What is the *general Welfare*?
Welfare Economic Perspectives on Prosperity and Equity

Charles F. Manski

Department of Economics and Institute for Policy Research
Northwestern University

First Public Version: January 2025
This Version: June 2025

Abstract

Researchers do not know what the framers of the United States Constitution intended when they wrote of the *general Welfare*. Nevertheless, economists can conjecture by specifying social welfare functions that aim to express the preferences of the population. Economists have often simplified utilitarian analysis of public policy by assuming that individuals have homogeneous, consequentialist, and self-centered preferences. In reality, individuals may hold heterogeneous private and social preferences. To enhance policy analysis, I argue that economists should specify social welfare functions that express the richness and variety of actual personal preferences over social states. The possibilities are vast. I focus on preferences for population prosperity and equity. There has been much controversy regarding interpretation of *equity*, a term that public discourse has used in vague and conflicting ways. Specifying social welfare functions that formally express different interpretations of equity may not eliminate disagreements, but it should clarify concepts of equity and reduce the inconsistencies that afflict verbal communication.

1. Introduction

A foundational objective of the Constitution of the United States is to promote the *general Welfare*. The Constitution uses the term twice. The Preamble states:

> "**We the People** of the United States, in Order to form a more perfect Union, establish Justice, insure domestic Tranquility, provide for the common defence, promote the general Welfare, and secure the Blessings of Liberty to ourselves and our Posterity, do ordain and establish this Constitution for the United States of America."

Article I, Section 8, Clause 1, sometimes called the *General Welfare Clause*, begins: "The Congress shall have Power To lay and collect Taxes, Duties, Imposts and Excises, to pay the Debts and provide for the common Defence and general Welfare of the United States."

The Constitutional premise that the United States should promote and provide for the *general Welfare* has rhetorical appeal, but it lacks substance. To make it meaningful requires a clear definition of the term. The Constitution contains no definition, nor have Acts of Congress and Supreme Court decisions clarified what constitutes the *general Welfare*. One might attempt to infer the intended meaning from parts of the Constitution such as the Bill of Rights, and from legislation that Congress has enacted and the Supreme Court has deemed constitutional. Inference might be straightforward if Americans were to hold a consensus perspective on what the nation should seek to achieve. It is evident, however, that the preferences of citizens have been heterogeneous to a considerable degree.

The vagueness of the Constitutional term *general Welfare* contrasts with the specificity of welfare economic study of public policy. Economists assume a particular social welfare function (SWF) or a well-defined class of SWFs. Research in public economics seeks to characterize the social welfare achieved by alternative feasible policies, aiming to find one that maximizes an SWF. Cost-benefit analysis is used to inform policy design and evaluation.

Economists studying policy choice in democracies strive to specify SWFs that in some manner express the values of society rather than the preferences of dictators. The U.S. Constitution begins with the words "We the People" and refers to the *general Welfare*. It does not endorse the April 2025 proclamation of

Donald Trump to journalists at the *Atlantic* that "I run the country and the world." [1]

Paul Samuelson placed responsibility for specification of the SWF on society rather than on the economist, writing (Samuelson, 1947, p. 220): "It is a legitimate exercise of economic analysis to examine the consequences of various value judgments, whether or not they are shared by the theorist." However, the vagueness of the term *general Welfare* and related terms has placed economists in the position of having to hypothesize SWFs rather than draw them transparently from society.

In my own research, I have studied policy choice with what I have called *pragmatic* social welfare functions. I have characterized pragmatic SWFs as ones motivated by (Manski, 2024, p. 58): "some combination of conjecture regarding societal values, empirical study of population preferences, and concern for analytical tractability." I have found the notion of a pragmatic SWF useful. I nevertheless worry that economists commonly hypothesize SWFs that are remote from the actual heterogeneous preferences of society. I have cautioned that economists sometimes use a veneer of pragmatism to conflate science with advocacy (Manski, 2011). That is, they may intentionally specify SWFs that justify policy conclusions they favor.

Economists have mainly studied *personalist* SWFs, ones that are functions only of the personal welfare (aka *utility*) of the members of a specified population. They have, moreover, commonly presumed utilitarian aggregation of interpersonally comparable cardinal utilities. My use of the term *personalist* SWF intentionally revises the word *welfarism* originated by Amartya Sen. He wrote (Sen, 1977, p. 1559): "The general approach of making no use of any information about the social states other than that of personal welfares generated in them may be called 'welfarism.' " I feel that *personalist* SWF expresses Sen's intended distinction more clearly than the word *welfarism*. The *Oxford English Dictionary* defines *personalism* as follows:[2] "A system of thought according to which reality has meaning only through the conscious minds of persons, ... or that reality consists of interacting persons."

[1] https://www.theatlantic.com/magazine/archive/2025/06/trump-second-term-comeback/682573/, accessed June 22, 2025.

[2] https://www.oed.com/dictionary/personalism_n?tab=meaning_and_use#30959752, accessed June 22, 2025.

Sen (1977) remarked (p. 1559): "welfarism as an approach to social decisions is very restrictive." Later, in his book *The Idea of Justice*, Sen (2009) took the stance of a moral philosopher who stands outside of society and offers ethical prescriptions for it. He belittled utilitarianism, writing that it presumes individuals aim to maximize "happiness," which he interpreted in a self-centered fashion.

I am puzzled by Sen's perspective that welfarism (i.e., personalist SWFs) "is very restrictive" and his interpretation of utilitarianism as presuming self-centered personal preferences. Although welfare economic policy analysis has often conjectured self-centered individuals, the basic economic concept of individual utility places no restrictions on what individuals seek to maximize. In principle, some individuals may be liberal moral philosophers, holding normative beliefs that society should adherence to notions of justice, fairness, and equity. Others may want to enhance inequalities in directions they favor and may advocate discrimination against members of certain groups. Each individual may hold a distinct view on the composition of the population—past, present, and future generations of humans and other forms of life—whose interests should be considered when measuring social welfare.

All of these possibilities may be embraced within the general concept of a personalist SWF and the particular one of utilitarian planning. The possibilities are too vast to explore entirely in one paper. To make the scope of discussion manageable, this paper considers personalist SWFs in which individuals may have preferences for population prosperity and equity. There has been much controversy regarding the meaning of *equity*. Public discourse has used the word in vague and conflicting ways. The related words *equality* and *fairness* have also been used loosely. I will view equality and fairness as synonyms for equity. Some writers may consider the three to be distinct concepts.

Economists speak of *distributional equity* when studying the population distribution of prosperity. Public economists speak of *horizontal equity* in taxation policy. Discourse on climate policy refers to *intergenerational equity*. Public health research uses the term *health equity* in various ways. Legal scholarship refers to *procedural equity*, using such terms as *equal treatment of equals* and *due process*. The U.S. Constitution refers to *due process* in the 5th and 14th Amendments, and it uses the term *equal protection*

of the laws in the 14th Amendment. Much recent debate regarding the combined term *diversity, equity, and inclusion* (*DEI*) has been vague about how each word should be interpreted.

Specifying SWFs that formally express different interpretations of equity may not eliminate disagreements about the types of equity that individuals value. However, I anticipate that it will reduce the misunderstandings and logical inconsistencies that presently afflict verbal communication. This is among the themes of the present paper.

Contrasting Welfare Economics and Moral Philosophy

Before proceeding, I think it important to contrast the usual mindsets of welfare economists and moral philosophers. With the exception of Sen and some others, welfare economists have mainly studied personalist SWFs. In contrast, a non-personalist perspective has been pervasive in moral philosophy. From the philosophical writings of the ancient Greeks through those of the Old and New Testaments through those of the Enlightenment to the present, philosophers have given themselves the authority and responsibility for interpreting the *general Welfare*. Philosophers have made normative arguments that an SWF *should* have certain properties, regardless of what the members of society think.

Philosophers have not agreed on what properties of SWFs warrant normative approval. Yet their literature shares some features. Philosophers often assert that certain deontological principles, such as procedural equity and justice, should supersede consequentialist ones. They often argue for lexicographic evaluation of policies, wherein society first restricts attention to policies deemed deontologically acceptable and only then considers the consequences of these policies. Lexicographic planning prohibits the quantitative weighing of welfare tradeoffs that economists usually recommend. Hildreth (1953) is a relatively rare instance of an economist who considered placing lexicographic ethical constraints on the space of feasible policies.[3]

[3] Using the letter S to denote the policy space, Hildreth wrote (p. 91):

> "It is possible that some ethical values which should be recognized in making social choices apply independently of individual preferences. In the present approach at least some of these values might be expressed as restrictions on the set of achievable alternatives and entered along with technical restrictions in the determination of S. Thus,

It is notable that philosophers often write as if values are binary rather than measurable or comparable. Thus, a policy may be classified as "just or unjust" or as "equitable or inequitable," without distinction of degrees of justice and equity; see Hsieh and Andersson (2021). It is also notable that philosophical discourse rarely engages concrete policy choices. Philosophers appear most comfortable discussing abstract general principles, using idealized examples to illustrate.[4]

Organization of the Paper

I focus on personalist welfare economics rather than the normative scholarship of moral philosophers. As prelude, Section 2 discusses devices that economists have used to circumvent or grossly simplify specification of SWFs. These include the Hicks (1939) introduction of the *new welfare economics* and the proposal by Kaldor (1939) and Hicks (1939) that economists use the artifice of Kaldor-Hicks efficiency to separate the study of Pareto efficiency and the distribution of personal welfare. They include the macroeconomic practice of using *representative-agent* models to avoid consideration of heterogeneity in personal preferences. Atkinson (2011), calling for "The Restoration of Welfare Economics" as a central subject of study by economists, referred to these devices as "avoidance strategies."

Section 3 discusses the common welfare economic practice of assuming that personal preferences are consequentialist and self-centered. These assumptions are not fundamental to welfare economics. Nevertheless, economists have often presumed that individuals only contemplate their personal circumstances and only value the consequences of actions, not the actions themselves. I use the utilitarian optimal income-tax theory pioneered by Mirrlees (1971) to illustrate. Thus research largely assumes that

---

> if certain social states were judged to be ethically undesirable independently of individual preferences, these states could be regarded as not achievable and could be excluded before the ordering based on individual preferences was applied."

[4] Research in jurisprudence has aspects in common with moral philosophy and economics. Some jurisprudence discusses broad principles of justice, equity, and rights in abstraction, as in moral philosophy. Other legal writing aims to be pragmatic, recognizing that reasonable construction and interpretation of particular statutes may require attention to subtle welfare tradeoffs. Research of the second type sometimes restricts itself to qualitative recognition of tradeoffs, but it sometimes proposes relatively precise rules for the practice of criminal or civil justice. Zamir and Medina (2008, 2010) discuss the tension between the various perspectives in legal research.

individuals have homogeneous preferences for (consumption, leisure) outcomes and differ only in their abilities to generate income. I also describe the use of stated-choice experiments to measure the distribution of health-related welfare in a heterogeneous population.

Section 4 opens discussion of personal preferences for equity, aiming to clarify important distinctions. I point out that it may be logically impossible to jointly achieve different types of equity. I call attention to multiple interpretations of the ideas of horizontal equity and equal treatment of equals. I consider two important substantive contexts, health equity and intergenerational equity. I observe that alternative perspectives on equity are embedded in different mechanisms for aggregating personal preferences into a personalist SWF.

Whereas the discussion in Sections 2 through 4 is verbal, Section 5 formally specifies SWFs that express preferences for population prosperity and equity. I focus on policies that present each member of the population with a choice set from which the member chooses an action. In this context, I formalize distributional preferences in which individual utility varies with the population distribution of choice sets, chosen actions, or realized personal welfare. I next specialize to utilitarian social welfare with individual utility separable in the distribution of opportunity. I present a simple analytical illustration and explain how to computationally approximate optimal utilitarian policy when analytical determination of the optimum is not feasible. I call attention to the complex problem of policy choice under uncertainty, when a planner has incomplete knowledge of the SWF of interest.

Section 6 concludes, reiterating themes developed throughout the paper.

2. Welfare Economic Attempts to Circumvent Specification of Personal Preferences

2.1. The New Welfare Economics and Kaldor-Hicks Efficiency

In the Introduction I quoted Samuelson (1947), who counseled economists that it is legitimate to perform policy analysis using welfare functions that express "various value judgments, whether or not they

are shared by the theorist." Economists have often sought to avoid basing analysis on explicit value judgements, Samuelson's perspective notwithstanding. Much economic research has used various devices to circumvent specification of an SWF.

One route was taken in the 1930s and 1940s by the economists who initiated study of the *new welfare economics* (Hicks, 1939). They did not want to specify a particular SWF, which would require taking a stand on interpersonal welfare tradeoffs. Hence, they retreated to the study of Pareto efficiency, the frontier of feasible policies that cannot be improved for all members of the population. A sequence of analyses relating perfect competition to Pareto efficiency under certain strong assumptions became labelled as the *Fundamental Theorems of Welfare Economics*. See Blaug (2007) for a historical review.

Restriction of attention to the Pareto frontier severely limited the ability of economists to study actual planning problems. Hildreth (1953) wrote (p. 82): "The new welfare economics has been primarily criticized for the narrowness of the range of questions to which it provides answers." He added (p. 91): "if we wish to go beyond the comparisons that are possible using only the principle of new welfare economics, the issue is not whether we can do so without making interpersonal comparisons of satisfactions. It is rather, what sorts of interpersonal comparisons are we willing to make." Chipman and Moore (1978) wrote (p. 548): "we shall argue that, judged in relation to its basic objective of enabling economists to make welfare prescriptions without having to make value judgments and, in particular, interpersonal comparisons of utility, the New Welfare Economics must be considered a failure."

Recognizing that the study of the Pareto frontier alone cannot suffice to inform policy choice, Kaldor (1939) and Hicks (1939) proposed that economists separate the study of Pareto efficiency and the distribution of personal welfare. Rather than consider the actual distribution of welfare achieved by a policy, they contemplated fictional redistributions that might in principle be achievable by transfers of wealth. It became common for economists to suggest the use of a hypothetical *lump-sum taxation* mechanism that

would levy a fixed tax on each individual, one whose magnitude cannot be altered by changes in individual behavior. See the discussion in Atkinson and Stiglitz (1980).[5]

Fictional redistribution has become a peripheral topic in economic theory, but the concept continues to be used in applied policy analysis. It is prominent in forms of cost-benefit analysis (CBA) that aggregate the monetary amount that each member of a population would be willing to pay for a specified change in policy relative to a given status quo (or, alternatively, the amount each person would be willing to pay to preserve the status quo). An individual's willingness to pay may be positive or negative depending on how a change in policy would affect the person. The methodology aggregates willingness to pay across the population and uses the result to evaluate a policy change.

To illustrate, I draw on my critique (Manski, 2015) of a review article by Domınguez-Rivera and Raphael (2015) of the use of this type of CBA to evaluate criminal-justice policies. I quote from their article at length not to criticize it, but rather because the writing is unusually clear and thoughtful. See Bar-Gill (2022) for related critique of CBA using willingness-to-pay concepts.

Using Willingness-to-Pay to Evaluate Criminal-Justice Policy

Domınguez-Rivera and Raphael (2015) call attention to some unpalatable features of measuring the social welfare of a policy by aggregate willingness-to-pay. They caution (p. 590):

> "Of course, there are limitations to this conceptual framework for public choice . . . To start, cost–benefit analysis and cost-effectiveness analysis provide a specific weighting (or social accounting) of the relative welfare of alternative groups in society that often conflicts with widely held beliefs regarding fairness and equity. This is a direct result of the use of money as the common metric used to place various benefits and costs on a common footing."

[5] Public economists have often viewed lump-sum taxation as an ideal approach to generation of tax revenue, calling such taxation "non-distortionary." This term is sometimes misinterpreted to mean that the tax does not affect individual behavior. However, a lump-sum tax necessarily affects behavior, as it changes personal wealth. Atkinson and Stiglitz (1980) called attention to this misinterpretation when they wrote (p. 469): "Even a lump-sum tax would have allocative effects. Yet (by definition), it is non-distortionary." Thus, the use of the word "non-distortionary" taxation is tautological, meaning only that an individual cannot alter (distort) the tax he must pay in a specified lump-sum tax assessment.

They observe that willingness to pay is positively associated with ability to pay and state that (p. 596): "This positive relationship between income and benefit and/or cost valuation ultimately results in greater weight being placed on the welfare of the well-to-do in cost-benefit calculations." They subsequently write that (p. 597): "the systematic tendency to place greater weight on the welfare of the wealthy is certainly of concern."

They consider deployment of police across poor and wealthy neighborhoods and write "A cost-effectiveness analysis would recommend reallocating police officers from the poor neighborhood to the wealthy neighborhood." They point out that this type of recommendation is common, stating (p. 600): "Examples of policy proposals that pass the benefit-cost test yet have questionable equity implications abound in many policy domains."

A reader contemplating the above and related excerpts from the article might conclude that CBA is a methodology invented by economists to serve the wealthiest segment of society. A reader might also anticipate that Dominguez-Rivera and Raphael, who repeatedly express concern with equity, would reject the methodology that they have described. To the contrary, they write (p. 601): "Nevertheless, there is a strong case to make for cost-benefit analysis as a principal input for policy making, equity concerns notwithstanding." To justify this, they rely on the concept of Kaldor-Hicks efficiency, writing (p. 601): "A policy proposal that would generate net positive benefits is one that has the potential to be a Pareto improvement if a mechanism exists to redistribute some of the net benefits to those who would otherwise be the policy losers."

As have many other economists, Dominguez-Rivera and Raphael suggest that society consider equity separately from CBA. They write (p. 590): "Responsible analysis requires . . . .a careful parallel analysis of the equity implications of policy alternatives." They later write (p. 628): "Given the unequal distribution of the cost of criminal victimization and anti-crime enforcement efforts, as well as the potential for perceived illegitimacy of the criminal justice system to undermine various public institutions, equity considerations deserve careful attention in all criminal justice policy choices." Yet they provide no guidance on how society might combine the willingness-to-pay type of CBA with equity considerations so as to make desirable policy decisions.

2.2. Representative-Agent Macroeconomics

Macroeconomists often seek a policy that maximizes the presented discounted value of Gross Domestic Product (GDP) or Gross World Product (GWP), without analysis of their distribution. To formalize this practice, they assume a fictitious *representative agent.* This construct, invoked to simplify analysis of economy-wide market interactions, eliminates consideration of heterogeneity in personal preferences. Macroeconomists not only assume a representative agent, but they further simplify analysis by supposing that the agent has only simple self-centered preferences. It typically is assumed that the agent has mathematically tractable preferences placing instantaneous utility on the magnitude of consumption of a single fictitious divisible commodity and that the agent maximizes the present-discounted value of this utility.

Efforts by macroeconomists to provide micro-foundations for representative agent models, beginning with Gorman (1953), have revealed that formal expression of heterogeneous personal preferences as the preferences of a fictitious representative agent is possible only when stringent simplifying assumptions hold. No claim has been made that these assumptions are realistic. Cherrier, Duarte, and Saïdi (2023) provide an informative historical review.

The prevalence of representative-agent models in macroeconomics has diminished in the 21st century, with increasing performance of research assuming various types of tractable heterogeneous-agent models. However, the focus of this recent work has not been to develop SWFs that realistically aggregate personal preferences. It has been to learn how different forms of heterogeneity may affect macroeconomic phenomena, including business cycles and growth. Rather than empirically learn the actual distribution of personal preferences, macroeconomists continue to assume that persons (or households) have tractable preferences placing instantaneous utility on the magnitude of consumption of a fictitious divisible commodity and that they maximize the present-discounted value of this utility. They continue to use

discounted GDP or GWP as a surrogate for the *general Welfare*, viewing the distribution of consumption as a separate concern. Again, the review article of Cherrier, Duarte, and Saïdi (2023) is informative.

3. Social Welfare Functions Assuming Consequentialist Self-Centered Preferences

This section discusses economic research that specifies an explicit personalist SWF. In principle, such a function can aggregate the welfare of individuals who hold heterogeneous consequentialist and deontological preferences over abstract social states. Arrow (1978) put it this way (p. 224):[6]

> What remains is the determination of the social ordering. On what data is it based? In particular, how does it relate to individual preferences over social states, what might be termed, "individual utilities." For purpose of this paper, I am accepting the viewpoint of the utilitarians and of welfare economics. It is assumed that each individual has some measure of the satisfaction he draws from each social state and that the social ordering is determined by the specification of these utilities for all possible social states."

In practice, economists rarely consider personal preferences over abstract social states. The convention has been to assume that preferences are consequentialist and self-centered, sometimes called egoism (Driver, 2022). Fleurbaey (2021) stated this in striking fashion, writing (p. 39):

> "It is standard in normative economics, as in political philosophy, to evaluate individual well-being on the basis of self-centered preferences, utility or advantage. Feelings of altruism, jealousy, etc. are ignored in order not to make the allocation of resources depend on the contingent distribution of benevolent and malevolent feelings among the population."

Fleurbaey did not argue that it is realistic to assume that personal preferences are self-centered. He only remarked that this assumption simplifies analysis.

I provide two prominent illustrations of research assuming that personal preferences are consequentialist and self-centered. Section 3.1 discusses utilitarian optimal income-tax theory. This has mainly assumed that individuals have homogeneous (consumption, leisure) preferences but heterogeneous

[6] Arrow's 1978 statement differs notably from his argument against utilitarian aggregation of utilities in his famous 1951 book. He wrote there that it (Arrow, 1951, p. 11): "seems to make no sense to add the utility of one individual, a psychic magnitude in his mind, with the utility of another individual." Comparison of the 1951 and 1978 statements illustrates that economists have vacillated on the meaningfulness of interpersonally comparable utilities.

abilities to generate labor income. Section 3.2 discusses utilitarian study by health economists of medical treatment. I describe the use of stated-choice data obtained from a sample of individuals to estimate the distribution of private health preferences in a population.

3.1. Optimal Income-Tax Theory

The seminal Mirrlees (1971) study of optimal income taxation has spawned a large body of pragmatic utilitarian analysis of tax policy. This body of work has mainly studied a simple static setting where individuals have homogeneous preferences for (consumption, leisure), utility increases in both goods, and individuals choose how much labor to supply.[7] Heterogeneity stems only from interpersonal differences in abilities to generate labor income, resulting in wage heterogeneity. A social planner is unable to tax wages directly but is able to tax gross income, which is wage multiplied by labor supply. Wages are assumed to be predetermined; thus, tax policy does not affect equilibrium wages.

The planner selects a tax schedule to maximize a utilitarian SWF, subject to a constraint that net tax revenue suffices to fund a predetermined level of government expenditure. The SWF used by the planner applies a specified monotone-concave transformation to the homogeneous personal utility function. This transformation expresses a type of preference for distributional equity. It motivates progressive income tax schedules that impose higher tax rates on persons with higher income and lower (or negative) rates on those with less income. The degree of concavity of the transformation, which affects the optimal income tax schedule and the redistribution achieved, has been a choice made by researchers rather than based on empirical analysis of population preferences for equity.[8]

---

[7] An early exception to the assumption of self-centered preference is Boskin and Sheshinski (1978). They assumed that personal welfare is a function not only of individual (consumption, leisure) but of an individual's position relative to the population distribution. A recent exception is Aronssohn and Johansson-Stenman (2023). They considered a society in which individuals are concerned with the mean of a transformation of disposable income in the population.

[8] Some research on optimal tax theory has proposed aggregation of a weighted average of cardinal utilities, the weights being determined by observable personal attributes. Saez and Stantcheva (2016) wrote (p. 24): "Weights directly capture society's concerns for fairness without being necessarily tied to individual utilities." Observe their reference

Mirrlees showed that, even with the many simplifying assumptions stated above, the structure of the optimal tax schedule is complex. The reason is that the tax schedule affects labor supply in a manner that depends on consumption-leisure preferences and abilities. The sensitivity of labor supply to taxes affects how much redistribution a society can accomplish. Atkinson and Stiglitz (1980) exposit Mirrlees (1971) and the contributions made soon after. Mirrlees (1997) gives a retrospective on his work and later developments. Kaplow (2024) is a recent review of the literature.

3.1.1. Absence of a Credible Basis for Specification of (Consumption-Leisure) Preferences

As a theorist, Mirrlees conjectured specific forms of (consumption, leisure) preferences, his aim being to simplify analysis of a complex mechanism design problem. He did not assert that his preference assumptions were realistic. Indeed, the fact that he titled his 1971 article "An Exploration in the Theory of Optimal Income Taxation" makes plain that he viewed his work as exploratory rather than as providing a basis for design of actual income tax policy. Further evidence of Mirrlees' recognition that he lacked empirical knowledge of personal preferences is found in the conclusion to his article, where he wrote (p. 207): "The examples discussed confirm, as one would expect, that the shape of the optimum earned-income tax schedule is rather sensitive to the distribution of skills within the population, and to the income-leisure preferences postulated. Neither is easy to estimate for real economies."

Credible specification of personal preferences is essential to evaluation of income tax policy. Utilitarian policy analysis should be based on actual rather than conjectured preferences if it strives to be relevant to the real world rather than only to serve as a challenge for economic theorists. In welfare-economic study of income taxation, the distribution of preferences matters in two ways. One is that it

---

to "society's concerns for fairness." This phrase appears to posit an external entity, perhaps a moral philosopher, who assigns weights to members of the population. In subsequent work, Stantcheva (2021) reported empirical findings on perceptions of the fairness of income and estate taxes elicited from respondents to a survey. The question asked respondents if they perceive aspects of the current tax structure to be "fair" or "unfair," without defining these words.

Reviewing research on optimal taxation, Fleurbay and Maniquet (2018) remarked that preference specifications often express values posited by researchers rather than the actual personal preferences of the population under study. They wrote (p. 1036): "Once one acknowledges that the choice of a particular utility measure is always strongly value laden, it is a small step to . . . . treat utilities as normative constructs." They then wrote that this "is probably the dominant view among optimal tax theorists."

determines the personalist SWF. The other is that prediction of labor-supply choices, which is necessary to predict tax revenue, requires knowledge of personal preferences.

It has long been known that economic theory does not predict the direction or magnitude of the response of labor supply to income taxation. To the contrary, it shows that a rational individual may respond in disparate ways. As tax rates increase, a person may decide to work less, work more, or not change labor supply at all. See Robbins (1930).

Modern labor economics envisions labor supply as a complex sequence of schooling, occupation, and work effort decisions made over the life course, perhaps with only bounded rationality. However, we need only consider a simple static scenario to see that a person may respond rationally to income taxes in disparate ways.

Suppose that a person with a predetermined wage and no unearned income allocates each day between paid work and the various non-paid activities that economists have traditionally called leisure. Let a proportional income tax reduce his wage by the prevailing tax rate, yielding his net wage. Assume that the person allocates time to maximize utility, which is an increasing function of net income and leisure.

Different utility functions imply different relationships between the tax rate and labor supply. For example, the labor supply implied by utility functions in the Constant-Elasticity-of-Substitution family increases or decreases with the tax rate depending on the elasticity of substitution. Other utility functions imply that labor supply is *backward-bending*. That is, hours worked may initially increase as net wage rises from zero but, above some threshold, decrease as net wage rises further. Still other utility functions yield more complex non-monotone relationships between net wage and labor supply. Stern (1986) describes a broad spectrum of possibilities.

Given that theory does not predict how income taxation affects labor supply, prediction requires empirical analysis. Robbins (1930) emphasized this, writing (p. 129): “we are left with the conclusion . . . . that any attempt to predict the effect of a change in the terms on which income is earned must proceed by inductive investigation of elasticities.” However, it was optimistic to think that empirical analysis would resolve the matter.

Fifty years after Robbins, Atkinson and Stiglitz (1980) wrote (p. 47): “Neither economic theory nor empirical evidence can provide a conclusive answer to the effect of income taxation on labour supply.” Concluding his detailed comparison of alternative utility and labor supply functions, Stern (1986) wrote (p. 173): “Our general conclusion must be in favour of diversity of functions and great caution in drawing policy conclusions on results based on a particular form.” Close to thirty years later, I reached this pessimistic conclusion after studying the problem of using data on the distribution of labor supply to identify the preference distribution (Manski, 2014, p. 146): “As I see it, we lack the knowledge of preferences necessary to credibly evaluate income tax policies.” Thus, I think it prudent to continue to regard optimal income tax theory as exploratory rather than as providing a credible basis for study of actual tax policy.

3.2. Optimal Choice of Medical Treatments

Health economists have viewed clinicians as utilitarian planners who treat populations of patients. A clinician observes certain covariates for each patient, who has some risk of illness. The objective is to maximize mean health-related utility. It is common to assume that care is individualistic, meaning that the care received by one patient may affect that person but does not affect others. This assumption is generally realistic when considering non-infectious diseases.

A common problem in clinical decision making is that treatments must be chosen with incomplete knowledge of their health outcomes. Health economists often assume that the clinician knows the objective probability distribution of personal outcomes that will occur if a patient with specified observed covariates is given a specified treatment; that is, the clinician has rational expectations. In this setting, the problem of optimizing utilitarian patient care has a simple solution: patients should be divided into groups having the same observed covariates and all patients in a group should be given the care that yields the highest within-group mean patient welfare. Patients with the same observed covariates should be treated uniformly.

Achievable utilitarian welfare weakly increases as more patient covariates are observed. Observing

more covariates enables a clinician to refine the probabilistic predictions of treatment outcomes on which decisions are based. Refining these predictions is beneficial if doing so affects optimal treatment choices. This important result has been discussed by Phelps and Mushlin (1988), Basu and Meltzer (2007), Manski (2013), and elsewhere. Manski, Mullahy, and Venkataramani (2023) provide a proof in the simple setting of choice between two treatments.

3.2.1. Evaluation of Health Welfare Using Stated-Choice Data in Hypothetical Scenarios

Whereas empirical studies of (consumption, leisure) preferences have generally used actual labor-supply data to perform revealed-preference analysis, economists studying health preferences have regularly used stated-choice data elicited from survey respondents. In empirical research of this type, a researcher poses multiple hypothetical choice settings to a person and asks the person to predict the choice he would make in each setting. If the respondents are a random sample of the population of interest, the data may be used to estimate the distribution of preferences in the population. Inference on personal preferences from data on stated choices has a long history in econometric analysis of discrete choice. Ben Akiva, McFadden, and Train (2019) provide a comprehensive review.

A practical advantage of stated-choice analysis is that the choice settings considered are not limited by what nature offers. A researcher can elicit predictions of behavior in a wide spectrum of hypothetical settings. A potentially serious issue is that interpretation of stated-choice data requires assumptions about the way that persons construe the scenarios posed and the cognitive processes they use when responding to questions. Manski (1999) offers a perspective on aspects of this subject.

Stated-choice data have been widely used in utilitarian assessment of medical treatments. The prevalent practice is to measure health-related cardinal utility on a scale called a quality-adjusted life year (QALY). Weinstein, Torrance, and McGuire (2009) describe the scale as follows (p. S5):

> "Health states must be valued on a scale where the value of being dead must be 0, because the absence of life is considered to be worth 0 QALYs. By convention, the upper end of the scale is defined as perfect health, with a value of 1. To permit aggregation of QALY changes, the value scale should have

> interval scale properties such that, for example, a gain from 0.2 to 0.4 is equally valuable as a gain from 0.6 to 0.8."

Although the basic QALY scale is [0, 1], health economists often consider the possibility that persons may view some adverse health states as having value less than 0, thus extending the scale.

To measure lifetime health utility under a specified policy, health economists commonly predict future longevity, assign a QALY level to each predicted future year alive, and sum the total. They typically do so without discounting the future. In this and other respects, QALY characterization of health utility makes questionable assumptions about health preferences to simplify analysis. Nevertheless, it makes a serious effort to empirically measure preference heterogeneity.[9]

4. Preferences for Equity

A vast array of possible preferences over social states opens when one entertains the possibility that individuals need not have purely consequentialist and self-centered preferences. Experimental economists has accumulated empirical evidence for various forms of *other-regarding* (or social) preferences; see Cooper and Kagel (2017) and Fehr and Charness (2025). Concluding their review article, Fehr and Charness call for recognition of social preferences in welfare economics, writing (p. 503):

> "it would be desirable to study the deeper implications of heterogenous social preferences for normative (public) economics. If individuals display altruistic or inequality averse distributional preferences, it does not make much sense to compute optimal policies on the basis of social welfare functions that assume that every individual only cares for his or her own consumption. Isn't economics, after all, built on a deep commitment to respect individuals' preferences? Likewise, if people care also for equality of opportunity, it appears of paramount importance to incorporate that notion into modern

[9] Especially prominent has been the preference elicitation approach known as EQ-5D, developed by the EuroQol Research Foundation; see Devlin and Brooks (2017). In the United Kingdom, the National Institute for Health and Care Excellence (NICE) uses estimates of mean QALY preferences obtained with EQ-5D to make recommendations regarding the cost effectiveness of treatments proposed for coverage in the government's National Health Service (NHS); see NICE (2013). The NHS use of mean QALY valuations to inform government decisions to pay for medical treatments is a remarkable instance of an actual social planner using an empirically-based form of utilitarian welfare analysis to make large-scale policy choices.

> welfare economics rather than computing optimal policies on the basis of a standard utilitarian welfare function that assumes that individuals only care for their own consumption."

Social preferences need not favor equity—some individuals may want to increase some inequalities. I will, however, focus on preference for equity, which has been a subject of intense public discussion and controversy. The controversy may persist in part because the word *equity* has been used in many distinct and often vague ways.

I aim to clarify the distinctions and sharpen the discourse. Section 4.1 observes that it is often impossible to jointly achieve different types of equity. Sections 4.2 and 4.3 discuss various interpretations of the ideas of horizontal equity and equal treatment of equals. Sections 4.4 and 4.5 focus on two substantive contexts, health equity and intergenerational equity. Section 4.6 remarks that alternative perspectives on equity are embedded in different mechanisms for aggregating personal preferences into SWFs.

4.1. Impossibility of Jointly Achieving Different Types of Equity

To begin, I call attention to the important fact that it is sometimes logically impossible to jointly achieve different types of equity.

A simple example occurs in medicine. Consider groups of persons who vary in their response to a given treatment of an illness, the treatment being more effective in curing some groups than others. Equalizing the rate at which these groups receive the treatment will yield disparities in their health outcomes. Contrariwise, equalizing health outcomes across groups will require disparities in the rate at which they receive the treatment. It is thus impossible to equalize both rates of treatment and health outcomes across groups. See Section 4.4 for further discussion.

Dominitz (2003) called attention to the impossibility of achieving different types of equity in a controversial aspect of criminal justice policy: racial profiling in traffic stops searching for illegal drugs. Racial profiling has been criticized as a deleterious inequity. Dominitz observed that profiling has been

measured in multiple ways: by comparison of race-specific search rates, drug find rates, thoroughness of search, rates of detention of the innocent, and rates of apprehension of the guilty. He showed that, if crime rates differ across races, it is logically impossible to simultaneously eliminate disparities in all of these respects. He cautioned (p. 415): "policy makers must decide whether to sacrifice equality of detention rates of the innocent, equality of apprehension rates of the guilty, or both, because they cannot be simultaneously satisfied."

4.2. Varieties of Horizontal Equity

In their textbook on public economics, Atkinson and Stiglitz (1980) distinguished several broad ways in which economists have used the term *horizontal equity* when considering taxation and other policies. They wrote (p. 293): "The principle of horizontal equity states that those who are in all relevant senses identical should be treated identically."[10] They juxtaposed three interpretations of horizontal equity that they had observed in the public economics literature:

(1) (p. 294): "horizontal equity is simply an implication of the more general principle of welfare maximization."

(2) (p. 294): "it is an independent principle of justice, which has to be set into the balance alongside maximization of welfare."

(3) (p. 295): "a restriction on instruments rather than as based on a comparison of distributions."

In interpretation (1), Atkinson and Stiglitz meant that analysis of policy choice assuming certain personalist SWFs and certain policy spaces yields the conclusion that homogeneous treatment of observationally identical persons maximizes social welfare. Interpretation (2) makes horizontal equity a non-personalist ethical consideration, which society may want to consider along with personalist welfare. Interpretation (3) views satisfaction of horizontal equity lexicographically, as an ethical mandate constraining the space of policies that society views as feasible.

[10] Considering people who are not identical in all relevant aspects, Atkinson and Stiglitz (1980) also observed that public economists studying taxation have referred to various forms of *vertical equity* and the related term *ability to pay*, advocating principles of *equal sacrifice* or *equal marginal sacrifice*. See Sections 11-4 and 13-1.

Atkinson and Stiglitz emphasized that the interpretations differ, writing (p. 295): "These three interpretations of horizontal equity are rather different. The first and second are concerned with the *results* of policy; the third is concerned with the *means* used to achieve the results."

4.3. Ex Ante and Ex Post Equal Treatment of Equals

The phrase *equal treatment of equals* is sometimes used as a synonym for horizontal equity. When studying policy choice under uncertainty, it is important to distinguish between ex ante and ex post equal treatment of equals.

In Manski (2009, 2024) and elsewhere, I have studied treatment diversification as an approach to coping with uncertainty regarding policy outcomes. A simple and instructive case is utilitarian allocation of a population of observationally identical individuals to two feasible treatments. Contemplating assignment of everyone to a single treatment, suppose that a planner does not know which treatment yields larger social welfare, but the planner can bound these welfare values. In this setting, I have shown that a specific fractional (aka diversified) allocation minimizes maximum regret.

I have called attention to the fact that diversification yields equal treatment of equals in the *ex-ante* sense that all members of the population have the same probability of receiving a particular treatment. Yet it violates equal treatment in the *ex-post* sense that different persons ultimately receive different treatments. Thus, equal treatment holds ex ante but not ex post.

Atkinson and Stiglitz (1980) recognized the distinction between ex-ante and ex-post equal treatment. They wrote (p. 296): "In some writing on welfare economics, it has been assumed that *ex ante* welfare is the natural welfare function; others would argue that the *ex ante* criterion is unacceptable and even, indeed, unconstitutional as a basis for taxation." They did not write more deeply on the subject because their book focused on policy choice in the absence of uncertainty about policy impacts. In contrast, uncertainty has been a central concern in my research on policy choice.

Manski (2009) observed that democratic societies usually adhere to the ex post sense of equal treatment. Americans with identical income, deductions, and exemptions are required to pay the same federal income tax. The Equal Protection clause in the 14th Amendment to the Constitution is held to mean that all persons in a jurisdiction are subject to the same laws, not that all persons have the same chance of being subject to different laws.

Nevertheless, some accepted policies yield equal treatment ex ante but not ex post. American examples include random tax audits, drug testing and airport screening, calls for jury service, and the Green Card and Vietnam draft lotteries. These policies have not been prompted by the desire to cope with uncertainty that motivates treatment diversification. Yet they indicate some willingness of society to accept policies that provide equal treatment ex ante but not ex post.

Acceptance of equal treatment ex ante but not ex post is especially notable in performance of randomized experiments, undertaken to learn about treatment response. Combining ex ante equal treatment with ex post unequal treatment is precisely what makes randomized experiments informative. Modern medical ethics permits randomization only under conditions of *clinical equipoise*; that is, when partial knowledge of treatment response prevents an ex ante determination that one treatment is superior to another.

4.4. Varying Perspectives on Health Equity

The term *health equity* has become familiar recently, with extraordinary breadth and vagueness of usage. Braverman (2006) describes the modern history of the term in public health research. The abstract to her article begins with the statement (p. 167): "There is little consensus about the meaning of the terms "'health disparities," "health inequalities," or "health equity." Among recent definitions by public health agencies, the U.S. Centers for Disease Control states: "Health equity is the state in which everyone has a fair and just opportunity to attain their highest level of health."[11] The U.S. National Cancer Institute states:

[11] https://www.cdc.gov/health-equity/what-is/index.html. Accessed May 24, 2025.

"A situation in which all people are given the chance to live as healthy a life as possible regardless of their race, ethnicity, sex, sexual orientation, disability, education, job, religion, language, where they live, or other factors."[12] The World Health Organization states: "Health is a fundamental human right. Health equity is achieved when everyone can attain their full potential for health and well-being." [13]

4.4.1. The Utilitarian Perspective

In Section 3.2, I remarked that utilitarian treatment choice by a clinician with rational expectations uses all available patient covariate information to maximize within-group mean personal welfare. Utilitarian treatment choice formally implements a perspective on the idea that medical decision making should be equitable: clinicians should do as well as possible for their patients, on average given what is known about them.

Utilitarian optimization does not imply that patients with different observed covariates should receive the same treatment or that they will experience the same health. The utilitarian sense of equity may be achieved with differences in treatments and health outcomes across groups of patients. Recall the discussion in Section 4.1, pointing out that it is impossible to equalize both rates of treatment and health outcomes across groups if groups of patients vary in their response to treatment of an illness.

4.4.2. Non-Utilitarian Perspectives on the Use of Measures of Race to Make Clinical Predictions

A growing segment of the medical community in the United States have deemed certain treatment and health disparities undesirable from non-utilitarian perspectives on health equity, particularly when the disparities are by race. An influential movement to remove race as a covariate in existing algorithms for medical risk prediction has developed.

---

[12] https://www.cancer.gov/publications/dictionaries/cancer-terms/def/health-equity. Accessed May 24, 2025.

[13] https://www.who.int/health-topics/health-equity#tab=tab_1. Accessed May 24, 2025.

Commentaries by Cerdeña *et al.* (2020), Vyas *et al.* (2020), and Briggs (2022) exemplify calls to cease the use of race as a covariate. Cerdeña *et al.* (2020) stated that (p. 1125): "race-based medicine . . . ., perpetuates health-care disparities." Vyas *et al.* (2020) called for a general reconsideration of the use of race in risk assessments. After documenting the use of race as a covariate in algorithms predicting patient outcomes in many fields of medicine, they asserted (p. 874): "Many of these race-adjusted algorithms guide decisions in ways that may direct more attention or resources to white patients than to members of racial and ethnic minorities. . . . . Given their potential to perpetuate or even amplify race-based health inequities, they merit thorough scrutiny." Briggs (2022) wrote (p. 2116):

> "while it is difficult to refute the central contention that optimal decision making requires the use of all covariates that are associated with outcome, the assumption that racial covariates, and their application within the medical arena, are sufficiently free from bias (structural, institutional or personal) misses the point of the underlying argument: that race is not the same as every other covariate in our arsenal. It is a covariate that is acting as a proxy for a wide range of other explanatory variables that could be genetic/biological but in many circumstances are more likely to be sociological/socioeconomic."

Leading institutions have recommended race-free risk prediction. A notable case is Delgado *et al.* (2021), which recommended removal of race as a predictor of kidney disease. This recommendation has since been implemented in major medical centers.

Manski (2022) questioned four assertions that have been advanced as arguments against the inclusion of race as a covariate in medical risk prediction. These assertions are: (i) race is a social, not biological, concept. (ii) there is no established causal link between race and the illness. (iii) using race may perpetuate or worsen racial health inequities. (iv) many persons are offended by the use of race in risk assessment.

I observed that assertions (i), (iii), and (iv) are empirical assertions that have been the subject of considerable controversy, with evidence being scant. I observed that assertions (i) and (ii) are not relevant to assessment of race as an informative predictor of illness. With the stated goal of making clinical decisions that would maximize utilitarian welfare, I concluded with this observation (p. 2113):

> "If an alternative perspective is to have a compelling foundation, it should explain why society should find it acceptable to make risk assessments using other patient characteristics that

> clinicians observe, but not race. It should explain why the social benefit of omitting race from risk assessment is sufficiently large that it exceeds the harm to the quality of patient care."

I made this statement in frustration because advocates of removing race from medical risk assessments have argued loosely. They have not made clear the logic of claiming that use of race as a predictor in medical risk assessment is not appropriate if race is a social construct. Nor have they made clear the logic of concern with causality, given that statistical association suffices for successful prediction within a population.

Empirical analysis of patient preferences has been lacking. In particular, it is not known to what extent patient populations would be willing to give up some accuracy in the medical predictions made for them, in order to mitigate the types of racial disparities that medical commentators have argued are undesirable.[14] Nor has there been much empirical study of how elimination of race as a predictor affects the magnitudes of the various types of disparities that commentators have deemed problematic. Manski, Mullahy, and Venkataramani (2023) elaborate on these themes.

4.5. Preferences for Intergenerational Equity

In Section 2.2, I observed that macroeconomists commonly use present discounted GDP or GWP as a surrogate for the *general Welfare*. In so doing, they ignore the distribution of income among persons who are currently alive. However, discounting the future takes an explicit stand on the distribution of income across generations, with the discount rate expressing preferences for intergenerational equity.

It is not feasible to elicit the preferences of future humans who are as yet unborn. Elicitation of preferences for intergenerational equity might in principle be feasible from the population who are alive

[14] Health economists have used stated-choice experiments to study personal preferences for various types of health equity in highly idealized settings. See, for example, Dolan (1998), Cookson *et al*. (2021), and Robson *et al*. (2024). The analyses in this body of work are remote from the concerns of medical commentators with the use of race to make clinical predictions.

currently, but that is not the practice.[15] Rather, the discount rate is a choice made by researchers. I use economic analysis of climate policy to illustrate.

4.5.1. Specification of the Discount Rate in Analysis of Climate Policy

Integrated assessment (IA) models enable quantitative evaluation of alternative climate policies. IA models provide long-run descriptions of the global economy, including the energy system and its role in economic production. They represent the climate and the links between the climatic effects of greenhouse gas (GHG) emissions and their impacts on the economy. IA models have become primary tools for comparing potential policies to reduce GHG emissions. A leading example is the Dynamic Integrated Climate Economy (DICE) model described in Nordhaus (2019).

In DICE and similar models, the economic losses from climate change are represented by damage functions that give the decreases in world-wide output resulting from increases in mean global temperature, as a proportional reduction or in dollar terms. Policy comparisons have been performed by considering a planner who seeks to make optimal trade-offs between the costs of carbon abatement and the global economic damages from climate change. The planner is assumed to face an optimal-control problem, the objective being to minimize the present discounted costs of abatement and damages over a time horizon. Thus, present discounted gross world product expresses social welfare.

The use of discounted world product to express welfare ignores the distribution of welfare among persons who are currently alive.[16] However, it does not ignore intergenerational equity. The specified

[15] While elicitation of the preferences of the current population for intergenerational equity has not been the practice, it is a central premise of *The Ministry for the Future* (Robinson, 2020). In his speculative novel, Robinson conjectures creation of a United Nations organization whose mission is (p. 16): "to advocate for the world's future generations of citizens, whose rights, as defined in the Universal Declaration of Human Rights, are as valid as our own."

[16] Discussions of climate policy often refer to distributional impacts, but only verbally. Consider, for example, a report on decarbonization policy published by a committee of the National Academies of Science, Engineering, and Medicine (2023). The Committee wrote that society should aim to achieve (p. 4): "a fair, equitable, and just 30-year transition" to decarbonization. It went on to write (p. 51):

> "the objective to reduce emissions to net zero by midcentury (or 50 percent by 2030) can be thought of as a constraint with the goal to minimize cost while maximizing desirable societal objectives of equity, employment, health, and public engagement."

discount rate quantifies the present value of the future costs and benefits resulting from climate policy. The appropriate discount rate has been a long-standing and contentious issue in climate economics. Controversy persists in part due to the fact that the choice of an appropriate discount rate is not only an empirical question concerning the state of the economy in the future. It is a normative matter concerning preferences for equity across future generations. Economists commonly use some version of the *Ramsey formula* (Ramsey, 1928) to expression the interplay of normative and empirical considerations in choosing a discount rate. See Dasgupta (2008).

The specified discount rate can be a highly consequential determinant of conclusions on optimal climate policy. This became plain in 2007 when Nordhaus (2007) and Stern (2007) reported the findings of studies using different discount rates. Stern used a relatively low rate and concluded that policy should seek to reduce GHG emissions aggressively and rapidly. Nordhaus used a relatively high rate and concluded that policy should act moderately and slowly.

4.6. Perspectives on Equity in Alternative Social Welfare Functions

To the extent that personal preferences are heterogeneous, optimal policy choice is necessarily sensitive to the preference aggregation mechanism specified in a personalist social welfare function. Different SWFs embody different perspectives on equity. Welfare economics has viewed the choice of the SWF as a normative meta-decision made by a social planner, who stands outside of society and acts in its behalf. Thus, welfare economics has in practice been only partially personalist.

To enrich the personalist aspect, economists could seek to learn what preferences individuals have over alternative SWFs. They could then conjecture a planner who decides how to aggregate preferences for the structure of the SWF. It thus appears that welfare economics cannot be completely personalist. An external

---

The Committee did not formalize the terms “fair, equitable, and just.” Nor did it assess what society should do if the “objectives of equity, employment, health, and public engagement” should be in tension with one another.

planner would have to specify a second-order welfare function to adjudicate heterogeneity in personal preferences for the SWF.

4.6.1. Harsanyi and Rawls As-If Consensus on the Social Welfare Function

Seeking to overcome the need to posit any external social planner, Harsanyi (1955) and Rawls (1971) conjured a thought experiment that they interpreted in competing ways, yielding dramatically different perspectives on equity. Harsanyi and Rawls both asserted that society as a whole would have consensus in favor of a particular SWF. However, they disagreed on what this consensus would be. For Harsanyi, it was utilitarian aggregation of interpersonally comparable cardinal utilities. For Rawls, it was maxmin evaluation of interpersonally comparable ordinal utilities.[17] I begin with the Rawls argument, which is more widely known, even though the Harsanyi proposal preceded it by about fifteen years.

Rawls argued that the SWF should be determined by a social contract. He maintained that this contract should express a consensus that he argued all rational people would accept in an *initial position*, characterized by a *veil of ignorance*. He wrote (p. 10):

> "the guiding idea is that the principles of justice for the basic structure of society are the object of the original agreement. They are the principles that free and rational persons concerned to further their own interests would accept in an initial position of equality."

He declared that he knew what principles free and rational persons would accept, writing (p. 13):

> "I shall maintain instead that the persons in the initial situation would choose two rather different principles: the first requires equality in the assignment of basic rights and duties, while the second holds that social and economic inequalities, for example inequalities of wealth and authority, are just only if they result in compensating benefits for everyone, and in particular for the least advantaged members of society."

Thus, Rawls assumed that personal welfare is ordinally comparable across individuals. and he argued that social welfare should be measured as the minimum personal welfare of all members of society.

[17] A distinction between the Harsanyi and Rawls arguments is that Harsanyi, but not Rawls, considered an individual's overall measurement of utility. Rawls conjectured a lexicographic process in which an individual first establishes certain principles of justice and then evaluates his material circumstances. He applied his maxmin argument regarding the SWF only to what he called *primary goods*, not the entirety of material circumstances.

Rawls did not originate the idea that all rational people would agree on the SWF in a hypothetical original position. Harsanyi (1955) posed a thought experiment of this type and reached a different conclusion. He argued that, not knowing their positions in society, individuals in the original position would place equal probability on realizing each possible position and would maximize expected utility. He thus concluded that all rational persons would accept a utilitarian SWF. Writing fifteen years later, Rawls barely acknowledged the precedent Harsanyi argument, mentioning Harsanyi by name only briefly in a footnote. Nevertheless, he attacked utilitarianism sharply.

The Harsanyi and Rawls arguments differ fundamentally in their perspectives on equity. Maximization of expected utility with equal probabilities of each possible initial position yields ex ante equal treatment of equals, but it provides no guarantee against extreme inequality in personal welfares ex post. Maximization of minimum utility guarantees against extreme deprivation in ex post minimum welfare, but it is insensitive to the distribution of welfare above the minimum.

Critics have questioned how one could know that all free and rational persons would accept either the Harsanyi or Rawls principles. In his review of the Rawls book, Arrow (1973) wrote (p. 247):

> "How do we know other peoples' welfare enough to apply a principle of justice?" . . . ."the criterion of universalizability may be impossible to achieve when people are really different, particularly when different life experiences mean that they can never have the same information."

He concluded by writing (p. 263):

> "To the extent that individuals are really individual, each an autonomous end in himself, to that extent they must be somewhat mysterious and inaccessible to each other. There cannot be any rule that is completely acceptable to all."

5. Social Welfare Functions Expressing Personal Preferences for Prosperity and Equity

I wrote at the outset that specification of SWFs expressing alternative interpretations of equity should reduce the misunderstandings and logical inconsistencies that presently afflict public discourse. I now proceed.

I begin abstractly in Section 5.1. This enables transparent presentation of basic principles, but some specificity is needed to express preferences for prosperity and equity. Section 5.2 focuses on policies that present each member of the population with a choice set from which the member chooses an action. I assume that individual utility has a self-centered component and a distributional component. The distributional component expresses personal preferences for population prosperity and equity. I distinguish preference for equality of choice sets (aka equal opportunity) and for equality of chosen actions.

Section 5.3 specializes further to formalize utilitarian social welfare when individual utility is additively separable in the population distribution of choice sets. Section 5.4 presents an illustration. Section 5.5 discusses computational approximation of optimal policy when it is not feasible to determine the optimum analytically. Section 5.6 considers policy choice under uncertainty, when a planner has incomplete knowledge of the SWF of interest.

5.1. Personalist Social Welfare in Abstraction

Welfare economists have sometimes posited that individuals have preferences over abstract social states. See, for example, Arrow (1951, 1978). A personalist SWF aggregates these preferences in some manner.

Let J denote the population under consideration. Let Φ denote the set of feasible policies and Z denote the set of feasible social states. For each policy $\varphi \in \Phi$, let $z_\varphi \in Z$ denote the social state attained under φ. Let each individual $j \in J$ have a preference ordering on Z, admitting a utility function $u_j(\cdot)$: $Z \to R$. If individual j were empowered to be the social planner, with the capacity to choose a policy unilaterally, j would choose one that maximizes $u_j(z_\varphi)$, $\varphi \in \Phi$.

A personalist SWF is a real function $W(\cdot)$ of the utilities of all members of J. For policy φ, personalist social welfare is $W[u_j(z_\varphi), j \in J]$. A planner with SWF $W(\cdot)$ chooses φ to maximize $W[u_j(z_\varphi), j \in J]$. Research in welfare economics has generally assumed that social welfare respects personal preferences, in the sense that $W(\cdot)$ is strictly monotone in each individual utility.

Research has differentiated SWFs in which the social preference ordering is invariant to ordinal transformations of individual utility from ones in which it is invariant to interpersonally comparable cardinal transformations. In the former case, the social welfare ordering of policies remains the same if one replaces $u_j(\cdot)$ by any strictly increasing function of $u_j(\cdot)$. In the latter case, the social welfare ordering remains the same if one replaces $u_j(\cdot)$ by $\alpha + \beta u_j(\cdot)$ for any real α and positive β.

In contrast with the literature discussed in Section 3, a central consideration in this paper is

Theme 1: The utility functions $u_j(\cdot)$, $j \in J$ need not be self-centered. ■

Individual utilities are self-centered if each social state z is a vector $z = (z_j, j \in J)$ and, for each j, $u_j(z)$ varies with z only through z_j. Personal preferences over abstract social states need not have this decomposable structure. In principle, $u_j(z)$ may vary with z in any manner.

It is important to recognize that making $W(\cdot)$ a personalist SWF does not imply that the individual utility functions $u_j(\cdot)$, $j \in J$ are themselves personalist. That is, $u_j(\cdot)$ may vary with z in ways other than through the values of $u_k(z)$, $k \in J$. The utility functions with distributional preferences to be introduced in Section 5.2 are not personalist.

5.2. Policies Imposing Individual Choice Sets

I now reduce the abstraction of Section 5.1. I suppose that a feasible policy presents each member of the population with a choice set. Under policy φ, individual $j \in J$ faces choice set $C_{\varphi j}$ and chooses action $c_{\varphi j} \in C_{\varphi j}$. Individuals have personal covariates $(x_k, k \in J)$ that do not vary with policy. These may, for example, include their demographic attributes and their utility functions. A social state is a vector $z_\varphi = (x_j, c_{\varphi j}, C_{\varphi j}; j \in J)$. The set of feasible social states is $Z = (z_\varphi, \varphi \in \Phi)$.

An example is income tax policy. Here φ is a tax structure, $C_{\varphi j}$ is the set of (consumption, leisure) pairs available to j under φ, and $c_{\varphi j}$ is the choice that j would make. Another example is medical treatment. Here φ is a health policy, $C_{\varphi j}$ is the set of treatments available to j under φ, and $c_{\varphi j}$ is the treatment that j would choose, perhaps in conjunction with a clinician.

Economists studying income tax and health policy commonly assume that personal preferences are self-centered. They also assume that utility is a function only of the choice $c_{\varphi j}$ that an individual makes, not the choice set $C_{\varphi j}$. Thus, $u_j(z_\varphi) = u_j(c_{\varphi j})$. These assumptions are restrictive. In principle, $u_j(\cdot)$ may be a function of the covariates, choices, and choice sets of all members of the population.

5.2.1. Distributional Preferences

Public discourse on policy typically concerns the distribution of circumstances across a population, not the circumstances of particular named individuals. For each policy φ, let $P(x, c_\varphi, C_\varphi)$ denote the population distribution of covariates, choices, and choice sets. With this notation, we may formalize *distributional preferences* by assuming that $u_j(\cdot)$ is a function of j's own choice and choice set $(c_{\varphi j}, C_{\varphi j})$, which are self-centered, and of the distribution $P(x, c_\varphi, C_\varphi)$. Thus,

$$(1) \qquad u_j(z_\varphi) = u_j[(c_{\varphi j}, C_{\varphi j}), P(x, c_\varphi, C_\varphi)].$$

Among utility functions of this form, considerable simplification in analysis of individual choice behavior occurs if $u_j(\cdot)$ is monotonically separable in its self-centered component. That is, let

$$(2) \qquad u_j(z_\varphi) = f_j[v_j(c_{\varphi j}, C_{\varphi j}), P(x, c_\varphi, C_\varphi)],$$

where $v_j(\cdot)$ and $f_j(\cdot)$ are real-valued and where $f_j(\cdot)$ is strictly increasing in $v_j(\cdot)$. Also let J be a population whose members interact at most infinitesimally, in the sense that the action chosen by any individual j does not affect the distribution $P(x, c_\varphi, C_\varphi)$. Given these conditions, we have

Theme 2: Policy analysis assuming personal preferences to be self-centered accurately portrays rational individual choice behavior under a specified policy even if individuals have distributional preferences. ■

The reasoning is that distribution $P(x, c_\varphi, C_\varphi)$ is invariant to the action chosen by j if interactions are at most infinitesmal. Hence, utility maximization by j implies that $c_{\varphi j}$ maximizes $v_j(d, C_{\varphi j})$ over $d \in C_{\varphi j}$. [18]

5.2.2. Interpretations of Equity

A utility function with distributional preferences may vary with the entire distribution $P(x, c_\varphi, C_\varphi)$. Many interpretations of equity can be formalized as specific properties of this distribution. Here I distinguish equality of opportunity, equality of chosen actions, and equality of realized personal welfare. For each interpretation, I distinguish complete equality and intergroup equality.

Equality of Opportunity

Complete equality of opportunity under policy φ means that all members of the population face the same choice set. Formally, $P(C_\varphi)$ is degenerate.

Complete equality of opportunity is a very strong property. Much recent public discourse has concerned weaker concepts of intergroup equality, for example racial or gender equality. To formalize these ideas, use the attributes x to classify the population into a set G of mutually exclusive and exhaustive attribute groups (X_g, $g \in G$). Let $P(C_\varphi|x \in X_g)$ denote the distribution of choice sets for persons with attributes in group X_g. I will interpret intergroup equality of opportunity under policy φ to mean that every group faces the same distribution of choice sets. That is, $P(C_\varphi|x \in X_g) = P(C_\varphi|x \in X_{g'})$ for all distinct groups g and g′.

[18] Although Theme 2 greatly simplifies analysis, I should caution that it is at odds with empirical evidence that individuals sometimes behave altruistically in settings where their private choices only infinitesimally affect the population distribution of choice sets and chosen actions.

I think it important to emphasize

Theme 3: Intergroup equality is consistent with any degree of intragroup variation in choice sets. ■

For example, a preference for racial equality of opportunity does not imply that all members of a given racial group face the same choice set. Racial equality is consistent with extreme inequality across individuals of each race.

Intergroup equality is equivalent to complete equality only in the limit case where G = J. Then each g ∈ G names a single individual. Hence, $P(C_\varphi|x \in X_g)$ is degenerate.

Equality of Chosen Actions

Complete equality of chosen actions under policy φ means that all members of the population make the same choice; thus, $P(c_\varphi)$ is degenerate. If individuals with heterogeneous preferences all face the same non-singleton choice set, they may make different choices. Thus, as illustrated earlier in the medical treatment example of Section 4.1,

Theme 4. It is not generally possible to achieve complete equality of both opportunity and chosen actions if individual preferences are heterogeneous and policy φ does not mandate one action. ■

To formalize intergroup equality, again classify the population into groups (X_g, g ∈ G). Let $P(c_\varphi|x \in X_g)$ denote the distribution of chosen actions for persons with attributes in group X_g. Intergroup equality of chosen actions under policy φ means that $P(c_\varphi|x \in X_g) = P(c_\varphi|x \in X_{g'})$ for all groups g and g′. Extending Theme 3, intergroup equality of chosen actions is consistent with any degree of intragroup variation in chosen actions.

Equality of Realized Personal Welfare

Choice sets and chosen actions are objective quantities that may be observable. Individual utilities are subjective and generally are not observable. Nevertheless, public discourse sometimes refers to equality of personal welfare. This concept is well-defined if utilities are interpersonally comparable. Then we may interpret complete equality of realized personal welfare under policy φ to mean that $P[u(z_\varphi)]$ is degenerate. Intergroup equality means that $P[u(z_\varphi)|x \in X_g] = P[u(z_\varphi)|x \in X_{g'}]$ for all groups g and g′.

5.3. Utilitarian Social Welfare with Individual Utility Separable in the Distribution of Opportunity

Policy discourse commonly expresses a preference for both prosperity and equity. Supposing that these concepts are well-defined, there may be close to consensus that one policy dominates another if the former increases both prosperity and equity relative to the latter. Policy comparison is subtle if neither dominates. Individuals may differ in their rankings of policies, depending on how they value prosperity and equity.

Adding to the specificity of Section 5.2, I now formalize prosperity-equity preferences under the assumption that individual utility functions are additively separable in the distribution of opportunity. I assume that the self-centered component of utility is a function of the chosen action but not the choice set. The distributional component is a function of the population distribution of covariates and choice sets but not the distribution of chosen actions or realized personal welfare. Thus, individual utility has the form

$$(3) \qquad u_j(z_\varphi) = v_j(c_{\varphi j}) + w_j[P(x, C_\varphi)],$$

where $w_j(\cdot)$ maps the distribution of covariates and choice sets into the real line. Observe that utility function (3) is not personalist. Individual j has a preference over the population distribution of choice sets, not the distribution of realized welfare.

If population members interact at most infinitesimally, utility maximization by individual j under policy φ implies that $c_{\varphi j}$ maximizes $v_j(d)$ over $d \in C_{\varphi j}$. Theme 2 observed that the choice is unaffected by

the distribution $P(x, C_\varphi)$. This does not imply that j would, if empowered to be the social planner, choose a policy to maximize $v_j(d)$ over $d \in C_{\varphi j}$, $\varphi \in \Phi$. A completely self-centered person would choose such a policy, but one with distributional preferences may not. I think it important to emphasize

Theme 5: If empowered to be the social planner, an individual with distributional preferences may choose to sacrifice some personal prosperity in order to enhance the population distribution of opportunities. ■

Utility function (3) has a relatively simple form, yet it is general enough to express a wide spectrum of personal preferences. The relative variability of functions $v_j(\cdot)$ and $w_j(\cdot)$ expresses the extent to which j values his own chosen action relative to the population distribution of opportunity. The form of $w_j(\cdot)$ expresses how j values population prosperity and equity.

Consider two policies φ and φ' such that $C_{\varphi k}$ is a superset of $C_{\varphi' k}$ for almost all $k \in J$. Thus, almost the entire population is more prosperous with policy φ than with φ'. If individual j is concerned only with population prosperity, without regard for equity, function $w_j(\cdot)$ should satisfy the inequality $w_j[P(x, C_\varphi)] > w_j[P(x, C_{\varphi'})]$. This inequality need not hold if j has a preference for equality of opportunity. A policy that enhances prosperity but reduces equality may be deemed worse than an alternative.

Loosely speaking, increasing equality of opportunity means that the spread of $P(x, C_\varphi)$ is less than that of $P(x, C_{\varphi'})$. There are many ways to measure the magnitude and spread of the distribution of opportunities. Hence, there are many ways that an individual might evaluate population prosperity and equity. Section 5.4 provides an illustration.

If individuals have interpersonally comparable utility functions of form (3), utilitarian social welfare under policy φ is

$$(4) \quad W[u_j(z_\varphi), j \in J] = E[v(c_\varphi)] + E\{w[P(x, C_\varphi)]\}.$$

The expectations are taken across the population. $E[v(c_\varphi)]$ is utilitarian welfare assuming that individuals are self-centered. $E\{w[P(x, C_\varphi)]\}$ is the mean distributional preference of the population. A utilitarian planner chooses a policy that maximizes (4) across $\varphi \in \Phi$, jointly considering the self-centered and distributional components of mean preferences.

5.4. Illustration of Optimal Utilitarian Policy: Provision of an Innovative Medical Treatment

In general, computation of optimal utilitarian policy is a challenging task, to be discussed in Section 5.5. Here I examine an illustrative setting that enables an instructive simple analysis. The illustration concerns the provision of an innovative medical treatment to the population.

Let a costless existing medical treatment A be available to all members of the population. Let policy φ make a new treatment B available to a randomly selected fraction φ of the population. Let treatment B be costless but the quantity be limited, enabling treatment of at most a fraction φ_{max} of the population. Thus, the set of feasible policies is the interval $\Phi = [0, \varphi_{max}]$.

Consider policy φ. Let individuals have utility functions of form (3). If only A is available to individual j, the choice set is $C_{\varphi j} = \{A\}$, implying that the self-centered component of utility is $v_j(A)$. If both A and B are available to j, the choice set is $C_{\varphi j} = \{A, B\}$, so self-centered utility for the chosen action is $\max[v_j(A), v_j(B)]$. Hence, mean self-centered utility is $E[v(c_\varphi)] = E[v(A)](1 - \varphi) + E\{\max[v(A), v(B)]\}\varphi$.

An individual may value population prosperity and equality of opportunity. I assume that $w_j[P(x, C_\varphi)] = \alpha_j\varphi - \beta_j\varphi(1 - \varphi)$. Here $\alpha_j \geq 0$ and $\beta_j \geq 0$ are person-specific parameters expressing the extent to which j values population prosperity and equity respectively. When $\alpha_j > 0$, $\alpha_j\varphi$ expresses a preference for greater population prosperity: utility increases linearly with the fraction of the population who face the larger choice set {A, B}. When $\beta_j > 0$, the quadratic function $-\beta_j\varphi(1 - \varphi)$ expresses a preference for greater population equality of opportunity. The function is maximized at the value zero when φ equals zero or one, each of which gives complete equality of opportunity. It is minimized at the value $-\beta_j/4$ when $\varphi = ½$, where inequality of opportunity is greatest.

With this utility specification, utilitarian social welfare is

$$(5)\quad W[u_j(z_\varphi), j \in J] = E[v(A)](1-\varphi) + E\{\max[v(A), v(B)]\}\varphi + E(\alpha)\varphi - E(\beta)\varphi(1-\varphi)$$
$$= E[v(A)] + \left[E\{\max[v(A), v(B)]\} - E[v(A)] + E(\alpha) - E(\beta)\right]\varphi + E(\beta)\varphi^2.$$

The term $E\{\max[v(A), v(B)]\} - E[v(A)]$ is non-negative, and it is positive except in the polar case where treatment A dominates B. Hence, social welfare increases with φ if $E(\alpha) - E(\beta) \geq 0$. It follows that the optimal policy is φ_{max} in this setting.

Now suppose that $E(\alpha) - E(\beta) < 0$. This inequality occurs if the mean valuation of equality of opportunity exceeds that of population prosperity. Social welfare remains an increasing function of φ and the optimal policy remains φ_{max} if $E(\alpha) - E(\beta)$ less negative than $E\{\max[v(A), v(B)]\} - E[v(A)]$ is positive.

The behavior of social welfare as a function of φ is more subtle if $E(\alpha) - E(\beta)$ more negative than $E\{\max[v(A), v(B)]\} - E[v(A)]$ is positive. Then the quadratic function (5) is minimized at a value of φ in the interval $[0, ½]$. Differentiating (5) with respect to φ yields the first order condition for minimum welfare, which occurs at the value

$$(6)\quad \varphi_{foc} = ½\left\{E(\beta) - E(\alpha) - [E\{\max[v(A), v(B)]\} - E[v(A)]]\right\}/E(\beta).$$

If $\varphi_{max} \leq \varphi_{foc}$, social welfare decreases with φ throughout the interval $[0, \varphi_{max}]$. Hence, the optimal policy is $\varphi = 0$, yielding social welfare $E[v(A)]$. If $\varphi_{max} > \varphi_{foc}$, social welfare decreases with φ in the interval $[0, \varphi_{foc}]$ and increases in the interval $[\varphi_{foc}, \varphi_{max}]$. The optimal policy is $\varphi = 0$ if (5) evaluated at φ_{max} is less than $E[v(A)\}$. It is $\varphi = \varphi_{max}$ if (5) evaluated at φ_{max} is greater than $E[v(A)]$.

5.5. Computational Approximation of Optimal Utilitarian Policy

The simplicity of the illustration in Section 5.4 should not obscure the fact that analytical determination of optimal utilitarian policy can be a challenging task. This became evident in the 1970s in research that sought to characterize optimal income tax policy under the idealized assumptions that individuals are self-centered and choose labor supply in static settings; see Atkinson and Stiglitz (1980). Study of optimal tax policy is yet more complex when individuals are presumed to face more realistic choice problems and to have distributional preferences. See Aronssohn and Johansson-Stenman (2023).

A tractable alternative to analytical study of optimal policy may be to approach the problem computationally. Monte Carlo simulation can approximate utilitarian welfare for a specified policy. Maximization of welfare over a relatively small finite set of policies can approximate maximization over a large set of policies.

Suppose that utilitarian welfare has form (4). To perform Monte Carlo simulation, one may draw a random sample of size N from J, say J_N. Specify a policy φ, assume observability of $[v_j(\cdot), w_j(\cdot), x_j, C_{\varphi j};\ j \in J_N]$, and assume rational choice behavior. Then one may determine the chosen action $c_{\varphi j}$ for each member of the sample J_N. The empirical averages $E_N[v(c_\varphi)]$, $P_N(x, C_\varphi)$, and $E_N\{w[P_N(x, C_\varphi)]\}$ approximate $E[v(c_\varphi)]$, $P(x, C_\varphi)$, and $E\{w[P(x, C_\varphi)]\}$ respectively. The Strong Law of Large Numbers implies that these empirical averages consistently estimate (4) as $N \to \infty$.

If the policy space Φ is finite and not too large, one may approximate the optimal utilitarian policy by solving the problem $\max_{\varphi \in \Phi} E_N[v(c_\varphi)] + E_N\{w[P_N(x, C_\varphi)]\}$. If Φ is too large for this maximization to be tractable, one may maximize over a finite subset of policies that aims to adequately approximate Φ.

5.6. Policy Choice under Uncertainty

The discussion of computation in Section 5.5 assumed observability of $[v_j(\cdot), w_j(\cdot), x_j, C_{\varphi j}]$ for a random sample of the population and for a tractable subset of Φ. In practice, this information is rarely available. I

called attention in Section 3.1.1 to the absence of a credible basis for specification of (consumption-leisure) preferences in studies of optimal income tax policy. Empirical knowledge of personal preferences is scant in many domains. How then to proceed?

A long term goal obviously should be to perform new empirical research that enhance the available knowledge of preferences. The prospect for meaningful progress are unclear, but the effort should be made.

In the near term, one may perform sensitivity analysis. Recall that Samuelson, 1947, p. 220) stated: "It is a legitimate exercise of economic analysis to examine the consequences of various value judgments, whether or not they are shared by the theorist." Thus, lacking empirical knowledge of the preferences of the population, welfare-economic research can conjecture alternative distributions $P[v(\cdot), w(\cdot), x, C_\varphi]$, $\varphi \in \Phi$, and determine optimal policy choice under each conjectured distribution.

In the near term, one may also study planning under uncertainty. I have studied this challenging problem in a program of research that began in the early 2000s, with focus on evaluation of policy by maximum regret. Manski (2024) provides a comprehensive exposition. This work has studied the decision problem of a planner who must choose a policy with incomplete knowledge of the SWF. The applications to date have addressed utilitarian planning assuming that individuals are self-centered. There is a pressing need to extend consideration to settings where individuals may have distributional preferences.

6. Conclusion

No external social planner exists in a democracy. Policy choices are made by a political process that need not maximize any social welfare function. In the United States, this process was established in the Constitution, which itself was created by a political process.

Researchers cannot definitively interpret what the framers of the Constitution had in mind when they wrote of the *general Welfare*. Nevertheless, welfare economics can contribute usefully to policy choice in democracies. Well-reasoned policy analysis is infeasible if the *general Welfare* remains an undefined concept. Study of policy choice with specified SWFs enables coherent analysis.

To be useful, I think it essential for welfare economics to move beyond conjecture of homogeneous, self-centered personal preferences. It should adequately express the richness and variety of preferences over social states. After reviewing relevant aspects of the literature on welfare economics in Sections 2 and 3, I distinguished various concepts of equity in Section 4. I then formalized a class of individual utility functions expressing distributional preferences in Section 5.

Section 5 emphasized five themes that hold for utility functions with these types of distributional preferences. These themes may not be novel, but I think that they warrant substantial new attention. In summary: (1) Utility functions need not be self-centered. (2) Policy analysis assuming personal preferences to be self-centered accurately portrays rational individual choice behavior under a specified policy even if individuals have distributional preferences. (3) Intergroup equality is consistent with any degree of intragroup variation in choice sets. (4) It is not generally possible to achieve complete equality of both opportunity and chosen actions if individual preferences are heterogeneous and a policy does not mandate one action. (5) If empowered to be the social planner, an individual with distributional preferences may choose to sacrifice some personal prosperity in order to enhance the population distribution of opportunities.

References

Aronssohn, T. and O. Johansson-Stenman (2023), "Optimal Taxation and Other-Regarding Preferences," https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4606987

Arrow, K. (1951), *Social Choice and Individual Values*, New York: Wiley.

Arrow, K. (1973), "Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice," *The Journal of Philosophy*, 70, 245-263.

Arrow, K. (1978), "Extended Sympathy and the Possibility of Social Choice," *Philosophia*, 7, 223-237.

Atkinson, A. (2011), "The Restoration of Welfare Economics," *American Economic Review Papers and Proceedings*, 101, 157-161.

Atkinson, A. and J. Stiglitz (1980), *Lectures on Public Economics*, Princeton, NJ: Princeton University Press, revised edition, 2015.

Bar Gill, O. (2021), "Willingness to Pay: A Welfarist Reassessment," *Yale Journal on Regulation*," 38, 503-542.

Basu, A. and D. Meltzer (2007), "Value of Information on Preference Heterogeneity and Individualized Care," *Medical Decision Making* 27, 112-27.

Ben Akiva, M., D. McFadden, and K. Train (2019), "Foundations of Stated Preference Elicitation: Consumer Behavior and Choice-based Conjoint Analysis," *Foundations and Trends® in Econometrics*, 10, 1-144.

Bentham, J. (1776), *A Fragment on Government*, Payne: London.

Blaug, M. (2007), "The Fundamental Theorems of Modern Welfare Economics, Historically Contemplated," *History of Political Economy*, 39, 185-207.

Boskin, M. and E. Sheshinski (1978), "Optimal Redistributive Taxation When Individual Welfare Depends Upon Relative Income," *Quarterly Journal of Economics*, 92, 589-601.

Braveman, P. (2006), "Health Disparities and Health Equity: Concepts and Measurement," *Annual Review of Public Health* 27, 167-194.

Briggs, A. (2022), "Healing the Past, Reimagining the Present, Investing in the Future: What Should Be the Role of Race as a Proxy Covariate in Health Economics Informed Health Care Policy?" *Health Economics.* 31, 2115-2119.

Cerdeña, J. M. Plaisime, and J. Tsai (2020), "From Race-Based to Race-Conscious Medicine: How Anti-Racist Uprisings Call Us to Act," *Lancet,* 396, 1125-1128.

Cherrier, B., P. Duarte, and A. Saïdi (2023), "Household Heterogeneity in Macroeconomic Models: A Historical Perspective," *European Economic Review*, 158, 1-16.

Chipman, J. and J. Moore (1978), "The New Welfare Economics 1939-1974," *International Economic Review*, 19, 547-584.

Cookson, R., S. Griffin, S., O. Norheim, and A. Culyer (2021), *Distributional Cost-Effectiveness Analysis*, Oxford: Oxford University Press.

Cooper, D. and J. Kagel (2017), "Other Regarding Preferences," *Handbook of Experimental Economics*, Vol. 2, edited by J. Kagel and A. Roth, 217-289.

Dasgupta, P. (2008), "Discounting Climate Change," *Journal of Risk and Uncertainty,* 37, 141-169.

Delgado, C. M. Baweja, D. Crews *et al.* (2021), "A Unifying approach for GFR Estimation: Recommendations of the NKF-ASN Task Force on Reassessing the Inclusion of Race in Diagnosing Kidney Disease," *Journal of the American Society of Nephrology,* 32, 2994-3015.

Devlin, N. and R. Brooks (2017), "EQ-5D and the EuroQol Group: Past, Present and Future," *Applied Health Economics and Health Policy*, 15, 127-137.

Dolan, P. (1998), "The Measurement of Individual utility and Social Welfare," *Journal of Health Economics*, 17, 39-52.

Domınguez-Rivera, P. and S. Raphael (2015), "The Role of the Cost-of-Crime Literature in Bridging the Gap Between Social Science Research and Policy Making: Potentials and Limitations," *Criminology & Public Policy,* 14, 589-632.

Dominitz, J. (2003), "How Do the Laws of Probability Constrain Legislative and Judicial Efforts to Stop Racial Profiling?" *American Law and Economics Review*, 5, 412-432.

Driver, J. (2022), "The History of Utilitarianism," Stanford Encyclopedia of Philosophy, https://plato.stanford.edu/archives/win2022/entries/utilitarianism-history/, accessed October 23, 2024.

Fehr, E. and G. Charness (2025), "Social Preferences: Fundamental Characteristics and Economic Consequences," *Journal of Economic Literature,* 63, 440-514.

Fleurbaey, M. (2021), "Normative Economics and Economic Justice," *Stanford Encyclopedia of Philosophy*.

Fleurbaey, M. and M. Maniquet (2018), "Optimal Taxation Theory and Principles of Fairness," *Journal of Economic Literature*, 56, 1029-1079.

Gorman, W. (1953), "Community Preference Fields," *Econometrica*, 63-80.

Harsanyi, J. (1955), "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, 63, 309-321.

Hicks, J. (1939), "The Foundations of Welfare Economics," *The Economic Journal*, 49, 696-712.

Hildreth, C. (1953), "Alternative Conditions for Social Orderings," *Econometrica*, 21, 81-94.

Hsieh, N. and H. Andersson (2021), "Incommensurable Values," *Stanford Encyclopedia of Philosophy*.

Kaldor, N. (1939), "Welfare Propositions in Economics and Interpersonal Comparisons of Utility," *The Economic Journal*, 49, 549-552.

Kaplow, L. (2024), "Optimal Income Taxation," *Journal of Economic Literature*, 62, 637-738.

Manski, C. (1999), "Analysis of Choice Expectations in Incomplete Scenarios," *Journal of Risk and Uncertainty*, 19, 49-66.

Manski, C. (2009), "Diversified Treatment under Ambiguity," *International Economic Review*, 50, 1013-1041.

Manski, C. (2011), "Policy Analysis with Incredible Certitude," *The Economic Journal*, 121, F261-F289.

Manski, C. (2013), "Diagnostic Testing and Treatment under Ambiguity: Using Decision Analysis to Inform Clinical Practice," *Proceedings of the National Academy of Sciences*, 110, 2064-2069.

Manski, C. (2015), "Narrow or Broad Cost-Benefit Analysis?" *Criminology and Public Policy*, 14, 647-651.

Manski, C. (2022), "Patient-Centered Appraisal of Race-Free Clinical Risk Assessment," *Health Economics,* 31, 2109-2114.

Manski, C. (2024), *Discourse on Social Planning under Uncertainty*, Cambridge: Cambridge University Press.

Manski, C., J. Mullahy, and A. Venkataramani (2023), "Using Measures of Race to Make Clinical Predictions: Decision Making, Patient Health, and Fairness," *Proceedings of the National Academy of Sciences*, 120, e2303370120, https://www.pnas.org/doi/10.1073/pnas.2303370120

Mirrlees J. (1971), "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies*, 38, 175-208.

Mirrlees, J. (1997), "Incentives and Information: The Economics of Carrots and Sticks," *Economic Journal*, 107, 1311-1329.

National Academies of Sciences, Engineering, and Medicine (2023), *Accelerating Decarbonization in the United States: Technology, Policy, and Societal Dimensions*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25931.

NICE (2013), *Guide to the Methods of Technology Appraisal 2013*, https://www.nice.org.uk/process/pmg9/chapter/the-appraisal-of-the-evidence-and-structured-decision-making, accessed December 15, 2024.

Nordhaus, W. (2007), "A Review of the *Stern Review on the Economics of Climate Change*," *Journal of Economic Literature*, 45, 686-702.

Nordhaus, W. (2019), "Climate Change: The Ultimate Challenge for Economics," *American Economic Review,* 109, 1991-2014.

Phelps, C. and A. Mushlin (1988), "Focusing Technology Assessment Using Medical Decision Theory," *Medical Decision Making* 8, 279-289.

Ramsey, F. (1928), "A Mathematical Theory of Saving," *Economic Journal*, 38, 543-559.

Rawls, J. (1971), *A Theory of Justice*, Cambridge: Harvard University Press.

Robbins, L. (1930), "On the Elasticity of Demand for Income in Terms of Effort," *Economica*, 29, 123-129.

Robinson, K. (2020), *The Ministry for the Future*, New York: Orbit.

Robson, M., O. O'Donnell, and T. Van Oorti (2024), "Aversion to Health Inequality--Pure, Income-related and Income-caused," *Journal of Health Economics*, 94, https://www.sciencedirect.com/science/article/pii/S0167629624000018

Samuelson, P., (1947), *Foundations of Economic Analysis*, Cambridge, MA: Harvard University Press.

Saez, E. and S. Stantcheva (2016), "Generalized Social Marginal Welfare Weights for Optimal Tax Theory," *American Economic Review*, 106, 24–45**.**

Sen, A. (1977), "On Weights and Measures: Informational Constraints in Social Welfare Analysis," *Econometrica*, 45, 1539-1572.

Sen, A. (2009), *The Idea of Justice*, Cambridge, MA: Harvard University Press.

Stantcheva, S. (2021), "Understanding Tax Policy: How Do People Reason?" *Quarterly Journal of Economics*, 136, 2309-2369.

Stern, N. (2007), *The Economics of Climate Change: The Stern Review,* Cambridge: Cambridge University Press.

Vyas, D., L. Eisenstein, and D. Jones (2020), "Hidden in Plain Sight—Reconsidering the Use of Race Correction in Clinical Algorithms," *New England Journal of Medicine,* 383, 874-882.

Weinstein, M., G. Torrance, and A. McGuire (2009), "QALYs: The Basics," *Value in Health*, 12, S5-S9.

Zamir, E. and B. Medina (2008), "Law, morality, and economics: integrating moral constraints with economic analysis of law," *California Law Review*, 96, 323-392.

Zamir, E. and B. Medina (2010), *Law, Economics, and Morality*, Oxford: Oxford University Press.