

## **John Roemer's Contributions to Distributive Justice Theory**

**Dick Arneson**

**For Roemer Conference, London. 2015**

John Roemer over the course of a long and distinguished career has made important contributions to areas of economic theory relevant to public policy choice and the assessment of institutions. In this presentation I focus on his contributions to theories of distributive justice, which seek to develop principles that will tell us when the distribution of benefits and burdens across persons is fair. Roemer's probing discussions of these issues are marked by great technical sophistication and a nuanced mastery of the economist's tool kit and sound instincts as to how these tools are best deployed to advance our understanding of ethical problems. My discussion will be marked by informality and even cartoon-like simplicity. I'm doing a water color sketch with broad brush strokes.

### **Roemer on Rawls and Dworkin.**

Roemer regards John Rawls and Ronald Dworkin as both advancing hypothetical social contract arguments to settle the appropriate content of principles of distributive justice and thereby to settle in what sense these principles should be egalitarian and how egalitarian they should be. So interpreted, both philosophers are following in the footsteps of the economist John Harsanyi, who in a famous essay tried to show that the maximization of the average level of welfare of people could be deduced as what individuals seeking to advance their expected utility would choose in a hypothetical social contract setting in which no one knows what social position he will occupy and must assume he is equally likely to be anybody (so the best he can do to maximize his expected utility is to choose a principle that requires maximizing the average of all persons in society). The philosophers are proceeding on the hunch that the philosophically favored interpretation of this social contract setting will result in the choice of some principle more egalitarian than utilitarianism. Moreover, the philosophers share with Harsanyi the belief that an appropriate interpretation of this hypothetical social contract construction will justify the principles that the imagined persons in the hypothetical decision problem would choose. That rational individuals in the correct interpretation of the hypothetical social contract would choose certain principles shows that we readers of the report of these imagined choices should accept these principles as morally binding.

Surveying the arguments, Roemer argues that when we do the math correctly, and work out the suggested models of rational choice precisely, the project founders. Harsanyi's construction does not exactly yield exactly any version of utilitarianism. This is roughly because utilitarianism is not well defined in the absence of stronger utility information, permitting interpersonal comparisons, than is present in Harsanyi's construction. But when one works out the implications of models close to Harsanyi's, modeling self-interested choice behind a veil of ignorance, one does not get anything close to egalitarian principles emerging from the construction. In a slogan, the assumptions that one is rationally pursuing one's interests via choice of principles of social regulation and is maximally ignorant lead to principles in the neighborhood of utilitarianism, not anything recognizably egalitarian.

Regarding Rawls and Dworkin, the verdict is more astringent. Rawls posits an original position in which the parties to the social contract are deprived of all particular information about themselves and know only such general facts as science and social

science deliver. The parties expect to inhabit a society that will be regulated by the principles chosen, so they can assume they presently exist, but they do not know their particular location in history. They are posited as seeking primary social goods, general purpose resources that will be useful for advancing any of many life plans they might want to pursue. As is well known, Rawls claims that the parties would choose a principle of equal liberties and an egalitarian principle regulating the distribution of social and economic inequalities; for our purposes the relevant bit is that institutions should maximize the social and economic primary goods expectations of the worst off social group.

Roemer in several publications challenges Rawls's philosophical rationale for making the veil of ignorance as thick as he makes it. Roemer also suggests that the decision problem faced by Rawls's imaginary social contractors cannot be solved given his stipulation of the terms of that decision problem. Roemer and Trannoy write: "Indeed, the only precise arguments that Rawls gives for the conclusion that the difference principle would be chosen in the original position . . . essentially state that the decision makers are extremely risk averse." But this is just an arbitrary stipulation on Rawls's part. Nothing in our idea of rational choice constrains the rational chooser to be risk-averse, rather than risk-neutral or risk-seeking. This is just settled for each individual by her actual preferences.

Roemer regards it as an unmitigated disaster for Rawls's hypothetical social contract approach that it begins to select and embrace plausibly egalitarian distributive principles only if you add to his decision problem a further assumption that the parties choosing principles to regulate social life in a society they expect to inhabit are highly risk averse. A risk averse person prefers to have the expected value of any lottery rather than the lottery.

#### **Risk aversion.**

Here I interject a comment. This verdict seems to me too sweeping. I think Rawls is on to something when he supposes that individuals in the individual position should be modeled as conservative in making choices under risk or uncertainty. I admit that Rawls is not clear on this point. He writes that when making a choice that will determine one's fundamental life prospects, determine whether one starts life with an adequate stock of resources to be able effectively to pursue any single life plan drawn from a reasonably capacious set of life plans, it is rational to choose as though one were extremely risk-averse as a psychological matter.

My suggestion is that what Rawls is gesturing toward here is the idea that rationality requires each of us to give extra weight to bad possible outcomes of choice when one's choice is between actions that will affect only oneself and each action one will have various possible outcomes and one can associate with each of these possible outcomes a probability that it will occur given one chooses the action in question and a value for oneself if it does occur. This rationality requirement follows from a natural extension of the prioritarian principle for justice policies with certainty of outcomes to choice of courses of action under risk together with the assumption that each person owes a duty of beneficence to himself for exactly the same reason that he owes a duty of beneficence that is impartial in its treatment of all affected people when one's actions can affect others besides oneself. Priority says that benefits matter more, the worse off are those to whom the benefits accrue. (Of course, when an individual faces a long string of

choices, each involving risk, none carrying risk of instant death and termination of the sequence, maximization of priority-weighted expected well-being coincides with straight maximization of expected well-being.

Priority as stated so far does not specify how much weight one ought to give to the consideration that the individual one might help is badly off, headed for low lifetime well-being absent one's possible intervention. The idea just gives us a family of views, not a specific proposal.

And although I have stated the view as a version of act consequentialism, I don't think that is crucial. Let the egalitarian /prioritarian beneficence principle be qualified by moral constraints and options, if you go in for that sort of thing. The main divide is between moral theories that do and those that don't embrace a strong principle of beneficence, and within the former group, those that do and don't prescribe putting a heavy thumb on the scale in favor of the worse off.

Consider the choice of an individual facing risk, when some possible outcomes of choice would leave the individual worse off by the measure of lifetime well-being. Consider a one person world inhabited just by Robinson Crusoe. If Robinson swims in the ocean he might enjoy himself or might be attacked by a shark and immediately die, whereas if he climbs trees lining the beach, he enjoys himself less but at worst falls onto the sand below without suffering serious injury. In thinking about what to do, Robinson should regard the different lives he might lead in the various scenarios depending on how the risk plays out as though they were lives of different persons, each of whom he should treat with equal concern. He should give greater priority to boosting the well-being of each his possible future selves, the worse off in well-being they would be. He should treat his possible future selves exactly as he would treat different persons whose lives are heading toward different lifetime levels of well-being, who could be helped or hurt by his choice of actions.

**{The rest of this section is an aside. You can skip to the next section.}**

Here's a toy example of a desert-catering prioritarian (DCP) justice ethic guiding policy. DCP says justice requires maximizing a function of aggregate individual well-being that assigns greater weight to gaining a well-being gain for an individual (1) the greater the gain, (2) the worse off in lifetime well-being the person would otherwise be, absent this gain, and (3) the more deserving the person is over the course of her life. In this model there is no production; there is a single all-purpose resource (manna) along with variable personal traits that enable people with varying efficiency to transform the resource into good life. The manna falls unevenly on people. There are two periods; in period 1, people enter a gambling casino and can play games of chance for small or large sums. Gambling is valued not as fun but for the winnings it might bring one. The individuals exit the casino with their resources altered by gambling luck. In period 2, each person consumes her resources. The justice agency has a choice of intervening either before the start of period 1 or period 2; only one intervention is possible. The agency has available some quantity of resources, which it can distribute to individuals so as to advance the cause of justice. If it intervenes before period 1, the agency has enough resources so that it can correct for bad brute luck and perfectly implement an equal opportunity ethic. If it intervenes before period 2, it has available a strictly smaller set of resources, and can transfer them to individuals to advance its conception of social justice. If the equal opportunity norm tells us to achieve equal opportunity at the highest possible

level, and if it is assumed that each person is entirely responsible for his gambling behavior, the solution is to intervene before period 1 with the larger set of resources, passing them out to equalize initial opportunities. Actual gambling behavior is not predictable by the justice agency.

I would suppose that the justice agency that subscribes to Rawlsian maximin of resource shares also chooses to intervene before period 1. It is not determined by the information provided so far what a utilitarian justice agency does; its aim is to maximize aggregate welfare, so it prefers more resources in people's hands and also prefers them in the hands of those who are more effective transformers of resources into personal welfare.

The DCP justice agency, anticipating considerable gambling behavior, does better to intervene at the start of period 2, transferring resources both to achieve better outcomes for those who would otherwise be worse off, after gambling losses, and also to reward those who abstain from gambling, which, being unproductive, is here just bad from a prioritarian perspective, whatever people's risk preferences happen to be. (Here I'm assuming welfare for the prioritarian is getting the items of an objective list not satisfying preferences.) In the ideal case, when each person is known to comply fully with moral requirements, the DCP agency prefers to intervene before period 1, seeking to maximize its social welfare function with the larger resource set. In this ideal case the agency knows that after resources are allocated, in period 1 individuals will just walk through the casino without gambling, and so the DCP social welfare function is realized to a greater degree than would be the case if it had instead intervened just prior to period 2.

#### **Roemer versus Dworkin.**

Roemer's verdict on Dworkin's theory of justice as fair insurance is also negative. He treats Dworkin's deployment of an insurance mechanism in the development of his theory as a version of a hypothetical social contract argument. On this interpretation, Dworkin is identifying egalitarian justice requirements, so far as compensation of individuals for bad circumstances they may face is concerned, with the insurance decisions that a rational person seeking to maximize her expected utility would make, against a background of equal external resources for all, against the possibility that she might be disabled or have low marketable talent, when she is ignorant of the likelihood that these possibilities will obtain.

I'm not sure this way of reading Dworkin is entirely accurate. Dworkin does not hope to derive egalitarianism from weaker normative premises by way of a hypothetical social contract. He starts with the assumption that the government (all of us in a political society acting collectively through the government) owes each of its citizens equal concern and respect. The government massively coerces those under its jurisdiction and claims to speak in their name. These facts together trigger a duty to treat all as equals, with equal concern and respect. This is a duty that does not lie on individuals acting privately, it falls uniquely on governments. The interpretive task that Dworkin sets himself is to develop the best, most defensible conception of what would be required for government to treat all with equal concern and respect. His first interpretive claim is that each individual has a nondelegable responsibility to develop her own views as to what is worthwhile in life and to pursue what she deems worthwhile as she chooses. Given this individual responsibility, the government's rightful task is to sustain a fair distribution of

resources and a fair framework for individual interaction. The question then becomes, what is that?

So Dworkin is not starting from precise weak normative premises and then deducing what they together imply by way of egalitarian justice requirements. He certainly does not have the idea of constraining the pursuit of self-interest by some notion of impartiality to generate egalitarian distribution principles. He starts with a vague egalitarian idea and engages in reflection aimed at choosing among its possible interpretations. --But I don't think any of this matters much, if at all, for the evaluation of Roemer's critique of Dworkin.

Dworkin's interpretive strategy seizes on the idea of fair insurance and holds tight to it for dear life. What is he thinking? In a nutshell, his idea is that when a noncrazy mentally competent person chooses under risk, her preferences appropriately dictate how much weight she should give to avoiding very bad outcomes versus obtaining very good ones. The ethic here is, let each person choose the lotteries she prefers, given she has a reasonable array of choices. The problem is that some important lotteries affecting life prospects play out without any individual having any choice about the matter. One is just born with more or less talent, and with or without disabilities. One is just born with a set of bodily traits that shape one's chances of later suffering the affliction of disease and bodily breakdown. One is just born with kin and family friends who bring it about that one has a better or worse start in life, higher or lower bank account wealth at the onset of adulthood. If people could somehow insure against disfavored outcomes in these primordial lotteries, the results would be ethically ideal. But this cannot be arranged. Hence, justice requires social arrangements that mimic the results that would have obtained had ideal insurance been possible. One must run imaginary insurance schemes and implement their results. Doing this, the government treats each of us with equal concern and respect. Justice is fair insurance.

One might worry that this ideal is incoherent or incomplete. To imagine an insurance market, one must characterize people's insurance decision problem, which requires inter alia that each person has a budget. What sets the initial budget? Also, there is the problem of extrapolating from choice under imaginary ideal circumstances to what is owed to people in actual nonideal circumstances. Perhaps with a fair initial allocation of resources each of us would self-insure and take out no health care insurance at all. It surely doesn't follow that when the initial allocation of resources is unfair, no social provision of health care insurance is what justice requires.

Dworkin has responses to the problems of identifying the budget for hypothetical insurance markets and characterizing the individuals who make purchases in them. One's initial wealth, in an ongoing society, depends on the gifts and inheritance one receives. Here two legitimate interests need to be balanced. One wants to safeguard against having a very poor initial position, with no wealth at the onset of adulthood. One also envisages oneself having acquired property legitimately over the course of one's life and wanting the freedom to expend it on gifts and bequests as one chooses just as one wants the freedom to spend one's money as one likes. Dworkin envisages a kind of hypothetical insurance market that balances these competing interests. This amounts to choosing a tax rate that takes some of what people get by gift and inheritance and uses it to augment the initial bank account wealth of those who have little. The fair initial distribution of wealth, which determines the budget for other hypothetical insurance

markets, is set in this way. The scheme protects us against having a very small resource share as we enter adult life at a cost of depriving us toward the end of our lives, in the event our lives have brought us accumulated wealth, of some of the resources we might have wished to bestow on others by gift or bequest.

This setup still looks incomplete. Dworkin's ideal assumes an ongoing ideal framework of interaction in which in each generation the initial distribution is fair and then people interact via market exchange. Imagining an ongoing Dworkinian ideal society does not settle the initial fair allocation, which is always presupposing some earlier fair initial allocation as required background. I suppose Dworkin could opt for a solution along the lines Robert Nozick once suggested, facing a similar indeterminacy. Since current property ownership proceeds from unfair starting points, let us have a once for all equal distribution of all resources (modified by Dworkin's insurance for talent and disability). From there the scheme would proceed as already described, with the fair ideal insurance (tax rate) for inheritance and gift setting initial fair allocations in each new generation. A further wrinkle is that Dworkin supposes we cannot sensibly imagine tailoring the strongly counterfactual hypothetical insurance schemes we are considering to individual hypothetical preference. We then get the view that justice for the sorts of problems we are considering requires something like social policies that mimic the insurance coverage that the average member of the community would have purchased with a fair share of resource on an insurance market tailored to be morally appropriate for the problem at hand.

Whatever we make of this, we must recognize that justice as fair insurance is not an attempt to organize society so that people's brute luck is equalized and their option luck is left unaltered. We use ordinary insurance to protect ourselves against bad option luck that might befall us, and hypothetical insurance markets would to some extent carry out the same function. Hence when criticisms of luck egalitarianism (understood as some ethic that proposes equalizing brute luck and not tampering with option luck outcomes) are brought against Dworkin's distributive justice ideal his response is that he is not a luck egalitarian. He tells us his favored ideal "is to make circumstances equal under some appropriate version of the envy test," and that the appropriate version of this test is "that people be made equal, so far as this is possible, in their opportunity to insure or provide against bad luck before it has occurred, or, if that is not possible, that people be awarded the compensation it is likely they would have insured to have if they had had that opportunity." To represent Dworkin's final view we should say something slightly different: that people should have fair initial shares and so fair opportunity to insure against bad luck before it occurs or to have the compensation they would have insured to have when the nature of the bad luck in question precludes reliance on actual insurance markets.

Roemer's criticisms of Dworkin take as their target his earlier views put forth in his 1981 articles. I have been urging that Dworkin's views have shifted over time. But I don't see that the shifts matter much given the character of Roemer's criticism. This is that the hypothetical insurance approach fails to deliver an intuitively acceptable egalitarian result. According to Roemer, a condition on acceptable egalitarian principles, or principles purporting to be egalitarian, is that they should recommend equalizing transfers from people who by sheer brute luck good fortune end up with larger resource holdings to people who by sheer bad brute luck end up with smaller resource holdings.

Transfers should go from advantaged to disadvantaged. Different egalitarianisms might be more or less demanding in the amounts of transfers required, but the tilt should be equalizing.

Dworkin's insurance markets do not meet this constraint when his scheme is applied to people who happen to have some possible nonperverse preferences over states of the world in which they end up, for example, blind or sighted. Since money might be able to do more for them if they are sighted, they might, when placed behind a veil of ignorance that bars them from knowledge of their actual disability status, hypothetically purchase insurance coverage that takes money from them if they turn out to be blind and gives them extra money if they turn out to be sighted.

This criticism would seem to apply also to Dworkin's proposal to set gift and inheritance tax rates and couple them to transfers augmenting people's initial holdings. Some people might prefer to face no tax that subtracts from the resources they have available for distribution as they choose, even at a cost of facing more meager initial resource holdings that might have been augmented by an ongoing tax and transfer scheme. They might even prefer a reverse tax that takes from them in the event they start out poor but compensates by giving extra to them later in life if they happen to be by then already rich. In principle one would want to set a policy that treats people in an individualized way, adjusted to their particular preferences, even if this would be in practice unfeasible. But what looks to be theoretically ideal for Dworkin for some cases again violates Roemer's minimal direction of transfer condition on acceptable egalitarianism.

Again, I interject a comment. Any distributive principle that balances the goals of maximizing the aggregate well-being of persons and assigning extra moral value to getting a gain for a person, the worse off she would otherwise be (any welfarist prioritarianism), will in some circumstances recommend reverse transfers of resources. If I am poor, and my holding resource X will yield Y units of well-being for me, if Bill Gates is a sufficiently good transformer of resources into well-being, transferring resources from me to Bill Gates, so he can enjoy an extra yacht, will be recommended by prioritarian justice. So I don't believe we should regard it as in itself morally bad that a moral theory recommends reverse transfers of resources (from disadvantaged to advantaged) in some possible states of the world.

The objection to Dworkin has to be that his theory recommends such reverse transfers in situations where we ought to find this morally unacceptable. Since Dworkin's theory takes resource allocation to be fundamental, his recommendations are opaque to assessment from a welfarist standpoint. I would say this already indicates something is awry; we should always in principle be looking beyond the allocation of resources to see what quality of lives people end up with, deploying those resources. The issue is welfarism versus resourcism.

Also, recall the risk aversion issue, which I claim is central to egalitarian justice. If I might turn out blind or sighted, and will have lower well-being when blind, then I should give extra weight, in making provision for my future, to the possible future in which I turn out to be blind.

**Roemer on equality of opportunity.**

Roemer's criticisms of Rawls and Dworkin are elegant and penetrating. I have raised criticisms, but these are not conversation stoppers. Roemer and Roemerians will have replies.

But Roemer's signal achievements in the area of distributive justice theory do not revolve around criticisms of others' views. Roemer has advanced brilliant ideas on the topic, what is the best understanding of the ideal of equality of opportunity, when the domain of resources is extended to include personal traits as resources. Consider the slogan: it is morally bad, unjust and unfair, if some are worse off than others through no fault or choice of their own. This slogan does not take a stand, on the further question, whether it is to some extent bad if some are worse off than others, independently of how the inequality arises. Nor does the slogan take a stand on the question, whether to some degree it is morally good, just and fair, if some are better off than others through their absence of fault or their positive deservingness. One version of the idea is that to the extent feasible, individuals should be fully compensated for inequalities, when they get the short end of the stick, that are due to their unchosen circumstances, and not compensated at all when the short-end inequalities that attach to them are due to their chosen actions, at least when their choices select from options that include reasonable alternatives. What formulations in this neighborhood amount to, and what plausibility they have, are so far unclear.

The problem of figuring out how one might implement equal opportunity is intertwined with the problem of achieving clarity of understanding about the notion. Offhand the problem might seem intractable. People's unchosen circumstances influence the option sets people face and their abilities to negotiate these option sets and choose effectively to gain good outcomes of interest to a society that aspires to distributive justice. People's choices and chosen actions interact with their circumstances to shape life outcomes. Being poor, for example, influences the options available and the outcomes I can reach. Being poor influences the quality of the education I obtain, which in turn influences my employment options. Being poor influences my preferences and so my choices. What I am responsible for is intertwined with what I am not responsible for.

Roemer makes a marvelously ingenious suggestion. Suppose we could identify all circumstances beyond a person's power to control that influence her choices that will affect how well off she will become. Even if we cannot do this, at any time we will have opinions as to how to make this identification. So we can (in principle) partition people into groups that are identical in these circumstances, or near enough identical. Call these types. Given any valuable objective of interest, such as education, income, or health, let us assume, perhaps not unrealistically, that we can identify a single dimension of effort that individuals can manifest that will increase their level of achievement of the objective. The effort displayed by an individual will reflect both her circumstances and that component of the determiner of effort which we suppose to be within the individual's control.

But by construction we have divided people into types identical in circumstances, so differences in effort among members of a single type reflect solely differences for which individuals can reasonably be held responsible. Roemer identifies the degree of one's effort with one's percentile rank of effort put forward among members of one's type. The degree of effort is the ethically relevant variable, for which it is reasonable to hold individuals responsible. The distribution of effort across types, and how one's effort



level compares to that of all others, is not within one's power to control, but where one sits in the distribution of effort within one's type can be reasonably deemed to be a suitable basis for individual responsibility.

Ideal equality of opportunity in this view would consist in equalizing outcomes for all individuals who are at the same centile of effort in their type. This generally cannot be done. For any given centile of individuals across types, and an assumed fixed budget available to the agency that seeks to improve matters by the standard of equal opportunity justice, we can maximize the achievement of advantage of the individual at that centile across the types. Again, it will generally be unfeasible to maximin in this way for every centile simultaneously. So according to Roemer a morally attractive equal opportunity policy will involve a compromise among the competing objectives of maximin in this way for each and every centile. Roemer has proposed maximizing the average of these maximin objectives.

In the enterprise as described so far, no attempt should be made to equalize the advantage levels of individuals whose degrees of effort are unequal. But this decision is not built into the framework Roemer provides. As he notes, equal opportunity as embraced in what has come to be called the luck egalitarian tradition combines an equalizing and a disequalizing aim. The luck egalitarian aims to reduce or remove inequalities due to circumstances and aims to let stand inequalities due to degree of effort (where "effort" might be variously interpreted in different versions of luck egalitarianism). These are distinct and independent aims. One could aim to prosecute either one without the other. Roemer eventually declares himself uncertain about how gung-ho one should be in pursuing the disequalizing aim. He provides a mathematical formulation of this agnostic stance in Roemer 2012 and Roemer and Trannoy 2013 (and elsewhere, for all I know). This recent statement is a generalization of the approach described in Roemer (1993, 1996, and 1998).

Roemer also countenances the thought that in different types of policy applications different values competing with equal opportunity might come into play and different versions of "equal opportunity" policies ought to be pursued. The branches of the ministry of justice charged with achieving advantages for people along the dimensions of education, health, work and income, and sexual fulfillment can appropriately follow very different algorithms in the name of equal opportunity broadly conceived.

Moreover, despite the metaphor of the "ministry of justice" in the previous paragraph, so far as I can see nothing in Roemer's abstract account requires the existence of a state, or only a state, as the agency responsible for fulfilling egalitarian justice aims. The equality of opportunity procedure that Roemer describes could in theory be carried out by conscientious individuals acting in concert under anarchy. (Each of us could put money into a pot, which would then provide the budget for volunteers who work to use the budget to advance equal opportunity.) Only a modest degree of political realism requires implicit reference to the modern bureaucratic state.

Speaking just for myself, I regard the framework for egalitarian policy analysis that Roemer has provided, and which I have only crudely characterized in the preceding paragraphs, enormously clarifying. Reading Roemer, one acquires a clear sense of what socialist equal opportunity principles demand, what values they presuppose, what tradeoffs and ambiguities and possible objections they require, involve, and invite. (I

borrow the useful phrase “socialist equality of opportunity” from G. A. Cohen 2009). The understanding thus acquired is understanding of moral principles and social justice, not merely some sort of applied engineering technique. Having been guilty of wallowing in vague formulations, I am grateful to Roemer (and to other philosophically minded economists and other social scientists) for using precise notions to pull out the plug of my bathtub. (Sorry, I am letting metaphor run amuck here. You get the point.)

Roemer tends to present his ideas on equality of opportunity as applications of distributive justice theories invented by philosophers. He presents himself sometimes as merely suggesting a mechanism for implementing the philosophers’ ideas. These modest descriptions wildly undervalue the creativity and interest of his proposals. These exemplify the philosophical value of the mathematical tools in the economists’ tool kit. When Roemer works to make more precise some vague philosophical formulation or slogan, we see more clearly what is really being proposed, or more usually, the several different proposals wrapped together and not distinguished in the vague formulations. We also see more clearly the content of the various ideas in play, and also their limits and the possible criticisms to which they are vulnerable. A slogan formulation can encompass claims that when made precise are revealed to be inconsistent, or compatible only under special conditions, or competing goals that cannot be maximized simultaneously, but instead call for tradeoffs. All of these possibilities turn up in the episodes that constitute the recent history of egalitarianism. Taking a warm bath in vague verbal formulations, one can fail to notice distinctions and implications that should be obvious.

In passing, we should note that the recent history of egalitarian justice thinking reveals quite a bit of back and forth between economists and philosophers. The sequence is not just Harsanyi to Rawls to Nozick to Dworkin to Cohen. Dworkin relies on the idea of envy-free distributions that emerges in the work of Duncan Foley and Hal Varian and others. Nozick is inspired by Hayek. The idea that economic exchange among private owners of resources might play a role in egalitarian justice theorizing can be traced back to discussions of market socialism among economists in the 1930s and perhaps further traced back to 19<sup>th</sup>-century writers including J. S. Mill’s *Principles of Political Economy*.

#### **Comment, criticism, and doubt.**

Roemer wavers in his statements as to what he intends to be offering us in developing the equal opportunity procedure. His official formulation is that he is providing a tool for social planners, so they can implement whatever opinions they happen to have on the question, for what features of their situations individuals should be held responsible, where this determination is to shape the compensation that is owed them. His model does not specify what is to count as an advantage for individuals, and does not tell us for what features of their situation individuals should be held responsible.

But clearly the ethical interest of carrying out the Roemer program depends on society’s having chosen reasonable rather than unreasonable interpretations of these matters. Roemer’s ethic says that people should not be held responsible (when what is at stake is the intrinsic rather than instrumental moral importance of deservingness) for that for which they really cannot reasonably be held responsible. The social planner’s opinions should correspond to what is objectively ethically correct. The same goes for determination of what counts as an advantage.

Roemer says he seeks a theory that will be “useful for policy makers. The Ministry of Health is interested in maximizing some distribution of life-expectancy or

quality-years of life, not subjective utility (e.g., it will try to discourage smoking, even though individuals may derive subjective pleasure from smoking); the Ministry of Education is interested in producing students who are well-educated, not happy students; and international lending institutions are interested in maximizing GDP per capital in countries, not the subjective utility of the country's politicians" (2012). But of course we should not care about the preferences of social planners and bureaucrats working in ministries of health and education except insofar as their preferences are good proxies for what will really advance the genuine well-being of the individuals who will be affected by their social policy choices. I don't say Roemer disagrees with the claim just made. I am interpreting his view here, not venturing criticism. (If Roemer is unwilling to identify the well-being of a person with the degree to which her subjective preferences are satisfied, I concur.)

My conjecture is that Roemer's agnostic statements reflect his underlying belief that what is ethically most important is reducing inequality between persons when the causes of that inequality are circumstances beyond their power to control. Even if society does not correctly identify all such circumstances, compensating for the ones it does correctly identify will be a step in the right direction.

In fact we don't have to guess at Roemer's underlying convictions. He tells us them. In early formulations of his equality of opportunity procedure, he states as the egalitarian goal the elimination of inequality due to circumstances and the toleration of inequalities due to degree of effort (insofar as it is feasible to carry out both goals together). In later formulations he gives decisive priority to the first goal. He declares, "I am militant on annihilating inequality due to circumstances; I am uncertain about how much inequality to allow due to differential effort" (Roemer 2012). One issue here is that we lack a theory of how much reward effort should receive. (This is certainly correct. It corresponds to the observation that in the prioritarian framework, if we modify priority for the worse off by adding priority for the more deserving, we have no good theory as to how much weight should be assigned to desert priority as compared to low well-being priority.)

Another issue is that we must not allow our norm of reward for effort to swamp or overpower our norm of leveling inequalities, which is not entirely reducible to a concern for reducing inequalities brought about by unchosen and uncourted circumstance. Following a suggestion made by G. A. Cohen, Roemer suggests that we should close the equal opportunity model by adding a principle that requires some degree of inequality aversion that is not conditioned on the genesis of the inequality in circumstance. Besides what Cohen calls socialist equality of opportunity we should embrace a norm of solidarity that instructs us to reduce the gap between worst off and best off. Maybe this solidarity norm can be interpreted so that it both puts a brake on the degree to which degree of effort can justify extra rewards for effort and sets an upper bound on allowable inequality no matter how the inequality is brought about.

Here I interject some comments. First, one might suppose that our distributive justice ethic ought to countenance some inequalities just on the ground that they would boost aggregate well-being, as any prioritarian ethic (less extreme than maximin) will do. Roemer does not disagree. His equality of opportunity ethic is flexible and moderate, in the sense that it is one principle among several to be included in our overall social justice doctrine. According to Roemer we might, and should, abate our wholehearted pursuit of

equality of opportunity when doing so would exact excessive costs in terms of privacy or other social values. On this ground he holds that a sensible health care ethic will not dictate policies that require health care providers to vary their treatment of potential patients according to their opinions of the deservingness status of these clients; we want to ensure the integrity of the doctor-patient relationship even though doing so reduces aggregate social welfare. Presumably he would say the same when equality of opportunity correctly construed can only be purchased at too high a cost in human well-being. Adding a Pareto proviso is not sufficient for this purpose. I myself tend to prefer a framework that puts the tradeoff between maximization of aggregate well-being and its fair distribution across persons front and center.

Once we countenance plural justice values, is our situation then hopeless? What justice requires depends on tradeoffs among component justice values and we lack an acceptable theory about how to assign weights to yield correct tradeoffs. I would say, not necessarily hopeless. We need to embark on an extended Frances Kamm type investigation. We should examine a wide variety of types of situations in which the components of our justice ethics will conflict, and seek to determine what intuitively seems plausible in each example and also seek to get a handle on what is driving our intuitions and whether they are being influenced by improper bias of some sort and, if so, purge our intuitive judgments of these improper influences. In carrying out this investigation we seek a Rawlsian ideally wide reflective equilibrium state, in which our general claims and our particular judgments fit together coherently, giving extra weight to our judgments about possible worlds close to the actual world. In carrying out this investigation we need to use our empirical knowledge about the biases and heuristics that affect judgment and choice, as a corrective and guide. Whatever the likely success of this enterprise, it is not any sort of defect in Roemer's social justice ethics that it leaves us with a problem of assigning weights to competing values. Any pluralistic theory will have to face this problem, and avoidance of the problem by adopting an incorrectly monistic theory is no solution.

Another comment: in trying to decide how much weight to give to deservingness considerations in our ethic, we need eventually to address the problem, what is the underlying basis of desert. If we envisage a series of exercises on the part of the egalitarian social planner, each one taking for granted some measure of "effort" such as labor force attachment or adherence to healthy lifestyle, we need to see these measures as proxy for some underlying X that is genuine desert. In my view this underlying desert basis is subjective, and is something like making conscientious efforts to orient one's will toward what (one conscientiously takes to be) is morally right and good. More on this below.

Another comment: It turns out that Roemer's equal opportunity ethic does not rebel against the influence of all unchosen circumstances in shaping people's fates. Not all determinants of fate are created equal.

Roemer distinguishes between two types of formative causal influences on people's life prospects: "the formation of preferences and aspirations in children" and the rest (Roemer and Trannoy 2013). According to Roemer, equality of opportunity ethically construed should seek to dampen or eliminate the impact of influences other than transmission of cultural values and aspirations to children by parents and other significant others. Religious values form one example of the circumstances that according to

Roemer should get a free pass. Roemer and Trannyoy: “If a coal miner loves coal-mining and instills in his child the desire to become a miner, this is a legitimate influence that does not call for compensation.” In particular, parental transmission of values that reduces the potential income of their children is not in itself grounds for compensation. (Roemer distinguishes sour grapes type mechanisms from pure parental influence. We should let stand only the latter, he holds.) One thought here is perhaps that if this influence called for compensation, that might in turn lead to social policy efforts to interfere with people’s freedom to raise their children, within broad limits, as they see fit, and Roemer might see this restriction of liberty as onerous.

But why? From an egalitarian welfarist perspective, Roemer’s examples are underdescribed. The mere fact that social environment influences might shape an individual’s preferences in ways that do not maximize his potential income is not in itself problematic. Consider Buddhist parents who teach children values that increase the children’s capacity for emotional self-regulation and also inspire a laid-back attitude that reduces potential income—the net effect on the individual’s well-being by any reasonable conception might be positive. But in other case, where parental and other social influences reduce children’s potential for achieving what is genuinely worthwhile, the socialization is regrettable and in principle should be compensated or ameliorated. Seeking social justice, we seek to bring about fair shares of genuinely good quality life for individuals, not shares of what some parents or guardians (perhaps innocently, perhaps culpably) mistakenly take to be genuinely good quality life.

At this point Roemer objects that if we believe parental socialization comes within the scope of equal opportunity justice, “then it is difficult to justify why we do not advocate raising children collectively” (Roemer and Trannyoy 2013) This is a leap. Tagging a very small number of particular persons with special responsibility to rear each child that is born may be generally speaking a very effective strategy for raising children well, and allowing biological parents standardly to play this guardian role may also be a generally very effective device to achieve childrearing aims any sensible ethic will espouse. But a presumption for parental prerogatives still allows lots of room for possible improving social policy and cultural norm reform.

Here the issue of resources-versus-welfare as the measure of individual’s condition for social justice purposes, along with the issue of whether to rely on objective or subjective conceptions of welfare, comes to the fore. I do not believe an adequate distributive justice ethic can remain neutral or agnostic on the proper distribuendum.

But this criticism, be it sound or unsound, does not come close to the heart of Roemer’s equal opportunity proposal. For all that has been said, the core of the Roemer equal opportunity procedure might be entirely defensible.

#### **Roemer’s Equality of Opportunity norm: a criticism.**

Although Roemer’s equality of opportunity procedure is brilliantly ingenious, it may be incorrect. More circumspectly, we perhaps should find his view implausible. To illustrate a possible difficulty, suppose there is wide dispersion in actual effort levels observed in some type. Suppose, for example, that in a type characterized by the features of being black females raised in impoverished circumstances (hereafter, BFRICs), a substantial fraction of type members exhibit very high actual effort levels. We might wonder if it is fair to regard these individuals at some particular centile level of their type as putting forth the same degree of effort, in any ethically interesting sense, as the

individuals in other types at the same centile of the distribution of actual effort of their type.

A possible conjecture is that society's decision as to what features of individuals' situations to count as circumstances beyond their power to control in this case is shown to be too crude by the anomalous relationship between type and actual effort expended that we have discovered in this case. But imagine that we have done very careful science and are very sure (not absolutely certain, of course) that the features we are calling circumstances do not just tend to correlate with actual effort expended but are causal forces that impede an individual from producing high effort. We can imagine our example occurs in the 23<sup>rd</sup> century and the science of human behavior has considerably progressed. Moreover, we are very sure that these features we are calling circumstances are indeed beyond the power of any individual to control. Furthermore, we are very sure that our catalog of features that we are classifying as circumstances is complete. No actual circumstance that impedes or promotes the production of actual effort by individuals is missing from the list of circumstances employed to define types in the population. So we are perplexed.

Now imagine that the range of actual effort levels displayed by the individuals comprising some disadvantaged type lies entirely above the range of actual effort levels displayed by individuals in advantaged types. In our example, the actual effort levels attained by BFRICs lie entirely above the actual effort levels displayed by, for example, white and Asian males raised in affluent households. We know that the circumstances of the BFRICS causally drag down their production of effort by comparison with the circumstances enjoyed by advantaged types. Yet the BFRICS at every centile of their effort distribution display very high effort.

The Roemer equality of opportunity procedure requires us to say that when individuals are completely segregated by type, so that all circumstances beyond anyone's power to control and exerting a causal influence on effort are filtered out, by definition the differences in effort expended by individuals within each type reflect differences in their autonomous choices of effort and so are appropriately are rewarded, with those within a type expending greater effort meriting greater reward. Each person is appropriately held accountable, to some extent, depending on how much reward we should assign to effort, for where she sits in the distribution of effort of her type. But the actual effort advanced varies with circumstances and so varies with type, and is not an appropriate trigger for deserved higher reward. The actual effort advanced by an individual is brought about by a promiscuous mix of circumstances and whatever it is that we should be seeking to reward as what makes an individual genuinely deserving. On this view, trying to bring about anything like equal reward for equal actual effort advanced is morally inappropriate.

However, our hypothetical example puts pressure on the assumptions underlying the equality of opportunity procedure. In particular, how can it be that when we have perfectly divided people into types so that the members of one type unambiguously are those who face causal hindrances to putting forth actual prosocial effort that members of another type do not face, we can ignore as morally irrelevant to just compensation differences in the aggregate effort displayed by types when it turns out that members of some disadvantaged types entirely outpace the members of advantaged types in actual effort put forward? In particular, our example puts pressure on the "assumption of

charity” that Roemer makes. This says that within any type, the distribution of the propensity to expend effort across individuals would be the same, were we able to factor out all the different circumstances which define different types. Maybe that is true. For all I know it might be true. But this is ultimately an empirical assumption which might prove false, and our hypothetical example stipulates conditions that provide considerable evidence that in the world of that example, the assumption is false.

Notice that it isn’t strictly true that the distribution of effort across types is beyond the power to control of any individual in the type. If I put forth more effort, the aggregate effort advanced by all members of my type shifts just a bit, and so the distribution of effort across types also shifts slightly. All members of a type might do the same.

The problem we are confronting is exhibited sharply if we imagine a different type of effort than the ones Roemer considers. In the Roemer equality of opportunity procedure, expenditure of effort by definition increases one’s attainment of some type of individual advantage, some component of individual well-being as we might think of it. For example, the objective might be good health through avoiding lung cancer and the effort for this example might be refraining from smoking. The fewer cigarettes one smokes, the greater one’s chances of avoiding becoming afflicted with lung cancer to one’s disadvantage.

But suppose we widen the pairing of effort and advantage. Consider the effort of putting forth sensible altruism—sacrifices of one’s own well-being for the sake of others when doing that achieves a better distribution of human well-being than would otherwise have occurred. We observe the actual level of sensible altruism that each person displays. Having divided the society accurately into types, we have successfully divided people into groups in such a way that we think we can reasonably hold people accountable for the degree of effort each expends, which is indicated by one’s centile rank in the distribution of sensible altruism expended by one’s type.

But again we (imagine we) find that the aggregate actual effort levels expended by some types are anomalous. It could be that circumstances beyond one’s power to control causally impede one’s attainment of some sorts of social objective and causally assists one’s attainments of other sorts. Circumstances that make it harder to devote years to one’s education might make it easier to display sensible altruism. But suppose again that we have investigated the causal effects of circumstances and in the cases we are considering they happen to be uniform. People whose circumstances impede their attainment of greater years of schooling are also impeded, and to a similar degree, in their manifesting sensible altruism. Yet BFRICs face less auspicious circumstances than others face and also exhibit greater actual levels of sensible altruism. We are even more perplexed.

Our perplexity is eased if we conjecture that what is going on is that BRFICs just behave more virtuously than others and are more virtuous. They are on the whole and on the average more morally deserving than members of other types, and relevantly for our purposes, more deserving than members of more advantaged types for whom producing actual effort is easier and hence not a good measure of true desert. In the hypothetical situation postulated, the Roemer equality of opportunity procedure will make judgments about appropriate compensation for unequal circumstances that are unfair and in particular skewed against the BFRICs, who are given insufficient credit for their virtue by the Roemer algorithm. The Roemer equality of opportunity procedure does not generally

successfully clarify and refine and implement the broad luck egalitarian principle that everyone should enjoy equal opportunity for advantages at the highest possible level. The idea is that if one behaves as meritoriously as it is (normatively not predictively) reasonable to expect someone to behave, given her circumstances, then one should not have worse prospects of gaining advantages than anyone else. This norm is not satisfied by the Roemer procedure in the hypothetical situation envisaged. But we don't know to what extent the hypothetical situation or some watered-down version of it actually obtains in our world.

What I have said about the sensible altruism extension of Roemer's procedure applies also to the applications he considers, where the objective is some kind of advantage for each individual and his advantage level increases with effort expended. Another way to put my point is that Roemer's ideas of "actual effort" and degrees of effort" cannot fill the roles assigned to them, in the general case. The effort that plausibly renders one genuinely deserving in such a way that one merits greater reward (intrinsically, not merely for instrumental carrot and stick reasons) is not the observable one dimensional measure of a type of behavior that helps an individual gain advantage differently understood in different policy domains.

It is instructive to compare the criticism being advanced here with a superficially similar criticism of Roemer's equal opportunity procedure once advanced by Brian Barry. Barry in effect proposes that what is morally relevant is not degree of effort expended but (at least sometimes) actual effort expended. His example is Asian youth, raised in families that value high achievement, and who are impelled by socialization to work hard at achievement. Their effort levels are far higher than those of suburban white American youth, raised in households that prize consumerism and gratification. The circumstances of socialization are beyond the individual's power to control, but Barry suggests that the Asian youth in the example should be rewarded for the painful effort they embrace.

I am not suggesting people should be rewarded for putting forth extra efforts whenever doing so is caused by circumstances beyond their power to control. I am suggesting that people should be rewarded for putting forth prosocial extra efforts when doing so is made (if anything) more, not less, difficult and painful by the ensemble of circumstances that are causal forces shaping their behavior, compared to others. Barry imagines a case in which those who put forward higher effort are being pulled by causal forces; my case is one where those who put forward higher effort are swimming against the causal forces impinging on them (or at least, against all the causal forces that a compatibilist theory tells us are responsibility-diminishing).

If the difficulty this discussion has identified holds up, then Ronald Dworkin has a *tu quoque* against Roemer's objections leveled against Dworkin's conception of justice as fair insurance. Recall that Roemer's objection was that in some circumstances Dworkin's egalitarian principles would implausibly imply that in the name of equality we should bring about reverse transfers of resources from disadvantaged to advantaged individuals. For example, fair insurance might require compensatory transfers from blind to sighted people. We are expected to conclude that fair insurance is an inadequate conception of justice, and I concur.

The objection against Roemer's equal opportunity procedure is that in some conditions it undervalues the deservingness of disadvantaged individuals and disadvantaged groups who rise above their circumstances and put forth high actual levels



of virtuous effort that indicate that their degree of effort and hence their true deservingness is very high. (Of course, my objection is symmetrical; Roemer's procedure may in some conditions misidentify the deserving rich rather than the deserving poor. This does not affect the point being made.) The moral significance of this failing depends on the weight we should give to compensation for true deservingness in the overall distributive justice norm. But so long as we give any weight at all to rewarding desert, even a small weight, Roemer's procedure does not capture the equal opportunity ethic it is seeking to represent in precise and operationalizable form.

The practical significance of the difficulty I have been trying to identify may be moot, for several reasons. One is that the claim is that in possible empirical states of the world, Roemer's norm gives implausible verdicts, and these possible states of the world might not in fact ever obtain. So one might hold that my issue is a nonissue. Another is that the notion of conscientious effort on which I am implicitly relying may be purely notional. By that I mean that whereas Roemer is careful to interpret the notion of effort so that in every application of the equal opportunity norm he considers, effort is one-dimensional and empirically observable, I am drifting toward interpreting effort as an indicator of the true state of one's soul, the orientation of one's will toward what one believes to be the right and the good, assessed by adjusting somehow for everything that makes it especially difficult and unpleasant for one to carry off this feat. Empirical indicators of effort are always to be viewed as rough proxies for this underlying orientation of the will. But how does one decide one has a good proxy measure for something one just has no clear idea how to measure even in principle? One is reminded of Jonathan Wolff's quip that measures of deservingness are either so fine-grained that it is unimaginable they could ever be applied or so coarse-grained that it is inconceivable that they could have any moral import. There is no viable middle ground here. (I believe this conclusion would be unduly skeptical, but this is a big and to my mind unresolved issue.)

In passing, note that Rawls's skepticism about desert, regarded as an element in a practical ethics to guide choice of social policy and individual action, might turn out to be prescient. Rawls wrote, "The precept which seems intuitively to come closest to rewarding moral desert is that of distribution according to effort, or perhaps better, conscientious effort. Once again, however, it seems clear that the effort a person is willing to make is influenced by his natural abilities and skills and the alternatives open to him. The better endowed are more likely, other things equal, to strive conscientiously, and there seems no way to discount for their greater good fortune. The idea of rewarding moral desert is impracticable." Notice, Rawls does not say there is no such thing as being truly deserving and undeserving. He says that for pragmatic reasons, social justice must ignore these matters.

It might be objected that it is wrong to regard Roemer's assumption of charity as empirically falsifiable. It is either a moral or a metaphysical claim. In either case, if we embrace it, no empirical states of the world could arise that would force us to give it up. We could always conjecture that contrary to our initial beliefs, the odd observations of actual observed effort levels across types indicate that our classification of individuals into types is inadequate.

This is a possible response, but one we should reject. As our science of human behavior improves, our confidence in our ability to classify individuals into types

increases, and then (if we should observe the possible phenomenon I have described), we are pushed toward finding it empirically unlikely that the BFRICs in our example are anything but more deep-down virtuous than others.

Another possible response is that to imagine that the science of human behavior and its causes improves is to imagine that we must be becoming ever more confident that all human choices are caused events and that some version of hard determinism must be correct. In that scenario, deservingness is never exemplified, and the part of our distributive justice ethics that prescribes responsiveness to differential deservingness becomes a dead letter. If you will, in this scenario we must conclude that everyone's deservingness score is the same, and not any sort of basis for differential reward. On this view our notion of human agency and human responsibility and accountability applies to the residual of human behavior whose causes are not adequately comprehended. When the causes of all human behaviors are in principle comprehended, our need for a theory of deservingness is nil. On this view the problem I have identified for Roemer never arises.

Here we are on the edge of deep waters, difficult issues. But it seems to me that in principle the advance of scientific understanding could leave room for some compatibilist understanding of human choices as both caused events and exercises of responsible agency. An alternative is that free will libertarianism will prove correct. This is a coherent if unlikely possibility, though if we try to characterize it in advance of the science that could make sense of it we will fall into incoherence (this would be akin to trying to describe non-Euclidean geometry before we have discovered it).

#### **Conclusion.**

So in my view the criticism of Roemer's equal opportunity procedure stands. But I have no idea how best to replace it. Roemer's contributions to social justice theory combine precision and insight in a welcome mixture. He has enriched our understanding of egalitarianism and—to mention a theme these comments have ignored—of the relationship between egalitarian theories of justice and traditional socialist ideals. For this we should all be grateful.