

April 24, 2009

“Responsibility in egalitarian theory”

by

John E. Roemer

Yale University

1. Introduction

The name of this conference, “Equality and responsibility in the social contract,” suggests two possible topics that I might address in these remarks. The first is how responsibility entered egalitarian theory in the last forty years, which was chiefly through the writings of John Rawls and Ronald Dworkin, who both relied deeply on contractarian arguments to advance their claims of what justice, or of what equality, consists in. The second topic is less high-minded: it is how we can justify egalitarian policy measures – more generally, what *degree* of equality we can justify as a goal of social policy -- given our actually existing social contracts in contemporary advanced democracies. I will attempt to address both of these topics in these remarks.

It may surprise you that I will argue that the contractarian arguments of both Rawls and Dworkin fail to justify the kind of equality they advocated. Indeed, the original position of Rawls and the hypothetical insurance market of Dworkin, which is his contractarian venue, both produce results that are undesirably unequal, if they are modeled correctly. This might seem to present a crisis for egalitarian theory, as so much of our present theory apparently relies on the work of Rawls and Dworkin. The crisis is this: if we find either Rawls’s original position or Dworkin’s hypothetical insurance market an attractive thought experiment, then must we not live with inegalitarian recommendations that I claim they produce? I will argue that we need not, because of one of their central premises: that individuals, or those denizens who represent them behind the veil of ignorance, are entirely self-interested. It is *this* assumption, I claim, which must be jettisoned, and replaced with a conception of solidarity.

2. Rawls’s original position

It is a dangerous undertaking to attempt to summarize Rawls’s argument in a few pages or minutes: there will be numerous objections that the various subtle qualifications

which Rawls gives through *Theory of Justice* have been ignored. Nevertheless, I think it is possible to do so. Some may object that I am forcing Rawls's complex thought into the economist's Procrustean bed. I make no attempt to deal, here, with the entire Rawlsian edifice, but only with the theory of economic distribution and how it is argued for.

For Rawls, people have life plans, which in economic theory would be called preferences. All life plans need certain common inputs, which Rawls calls primary goods. Rawls assumes that an index of primary goods can be created, which can be used to aggregate any bundle of primary goods into a single number: that is, certain weights are attached to each good. Once we have such an index, we can say that one bundle of primary goods, x , is greater than another, y , when the weighted sum of the goods in x is greater than the weighted sum of the goods in y , using the indexing system as weights. The important assumption is that *every life plan is enhanced by having a greater amount of primary goods*, so interpreted. It is a heroic assumption that such an index can be constructed, but I will not dwell on this point. Indeed, if such an index exists, then an economist would say that every person's life plan can be *represented* by a utility function that is just the indexed sum of the primary goods that he consumes or enjoys.

Rawls views the many circumstances of a person's situation as morally arbitrary: it is a matter of luck what the birth lottery gives you, in the way of the family to which you are assigned, the talents and handicaps that you have, and the income-earning capacity that these circumstances will eventually engender. In the original position, there is a soul who represents each person, and these souls are to agree on a distributive rule for primary goods. They do not know the plans of life (preferences) of the persons whom they represent, but they do know that every person wants more primary goods rather than fewer – as aggregated by the index. Not only do they not know the life plan of their person, but they do not know the actual distribution of life plans in the actual world. The souls in the original position are assumed to be interested only in the primary goods their principals receive: they are self-interested.

The veil of ignorance which filters out all these facts about the world is constructed in order to shield the souls from knowledge of facts which describe *morally arbitrary* events. The morally arbitrary events in question are the actual circumstances and endowments of individual persons. But why shield the souls from knowledge of the

life plans of their persons? Rawls never argues that a person's life plan is morally arbitrary: in fact, his view is quite the opposite, that persons are responsible for their plans of life. Rawls writes, in 1982, "The third respect in which citizens are regarded as free is that they are regarded as capable of taking responsibility for their ends and this affects how their various claims are assessed...citizens are thought capable of adjusting their aims and aspirations in the light of what they can reasonably expect to provide for." But if persons are responsible for their 'ends,' which I take to be their plans of life, why should they not know these life plans in the original position? After all, the original position is meant only to obscure knowledge of morally arbitrary features. And why should they not know, in addition, the distribution of life plans among the whole society? Is this not a fact about the world, a fact that is surely not morally arbitrary?

Rawls has, in fact, constructed the original position so that all souls are identical: they each know only that they wish to maximize the index of primary goods of their person, and they know nothing about probabilities or preferences. So what goes on behind the veil of ignorance cannot be described as a contract – it is in fact a decision problem of a single agent. (In fact Rawls recognizes this; he writes "...it is clear that since the differences among the parties are unknown to them, and everyone is equally rational and similarly situated, each is convinced by the same arguments. Therefore, we can view the agreement in the original position from the standpoint of one person selected at random. (TJ, 120)") The agent simply wants to maximize the primary goods going to his person. I must emphasize that there is no altruism here: Rawls assumes that each soul represents only the interests of its person.

Not knowing any probabilities or even the preferences over risk that their persons have, there is really no accepted way of solving this decision problem. The veil of ignorance is just too thick to use decision theory. Nevertheless, Rawls gives three arguments for why the soul will choose the difference principle, which is the distributive rule that chooses that distribution of primary goods which awards the largest index of primary goods to the person who receives the smallest such index – hence, the terminology of maximin. The obvious way to justify this choice is that the soul is hyper-risk averse. It is possible that the person whom it represents will be a member of

that worst-off class, and hence a hyper-risk averse soul would choose to make that class as well off as possible.

Although Rawls equivocates, and sometimes claims that he is *not* relying on hyper risk aversion, it is very clear that, indeed, he is. He writes that choosing the maximin distribution would not *generally* be a good choice when facing risk, but that there are three features of the original position which make it so. The first feature is that the probabilities of having various life plans and circumstances are unknown, so there are no probabilistic computations one can make. This begs the question: as I have said, the probabilities *should* be known, for they are facts about the world, not morally arbitrary. The second feature is that “the person choosing has a conception of the good such that he cares very little, if anything, for what he might gain above the minimum stipend that he can, in fact, be sure of by following the maximin rule. It is not worthwhile for him to take a chance for the sake of a further advantage, especially when it may turn out that he loses much that is important to him.” This is, purely and simply, a statement of extreme risk aversion. Is this a general description of human psychology? Surely not. The third feature is that “the rejected alternatives [i.e., non-maximin distributions] have outcomes that one can hardly accept. The situation involves grave risks.” But this, too, is simply a statement of extreme risk aversion. Why can ‘one hardly accept’ an alternative to maximin? Only because one is extremely risk averse, and pays attention only to the event that one ends up in the worst-off class. These arguments all appear on p. 134 of TJ (1999 edition).

Some political philosophers respond to these criticisms of Rawls’s original-position formulation by saying that, after all, the construction was not important for his view, that it was just an example of how the view might be justified. This an incredible claim. Rawls’s formulation that his theory is ‘justice is fairness’ relies entirely on the original position: Rawlsian justice is an instance of fairness *precisely* because it is a contractarian outcome in an original position which shields persons from knowledge of morally arbitrary facts. (Rawls writes, “The OP is, one might say, the appropriate initial status quo, and thus the fundamental agreements reached in it are fair. This explains the propriety of the name ‘justice as fairness’: it conveys the idea that the principles of justice are agreed to in an initial situation that is fair. (TJ, 11)”) If one excises the

original position from *Theory of Justice*, one loses as well that claim that justice is fairness.

If one accepts these criticisms, then a natural move would be to reconstruct the original position, endowing the souls with knowledge of the life plans of their persons, of the distribution of life plans, and of the distribution of circumstances, although *not* the knowledge of their own person's circumstances, which are morally arbitrary. This is essentially the construction that John Harsanyi made in 1953, and he *thought* that he had deduced that the social contract would be utilitarian – that is, that it would distribute wealth and other resources to maximize total welfare in society. But he misinterpreted his result: what he in fact showed is that, under assumptions that I think are reasonable, the social contract would call for maximizing some weighted sum of the von Neumann Morgenstern utilities of persons. It would take me too far afield to explain why this is not equivalent to utilitarianism; for this, consult Roemer (1996, chapter 4). Recently, Juan Moreno-Ternero (2008) and I have offered a way to complete Harsanyi's argument. The quite unattractive result is that, unless persons are very risk averse, the original position can recommend distributions of wealth which are *perverse*, in the sense of giving more wealth to talented people than to untalented ones. So the original position, amended à la Harsanyi, and further refined according to Moreno-Ternero and me, produces recommendations that any egalitarian would condemn, *unless* persons are very risk averse. I do not believe there is a valid argument for claiming persons actually possess the degrees of risk aversion required to avoid this result. Furthermore, if the original position, so constructed to filter out information about morally arbitrary events, produces this outcome when individuals are *not* extremely risk averse, is that result not fair according to Rawls, in that situation? The egalitarian would seem to have to choose between fairness and equality.

3. Dworkin's hypothetical insurance market

I have argued that Rawls inserted the issue of responsibility into his theory in only a vague and inconsistent way. He argued that persons are responsible for their life plans, but this was not reflected in his contractarian venue, where souls do not know the life plans of the persons they represent. Amartya Sen offered a critique of Rawls in his 1980

lecture, "Equality of what?" but it did not focus upon this problem, but rather on Rawls's choice of the index of primary goods as the maximand for souls, rather than upon an index of functionings. One way of looking at Sen's critique is to observe that he is contesting Rawls's assumption that a single set of weights is appropriate to create the index of primary goods for all persons. People have different ways of combining primary goods into functionings, Sen would say, based on their physical and psychological endowments, and it is the vectors of functionings that people have that should be the concern of justice, not the vectors of primary goods. This criticism would be irrelevant if every person aggregated primary goods into functionings in the same way: and so for Sen's critique to be interesting, it must presuppose that there is no single index of primary goods which can be applied. Sen later introduced a concept of freedom and responsibility into his approach, by saying that it is the *set* of vectors of functionings available to a person which is ethically salient: he called this set the person's *capability*. Justice requires us to equalize capabilities, but the particular choice of a vector of functionings from the capability is the person's responsibility, and is no concern of justice. This theory requires us to be able to *compare sets*: when is *A*'s capability greater than or better than *B*'s? Sen never offered a solution to this problem, saying simply that in some cases we can obviously make this kind of judgment, for example, when one person's capability contains another's. I do not believe Sen made much headway on the responsibility question, although I do think proposing his objective-list theory of justice was a worthwhile contribution. I believe we can go along way towards evaluating the justice of actual distributions looking at the degrees of various functionings which those distributions enable in people.

In 1981, Ronald Dworkin published two papers, which advanced the discussion in a major way. He posed the question, what should an egalitarian seek to equalize, and argued in the first paper that equalizing welfare was an unattractive option; in the second paper, he argued the right thing to equalize was resource bundles, and he proposed how to do so. For Dworkin, persons have circumstances, which include their environments and internal talents and handicaps, and preferences or ambitions, what Rawls would have called their life plans. Circumstances are resources, and should be equalized across persons: but people are responsible for the exercising their preferences.

Dworkin distinguishes between brute luck and option luck. Brute luck is an event affecting the person against which he could take no precaution, against which he could not insure. Option luck is a stochastic event a person chooses to face. Thus if my house burns down and I chose not to buy fire insurance, that is an instance of option luck. If I am hit by lightning while hiking, and the weather forecast predicted fine weather, and there was, furthermore, no lightning insurance available, that is brute luck. If the weather forecast was stormy, and I chose to hike, the lightning strike is option luck. Dworkin argues that brute luck is unfair, while option luck is fair, precisely because a person is responsible for his choices. Perhaps the most important instance of brute luck is the birth lottery, which assigns circumstances – families and genetic constitutions – to persons. No insurance was available to the individual in this case.

If all resources were transferable, then equality of resources would simply mandate that an equal-division of resources be mandated, and then people could trade to a competitive equilibrium. This would produce a Pareto efficient allocation, in which no one envied any other person, in the sense of desiring the other's bundle. This is exactly the case of the auction on the desert island that the shipwrecked travelers face in Dworkin's second article. Unfortunately, many important resources are not transferable: families into which one is born, talents, and handicaps. Dworkin's ingenious proposal is to transform the brute luck of the birth lottery, which allocates these resources, into option luck, by use of a hypothetical insurance market.

The venue for this market is an analog to Rawls's original position. In this venue, each person is represented by a soul. The soul knows the preferences of her person, including her preferences over risk. She does not know the circumstances of her person, but she does know the distribution of circumstances in the world. In Rawlsian parlance, she knows exactly those things her person is responsible for (preferences) and facts about the world (the distribution of circumstances). She does not know what is morally arbitrary, namely, her own circumstances, or as Dworkin would also say, resource endowment. In this situation, an insurance market will operate, where each soul will have the same amount of fictional money to purchase insurance -- this equality is key. Insurance can be purchased to pay out contingent upon the realization of the birth lottery. Thus, souls can insure against their persons having bad luck in the birth lottery, and in so

insuring, they use the risk preferences of their particular persons. This seems to be a perfect model for Dworkin's theory: for the original position shields persons precisely of morally arbitrary facts and holds them responsible for their choices.

Unfortunately, Dworkin's discussion of how this insurance market would operate is clumsy and ad hoc. He should have consulted an economist at this point, because there is a beautiful theory of insurance markets, due to Kenneth Arrow, which can be applied to Dworkin's problem. Indeed, I have applied to the standard model of insurance to Dworkin's veil of ignorance construction, and it leads to the following result: unless the individuals are sufficiently risk averse then the consequence of Dworkinian insurance will be to transfer resources from the disabled to the able.

The intuition for this unhappy result is as follows. Suppose that there are two people in the world, Andrea and Bob. Andrea is talented and Bob is handicapped. 'Talent' in this example is the ability to convert resources (wealth) into some desired outcome – say functioning or welfare. Talent is the resource that is distributed by the brute luck of the birth lottery. The talented person transforms wealth more efficiently into welfare than the disabled person.

In the Dworkinian insurance market, there are two souls, one representing Andrea and the other Bob: call them Alpha and Beta. Each knows the risk preferences of its person, but not whether its person is the talented or the disabled one – that is, the souls do not know how the birth lottery plays out. Alpha and Beta have the opportunity to take out insurance against being born disabled (i.e., as Bob). An insurance contract will take the form: "If I am born as Andrea, then I will pay x dollars to Bob; if I am born Bob, then I will receive y dollars from Andrea." Equilibrium in the insurance market exists when the two contracts made by Alpha and Beta match, in the sense that, in both states of the world, demands and supplies of the transfers are equal. Now there are two contrasting forces at play: on the one hand, the souls wish to protect themselves against having too little wealth in the state in which they are born as Bob; on the other hand, they wish to take advantage of their superior ability to transform the resource into welfare when they are born as Andrea. Which of these two forces dominates depends upon the degree of risk aversion of Andrea and Bob (and hence, of Alpha and Beta). *If Andrea and Bob are insufficiently risk averse, they will write contracts so that the soul born as Andrea*

receives the majority of the wealth. This is a perverse result, from the resource-egalitarian viewpoint. Thus, whether the Dworkinian insurance market realizes Dworkin's aim, of transferring wealth resources to those born with insufficient internal resources, is entirely contingent on how risk averse actual persons are. The insurance contracts can result in perverse outcomes if Andrea and Bob are insufficiently risk averse.

4. What's wrong with the original position?

Rawls's project was to produce a recommendation of equality (the difference principle) from two assumptions only: rationality and self-interest. Dworkin also believed that, from an assumption of self-interest, hypothetical insurance markets would suffice to allocate transferable wealth in a way to properly compensate persons for disadvantaged circumstances which were non-transferable. If egalitarians find the *gedankenexperiment* of the original position attractive, they are faced with a problem: should the original position be discarded as a tool, or should equality be abandoned as a desideratum?

I have puzzled over this quandary for some years. Nor am I the only one: Serge-Christophe Kolm and Brian Barry have both understood these issues, if not in precisely the way I have explicated them, and have come to similar conclusions. My solution is to challenge the premise of self-interest. Rawls's grand goal is unachievable: rationality and pure self-interest will not (as far as I can see) produce equality as the social contract. One needs in addition an assumption of solidarity. Kolm states it differently: he argues that *reciprocity* is what is required (which may be not exactly the same thing as solidarity). Some kind of social glue is needed.

If we posit that that social glue is embedded in the preferences of individuals, then it is not difficult to construct an original position which will produce an egalitarian social contract. But once solidarity is postulated, one does not need the original position! One can pass directly from solidarity to equality – or, to be precise, to a degree of equality consistent with the degree of solidarity that people feel.

My view is that Rawls should have skipped the whole original position construction, and simply argued that because the assignment of families and talents to persons is morally arbitrary, compensation from the advantaged to the disadvantaged is

required. The moral arbitrariness of the distribution of these resources must be taken as a premise. Dworkin's distinction between choice and circumstance is certainly key: to some degree at least, inequality which is due to the exercise of choice is ethically all right. How can one formulate the problem?

5. Equality of opportunity

In 1989, G.A. Cohen (1989) and Richard Arneson(1989) each published articles critiquing Dworkin – but not along the lines I have presented. (I first published my critique of the hypothetical insurance market in 1985.) Both praised Dworkin for introducing responsibility into egalitarian theory. Arneson argued that Dworkin was right to move away from equality of welfare, because of the expensive taste argument, but that instead of moving to equality of resources, he should have opted for equality of opportunity for welfare. Arneson then attempted to provide a definition of what it means to equalize opportunities for welfare. Cohen argued that Dworkin had misplaced the cut between preferences and circumstances: there are certain preferences a person should not be held responsible for, even he is glad he has them. Think of Sen's tamed housewife and battered slave, persons who adopt insufficiently ambitious preferences because of cognitive dissonance or social pressure, in order to remain sane in their poorly resourced environments. Dworkin only discounted preferences when they are addictions or cravings – preferences the person would rather not have. But many preferences are themselves *induced* by unjust circumstances, and we are not fully compensating persons who hold them if we hold them responsible for their choices. Cohen called for 'equal access to advantage,' something sounding not too different from Arneson's equality of opportunity for welfare. Cohen has never made his suggestion more precise.

Inspired by this debate, I proposed (first, in Roemer (1993)) an approach which I view as an attempt to carry out Arneson's recommendation. I will outline this approach and then give some examples of its application, reply to some critiques, and issue some caveats about the scope of its application.

The language of my proposal contains five words: objective, circumstances, type, effort, and policy. We suppose a planner who desires to equalize opportunities, among

a given population, for the acquisition of some objective: Cohen would call the objective ‘advantage,’ and Arneson would call it ‘welfare,’ but I wish to be less grandiose, and to think of it as something more observable and limited, such as life expectancy, or wage-earning capacity, or income. These objectives are things that policy makers are concerned about; they are less encompassing and subjective than Cohen’s advantage or Arneson’s welfare. Individuals each possess circumstances, and have the capacity to apply effort. The value of the objective which a person will achieve is taken to be a function $u(C, e, \varphi)$ of his circumstances (C), his effort (e), and the policy of the planner (φ). I emphasize that u is *not* a utility function: persons may be differentially interested in realizing the objective at hand, and therefore will differentially exert effort to accomplish that end. For instance, if u is life expectancy, one example of low effort is smoking, because it is a voluntary activity (let us suppose) which reduces the value of that objective. My planner, however, is – in this case-- the minister of health, and she is interested in equalizing opportunities for life expectancy in her society.

For the sake of practicality, I partition the set of circumstances into a finite number of elements, which I call *types*. A type is the set of people with similar circumstances. Circumstances, as for Dworkin, are taken to be aspects of persons’ environments which are beyond their control and which influence their ability to acquire the objective.

Continuing the life-expectancy example, we would define effort as an index of the choices that people make which influence their life expectancy – perhaps an index including whether or not they smoke, and their eating and exercise habits. We can construct this index by observing the relationship of these behaviors and circumstances to age of death in our population. This will give us the function u . Now the set of policies Φ is a set of interventions available to health ministry: health insurance, spending on education about the links between effort and mortality, income transfers, and so on.

For any policy φ that is chosen, there will result a specific distribution of effort in each social type. It is important to note that these distributions of effort may be very different across types. For instance, if we take socio-economic status as a circumstance, it may be the case that the distribution of effort is much worse in poor types than rich

types. In particular, the distribution of effort is itself a *circumstance*: it is beyond the control of any individual, and it is influenced by the other circumstances.

Equality of opportunity is achieved, I say, when the degree of acquisition of the objective is independent of a person's circumstances, but responsive to his degree of effort. This is the interpretation of Dworkin's idea: people should be indemnified against the effects of their circumstances, but held responsible for their efforts. The tricky part is that the distribution of effort is, as I said, itself a function of circumstances. So how can the Dworkinian idea be implemented?

We can easily compare the efforts of people within the same type: simply observe which one has the higher index of the choices we call effort. But how can we compare the efforts of two people in different types? We require a measure of the degree of effort which sterilizes the index of effort of circumstances upon it. Such an index is the *rank* of the person in the effort distribution of her type. Thus, I declare two people, in different types, to have expended the same *degree of effort*, if they lie at the same centile of the effort distributions of their types.

Now by definition, effort is defined as that constellation of choices which, if increased, will increase the value of the objective, within any type. It follows that the individual at the p^{th} centile of the effort distribution of his type, will also be the person at p^{th} centile of the distribution of the objective in his type (since circumstances and policy are fixed for a type). Therefore, equality of opportunity for the objective is achieved precisely when the distributions of the objective are identical across types. For this is the result if we equalize the objectives at every centile of their distribution, across types. That is, the planner desires to choose the policy to equalize the *distribution* of the objective for every type.

To summarize, the *equality* part of equal opportunity is that regardless of one's circumstances, one will face the same distribution of the objective – that is, one will have the same probability of achieving given levels of acquisition of the objective, regardless of type. The *opportunity* part is that we do not seek to equalize the value of the objective across all persons, but only to equalize the distributions of the objective across types. Greater effort will bring about a greater degree of acquisition of the objective within type.

Thus, a person is not responsible for his circumstances, but is responsible for his choice of effort.

In practice, it is almost always impossible to find a policy that will equalize perfectly the distributions of the objective across types. One must, therefore, adopt some second-best approach. A number of such approaches have been proposed. The simplest one is to choose that policy that maximizes the minimum of the mean value of the objective, across types. This sounds very much like Rawls's difference principle – except here, it is applied only to the aggregates called types, rather than to individuals. There are other second-best measures that I will not discuss, as this would involve an uncomfortable intrusion of mathematics into the presentation.

I will present two applications of this approach. The first takes the population to be American males, the objective to be wage-earning capacity, circumstances to be characteristics of the individual's family background (parental education, parental income, etc.). To be specific, in the computation I describe, my co-author, Julian Betts, and I chose the education of the mother to be the single characteristic of type, and we partitioned the population of young men into four types: mother had less than 8 years of education, mother had more than 8 years but less than 12 years, mother had 12 years, and mother had more than 12 years. The policies are educational finance policies, of the form (x_1, x_2, x_3, x_4) where we take a given total budget for K-12 education, and assign a different amount of expenditure to each of the four types. In the US, because we have local financing of K-12 education, there is much variation in per-pupil spending around the country, and we can use available data sets to estimate how adult wages vary as a function of per-pupil expenditure and type. Using these data sets, we compute the required relationships, and then we solve the following problem: Find the vector of type-expenditures $x = (x_1, x_2, x_3, x_4)$ that makes the distribution of wages across the four types as close as possible to identical. (We use a somewhat more complex measure than maximizing the minimum of the mean wages across types.) The result is presented in table 1:

Table 1: EOp allocation of investment with per capita budget of \$2500 per student per annum

Parental Ed'n	< 8 years	8 < ed < 12 yrs	12 yrs	> 12 yrs
EOp investment	\$5360	\$3620	\$1880	\$1100

From: Betts and Roemer (2008)

John Roemer 4/24/09 10:58 AM

Comment: Slide 1

The young men in the sample were in secondary school in the late 1960s, and we took the annual per-pupil expenditure to be \$2500, the US national average at that time. The equal-opportunity allocation would have been to spend approximately five times as much per pupil on the most disadvantaged type as on the most advantaged type. Note that this allocation is very different from the 'equal resource' allocation of \$2500 per pupil of all types. Indeed, even achieving that equal-resource allocation would have been a considerable improvement over the existing US allocation in 1970, where richer municipalities spend more on their pupils than poor ones. The equal-opportunity allocation is highly skewed towards the disadvantaged.

Moreover, we predict that if educational resources had been allocated in this manner, then the average wage would have risen by 2.6% compared to what it in fact was. Thus, equalizing opportunities, in this case, increases total output. This happy result will not always occur with equal-opportunity policies.

John Roemer 4/24/09 10:58 AM

Comment: Slide 2

My second example is a hypothetical health example. Suppose there is a society with two types, Poor and Rich; 25% of the population is Poor and 75% is Rich. There are two diseases, tuberculosis and cancer. Both types are susceptible to cancer, but only the Poor will contract tuberculosis. The probability of contracting cancer is the same for the Poor and the Rich, as a function of the quality of life-style (which is effort), which is:

$$s^{CP}(q) = s^{CR}(q) = 1 - \frac{2q}{3}.$$

A Poor person's probability of contracting TB as a function of life-style quality is:

$$s^{TP}(q) = 1 - \frac{q}{3}.$$

The life expectancy of a Rich person is:

70, if cancer is not contracted, and

John Roemer 4/24/09 10:59 AM

Comment: Slide 4

$$60 + 10 \frac{x_c - 1}{x_c + 1}, \text{ if cancer is contracted, and } x_c \text{ is spent on its treatment.}$$

Suppose the life expectancy of a Poor person is

70 if neither disease is contracted,

$$60 + 10 \frac{x_c - 1}{x_c + 1} \text{ if cancer is contracted and } x_c \text{ is spent on its treatment, and}$$

$$50 + 20 \frac{.1x_t - 1}{.1x_t + 1} \text{ if tuberculosis is contracted and } x_t \text{ is spent on its treatment.}$$

Thus, the Poor will have a life expectancy of only 30 if TB is contracted and not treated, and TB is hard to treat, so life expectancy does not increase very rapidly with expenditures on the disease.

Now suppose that the Poor have life-styles whose qualities q are uniformly distributed on the interval $[0,1]$, while the Rich have life-style qualities that are uniformly distributed on the interval $[0.5, 1.5]$. Thus, the Rich have healthier life styles, on average. Suppose that the health-care budget of the country is \$3000 per capita. Suppose a *policy* is a vector of expenditures (x_c, x_t) which prescribes how much will be spent on each case of cancer and each case of TB. Note these expenditures are, ex hypothesi, independent of the type of person; treatment is *horizontally equitable*, in the sense that each case is treated in the same way, regardless of the type of person, *and regardless of the life-style of that person*. Of course, given the knowledge frequencies of the disease in the population as a function of life-style, and of the distributions of life-style of the two types, the Ministry can compute the set of feasible policies. The problem then is to choose the policy which maximizes the minimum life expectancy across the two types.

The solution turns out to be:

$$x_c^{EOp} = \$250, \quad x_{TB}^{EOp} = \$13,900;$$

much more is spent on a case of TB than on a case of cancer.

The life expectancies of the two types as a function of their life-styles turn out, given these expenditures, to be those shown in Figure 1. We see that, of course, life expectancy in both types remains an increasing function of life-style quality, for that affects the probability of contracting the diseases. Moreover, even at the EOp solution,

John Roemer 4/24/09 10:59 AM

Comment: Slide 5

John Roemer 4/24/09 11:00 AM

Comment: Slide 6

the Rich still have higher life expectancy than the poor at every degree of life-style quality. Why does the life-expectancy of the Poor remain considerably below that of the Rich? Because we do not discriminate between the poor and the Rich in treatment. There is no other feasible policy which would raise the average life expectancy of the Poor above what it is at this policy, *given* that we maintain horizontal equity. *If* we wished to discriminate between the Poor and Rich in cancer treatment, we could raise the life expectancy of the Poor considerably higher, at the cost of reducing the life expectancy of the Rich. But I have chosen the set of policies to disallow this kind of discrimination.

Let me contrast the EOp policy with the *utilitarian* policy, which is the policy from the same feasible set that maximizes the average life expectancy in the entire population. This policy turns out to be

$$x_C^{U^t} = \$2520, \quad x_{TB}^{U^t} = \$9350.$$

We now spend much more on cancer than before, and correspondingly less on TB. Notice that the life expectancies of the two types are much farther apart under the utilitarian policy: we have a higher average life expectancy in the population at the cost of greater inequality of life expectancy across types. Virtually all health ministries today, were they to face this kind of problem, would choose the utilitarian policy, because that is the dominant ideology in health economics, as in economics quite generally.

Let me summarize a number of features of these two examples.

1. Effort is treated as the residual explanandum of the objective, once circumstances are delineated. If we wish to predict the distribution of the objective within each type, we will require a theory of how the policy chosen will influence the distributions of effort, and hence the distributions of the objective.
2. The policy itself can be type-blind and effort-blind in its application, as in the health example. This is a matter of how the policy space is chosen. We *can* make the policy applied depend on the individual's type, but we *need* not. Similarly, we can make the policy applied depend upon the individual's effort, but we need not. In both examples given, the policy did not depend upon individual effort; in the education example, it does depend upon the individual's type.

John Roemer 4/24/09 11:02 AM

Comment: Slide 7

3. Equality of opportunity attempts to equalize the objective across types while utilitarianism attempts to maximize the average of the objective in the aggregate. Thus EOp policies will deliver *lower average accomplishment* than utilitarian ones and utilitarian policies will be *less equal across types* than equal-opportunity ones. This is *by definition* of what it means to maximize each of the two social objectives subject to constraint.

4. One can design equal-opportunity policy to be consonant with a society's views concerning responsibility by choosing the set of circumstances to reflect those views. The more characteristics of the person's environment we include in this set, the less will performance be attributed to effort. In the limit, if we choose each individual to comprise a single type, then there is no scope for effort, and the equal opportunity converges to the Rawlsian maximin across all persons. At the other extreme, if we choose all persons to be of one type, then the policy is simply utilitarian: maximize the average value of the objective tout court. In general, the EOp policy is less redistributive than complete maximin and more redistributive than utilitarianism.

6. Critiques and caveats

Some writers have argued that to implement the equal-opportunity policy requires an invasion of privacy, which is required to ascertain the effort that individuals have expended. (One envisions grilling a lung-cancer victim about his smoking behavior before treating him.) The examples show this is false: in neither example is it necessary to ascertain the effort of individuals before applying the policy. Indeed, in the health example, it is not even necessary to ascertain their type! The planning ministry does need to know the effort response to policies by type, in the society: but this can be ascertained with studies on small representative samples of the population. Formally, it is only *distributions* of effort and objective with which the planner is concerned, not individual responses.

Some critics say that effort, in the sense of actions for which individuals should be held responsible, is a useless concept when it comes to something so vital as health, or even education. These critics say that relevant actions in these spheres are socially determined, not voluntary. My response is that they are both socially determined *and*

voluntary. When the price of tobacco is sufficiently high, fewer people smoke: thus, smoking is (to some degree at least) an act of choice. More educated people typically smoke less: so smoking is (to some degree at least) socially determined -- that is determined by circumstances. In any case, the equal-opportunity theory is *agnostic* concerning the *true* role of voluntary effort versus circumstances in determining behavior: rather than solving this metaphysical problem, it provides a policy which is *consonant* with any particular view one might delineate concerning responsibility and circumstance.

Some critics focus upon the *inequality* (within types) that results from application of the EOp policy: but one may just as well focus upon the *equality* (across types) which it implements. Indeed, the EOp educational finance policies I presented are radically more egalitarian than the finance policies of any actual country, in terms of compensation for social disadvantage.

Now, to some caveats. What is the scope of equal-opportunity policy? Should the Ministry of Sports require professional basketball teams to admit a certain number of very high effort, short players (that is, should height be considered a circumstance in this case)? Obviously the answer is no. Equal-opportunity policy considers only the welfare of the people who are competing for some social advantage in a world of scarce resources. The examples I have given do not take into account the effect of this policy on the general welfare of the entire population. In the case of professional basketball, there is a very small number of players, and millions of fans: the welfare of the fans should count more than the welfare of the players. Assuming short players on a team would lead to an inferior game, the fans' welfare trumps equalizing opportunities among players.

Consider a more serious example. Should we admit some socially disadvantaged but inferior surgeons to the surgery profession? Fairness among surgeons would perhaps say yes; consideration of the welfare of the patients would say no.

Richard Arneson's response to such problems is to say that it is a mistake to apply equal-opportunity policy in discrete spheres, as I have done here: rather, one should aim to equalize opportunities for welfare over the entire population. Thus, we would, for example, limit the extent to which we license high-effort surgeons from disadvantaged

backgrounds, because doing so would reduce welfare for others. I have no objection to Arneson's proposal except that it is unworkable. Practicality requires the discrete-sphere approach, and so we must be content to deal with the kind of conflict I have described in a somewhat ad hoc way.

I would propose a general rule of thumb: when we are considering individuals who are being trained or educated, I would tend to favor applying equal-opportunity principles. When we consider people competing for professions or occupations in society, I would stress merit. Thus, I would have an equal-opportunity admissions process for medical schools, but would apply strict merit principles in licensing would-be surgeons. I would admit a certain number of short basketball players to high school teams, which have an educational function, but I would consider only basketball prowess for recruitment to professional teams.

I believe this rule of thumb has general support among some citizenries, at least in the United States. In the debate over affirmative action admissions to universities in the US, many objected to counting race as a circumstance, who do not object to counting socio-economic circumstances. Thus the University of California and the University of Texas, both public institutions, some years ago replaced racial preferences with preferences for socio-economic disadvantage: of course, African-Americans and Hispanics count highly in the economically most disadvantaged type. Citizens, in this case, did not object to an equal-opportunity policy, but they objected to counting race as a circumstance, because it sometimes did not correlate well with economic disadvantage. (Indeed, those opponents argued that many of the beneficiaries of the racial affirmative action policies were upper-middle class African Americans.) But with respect to hiring, citizens objected to any kind of affirmative action: here, they expressed the opinion that hiring should be meritocratic. This is the cut between education and training, on the one hand, and competition for occupations, on the other, that I suggested above.

7. How much inequality is due to circumstances?

Can we say how much inequality in income, in a country, is due to circumstances, and how much to differential effort? Obviously, it depends on how complete a list of circumstances one makes. One could go 'all the way down,' and say that the individual

is responsible for nothing, all her actions being determined by prior environmental stimuli and her genetically determined physical make-up. But if one is a philosophical compatibilist, and reserves a space for responsible choice, despite holding a materialist view of human behavior, the question is sensible.

It is, however, generally impossible to answer because of a poverty of appropriate data sets. Suppose one counts the following six characteristics as circumstances for the individual: his parents' income when he was a child, his parents' education, his IQ taken when he was a child, the number of siblings he had, whether he grew up in a single or dual parent household, the region of the country he grew up in, and his race/ethnicity. (I have not mentioned sex: I would do one equal-opportunity calculation for each sex.) If each of these circumstances can take on let us say three possible values, then we have $3^7 = 2187$ types. If we would like to have at least 100 persons in each type, for statistical significance, we would need a sample of millions. We could then see how much the distribution of earnings varied across the types.

There are hardly any such data sets today, but there is one – in Sweden, which collects all these data on young men when they enter military service. Anders Bjorklund, Markus Jantti, and I are using this data set to compute the effect of circumstances on earnings of young men. Using a decomposition method, we have computed the fraction of the Gini coefficient of earnings that can be attributed to each characteristic. The result is presented in Table 3:

Table 3 Relative contribution of six characteristics to the earnings Gini coefficient

Parental Income Type	7.9%
Parental Educ. Type	0.5%
IQ type	12.5%
Sibling type	1.0%
Family type	1.6%
BMI type	0.9%
Residual	75.6%

John Roemer 4/24/09 11:03 AM

Comment: Slide 8

One could argue that the body-mass index (BMI) type reflects both choice and circumstance; I believe that IQ type, if IQ is taken at a young age, is entirely a circumstance, since both nature and nurture are circumstances. We see that an astonishing 75.6% of the Gini coefficient is not explained by circumstances – it is, by definition, due to differential effort. There are caveats. First, the above calculation has *not* taken into account the effect of circumstances on effort, which must be done. We intend to do that, and doing so will certainly decrease the role of the residual somewhat. Secondly, some of what the residual explains is surely due not to effort but to luck – luck of an episodic kind, which is not captured by any of the circumstances. Needless to say, episodic luck is extremely hard to measure. Thirdly, these are Swedish data, and the Scandinavian countries are the most advanced in the world, with respect to eliminating the effect of circumstances on earnings. Indeed, note the very minor role of parental education in explaining earnings. The fact that the most important circumstance in determining earnings inequality is IQ indicates that Sweden is still, as one would expect, a meritocracy to a significant degree.

If, indeed, you accept our list of circumstances as fairly complete, capturing almost all of what a modern citizenry would call morally arbitrary features of persons, then the implication of table 3 is that a society with a Gini coefficient not much lower than Sweden's would be distributively just.

If we could calculate the same decomposition for a developing country, the circumstances would explain much more of the Gini coefficient. Indeed, Ferreira et al have made such calculations for a number of Latin American countries, using many fewer types than we used with the Swedish data: they report that between 30 and 50% of income inequality is due to circumstances there, which is a lower bound on the true amount, because of the limited number of circumstances they used. **Some examples**

In conclusion, I propose that a good measure of economic development is the degree to which earnings inequality is due to effort, not circumstances. A country would count as developed to the extent that it had eliminated the effect of morally arbitrary factors on earnings. In other words, economic development means approaching distributive justice, in the sense that contemporary egalitarian theory has defined it.

References

- Arneson, R. 1989. "Equality and equal opportunity for welfare," *Philosophical Studies*
- Cohen, G.A. 1989. "On the currency of egalitarian justice," *Ethics*
- Dworkin, R. 1981a. "What is equality? Part 1: Equality of welfare," *Phil. & Public Affairs*
- Dworkin, R. 1981a. "What is equality? Part 2: Equality of resources," *Phil. & Public Affairs*
- Moreno-Ternero, J. and J. Roemer, 2008. "The veil of ignorance violates priority," *Economics & Philosophy*
- Rawls, J. 1971. *A theory of justice*, Harvard Univ. Press
- Roemer, J. 1985. "Equality of talent," *Economics & Philosophy*
- Roemer, J. 1993. "A pragmatic approach to responsibility for the egalitarian planner," *Philosophy & Public Affairs*

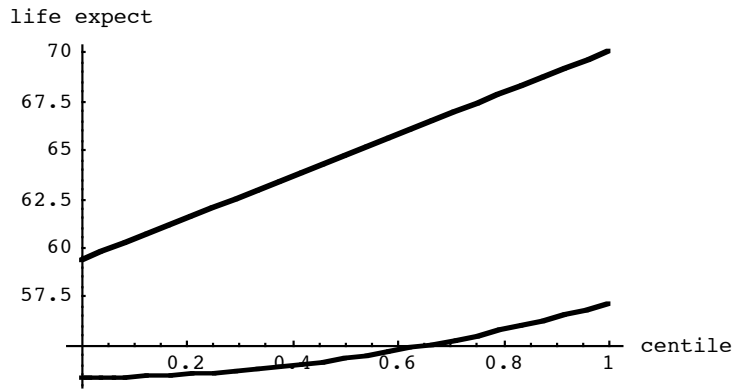


Figure 1 Life expectancy as a function of effort at the equal-opportunity optimum

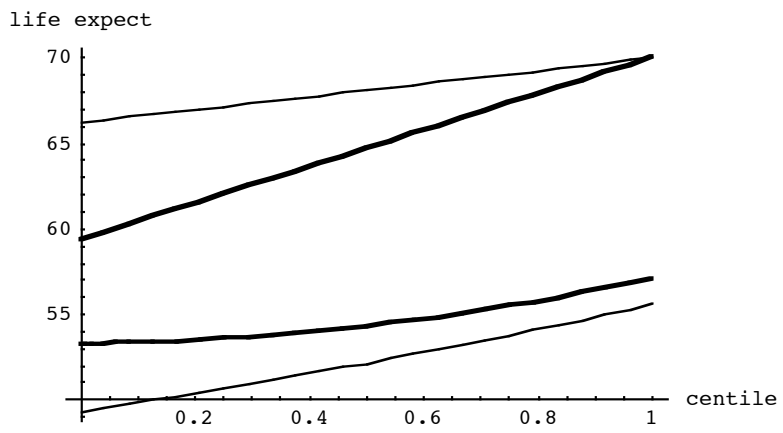


Figure 2 Life expectancy as a function of effort: EOp and Utilitarian