions. Economists generally employ the idea of a “benevolent social planner” when working with social welfare functions. One might imagine that such an assumedly benevolent dictatorship is not essential, and that it is possible to formulate a social welfare function based on some compromise among different points of view. However, the Arrow Theorem reveals that there is no compromising mechanism that preserves the property that if option $A$ is chosen over option $B$ when option $C$ is under consideration, $A$ will still be chosen over $B$ when $C$ is not under consideration (Arrow, 1963, pp. 97-100; Kemp and Ng, 1976). Thus it is impossible to design a non-dictatorial mechanism for the aggregation of individual preferences that would provide the aura of correctness that comes from consistency.

The social welfare function that a benevolent social planner consults reveals the relative weights that the planner attaches to dollars of benefits and costs for different persons. If the social planner decrees that maximizing social welfare is achieved by maximizing aggregate income measured in dollars, then the social welfare function is equivalent to the benefit-cost criterion. If the social planner uses some other criterion, then the social welfare function provides a way of aggregating consequences for different persons that goes beyond simply adding up dollars across people. Although social welfare functions do not measure utility, because utility cannot be measured reliably, changes in measured social welfare reveal the changes in social utility that the capable benevolent social planner attaches to changes in income for different people, providing implied social weights for income changes for different people.

Evaluating potential policy changes by a social welfare function is coherent but has a troublesome logical consequence. Using the weighted sum of income changes as the criterion for a change being an improvement implies that any transfer of money from one person to another is an improvement if the person who receives the money has a greater social weight for income changes than the person from whom the money is transferred. While such transfers increase social welfare as measured by the social welfare function, they undermine incentives to
act efficiently, because the argument for such transfers continues to the point of full equality of the marginal utility of money. If all such transfers were to be made, then people would have an incentive to be productive only to the extent that they valued the opportunity to benefit others. Utilitarians address this incentive effect by considering a transfer worthwhile only if it increases the sum of utilities after taking account of incentive effects. In fact, this is the perspective of optimal taxation theory (Vickrey, 1945; Mirrlees, 1971).

Empirically, one of the factors that seems to affect the marginal utility of money is whether the change in available income is an increase or a decrease (Tversky and Kahneman, 1991). This led Martin Bailey (2001, p. 83) to suggest that, in evaluating proposed policies, losses should be given three times the weight of gains.

An alternative approach to the concern that transfers will end only when marginal utilities of income are equalized is to move from Utilitarianism to constrained Utilitarianism. Under constrained Utilitarianism, some state actions are simply off-limits because of rights, constitutional provisions, and other rules. The agenda of permissible changes is not the set of all conceivable changes but rather the set of all changes that do not conflict with prohibitions on taking private property without compensation, on subjecting different people to different tax schedules, and so on. The Utilitarian criterion is applied only after establishing that such constraints do not preclude a proposed project. Project evaluation therefore becomes a two-step process. The proposed project passes the first step only if it does not violate any restriction on how people may be treated. Upon passing the first step, the proposed project is accepted in the second step if it increases aggregate utility, as measured by a social welfare function. Similarly, if a policy is to be selected from a broad range of possibilities, then all policies that violate restrictions on how people may be treated are excluded in the first step, and the remaining policy with the greatest aggregate utility is selected in the second step.

The Utilitarian approach to aggregating consequences for different persons is to sum changes in individual utility. Social welfare functions admit a wide range of aggregation possibilities: any non-decreasing function of the indifference curves attained by individuals is a potential social welfare function. This flexibility led economists in the 1970s to respond in terms of social welfare functions to ideas offered by John Rawls. Rawls argued that from behind a veil of
ignorance, a person appraising alternative social institutions would appraise them in terms of the prospects that they offered for the least advantaged members of society (Rawls, 1999, p. 65). Economists responded that, in terms of economics, this can be described as a social welfare function that is not the sum of individual utilities but rather the minimum of the individual utilities. Economists went on to point out that there is a continuum of possibilities for which the Utilitarian and Rawlsian social welfare functions are extremes. If the social welfare function is expressed as \((\text{the sum of } U_i^\alpha)^{1/\alpha}\), where \(U_i\) is individual utility, then the Utilitarian case is the case where \(\alpha\) is 1, and the Rawlsian case is the limit as \(\alpha\) approaches minus infinity (Atkinson, 1973; Feldstein, 1973).

For many years, economists eschewed any analysis involving the summation of individual utility, because of the impossibility of making objectively meaningful inter-personal comparisons of utility. The taboo on such comparisons began to break down in about 1970, when economists began to weight changes in income by estimates of the marginal utility of money. The most common weighting factor used was the reciprocal of income. Economists did so in spite of Samuelson’s (1950) warning that no indicator based on a weighted sum of incomes could be valid if relative prices changed and people differed in preferences or incomes, because this practice represented the closest available approximation to the Utilitarian criterion of maximizing the sum of individual utilities.

*Justifying coercion through consensus and revealed preferences*

The approach to justifying coercion that has become popular among public choice scholars during the second half of the 20th century builds upon the contractarian proposals by Thomas Hobbes and John Locke. The now customary justification, proposed by James Buchanan and Gordon Tullock (1962) in *The Calculus of Consent*, involves the claim that people would agree to the proposed coercion behind a veil of ignorance, that is, if they did not know the positions in which they would find themselves when the rules were to be applied. Their perspective is similar to the perspective of John Rawls (1958), except that Rawls sought a recipe by which an individual might discern what is just, rather than a justification of coercion. And while Buchanan and Tullock described a mechanism by which people could agree on common goals, Rawls argued that he had identified the principles that people would chose behind the veil of ignorance. The principle that has drawn the most controversy is Rawls’ claim that people would
condone inequalities if and only if those inequalities benefit the least-advantaged members of society. In contrast, John Harsanyi (1953) proposed that people would choose Utilitarianism if they did not know the positions they would occupy when the rules were applied.

Although the idea of coercively enforcing what people would agree to behind a veil of ignorance is coherent, it is vulnerable to a criticism offered by Bruce Ackerman (1980, pp. 327-336): When a person says, “There is no way that I would have agreed to that from a position of uncertainty about my position,” it is unsatisfying to have that person silenced with a response of, “I don’t need to listen to anything you say, because now you are not in that position of uncertainty about how you would be affected.” Since it is not actually possible to step behind the veil of ignorance, any speculation about the decisions that people would make behind this veil must remain speculation. Supporters of Rawls’ position might respond to Ackerman’s challenge by pointing out that Rawls was not seeking to justify coercive government action. Rather, he was describing a thought process by which a person, or a group in dialog, might discern what is just. However, the coherence of giving such an answer to Ackerman is undercut by the fact that Rawls viewed political power always as coercive power. But regardless of what people might decide behind a veil of ignorance, any effort to maximize the sum of utilities an adequate justification for coercion? Bentham and John Stuart Mill would have said so, but there may be few who would join them.

While dialog may lead to substantial agreement among scholars as to what would maximize the sum of utilities, an alternative justification of coercion, in the spirit of Buchanan and Tullock (1962), is that ultimately all governmental authority, including the use of coercion to enforce legally established rules, can be justified only through a constitutional framework that has been established through democratic process. This perspective is augmented by connections between democratic process and efficient choice. For situations in which voters act as disinterested judges and all voters have the same probability (greater than ½) of choosing the better of two options, Condorcet (1785) pointed out that the option chosen by the majority is more likely to be the better option than the one chosen by the minority. If voters are regarded as self-interested advocates rather than disinterested judges, then the corresponding theorem proposed by Bowen (1943) states that if voters’ interests with respect to some proposal are distributed symmetrically
around their mean value, then that mean value is positive if and only if there is majority support for the proposal. In both situations, the preferences that are revealed through the democratic process are offered as justification for choosing, and coercively enforcing, one option over the other, even when the winning option does not have unanimous support.

How might governments justify coercive political action when intensities of preferences are not distributed symmetrically? For many years, economists accepted the suggestion of Paul Samuelson (1954) that it is not possible to motivate people to report their preferences for public goods truthfully. However, the “demand-revealing process” or “pivotal mechanism” invented by Edward Clarke (1971; see also Tideman and Tullock, 1976) charges people the marginal costs of accommodating their preferences and thereby gives them an incentive to make truthful reports of their preferences, in dollars. If, as suggested by I.J. Good (1977), one uses estimates of the marginal utility of money as weights in Clarke’s demand-revealing process, then it is possible to select public policies on the basis of what appears to generate the greatest utility.

However, using the demand-revealing process augmented by weights that reflect estimates of people’s marginal utilities of money does not lead to possibilities for making everyone better off. Any deviation from unanimous agreement is an indication that the policy to be implemented will make those not in agreement worse off. The demand-revealing process identifies changes that would make everyone better off if compensation could be paid, but a practice of providing the compensation that would make everyone actually better off would destroy the incentive to reveal the information necessary to identify them. A similar constrained capacity to estimate what maximizes utility arises when, as suggested by Tideman and Plassmann (2017), one adds I.J. Good’s individual weights to any of the other mechanisms that operate in the spirit of the demand-revealing process. Thus while these decision mechanisms are able to identify options that are Pareto efficient, they do not provide justification, for those who benefit from what is efficient, to impose their will coercively on those who disagree. Declining to act on the basis of a showing of efficiency is consistent with John Stuart Mill’s argument that people have an obligation to act morally but no obligation to act efficiently, so that the possibility that an inefficiency will occur does not justify coercion.
An alternative path to justifying coercion with regard to efficient outcomes can be constructed in the spirit of Charles Tiebout’s (1956) insight that, rather than trying to find acceptable policies for all, it may be better to let citizens with similar preferences form their own communities, so that all citizens approve of the policies within their communities. If people have rights to leave the communities of which they are members and have opportunities to form new ones with those who find one another mutually compatible, then those who want to impose rules coercively on fellow citizens who ask, “What right do you have?” can justly answer, “If you don’t like it, form your own polity.”

The Tieboutian construct of competing communities requires that those who disagree with coercive rules be able to object; it therefore does not provide an answer to the challenge, “What about the concerns of those who cannot complain today because they have not been born yet?” This objection brings us back to the Utilitarian justification for state action. The universalistic Utilitarian must acknowledge that those who are yet to be born count no less than those who are alive. Thus a Utilitarian should seek to maximize the sum of all future utility of all persons, present and future. There can be discounting for uncertainty. There can be discounting of dollars for their exchangeability for more future dollars. There can be discounting of dollars for being received by those with a lower marginal utility of money. But there should be no discounting of utility for being in the future; doing so would violate the fundamental Aristotelian ideal that fairness requires that equals be treated equally. There is no justification for treating temporal distance any differently than spatial distance. True Utilitarianism requires treating all people equally, no matter how far apart they live in time or space. Of course, the unknowability of the distant future makes such an effort problematic. Still, it is possible to specify that there should be a one-for-one trade-off between the utility of any person today and any other person at any other time. The rights-constrained Utilitarian would only add that one must protect everyone’s rights before the trade-offs are made.

Conclusion

Three different applications of Utilitarian thought have affected the thinking of economists—first, the positive notion that people act so as to maximize their individual subjective utilities, second, the normative idea that good people treat their neighbors as they treat themselves, and
third, the normative idea that the goal of maximizing a society’s utility justifies coercing people in accordance with the state’s rules and policies.

The notion that people act so as to maximize their individual subjective utilities has been the backbone of economic analysis since the writings of Stanley Jevons (1871) and Carl Menger (1871). The predictions of this framework have been tested and confirmed empirically many times, and there remains little doubt that the idea that people seek pleasures and shun pain leads to a useful framework for describing the decision processes of individuals.

Similarly, the ideas of treating our neighbors as we treat ourselves and bringing happiness to others describe behavior that many people consider highly laudable. Thus Utilitarianism is a wonderful goal for individual decisions about how to allocate the things that are uncontestedly theirs.

Finally, the public-spiritedness of Utilitarians in seeking to apply that same rule to state action deserves respect. But a comparison of the increases in utility of some persons with the decrease in utility of others, even if it be accurate (which can never be verified) is not a sufficient basis for the coercive imposition of that which is seen as increasing aggregate utility. Proposed state actions must be sifted through a filter for consistency with individual rights before the unfortunately imperfect criterion of increasing the sum of utility is applied. It is consistent with justice to require citizens to accept losses on the basis of a Utilitarian criterion if they have agreed to a constitution that supports this, but attempting to justify coercion by an appeal to Utilitarianism without such constitutional support will not provide the basis for an adequate response to the objections of dissenting citizens.
References:


Bentham, J. (1816), Anarchical Fallacies, Published in French as Tactique des Assemblées législatives, suivie d’un traité des sophismes politiques, Geneve: Dumont.


Condorcet, Jean-Antoine-Nicolas de Caritat, marquis de (1785), Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix, Paris, Imprimerie Royale.


Dupuit, J. (1844), De la mesure de l’utilité des travaux publics, Annales des ponts et chaussées, Second series, 8.


http://faculty.som.yale.edu/ShaneFrederick/thesis_s.frederick.pdf).


Endnotes

* Corresponding author.

1 The demand-revealing process motivates individuals to report their monetary gains and losses, but only if there are no compensation payments. See Clarke (1971) and Tideman and Tullock (1976).

2 Early analyses equivalent to cost-benefit analysis can be traced as far back as to Dupuit (1844).

3 Bergson conveyed this to one of the authors (Tideman) in a conversation in about 1971.

4 The alternative to assuming a benevolent social planner is to assume the existence of a representative agent, thereby bypassing and obscuring the social decision making problem altogether.

5 The explanation of why this is what one should take from the Arrow Theorem can be found in Tideman (2006).


7 See Rawls (1991), p. 216. Rawls qualifies the exercise of political power as “proper and hence justifiable only when it is exercised in accordance with a constitution the essentials of which all citizens may reasonably be expected to endorse in the light of principles and ideals acceptable to them as reasonable and rational” (ibid, p. 217). For further discussion, see Friedman (2003), Chapter 8.

8 For analyses and discussions of Condorcet’s jury theorem, see Black (1958) and McLean and Hewitt (1994).

9 For a discussion of this proposition, see Frederick (1999).