

Chapter 4

No Exit: Framing the Problem of Justice

Tom G. Palmer

Recent years have seen a marked retreat from cosmopolitan liberalism in moral and political theory. What is remarkable is that that retreat has been carried out under the banner of liberalism. Many modern “liberals” (in contrast to “classical” liberals) have taken as their touchstone adherence to the redistributive welfare state. The obligations and benefits allocated by welfare states are normally limited to particular groups of subjects or citizens; they are not universal.¹

John Rawls has prepared the field for much of that retreat. In his books and various articles Rawls has combined powerful sources of legitimacy for institutions and practices—mutual advantage, fairness, and agreement—into one account: justice as fairness. When he stipulates that the principles of a just society “are those a person would choose for the design of a society in which his enemy is to assign him his place,” Rawls builds in a highly illiberal premise into his theory of justice by foreclosing the exit option.² What was implicit in *A Theory of Justice* is made explicit in *Political Liberalism*: in setting out “the context of a social contract,” Rawls insists that “membership in our society is given, that we would not know what we would have been like had we not belonged to it Since membership in their society is given, there is no question of the parties comparing the attractions of other societies.”³ Rawls finds that for his construction to work it must be restricted to “a structure of basic institutions we enter only by birth and exit only by death”;⁴ strikingly, “the question of our entering another society does not arise.”⁵ Rawls has attempted to generate allegedly liberal outcomes from a choice situation that is highly illiberal. Cosmopolitanism, universalism, and internationalism—defining characteristics of traditional liberalism—disappear in the Rawlsian enterprise. That disappearance represents a general trend in contemporary political theory. Communitarians, socialists, welfare statistes, nationalists (and the various combinations among those) have been drawn away from cosmopolitanism for many of the very same reasons that Rawls posits “a structure of basic institutions we enter only by birth and exit only by death.” The case for preferring coercion over free association rests on an assumption that is both hardly self-evident and incompatible with core elements of

¹ Rights theorists traditionally distinguished “connate” and “adventitious” obligations: the former are universal and are justified on grounds of universal features of human life or sociality, whereas the latter are particular and accordingly require particular justifications, albeit according to universal principles. See Samuel Pufendorf, *The Political Writings of Samuel Pufendorf*, ed. Craig L. Carr, trans. Michael J. Seidler, (Oxford, 1994), esp. p. 51.

² John Rawls, *A Theory of Justice* (Cambridge, Mass., 1971), p. 152.

³ Rawls (1993a), p. 276.

⁴ *Ibid.*, pp. 135–6.

⁵ *Ibid.*, p. 277.

the liberal tradition. “Fairness” is a solution to a problem that was slipped into the specification of the choice situation by fiat, by arbitrarily excluding well established, well understood, and widely employed principles of free association from the options available to choosers.

The sleight-of-hand involved is remarkably similar to that typically involved in specification of the choice situation governing the provision of public goods. Once a decision has been made to produce a good on a non-exclusive basis, it is then asserted that the good cannot be produced through voluntary, uncoerced cooperation. Or the demonstration begins by assuming the existence of a good from which consumers cannot be excluded (or can only be excluded at some cost), when the problem is to produce such goods in the first place. Assuming that the good exists is hardly a solution to the problem of how to produce it. Similarly, by excluding exit or choice among options *as* an option, the problem of distribution of rights and obligations is converted into a pure bargaining game, which sets the stage for Rawls’s voluminous writings and the many jots and tittles added by his followers. By eliminating such options, Rawls does not solve a problem of justice or fair division; he creates it.

What follows is a brief excursion through the Rawlsian choice situation, setting out several unresolved problems and showing the significance of a quite problematic assumption that has received too little attention.

Mutual Advantage and Distributive Justice

In order to elicit agreement to the basic structure of society, Rawls argues, all groups, including the worst off, must benefit from the scheme of cooperation that governs a society: “since everyone’s well-being depends upon a scheme of cooperation without which no one could have a satisfactory life, the division of advantages should be such as to draw forth the willing cooperation of everyone taking part in it, including those less well situated.”⁶ Rawls stands firmly in the tradition of social contract theory in seeking to ground the social contract on mutual advantage. All must expect to benefit, in order for all to be obligated.

Rawls believes that people seeking their mutual advantage behind a veil of ignorance would insist on a benchmark of equality in the distribution of resources, departures from which would require justification. As Rawls states, “from the standpoint of one person selected arbitrarily”:

Since it is not reasonable for him to expect more than an equal share in the division of social goods, and since it is not rational for him to agree to less, the sensible thing for him to do is to acknowledge as the first principle of justice one requiring an equal distribution. Indeed, this principle is so obvious that we would expect it to occur to anyone immediately.

Thus, the parties start with a principle establishing equal liberty for all, including equality of opportunity, as well as an equal distribution of income and wealth. But there is no reason why this acknowledgment should be final. If there are inequalities in the basic

⁶ Rawls, *A Theory of Justice*, p. 15.

structure that work to make everyone better off in comparison with the benchmark of initial equality, why not permit them?⁷

The reason that inequalities may be to the advantage of those who end up with less, is that if “these inequalities set up various incentives which succeed in eliciting more productive efforts, a person in the original position may look upon them as necessary to cover the costs of training and to encourage effective performance.”⁸ That leads Rawls to the formulation of his “difference principle,” according to which “social and economic inequalities are to be arranged so that they are both: (a) to the greatest benefit of the least advantaged, consistent with the just savings principle, and (b) attached to offices and positions open to all under conditions of fair equality of opportunity.”⁹

How Much Inequality does the Difference Principle License?

In an early review of *A Theory of Justice*, Thomas C. Grey raised a serious objection to the justification of the difference principle, suggesting that it suffers from “a

⁷ Ibid., pp. 150–51. Note that Rawls uses the term “reasonable” to restrict claims for more than an equal share and the term “rational” to preclude acquiescence to less than an equal share. In *A Theory of Justice*, Rawls claims, in effect, that the “reasonable” can be deduced from the “rational,” understood in terms of maximization of interest satisfaction: “The theory of justice is a part, perhaps the most significant part, of the theory of rational choice” (p. 16). That claim is retracted in his later book, *Political Liberalism*, in which he argues that his earlier statement was mistaken, that it did not accurately express his intentions: “There is no thought of deriving those principles from the concept of rationality as the sole normative concept” (p. 53). Rawls cites approvingly T. M. Scanlon’s essay, “Contractualism and Utilitarianism” (T. M. Scanlon, “Contractualism and Utilitarianism,” in Amartya Sen and Bernard Williams (eds), *Utilitarianism and Beyond* (Cambridge, 1982), pp. 103–28, in which Scanlon argues that the contractualist enterprise is properly understood as driven by “the desire to find and agree on principles which no one who had this desire could reasonably reject” (p. 111) and that that “moral argument” (p. 111) is to be distinguished from the rational calculation of advantage. Scanlon concludes his distinction between contractualist views driven by the “rational” and those driven by the “reasonable” as follows: “On one view, concern with protection is fundamental, and general agreement becomes relevant as a means or a necessary condition for securing this protection. On the other, contractualist view, the desire for protection is an important factor determining the content of morality because it determines what can reasonably be agreed to. But the idea of general agreement does not arise as a means of securing protection. It is, in a more fundamental sense, what morality is all about” (p. 128). After rejecting his earlier formulation of the theory of justice as a part of the theory of rational choice, Rawls claims, in *Political Liberalism*: “I believe that the text of *Theory* as a whole supports this interpretation” (p. 53). The distinction in the text cited Rawls, *A Theory of Justice*, pp. 150–51, in which “reasonable” limits claims for more and “rational” limits acquiescence in less, supports this interpretation, but it still leaves problems in the justification for inequality in Rawls’s theory, as we shall see later, and may leave unjustified certain central Rawlsian principles.

⁸ Rawls, *A Theory of Justice*, p. 151.

⁹ Ibid., p. 302.

psychological or moral if not a logical inconsistency.”¹⁰ The inconsistency concerns the combination of the insistence of a baseline of equality in the distribution of wealth, arising from the alleged “moral irrelevance” of natural or inherited talents or capabilities, and the admission of incentives in the justification of divergence from equality and the establishment of a regime of inequality in the distribution of wealth.

Grey argues that that justification suffers from a “failure to establish any limits of justice on the bargaining power of those with more than average productive abilities. The standard is in principle indeterminate between complete equality of income, on the one hand, and income wholly determined by productivity, on the other.”¹¹ As Grey notes, “The principle apparently contemplates some extra payment to those with scarce or special skills ... [in the belief that] paying differential incomes to those with extra skills will bring forth extra production from them.”¹²

How can I justly bargain for more than an equal share by threatening to withhold my scarce talent? It must be because I have some special claim on these talents and their fruits. But if they are considered a social asset—as Rawls apparently regards them—I can have no such special claim. On the other hand, if my talents and the fruits of the use I choose to make of them belong to me, then there can be no justified coercion of me if I do not choose to share them with others. The inequalities which Rawls would permit suggest some acquiescence in this view. But those who press this view to its limits will reject the notion that greater equality of income is a legitimate aim of public policy.¹³

Let us press that view to its limits: the difference principle either results in no deviation from voluntary market outcomes, as the allowance of incentives entails the recognition of claims by possessors of natural assets and consequently of bargaining rights, *or*, if the difference principle *is* to realize a redistributive aim, it requires both premises and outcomes that are highly objectionable and counterintuitive, and certainly highly *illiberal*.¹⁴ Rawls’s attempt to justify a redistributive state can only

¹⁰ Thomas C. Grey, “The First Virtue,” *Stanford Law Review*, 25 (1973): 286–327, p. 325.

¹¹ *Ibid.*, p. 322.

¹² *Ibid.*, p. 322.

¹³ *Ibid.*, p. 325.

¹⁴ The fundamental issues I will deal with were first raised by Grey, “The First Virtue”; others who have trod this ground, in one way or another, are the following: Robert Nozick, *Anarchy, State, and Utopia* (New York, 1974), esp. pp. 183–97; Jan Narveson, “A Puzzle about Economic Justice in Rawls’s Theory,” *Social Theory and Practice*, 4 (1976): 1–27 and, by the same author, “Rawls on Equal Distribution of Wealth,” *Philosophia*, 17 (1977): 281–92. Also see on these issues Donagan, Alan, *Morality, Property, and Slavery. The Lindley Lecture 1980* (Lawrence, Kans, 1981); Brian Barry, *Theories of Justice* (London 1989); and G. A. Cohen, “Incentives, Inequality, and Community,” in Grethe B. Peterson (ed.), *The Tanner Lectures on Human Values, Volume Thirteen* (Salt Lake City, 1992), pp. 263–329.

I intend to draw on some of their arguments, but also to extend the critique further and to connect these problems with the problem of the membership of the contracting group, that is with Rawls’s “closure” move discussed above. Barry notes that “Rawls has never, it seems to me, addressed himself to this line of attack” (p. 234). I believe that, although he

succeed by rejecting such fundamental liberal precepts as the right of exit and the universality of principles of justice.

A problem ancillary to the problem that Grey identifies as the indeterminacy of the difference principle “between complete equality of income, on the one hand, and income wholly determined by productivity, on the other,” is, What is the baseline of equality in the distribution of wealth from which inequality is to be measured? Given that there is some wealth in existence at any given time, are we to say that inequalities are to be measured from the base currently existing at time t_1 (Wt_1), with Wt_1 to be shared out precisely equally? Setting aside the perhaps insuperable difficulties in determining how much inequality (not too much and not too little) is necessary to satisfy the difference principle from any base we can ask why the total wealth later on, say at t_2 , should not be shared out equally. How are we to determine the baseline of equal distribution of wealth from which inequality is to be justified? Why use Wt_1 rather than Wt_2 as the baseline of equal wealth distribution? Given that production is a process that takes place over time, why should any one of the periods $Wt_1 \dots \infty$ be subject to strict equalization, but not others? There seems little in the specification of the original position to lend us guidance; certainly the discussion in *A Theory of Justice* of “intergenerational justice”¹⁵ is no guide.

Besides those problems of determining the baseline and the allowable divergences therefrom, there is the far more serious problem of determining just why some are allowed to demand more than are others. Under a principle of strictly equal distribution of total product, one who produces more would be expected to produce for the benefit of others, as well as for herself, since out of a population of k she would receive only $1/k$ of the increase in productivity due to her additional efforts. But Rawls specifically insists that we cannot expect individuals to serve one another in that way:

One might think that ideally individuals should want to serve one another. But since the parties are assumed not to take an interest in one another’s interests, their acceptance of these inequalities is only the acceptance of the relations in which men stand in the circumstances of justice. They have no grounds for complaining of one another’s motives. A person in the original position would, therefore, concede the justice of these inequalities.¹⁶

What is being asserted, then, is that those who insist on being paid more to produce more are allowed to demand a larger share of the total product than others receive. What they are allowed to say, in effect, is that they will not produce anything, or they will deliberately produce less, unless they are allowed to demand (and to receive) a special incentive in the form of an unequal share, that is, unless they are allowed to make an “unreasonable” claim.¹⁷

does not address those criticisms directly, Rawls’s move to close entry and exit options to the contracting parties can be understood as his response to problems of this sort.

¹⁵ Rawls, *A Theory of Justice*, pp. 284–93.

¹⁶ *Ibid.*, p. 151.

¹⁷ Recall that Rawls writes that “it is not reasonable for him to expect more than an equal share in the division of social goods” (Rawls, *A Theory of Justice* p. 150). Narveson

The very idea of incentives for greater production (as opposed to covering the costs of training and production, which is fully consistent with equal distribution of consumer goods, so long as no residual remains over the costs of production)¹⁸ is an admission of a *right* of the better endowed, or of those who are simply better at bargaining (which need not entail any absolute advantage in productive capacity), to demand a larger share of the product. If a person must be *allowed* to “hold out” for more, it can only be because she has a right to do so. Having a right *not* to work entails a right to bargain for a larger share of the product in exchange for working; if she has a right to employ or to withhold her time, labor, and endowments, then she has a liberty right to them (to employ them) and a claim right (to exclude others from employing them against her will). Rawls, however, has committed himself to the position that such factors as intelligence, strength, cleverness, and even motivation (and *certainly* bargaining ability), are “undeserved” and that they are therefore morally irrelevant in the decision of a fundamentally moral problem, the distribution of goods in society.¹⁹ Yet by introducing incentives for greater production, he tacitly allows those “morally irrelevant” factors to play a role in the distribution of goods. The difference principle seems to allow either voluntary market distribution or completely equal distribution; nothing in between complete equality and completely voluntary arrangements (which may, of course, also be equal, depending on the voluntary choices of the cooperating agents) seems to be derivable from Rawls’s arguments. There is a fundamental contradiction in Rawls’s enterprise.

Exile to the Production Frontier

Rawlsian justice faces a problem: either it amounts to classical liberal voluntarism, with no fairness-based coercive redistribution of income, or it demands complete equality in the distribution of wealth. The latter route leads Rawls away from traditional liberalism and universalism, for if the difference principle were taken seriously as a redistributive principle, it would entail an enforceable obligation to produce at one’s production frontier, that means, work could be made compulsory on the principle of conscript labor. As Alan Donagan notes, the difference principle “is a principle of servitude: not only does it expropriate to a common pool everything

responds to the claim that those who receive less “have no grounds for complaining of one another’s motives” as follows: “they *do* have a ground for ‘complaining of one another’s motives.’ The ground is that they are likely to end up with less, with a smaller than equal share, if people are permitted to be motivated by self-interest (say) instead of by equality. Indeed, Rawls’s argument here really seems to concede my basic point, which is that on this showing, the motive of justice would always direct one to sharing equally with one’s fellows” (p. 287). Cohen claims that “when true to itself, Rawlsian justice condemns such incentives, and ... no society whose members are *themselves* unambiguously committed to the difference principle need use special incentives to motivate talented producers” (Cohen, “Incentives, Inequality, and Community,” p. 310).

¹⁸ See Narveson, “Rawls on Equal Distribution of Wealth,” pp. 285–6.

¹⁹ Rawls, *A Theory of Justice*, p. 102: “No one deserves his greater natural capacity nor merits a more favorable starting place in society.”

above the minimum to the social product, returning only what is necessary to ensure that the difference will continue to be produced, but it also requires a social structure that takes away the option of idleness when one's own needs and the demands of beneficence have been met."²⁰ Rawls's scheme, to be made operational, would entail a duty to work up to the point where the worst off have been made as well off as possible, as "social and economic inequalities are to be arranged ... *to the greatest benefit of the least advantaged*"²¹ (emphasis added).

Rawls is not saved by his first principle of justice (equal liberties) from the unpleasant conclusion that forced labor may be justified to enforce the universal obligation to produce one's maximum product so as to enable redistribution to the greatest benefit of the worst off group, because the only liberties he includes in the "basic liberties" are "political" and, in Donagan's words, "*these basic liberties would not directly exclude the conscription of labor in time of peace.*"²² What *could* save the principle, according to Brian Barry, is the second half of Rawls's second principle, which Rawls states as follows:

Social and economic inequalities are to be arranged so that they are both (a) to the greatest benefit of the least advantaged and (b) attached to offices and positions open to all under conditions of fair equality of opportunity.²³

Part (b) of the second principle above, Barry claims, "looks like a pretty strong statement of the importance of occupational choice, and I think that, even if Rawls does not list it among the 'basic liberties,' the best interpretation of his theory is that he regards it as having priority over the difference principle."²⁴ So construed, the

²⁰ Donagan, p. 21. Rawls, *A Theory of Justice* characterizes the Kantian principle of treating persons as ends in themselves as follows: "To regard persons as ends in themselves in the basic design of society is to agree to forgo those gains which do not contribute to their representative expectations. By contrast, to regard persons as means is to be prepared to impose upon them lower prospects for the sake of the higher expectations of others" (p. 180). Donagan does not directly raise the following point in his critique, but his approach suggests the further specification that, if leisure (or even simple idleness) is a good, and the attainment of an additional increment of it a gain, which seem plausible enough assumptions, then parties to the contract must forgo leisure or idleness that does "not contribute" to the representative expectations of others less well situated than they are. That would imply a duty to produce until, at least, the well being of the least well situated is maximized.

²¹ Rawls, *A Theory of Justice*, p. 302; Rawls adds "consistent with the just savings principle."

²² Donagan, p. 19. Nor, it might be added, in time of war.

²³ Rawls, *A Theory of Justice*, p. 83.

²⁴ Barry, p. 399. Donagan considers and rejects this interpretation of the significance of fair conditions of equality of opportunity, noting *en passant* that "the word 'opportunity' is less than felicitous when the inequalities are burdens rather than rewards" (p. 19). It is also a bit difficult to understand how an inequality of liberty or of opportunity would enhance the liberty or opportunity of those with the lesser liberty or opportunity; if liberty is precisely the sort of thing that can be distributed precisely equally, how can diminishing the liberty of one increase the liberty of another? Rawls notes this problem obliquely when he distinguishes between "liberty" and the "value of liberty." As Rawls notes (*A Theory of Justice*, p. 204),

constraints placed upon the difference principle by the second half of the second principle make the limitation on inequality of wealth (or income) otiose, because one can always claim that one wants (or “needs”) a larger than equal share of the product, and back up the claim by a threat not to produce if one does not receive a larger share. That threat would be justified by Barry’s construal of part (b) of the second principle. There is a further problem: even if conscripted labor services *were* ruled out by the principle of open offices and fair opportunity (something that is *not* entirely clear from the priorities of principles that Rawls lays out, but which seems a fair enough move to make in defense of the Rawlsian approach), the principle of fair equality of opportunity would *not* be *inconsistent* with taxation of labor at its highest assessed use, as land and other assets are taxed.

To understand how such a capacity tax would work, consider the case of taxation of the productive asset of land. Land is typically taxed on the basis of an assessment of its most highly remunerative employment (measured in money), regardless of whether it is actually being employed in that way. In the same way, a capacity tax could be assessed based on what one “could earn” even if one were not in fact earning that at the moment because one placed a high value on leisure or simply because one does not like the occupation that would produce the highest income.²⁵ That would not run afoul of the “principle of fair opportunity,” even under its most generously liberal construal of it, although the effects of such a capacity tax would

“liberty and the worth of liberty are distinguished as follows: liberty is represented by the complete system of the liberties of equal citizenship, while the worth of liberty to persons and groups is proportional to their capacity to advance their ends within the framework the system defines. Freedom as equal liberty is the same for all; the question of compensating for a lesser than equal liberty does not arise. But the worth of liberty is not the same for everyone.” On the other hand, if opportunities are differentiated from simple liberty (for instance, as in the sense in which we say, “A had more opportunities than B,” even though they were equally free to take advantage of what opportunities they did face), then we could arrange *inequalities* in opportunities so that “an inequality of opportunity . . . [will] enhance the opportunities of those with the lesser opportunity” (*A Theory of Justice*, p. 303). Why would such an allowable inequality of opportunity not extend to limiting one’s access to choice of occupation? (“Affirmative action” requirements and mandatory hiring quotas, by limiting the opportunities of some groups in competition with others, limit the opportunities of those who are excluded or who are not included in the quotas; the excluded certainly no longer enjoy “equality of opportunity.”) And if the principle is admitted that choice can be limited in this way, then could it be limited to one option, and one only? If so, that would be equivalent to a mandate to pursue the one remaining option, which result would lead us from the principle of fair equality of opportunity, construed as differentiated from equality of liberty, *not* to a defense of labor market freedom against coercion, but to a justification of inequalities in choice options in labor markets, with the choice set reduced to one, all in pursuit of advancing the opportunities (or perhaps the value of the opportunities) of the least advantaged.

²⁵ For proposals along these lines, see Richard Arneson, “Lockean Self-Ownership: Towards a Demolition,” *Political Studies*, 39 (1991): 36–54 and “Property Rights in Persons,” *Social Philosophy and Policy*, 9/1 (1992): 201–30. Also confer James Sterba, “From Liberty to Welfare,” *Ethics*, 105/1 (October 1994): 64–98.

be difficult to differentiate from chattel slavery.²⁶ If such a capacity tax were deemed to maximize the position of the least advantaged, then it would seem to be demanded by the difference principle; and there must surely be some wealth producers (or even classes of them) who are producing less wealth than they *could* produce; were they to produce more of those goods there would be a larger surplus available for distribution to the least well off; hence, by not producing what they *could* produce, such “slackers” are diminishing the prospects of the least well off, relative to what the least well off could enjoy were the “slackers” to produce more.

Who is Everyone?

The considerations above go to the question of whether Rawls has developed an authentically *liberal* theory of justice. If forced labor were the outcome, it would be hard to see how that could be squared with the tradition of liberalism, which has, after all, the concept of liberty in its very name.

One could attempt to make a system of illiberal coercion universal, but Rawls does not take that approach. In the process, he abandons cosmopolitanism, as well. The reason is embedded in the difference principle itself, for when Rawls refers to “inequalities in the basic structure that work to make everyone better off in comparison with the benchmark of equality,” (*A Theory Of Justice*, p.151) he leaves ambiguous the question of whom he is designating by the term “everyone.” Does it mean every party to a transaction, or every party to the social contract, or does it mean literally every human being, or even every rational creature? Who is covered? Presumably, Rawls intends every member of the basic cooperating group governed by the contract and that is invariably cashed out to mean the subjects of a particular nation state. Allan Gibbard, in his review of Brian Barry’s *Theories of Justice*,²⁷ addresses the issue of the membership of the group over which the difference principle is to extend. In response to Brian Barry’s challenge, on grounds of “impartiality,” to extend the principle to the entirety of humanity, Gibbard draws out of Rawls’s work an approach that does not “hover uneasily . . . between impartiality and mutual advantage,” as Barry claims, but that is “intrinsically reciprocal.”²⁸ Gibbard imagines a “well-off person” putting to Rawls the following question: “Why limit myself in pursuit of my own advantage?”²⁹ Gibbard answers for Rawls as follows:

²⁶ This lesson was learned by the indigenous inhabitants of Africa and South America, among other places, when European states colonized them and, perhaps in line with the principle that taxes must be paid for benefits received, imposed taxes that could be paid, of course, only in European money. The only way for most of the indigenous inhabitants to get this European money was to work in plantations and mines owned by Europeans. See Parker T. Moon, *Imperialism and World Politics* (New York, 1926) for a graphic and detailed description of colonial taxing policies in Africa.

²⁷ Allan Gibbard, “Constructing Justice,” *Philosophy and Public Affairs*, 20 (Summer 1991): 264–79.

²⁸ *Ibid.*, p. 266.

²⁹ *Ibid.*, p. 269.

Rawls, in effect, gives this answer: “You have what you have only because others constrain themselves, in ways that make for a fair cooperative venture for mutual advantage. Constrain yourself by those rules in return, and you give them fair return for what they give you.”³⁰

Gibbard offers a thought experiment to clarify the notion of reciprocity: imagine that “each person sprang into existence on a separate island, adult and able, and each was protected by the water from threats and takings” and that the islands differ in fertility.³¹ Gibbard concludes that

from the bare assumption that fertility is morally arbitrary, no obligation to share follows. The lucky ones could still admit that their luck is morally arbitrary, and still ask “Why share?” One answer that they could not be given is that sharing pays others back for their cooperation or their restraint. No one has cooperated and no one has restrained himself, and so there is nothing to pay back. Motives of fair reciprocity, then, would not lead the lucky ones to share, even though they freely admitted that their luck was morally arbitrary.³²

Rawls takes up that theme in *Political Liberalism*, endorsing Gibbard’s approach and maintaining that “the idea of reciprocity is not the idea of mutual advantage.”³³ Gibbard’s attempted solution comes at a high cost, however, in the form of contradicting openly another basic Rawlsian tenet. Among the “formal constraints of the concept of right” in *A Theory of Justice* are the requirements that conceptions of justice must be general, universal, public, capable of imposing a transitive ordering on conflicting claims, and final.³⁴ In his elaboration of the penultimate item on his list of formal constraints, the requirement that a concept of right create a transitive ordering of competing claims, Rawls notes that “it is to avoid the appeal to force and cunning that the principles of right and justice are accepted. Thus I assume that to each according to his threat advantage is not a conception of justice.”³⁵ But a consequence of the very definition of an “intrinsically reciprocating group,” to which Gibbard and Rawls limit the range of the principles of justice, is that they rest the entire theory of justice on threat advantage, on the possibility of the employment of force and cunning. Those who can make credible threats of harm are accorded membership in intrinsically reciprocating groups and are entitled to be treated “fairly,” while those who cannot make credible threats do not deserve to be treated justly.

Gibbard sets up his thought experiment by specifying that among the islanders “each was protected by the water from threats and takings.” In his scenario there is no *reciprocity* of restraint, because the islanders have no intercourse whatsoever; the islanders need not even know of each other’s existence. There is neither mutual advantage nor reciprocity. But if the parties *are* capable of harming each other it

³⁰ Ibid., p. 269.

³¹ Ibid., p. 269.

³² Ibid., p. 269.

³³ John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993), p. 17.

³⁴ Rawls, *A Theory of Justice*, pp. 130–36.

³⁵ Ibid., p. 134.

might reasonably be said that they *do* reciprocate simply by *abstaining* from aggressive “threats and takings”; if not killing or stealing qualifies as “cooperation,” then, Gibbard’s argument maintains, the difference principle *does* extend over all of the islanders—lucky and unlucky—as a group and inequalities among them must be specially justified (if, that is, such inequalities can be justified at all, given the problems with “incentives” noted above). Gibbard and Rawls define “intrinsically reciprocal” groups by determining whether the members of the groups have the capacity to harm one another, which seems more congenial to starting with a Hobbesian threat advantage (which Rawls dismisses from his construction) than with the idea of people joined voluntarily together in an ongoing project of social cooperation.

There is an ancillary problem: if “everyone” (however defined) must be made better off, what does “better off” mean? Jan Narveson distinguishes two possible meanings of the phrase:

D1. “Everyone is better off” = “No one is worse off” (persons outside the transaction may be unaffected, compatibly with the requirement)

D2. “Everyone is better off” = “Each person’s situation is improved by comparison with what it was before the transaction” (persons outside the transaction must be favorably affected, to meet the requirement)³⁶

As Narveson notes, interpreting the phrase in the first way would make the difference principle otiose, and interpreting it in the second way makes it indeterminate, for the reasons that Thomas Grey offered. What “make everyone better off” must mean, if the difference principle is not to be rendered otiose (as Narveson’s interpretation “D1” would do), is that some people are to be given a veto power over the mutually advantageous interactions of others. Thus, even though C is *not hurt* by A’s and B’s voluntary cooperation, C is able to insist forcibly that A and B not be allowed to cooperate unless C receive a payment for not exercising her veto; the payment would come in the form of some share (in principle indeterminate) of the results of A’s and B’s cooperation. That means that C has a kind of ownership claim over A and B; she would determine how A and B shall use their resources and demand a payment from them for the privilege of being allowed to cooperate together.³⁷

³⁶ Narveson, p. 9.

³⁷ If my analysis of Gibbard’s formulation of the principle of reciprocity is correct, and the group over which the principles of justice are to apply is apparently defined by whether they can hurt each other through “threats or takings,” then this begins to sound more like a classic “protection racket,” in which parties must pay third parties to be able to cooperate in peace, than like a cooperative venture. Gibbard, “Constructing Justice,” indicates that something of the sort must be the case when he notes, “It is not only cooperation that is at stake in questions of justice, but mutual restraint: nonaggression and respect for a system of property rights. Just people forgo chances to seize advantage, and the idea of Justice as Fair Reciprocity is that they forgo these chances in return for others’ voluntary support of just arrangements. Now a coalition that withdraws from society renounces any claim to justice from those who remain” (p. 272). If those who attempt to withdraw have *no claim to justice*,

If it were to be responded that Rawls only intends for the difference principle to be embedded in the “basic structure” and not to be applied to individual transactions, one can point out that *all* the transactions to which it might be applied necessarily *are* individual transactions (there being no abstract or universal transactions), and that the principle must be applied over some temporal horizon, whether transaction-by-transaction (regardless of the duration of the transaction), or on a daily, monthly, yearly, or some other basis. The difference principle may be *articulated* at the stage of determining the basic structure in the Original Position, but it is *exercised* upon the transactions that take place after the basic structure has been determined. Which time horizon is chosen, from minute-by-minute to every hundred years, is not relevant to the general point I am making. One group (C) is given a right of veto over the voluntary transactions of others (A and B) unless they (C) get a cut, and that amounts to holding a claim right or a power against the members of A and B, by being able to stop them by force from transacting.³⁸ If mutual advantage defines socially cooperative groups, and therefore the membership of the social contract, one can legitimately ask why a group cannot separate out if they believe that *their* mutual advantage would lie in a contract governing distribution of a joint product produced *without* the cooperation of others who might produce a smaller share and then demand and enforce a share of the product disproportionate to their contribution.³⁹ Allowing such withdrawal would undercut the difference principle by allowing the exit of those who believe that they could do better without having their product redistributed to others.

It is for that reason—to salvage the difference principle—that Rawls must “close the door” of exit, by insisting that entry into society “is only by birth and exit from it

then those who remain are indeed free to attempt to enslave the withdrawing members, as Gibbard notes. Perhaps a more truly liberal approach would be to construe “cooperation” such that withdrawal from it need not mean dropping all restraint from aggression against persons and property; withdrawal would be simply an attempt to get away from other aspects of a cooperative scheme that are not to the interest or liking of the withdrawing parties. The proper reciprocal response to mere nonaggression, it seems, is simply nonaggression, rather than equalization of assets subject to the difference principle; if that is so—if withdrawal from a group need not entail complete withdrawal from all considerations of morality and justice, as it does for Gibbard—then why cannot any group of cooperators withdraw from the reciprocal scheme governed by the difference principle, but remain in the more limited “cooperative” relationship of not aggressing against one another? This would make the difference principle a truly voluntary principle among full cooperators, defined by those who actually consent to remain within the “fully cooperative” group. Just as the members of a club of equals might arrive at a principle of equal distribution, but are not constrained to do so if they agree to another principle, those who agree to cooperate in Rawls’s favored sense would be bound to his principles of distributive justice, while others would not.

³⁸ Rawls, *A Theory of Justice*, p. 282: “Because they start from equal shares, those who benefit least (taking equal division as the benchmark) have, so to speak, a veto. And thus the parties arrive at the difference principle.”

³⁹ Compare Barry, p. 243: “It is fairly clear that allowing the withdrawal of whole blocs from the society to define alternative nonagreement points would pose severe dangers to the claim that the difference principle is better for (almost) everyone than they could do in the absence of cooperation.”

is only by death.” If individuals were to be able to withdraw and form other groups at will, the core problems of fair distribution with which Rawls is concerned would effectively disappear. The closure move does not so much solve a problem of justice and fair division, as create it.

Joint Products and the Problem of Distribution

The Rawlsian enterprise rests on the idea of the “fair distribution” of a joint product. Before discussing “fair distribution,” an examination of the idea of a “joint product” is in order. Social cooperation may be valuable for a number of reasons, but surely prominent (if not foremost) among them is because cooperation results in a product greater than would be possible in its absence. The problem of justice is (at least principally) a problem of how that joint product is shared, allocated, or distributed, or of how the rewards of cooperative activity are imputed to those who have contributed to such activity. As Russell Hardin poses the problem, “If the bulk of what we enjoy is the result of collective endeavors in our society, then the bulk of what we have is up for redistribution.”⁴⁰

Economic science has something to say about the distribution of joint products among cooperators. In general, when markets exist (meaning, in this context, alternative uses of goods and free choice of factor owners among alternative employments of their factors) and factors of production are complementary and non-specific, meaning that they can be combined in various ways and put to a range of productive uses (for example, land for farming, recreation, or housing), then the proportion of the anticipated final value of the joint product imputed to the factors, that is, the *prices* paid for the factors, will tend to reflect their marginal value product (MVP). (To be more precise, what is imputed to the factors of production is the *discounted* marginal value product (DMVP) with the discount reflecting preference accorded to the time between the payment of the factor and the realized value or sale price of the final product.) When two or more factors of production are complementary but purely product-specific, then no market (in the sense of alternative uses of the factors, which entails the existence of opportunity costs) exists, meaning that a

⁴⁰ Russell Hardin, *Morality Within the Limits of Reason* (Chicago, 1988), p. 134. Rawls is even more strenuous in setting the non-agreement point and therefore the inclusiveness of what is up for distribution: “The difference principle represents, in effect, an agreement to regard the distribution of natural talents as a common asset and to share in the benefits of this distribution whatever it turns out to be” (Rawls, *A Theory of Justice*, p. 101). It seems that *everything*, and not merely “the bulk of what we enjoy,” viz. what we enjoy in excess of what we could have in total isolation from others, is up for distribution. See also David Gauthier, *Morals by Agreement* (Oxford, 1986), p. 254: “In seeking to treat persons as pure beings, freed from the arbitrariness of their individuating characteristics, Rawls succeeds in treating persons only as social instruments. In denying to each person a right to his individual assets, Rawls succeeds in treating persons only as social assets. In denying to each person a right to his individual persons, Rawls is led to collectivize those assets.” Thus, Rawls’s own account suffers from the same defect on the basis of which he rejects utilitarianism, namely that it “does not respect the distinction between persons” (Rawls, *A Theory of Justice*, p. 27).

determinate marginal value product cannot be imputed to the factors in the form of a price or dividend paid; all that can be determined is a joint MVP, namely the MVP of the combination of the two factors. (When the factors can only be combined in fixed ratios, then there is no way to determine marginal value product at all.) That is to say, the MVP to be imputed to the two factors together is equal to the total product (TP) minus the sum of the MVPs of all other factors. The proportion of the residual joint MVP of the specific factors to be shared by the factor owners will be purely a result of bargaining and cannot be determined by reference to marginal product. It is precisely the lack of an alternative employment or allocation of the factor that results in an indeterminate distribution of imputed value over which the factor owners must bargain or arrive at some salient or “reasonable” distributive outcome or principle.⁴¹

One way to *create* a situation of bargaining over the distribution of a joint product is to eliminate all alternative uses of the factor inputs by not allowing factor possessors to withdraw some or all of the factors and employ them elsewhere or not at all. If alternative uses disappear—perhaps due to radical changes in technology—or are cut off by fiat, the result is an indeterminate range of distributions (or imputation) of the value of the joint product to the factors in the form of the prices that they can command.

In effect, what Rawls and his followers are arguing is that “society” is a joint product the distribution of the value of which cannot be allowed to result from market principles of voluntary choice among alternatives, principles which normally result in a tendency for factor owners to receive their marginal value products.⁴² The

⁴¹ Note that a similar problem exists in the case of bilateral monopoly, in which there is both only one seller (“monopoly”) and only one buyer (“monopsony”), meaning that only a range of prices could be theoretically established, namely between the lowest price that would be accepted by the seller and the highest price that would be paid by the buyer. Precisely where the actual price would fall between these two boundaries is purely a matter of bargaining. As Richard Posner notes, “When a monopolistic supplier of labor confronts a monopsonistic buyer of labor, the exact price and quantity that will be set depend upon the parties’ relative skills at bargaining, ability to use intimidation or bring political pressures to bear, and perhaps other factors” (Richard Posner, *Economic Analysis of Law* (Boston, 1972), p. 136).

⁴² If the idea of the margin is rejected as entirely inappropriate to the analysis of justice, perhaps because individual contributions to the “joint product” in question (the products of social cooperation) cannot in any way be identified, then the basis of the difference principle itself is altogether obliterated. As Robert Nozick notes, “When it is necessary to provide incentives to some to perform their productive activities, there is no talk of a joint social product from which no individual’s contribution can be disentangled. If the product was all that inextricably joint, it couldn’t be known that the extra incentives were going to the crucial persons; and it couldn’t be known that the additional product produced by these now motivated people is greater than the expenditure to them in incentives” (Nozick, p. 189). Knut Wicksell, in his “A New Principle of Just Taxation”, similarly argued against Mill’s principles of distributive justice, for even when considering activities of the state (rather than all social activities, as Rawls does), “The discussion is almost always concerned with this or that change in the scope of the State’s operations, this or that extension or (much more rarely) contraction of the separate branches of public activity” (Knut Wicksell, “A New Principle of Just Taxation” in Richard A. Musgrave, and Alan T. Peacock (eds), *Classics in the Theory of*

reason is that no alternative allocation of the factors, including our labor power, can be considered that will generate reservation prices capable of establishing at least a set of highest and lowest exchange ratios (or set of baselines) within which the value product is to be determined. It then follows that, as Hardin notes, “the bulk of what we have is up for redistribution.”⁴³ This notion of a joint product attributable to factors of production that are *specific* (that is, for which there is no market or range of uses among or within which they may be allocated) is crucial to the enterprise of generating fair principles of distribution. It is the *closing off* of alternative uses by fiat that results in a pure bargaining or agreement problem.

One could argue that I am guilty here of the fallacy of affirming the consequent:

(A) If two or more factors of production have no alternative employments, then there is a bargaining (or rational agreement) problem.

(B) There is a bargaining (or rational agreement) problem.

(C) Therefore, there are two or more factors of production that have no alternative employments.

This would be a fallacious argument. Rawls has, however, saved me from falling into fallacy by specifically noting that, in order to avoid admitting market principles of voluntary agreement into “the political,” entry and exit options (that is alternative employments of our resources, most notably of our persons) are to be shut off.

In *Political Liberalism*, Rawls straightforwardly acknowledges that freedom of entry and exit do generate “measures of worth” and therefore “a basis for contractual calculations” (other than bargaining, salience, or some other form of agreement). But he claims that that applies only to “particular agreements” and not to “the political.” In the case of association,

one simply notes how the venture or association would fare without that person’s joining, and the difference measures their worth to the venture or association. The attractiveness of joining to the individuals is ascertained by a comparison with their opportunities. Thus particular agreements are reached in the context of existing and foreseeable configurations

Public Finance (2nd edn, New York, 1994), p. 78). Wicksell defends the marginal principle against Mill’s claim that “Government must be regarded as so pre-eminently a concern of all, that to determine who are most interested in it is of no real importance,” which means that allocation of burdens on the basis of benefits is beside the point. (quoted by Wicksell on p. 78)

⁴³ Note that Hardin does not write “distribution,” but “redistribution,” implying that assets have already been allocated and must now be *reallocated*. If the initial distribution occurred in compliance with the rules of justice, and then the results were judged to merit redistribution, then the rules of justice that resulted in the initial distribution are not action-guiding, in the sense of ensuring that when they are followed the parties need not fear coercive sanctions; the redistribution principle is incompatible with the rule of law. Rawls asserts otherwise, but his characterization of redistributive taxation could just as easily be applied to rape, muggings, or lynchings: “Taxes and restrictions are all in principle foreseeable, and holdings are acquired on the known condition that certain transfers and redistributions will be made” (Rawls, *Political Liberalism*), p. 283.

of relationships within the basic structure; and it is these configurations that provide a basis for contractual calculations.

The context of a social contract is strikingly different, and must allow for three facts, among others: namely, that membership in our society is given, that we cannot know what we would have been like had we not belonged to it (perhaps the thought itself lacks sense), and that society as a whole has no ends or ordering of ends in the way that associations and individuals do. The bearing of these facts is clear once we try to regard the social contract as an ordinary agreement and ask how deliberations leading up to it would proceed. Since membership in their society is given, there is no question of the parties comparing the attractions of other societies. Moreover, there is no way to identify someone's potential contribution to society who is not yet a member of it; for this potentiality cannot be known and is, in any case, irrelevant to their present situation.⁴⁴

Rawls acknowledges that free allocation of resources among alternative employments generates "measures of worth" but states that "there is no way to identify someone's potential contribution to society who is not yet a member of it." That is patently false; many people *do* make decisions to emigrate from one society to another, sometimes more than once, and they do so on the basis of the kinds of estimations that Rawls asserts are not possible.

In order to avoid market solutions to the problems he intends to treat, Rawls must stipulate that "membership in society" is "given," that is, unalterable. It is a stipulation that is contradicted by the experience of millions of persons. In addition to being an unhelpful characterization of the problem of political choice, the stipulation that parties cannot choose "membership in our society" does not help us to *solve* problems of "distributive justice," for it *generates* them instead.

Fairness and inescapable public goods

Rawls places great weight in the generation of obligations on what he calls "the principle of fairness." According to Rawls,

The main idea is that when a number of persons engage in a mutually advantageous cooperative venture according to rules, and thus restrict their liberty in ways necessary to yield advantages for all, those who have submitted to these restrictions have a right to a similar acquiescence on the part of those who have benefited from their submission. We are not to gain from the cooperative labors of others without doing our fair share.⁴⁵

Rawls even claims, quite implausibly, that "All obligations arise in this way."⁴⁶

⁴⁴ Rawls, *Political Liberalism*, pp. 275–6.

⁴⁵ Rawls, *A Theory of Justice*, pp. 111–12. Rawls acknowledges a debt to H. L. A. Hart's essay, "Are There Any Natural Rights?" (H. L. A. Hart, "Are There Any Natural Rights?," *Philosophical Review*, 64 (1955): 175–91, reprinted in Jeremy Waldron (ed.), *Theories of Rights* (Oxford, 1984), pp. 77–109).

⁴⁶ Rawls, *A Theory of Justice*, p. 112. Rawls distinguishes "obligations" from "natural duties"; the former "arise as a result of our voluntary acts" and are "normally owed to definite individuals, namely, those who are cooperating together to maintain the arrangement in

The cooperative venture with which Rawls is concerned in setting out his theory of justice as fairness is “the basic structure of society, or more exactly, the way in which major social institutions distribute fundamental rights and duties and determine the division of advantages from social cooperation.”⁴⁷ It is significant that he refers to the basic structure of society as “the arrangement of major social institutions into *one* scheme of cooperation”⁴⁸ (emphasis added). Rawls is a very careful writer, and I believe that in *A Theory of Justice* he had, even if not in a fully articulated way, the idea that the option of comparing alternative “arrangement[s] of social institutions” would not be allowed to play any role in agreement on the principles of social cooperation. In effect, by insisting on that restriction, he conjured up the very problem that his work was written to solve.

The social cooperation made possible by the basic structure is to the advantage of all members of society, and that advantage is demonstrated through their agreement to the basic structure governing the division of the advantages: “the guiding idea is that the principles of justice for the basic structure of society are the object of the original agreement.”⁴⁹ The choice situation, as well as the agreement itself, are understood “as a purely hypothetical situation characterized so as to lead to a certain conception of justice.”⁵⁰ Robert Nozick has responded to that assertion of the principle of fairness with a counterargument and a counter-assertion. The counterargument concerns the coherence of Hart’s general argument about the generation of special obligations,⁵¹ and the counter-assertion is as follows:

On the face of it, enforcing the principle of fairness is objectionable. You may not decide to give me something, for example a book, and then grab money from me to pay for it, even if I have nothing better to spend the money on.⁵²

question” (p. 113), while the latter “apply to us without regard to our voluntary acts” (p. 114) and “are owed not only to definite individuals, say to those cooperating together in a particular social arrangement, but to persons generally” (p. 115).

⁴⁷ Ibid., p. 7.

⁴⁸ Ibid., p. 54.

⁴⁹ Ibid., p. 11.

⁵⁰ Ibid., p. 12.

⁵¹ The counterargument is that, in Hart’s argument, what Nozick refers to as “the point” of special rights to force others to behave as they are obligated can only be illustrated against a background of a general right not to be forced; this would, argues Nozick, entail that “only against a background of *permitted* forcing can we understand the point of *general* rights,” which would undercut Hart’s claim (Hart, “Are There Any Natural Rights?,” p. 90) that “in the case of special rights as well as of general rights recognition of them implies the recognition of the equal right of all men to be free” (Nozick, p. 92). The general force of Nozick’s response is that “rights of forcing” are not strictly deducible as corollaries from obligations, even of the strong sort (distinguished from moral duties) that Hart is describing. An alternative account of special obligations that entails no corollary “right of forcing” is to be found in Randy Barnett (1986a and 1986b). In this “consent theory of contract” such obligations are to respect property titles that have been transferred by contractual means, rather than to perform certain actions; in the case of nonfeasance the remedy is not specific performance, entailing a “right to force,” but transference of a property claim, for example, a non-performance bond.

⁵² Nozick, p. 95.

On its face, Nozick's counter assertion seems quite plausible, at least for most situations, especially considering the experience so many have had in airports with disciples of messianic religious or political leaders thrusting books or flowers into their hands and then insisting on "contributions." We cannot admit a theory of obligations that allows people gratuitously and unilaterally to impose benefits on others and thereby to create obligations with corollary "rights to force" on the part of the unilateral benefit-imposers.⁵³

But perhaps Rawls and others in the tradition of "hypothetical" social contract have only a certain class of benefits in mind: those from which one "cannot" escape. They may readily admit that the pressing of a flower into the hand of a traveling passenger does not generate an enforceable obligation on the part of the recipient to pay her "fair share" of the costs of producing and distributing the flowers. These goods are private goods, the enjoyment of which is best left to the voluntary choices of producers and consumers; that is so because such benefits are easy to "escape."⁵⁴

A significant literature has emerged defending the fairness principle on the grounds of the inescapability or presumptive benefits of public goods. Richard Arneson, for example, admits the force of Nozick's objections to the principle of fairness but argues that it can be "revised" such that obligations and corollary rights to force only arise in the case of "particular types of benefits"⁵⁵ that satisfy certain carefully specified criteria: the benefits must derive from the provision of "pure public goods," characterized by jointness of consumption, nonexcludability, and equal consumption by all members of the group.⁵⁶ In such cases, to allow "free riders to enjoy the benefit of the scheme without helping defray its cost ... [is] often morally repugnant."⁵⁷

Arneson carefully specifies the cases in which free-riders are to be distinguished or selected out from the class of potential beneficiaries, to wit, with respect to a

⁵³ It may be worth nothing that benefit-based obligations are infiltrating contract law itself, and not just political theory, as P. S. Atiyah, *The Rise and Fall of Freedom of Contract* (Oxford, 1979), pp. 764–70 shows.

⁵⁴ The language of "escaping" is borrowed from Albert Hirschman's treatment of public goods: "The distinguishing characteristic of these goods is not only that they *can* be consumed by anyone, but that there is *no escape* from consuming them unless one were to leave the community by which they were provided" (Albert Hirschman, *Exit, Voice, and Loyalty* (Cambridge, Mass., 1970), p. 101).

⁵⁵ Richard Arneson, "The Principle of Fairness and Free-Rider Problems," *Ethics*, 92 (July 1982): 616–33, p. 618.

⁵⁶ *Ibid.*, p. 618. Limiting the class of benefits to "pure public goods" obviates the necessity for voluntary acceptance, asserts Arneson, for, "once a pure public good is supplied to a group of persons, there cannot really be any voluntary acceptance or enjoyment of the benefit by individual consumers. One cannot voluntarily accept a good one cannot voluntarily reject" (p. 619). Attention should be paid to the emphasized terms: "*once* a pure public good is supplied"; and "a good one *cannot voluntarily reject*." The assumptions that the good has already been supplied and that one cannot escape it are crucial—and highly problematic—moves in the argument, for they foreclose the option of pre-contract excludability; they assume away the problem of *producing* goods by assuming that they already exist; and they may limit the applicability of the principle to a very tiny class of "inherited" goods that are found (assuming no investment in the finding) and not produced.

⁵⁷ *Ibid.*, p. 621.

group of beneficiaries “G” of a good “B,” it is specified: (a) that B is genuinely collective with respect to G; (b) that the benefits to each member of G are greater than her “fair share” of the costs of supplying them;⁵⁸ (c) that “no beneficiary who has a disinterested motive for not contributing ... is required to contribute”;⁵⁹ (d) that it is unfeasible to provide B through the supply of private benefits to “each

⁵⁸ Note that application of this criteria to providers of collective goods could create havoc for, among others, the airline, cinema, and hotel industries, which engage in “discriminatory pricing” as part of a profit-maximizing strategy, charging often widely different rates for physically identical benefits, such as a hotel room or a seat in an airplane or a cinema. (The losers in such schemes are often professional business travelers, who must book in advance and spend a Saturday night away from family, in the case of the travel industry, and young to middle-aged movie goers, who are charged more than children, students, and senior citizens.) The existence of the hotel facilities (corridors, elevators, security, and so on) and the fact of the plane’s flight and the movie’s showing, regardless of whether it is empty, half-full, or full, are clearly collective goods with respect to those staying in rooms or sitting in seats; however, discriminatory pricing, that is, a clear case of “unfair shares,” are among the most effective means of providing such collective goods by charging the most to those who receive the most benefit from the nonrivalrous good, as measured by willingness to pay.

⁵⁹ Arneson explains in a footnote that that criterion is “intended to ensure that the principle of fairness will not lay obligations upon those who are genuinely conscientious objectors to the scheme for supplying public goods. My understanding of the requirement is that, in order to have a disinterested motive, the beliefs which give rise to the motive cannot be acquired or sustained in a culpably irrational fashion.” How one would determine “the beliefs that give rise to the motive” is left quite unclear, except for his explanation (Arneson, “The Principle of Fairness and Free-Rider Problem”, p. 632) that “If Jones has a deeply entrenched belief grossly at variance with the facts, and this counts as negligent or culpable ignorance, his obligation stands. Just having bizarre beliefs about the origins of the collective benefits one enjoys does not relieve one of the obligation to pay one’s fair share.” Thus, it might seem that the Amish belief that the Social Security System in the United States is a retirement plan (given current actuarial facts, this certainly may qualify as a “bizarre belief” by itself) and their view that divine providence requires their abstention from commercial insurance and retirement plans because participation evidences a lack of faith in providence, “does not relieve [them] of the obligation to pay [their] fair share.” That entails, then, a right on behalf of admitted beneficiaries to override *stated* preferences of others in pursuit of their own benefit; as David Schmidtz remarks of what he calls “honest holdouts” and Arneson refers to as “genuinely conscientious objectors,” after discussing the difficulties of distinguishing free riders who are strategically passing themselves off as honest holdouts: “The prospect of justifying institutionalized coercion as a solution to public goods problems depends upon whether there is a prior justification for engaging in actions that can be expected to force some people (perhaps a minority, but not necessarily so) to help pay for other people’s projects. Without this prior justification, the argument begs a most important question: Coercion may be necessary to force honest holdouts to pay for other people’s projects, but what exactly is it that makes pursuing this goal permissible in the first place? Merely pointing out that coercing people is the most efficient way for others to get what they want does not even begin to discharge the burden of proof” (David Schmidtz, *The Limits of Government: An Essay on the Public Goods Argument* (Oxford, 1991), p. 84).

beneficiary” sufficient to induce her contribution;⁶⁰ (e) that each person finds her “fair share of the costs” to entail a disutility;⁶¹ (f) that “no member’s choice is made under the expectation that it will influence any other member’s choice”;⁶² (g) that “No single member of G ... [nor] any coalition of a few members of G find it possible to divide the costs of B among the members of the coalition so that each member of the coalition will find the benefits of B to him outweigh the cost to him of contributing toward the supply of B according to the terms of the coalition. A large number of persons must contribute toward the supply of B if the benefits each receives are to overbalance the cost of each one’s contribution.”⁶³

⁶⁰ That would eliminate the enormous category of public goods that can be provided through “tie-ins” with private goods. See for instance, Daniel Klein, “Tie-Ins and the Market Provision of Public Goods,” *Harvard Journal of Law and Public Policy*, 10 (1987): 451–74 and for a classic case study Ronald Coase, “The Lighthouse in Economics,” *Journal of Law and Economics*, 17 (October 1974): 357–76.

⁶¹ That seems to exclude cases where altruists might derive pleasure from the sheer knowledge that they were helping to bring about a pure public good. How we would know what the motivations of contributors are is not explained, and may render this criterion non-operational.

⁶² That criterion is a bit puzzling. Arneson (in Arneson, “The Principle of Fairness and Free-Rider Problems”) offers in a footnote the explanation that “No single individual’s decision is expected to influence the decisions of others, but note that this is compatible with individuals basing their choices on expectations about what the aggregate of others will decide.” If the aggregate of others is intended to mean unanimity-minus-yourself, it would mean that expectation of unanimous contribution would be an acceptable motivation to contribute, but the knowledge that another has agreed to contribute on the condition that you do would not be an acceptable motivation, presumably because, unless it is a two person game, there might still be free-riders who are outside of the agreement. Such a requirement would eliminate those solutions to collective goods problems set forth in Earl R. Brubaker: “it seems worth proposing an additional rival to the free-rider hypothesis, and for convenience it might be called the golden rule of revelation. It asserts that under pre-contract group excludability the dominant tendency will be for each individual to reveal accurately his preference for a collective good provided that he has some assurance that others will match his offer in amounts he perceives as appropriate” (Earl R. Brubaker, “Free Ride, Free Revelation, or Golden Rule?,” *Journal of Law and Economics*, 18 (April 1975): 147–61, reprinted in Tyler Cowen (ed.), *The Theory of Market Failure* (Fairfax, Va., 1988), pp. 99–100). That thesis has been expanded and provided additional evidence from experimental economics by Schmidt; see especially chapters 4 and 5 on “conditionally binding assurance contracts.” Such contracts would only meet Arneson’s criterion “f” if the group of contractors included each and every beneficiary of the pure public good and the decision were made simultaneously and with no stepwise movement toward unanimity. That strikes me as a strange and unreasonable limitation on the provision of public goods.

⁶³ That seems to rule out voluntary solutions to public goods problems that exhibit the “straddle payoff” considered by de Jasay (in Anthony de Jasay, *Social Contract, Free Ride: A Study of the Public Goods Problem* (Oxford, 1989)), under which “under certain conditions being a ‘sucker’ is superior to having the exchange regime” (p. 144), a condition found more commonly than one would expect were one looking through lenses ground to Paul Samuelson’s specifications. See, for example, the study of voluntary and non-patent protected industrial and technological innovation in Tom G. Palmer, “Intellectual Property: A Non-Posnerian Law and

Under those conditions, according to Arneson, “each person who benefits from the cooperative scheme supplying B can correctly reason as follows: either other persons will contribute sufficient amounts to assure continued provision of B, or they will not. In either case, the individual is better off if he does not contribute If this reasoning induces an individual not to contribute, he counts as a free rider.”⁶⁴ Free riding is, according to Arneson, unfair and therefore activates the “rights to force” entailed by the fairness principle.

Not only has the range of applicability of the principle been limited to a very small set (perhaps the null set) of goods, leaving obligations to do our “fair share” for all of the rest unjustified (at least by that principle), but perhaps more importantly, he has made it very difficult to know precisely whether any good actually meets those criteria. That is so not only because it requires asking questions about motivations and “reasoning,” which are notoriously difficult to ascertain, but because if the good *is already produced* by the state and funded coercively on the basis of the principle of fairness, we cannot know whether it *could* have been produced through discriminatory pricing, tie-ins, altruism, conditionally binding assurance contracts (for example, “matching pledges”), or “straddle payoff” strategies.⁶⁵ Especially considering that such factors as technology, motivation, and expectations figure in Arneson’s criteria, it is difficult to imagine how we would determine whether a good should be provided collectively or privately, in the absence of the competition made possible by voluntary choice,⁶⁶ and, if collectively, whether it satisfies the criteria set out by Arneson and that we should therefore conclude that the burdens of its production should be shared on the basis of the principle of fairness.⁶⁷

Economics Approach,” *Hamline Law Review*, 12 (Spring 1989): 261–304, reprinted in Adam Moore (ed.), *Intellectual Property: Moral, Legal, and International Dilemmas* (New York, 1998), pp. 179–224; Jack Hirschleifer, “The Private and Social Value of Information and the Reward to Inventive Activity,” *American Economic Review*, 61 (1971): 561–73; and George Bittlingmayer, “Property Rights, Progress, and the Aircraft Patent Agreement,” *Journal of Law and Economics*, 76 (1986): 227–48.

⁶⁴ Arneson, “The Principle of Fairness and Free-Rider Problems,” p. 622.

⁶⁵ Actually, the last would, apparently, be considered “unfair” and hence unjust to those who produce the good. Contrast with Arneson’s conclusion of injustice the interesting discussion in Jasay, *Social Contract, Free Ride: A Study of the Public Goods Problem*, chapters 9 and 10. Jasay reaches rather different conclusions from Arneson. “Straddle payoff” games have structures similar to “battle-of-the-sexes” coordination games; for a description of the latter, see Jean Hampton, *Hobbes and the Social Contract Tradition* (Cambridge, 1986), pp. 151–61.

⁶⁶ See F. A. Hayek, “Competition as a Discovery Procedure,” in F. A. Hayek, *New Studies in Philosophy, Politics, Economics and the History of Ideas* (Chicago, 1978), pp. 179–90.

⁶⁷ The criteria might even limit the class of goods subject to the principle to those *already* produced (or in a sense inherited), and since the good is *already produced*, there is no problem of allocating the burdens of its production. What might arise, however, would be problems of equal or universal access to such already-produced collective goods. An example might be the so-called Royal Roads of England, so-called not because the crown produced them but because the crown seized them (they were produced centuries before by the Romans). Issues of access to collective goods are discussed in Carol Rose, “The Comedy of the Commons:

We should remember, further, that the difference between a collective good and a private good is not an inherent quality of the good as such, at least for most conceivable goods, namely those for which the publicness is not definitionally a part of the good (as publicness would be, for example, for the good of “republicanism”).⁶⁸ As Kenneth Goldin has argued:

The evidence suggests that we are not faced with a set of goods and services which have the inherent characteristics of public goods. Rather, we are faced with an unavoidable choice regarding every good or service: shall everyone have equal access to that service (in which case the service will be similar to a public good) or shall the service be available selectively: to some, but not to others? In practice, public goods theory is often used in such a way that one overlooks this important choice problem.⁶⁹

Thus, collective or private provision may be a matter of choice, and not of the inherent characteristics of the good as such. The choice may indeed be a response to high costs of exclusion or of contractual mechanisms in some cases, but it is *no less a choice* for that.⁷⁰ To say that the publicness of the good entails an obligation is to say that *the choice to produce it publicly* is what determines the obligation, which simply means that the *choice* of the majority can obligate the minority (or vice versa), thus undermining the claim that it is a problem of fairness *simpliciter* that generates obligations and rights to force non-contributors.⁷¹ To take the constantly raised “hardest” case, the current political system of nation states, for example, conjures up “units” of “national” defense when it is not at all clear that citizens of, say, Hawaii and New Hampshire are naturally “enjoying the same good of defense,” such that citizens of one may be coerced on the basis of the principle of fairness into paying

Custom, Commerce, and Inherently Public Property,” *University of Chicago Law Review*, 53/3 (Summer 1986): 711–81. I disagree, however, with the range of goods to which Rose’s reasoning might apply: I do not see how her arguments apply to goods that must be produced, rather than merely found without any opportunity cost incurred in the finding.

⁶⁸ To define the good as including state (or “public”) action is to rig the game; goods such as community, which people may actually desire, should not be defined in terms of how they are produced, as doing so renders the conclusion that they can only be produced in that way merely trivially true.

⁶⁹ Kenneth D. Goldin, “Equal Access vs. Selective Access: A Critique of Public Goods Theory,” *Public Choice*, 29 (Spring 1977): 53–71, p. 68.

⁷⁰ If there are good reasons for collective provision, then there are benefits to the parties that can be offered to induce holdouts to become participants. If the response is that the beneficiaries of public provision don’t *want* to forgo some of their benefits in order to induce others to cooperate, then one response may be that they have no right to force others to “cooperate” merely in order to gratify *their own* desires at the lowest cost.

⁷¹ Jasay notes that it is not “logistics” but “social choice” that determines which goods will be provided publicly: “Instead of being public because it was non-excludable, a good became non-excluded because the ‘public’ wanted it so” (Jasay, *Social Contract, Free Ride: A Study of the Public Goods Problem*, p. 156). See also the treatment in Antony de Jasay, *Justice and Its Surroundings* (Indianapolis, 2002).

for the defense of the other.⁷² Arneson attempts to meet Nozick's objection to the fairness principle by arguing that "a principle that Nozick cannot disavow without disavowing central commitments of his political philosophy requires acceptance of a revised principle of fairness."⁷³ Arneson's argument is that, if appropriation of property is to be justified, not on the basis of universal consent, but on the basis of the benefits generated even for those who face diminished opportunities of common use, then the principle of benefit can certainly be used to generate obligations without voluntary consent. While it is not clear that Nozick would agree with this argument, there is some evidence in the text of Locke to support a Lockean "fairness principle."⁷⁴ It is true that there is an implicit appeal to fairness at work as a part of Locke's account of political obligation,⁷⁵ but it is less clear that his argument for rights of appropriators to protect their justly appropriated property and exclude others from its enjoyment must rest solely or even partially on the *benefits* to the excluded; Locke argues strenuously that they are not *harmed* by this appropriation, but that they are equally *benefited* is not necessary, since Locke is seeking to avoid entirely the problem of "Consent of all the Commoners": "And the taking of this or that part, does not depend on the express consent of all the Commoners."⁷⁶ Indeed, one can argue not only that the Lockean enterprise is not about demonstrating the benefits of appropriation to all the nonappropriators, but that it is the benefit

⁷² Of course, it should also be noted that one person's benefit may be another person's bane. As Albert Hirschman notes, "[H]e who says public goods says public evils. The latter result not only from universally sensed inadequacies in the supply of public goods, but from the fact that what is a public good for some—say, a plentiful supply of police dogs and atomic bombs—may well be judged a public evil by others in the same community" (Hirshman, p. 101).

⁷³ Arneson, "The Principle of Fairness and Free-Rider Problems," p. 617.

⁷⁴ As Locke argues in the *Second Treatise* of the transfer of rights of enforcement of the law of nature from the individual to the political society: "For being now in a new State, wherein he is to enjoy many Conveniences from the labour, assistance, and society of others in the same Community, as well as protection from its whole strength; he is to part with as much of his natural liberty in providing for himself, as the good, prosperity, and safety of the society shall require: which is not only necessary, but just; since the other Members of the Society do the like" (John Locke, *Two Treatises of Government*, Peter Laslett (ed.) (Cambridge, 1988) II, §130, p. 353).

⁷⁵ See the treatment of this issue generally from a Lockean perspective in Simmons (1993), pp. 248–60.

⁷⁶ Locke (1988), II, §28, p. 289. Locke's contractarian argument is limited to political institutions, and not to moral institutions or practices generally. He is a "state contractarian" and not a "moral contractarian", to use Jean Hampton's terminology in her essay "Two Faces of Contractarian Thought" (in Peter Vallentyne (ed.), *Contractarianism and Rational Choice* (Cambridge, 1991), pp. 31–55). Locke, in fact, explicitly *rejects* "moral contractarianism" by insisting that the principle of property in one's person and subsequent appropriation of unowned resources "do not depend on the express consent of all the Commoners." Contractual agreement is used only to justify—and limit—political institutions. The realms of justice, morality, and the state for Locke are not, as they seem to be for Rawls, virtually co-extensive. As Hampton remarks, "Interestingly, Rawls does not even question the state's legitimacy in *A Theory of Justice*" (p. 32).

to *appropriators* that is the central issue; Locke shows that “commoners” should *become* appropriators both in their own interest and, of great if not equal importance, in order to secure freedom:

[Y]et being given for the use of Men, there must of necessity be a means *to appropriate* them some way or other before they can be of any use, or at all beneficial to any particular Man. The Fruit, or Venison, which nourishes the wild *Indian*, who knows no Inclosure, and is still a Tenant in common, must be his, and so his, *i.e.*, a part of him, that another can no longer have any right to it, before it can do him any good for the support of his Life.⁷⁷

The key phrase is “a part of him,” for “Though the Earth, and all inferior Creatures be common to all Men, yet every Man has a Property in his own Person. This no Body has any Right to but himself.”⁷⁸ Pace Rawls, our fundamental rights are to our bodies, and those are rights that are in no way held in common with others (that is, each person owns her own body and only her body, rather than a share in common rights to the use of all bodies). Without property in corporeal objects there is no way to delineate for embodied persons the spheres of action that constitute individual liberty, within which it would be unjust to initiate violence.⁷⁹

The generation of benefits for others need not figure prominently in a Lockean justification of rights of appropriators, and therefore endorsing a Lockean defense of property need not entail endorsement of the principle of benefit-based obligations. Arneson’s appeal to the text of Locke fails to provide support for his revised principle of fairness.

To deal with the problem of “honest holdouts,” George Klosko has sought to ground fairness-based obligations by identifying “presumptively beneficial” goods, which “must be desired by rational individuals regardless of whatever else they desire, though even this account presupposes a background of generally accepted values and beliefs.”⁸⁰ Identifying such goods, Klosko argues, negates the validity of claims on the part of non-contributors to be “honest holdouts” or “genuinely conscientious objectors.” According to Klosko:

There can be little doubt that the citizens of many modern states enjoy nonexcludable presumptive goods that depend upon the cooperative efforts of their fellows. Among such goods are ... national defense, protection from environmental hazards, and public health assurances. I think it is clear that individuals are obligated to contribute to the provision of these goods ... the principle of fairness presents an attractive account of the nature of their obligation.⁸¹

⁷⁷ Locke II, §26, p. 287.

⁷⁸ *Ibid.* II, §27, p. 287.

⁷⁹ For one Lockean grounding of natural property rights in property in one’s person, see Samuel C. Wheeler, III, “Natural Property Rights as Body Rights,” *Noûs*, 14/2 (1980): 171–94.

⁸⁰ George Klosko, “The Principle of Fairness and Political Obligation,” *Ethics* 97 (January 1987): 353–62, p. 353.

⁸¹ *Ibid.*, p. 356.

Klosko *asserts* that the principle of fairness provides an “attractive account” of the obligation he claims, but he gives no *argument* for it. The closest that he comes is a citation of a series of studies, mainly of American subjects, purporting to show that “people’s feelings that they are obligated to bear various burdens are significantly affected by their assessment of the burdens that other individuals are bearing or willing to bear.”⁸² This may show that some—or even most—Americans believe in the principle of fairness and “doing one’s share” in paying taxes, but it is hardly an *argument* for “doing one’s fair share” being obligatory or coercively enforceable. It may simply reflect the fact of decades of state control of education, and not of some universally (or even widely) intuited truth.

The dependence of “publicness” on choice, rather than on inherent characteristics of goods, leads us to the conclusion that the principle of fairness does not apply to goods from which we *cannot* escape, but to goods from which we will not be *allowed* to escape. In Arneson’s terms, “once a public good is supplied,” that means, after a decision has been made by some group with the power to make such decisions, then “One cannot voluntarily accept a good one cannot voluntarily reject.”⁸³ The “principle of fairness” depends on closure of exit options, on a *choice* by some not to allow others to evade alleged benefits.

A Hypothetical Contract with People you Cannot Escape

It is central to the Rawlsian social contractarian enterprise that parties to the contract not be allowed to escape from either its benefits or its burdens, and that the terms of the social contract must reflect that inescapability. The consequences of that inescapability are significant. If, for example, you cannot escape from your enemies in order to associate only with your friends, or at least with those who bear you no ill will, then it makes sense to establish rules of a rather different cast from those you would establish if you *could* escape those who are liable to make various coercive demands on you. Rawls raises that very scenario in *A Theory of Justice*,

⁸² Ibid., p. 357.

⁸³ Arneson, “The Principle of Fairness and Free-Rider Problems,” p. 619. There may be some externalities from which we cannot, in fact, escape that may pose fairness problems; they are negative externalities—such as air or noise pollution—for which there are threshold effects, for example two people enjoying wood fires or two people playing rock music is not a negative externality, but three people doing so is. A property rights approach can solve the problem of such externalities by enforcing rights against coercive interference with our property rights (for instance tort law) and by allowing contractual internalization of the externalities, but those activities below the threshold necessary to produce a rights-violating externality could be considered allocatable on the basis of fair access (for example by taking turns or auctioning off the rights), since strong Kantian-like solutions (“nobody may enjoy log fires, because if everybody did it would be disastrous”) are clearly sub-optimal, for *some* log-burning or rock-music playing below the rights-violating-externality-producing threshold produces goods for some without harming others. Griffin (in James Griffin, “Some Problems of Fairness,” *Ethics* 96 (October 1985): 100–118) suggests fairness-based solutions to such cases.

[T]he two principles are those a person would choose for the design of a society in which his enemy is to assign him his place.⁸⁴

Perhaps a better way to “protect themselves against such a contingency” would be to have exit rights from the start so that parties could escape from situations in which their enemies might assign them their places. What is remarkable about Rawls’s formulation is that the supposition that our enemies will assign us our places in society, even if only an “as if” constraint in the choice situation, implies a concern for the strategic interaction of rationally maximizing parties with conflicting interests and no preexisting moral restraint, something that is denied in Rawls’s work. Since we know that the constraints on the original position entail complete symmetry of interests, knowledge, and the like, we “can view the choice in the original position from the standpoint of one person selected at random.”⁸⁵ The consequence of that symmetry and of reduction to the situation of a single individual choosing among principles is that “the parties have no basis for bargaining in the usual sense.”⁸⁶ The ignorance of the parties

makes possible a unanimous choice of a particular conception of justice. Without these limitations on knowledge the bargaining problem of the original position would be hopelessly complicated. Even if theoretically a solution were to exist, we would not, at present anyway, be able to determine it.⁸⁷

Yet Rawls supposes that a party should adopt a conservative strategy, “reasoning as if” the party were coping with a situation in which “his enemy is to assign him his place.” The allegedly “hopelessly complicated bargaining problem,” contrary to Rawls’s assertion, is amenable to solution by specifying a suitable set of initial endowments (along the lines of Lockean self proprietorship) and allowing free exit rights in accordance with principles of voluntary association. (That would be to take the proposal of David Gauthier in *Morals by Agreement*⁸⁸ a step further, not only establishing a non-agreement point to serve as a baseline for distribution of the cooperative surplus over non-agreement, but introducing as well the very realistic option of choosing from among a variety of contracting groups.)

In his later work, Rawls spells out in detail why in the choice situation one can not be *allowed* to escape and how one should reason to principles accordingly. In *Political Liberalism*, Rawls takes great pains to differentiate “a well-ordered democratic society” from a “community” and from an “association.”⁸⁹ The form of cooperative endeavor about which Rawls is writing “is closed ... in that entry into it is only by birth and exit from it is only by death.”⁹⁰ Further, “we are not seen as

⁸⁴ Rawls, *A Theory of Justice*, p. 152. Rawls goes on (p. 153) to assert that the parties “should not reason from false premises,” but that the parties are still to reason as if they “were forced to protect themselves against such a contingency.”

⁸⁵ *Ibid.*, p. 139.

⁸⁶ *Ibid.*, p. 139.

⁸⁷ *Ibid.*, p. 140.

⁸⁸ David Gauthier, *Morals by Agreement* (Oxford, 1986).

⁸⁹ Rawls, *Political Liberalism*, p. 40.

⁹⁰ *Ibid.*, pp. 40–41.

joining society at the age of reason, as we might join an association, but as being born into society where we will lead a complete life.”⁹¹

Justice and Exit Rights

Rawls rejects the exit option in the specification of the choice situation for two primary reasons. First, exit rights cannot be cited to warrant the justice of an arrangement. Thus, “While the principles adopted will no doubt allow for emigration (subject to suitable qualifications), they will not permit arrangements that would be just only if emigration were allowed.”⁹² Rawls writes: “Political society is closed: we come to be within it and we do not, and indeed cannot, enter or leave it voluntarily.”⁹³ The reason is that “the right of emigration does not make the acceptance of political authority voluntary in the way that freedom of thought and liberty of conscience make the acceptance of ecclesiastical authority voluntary.”⁹⁴ Thus, “The political is distinct from the associational, which is voluntary in ways that the political is not.”⁹⁵

Rawls may have in his crosshairs the theory of John Locke, who offered as an argument for an obligation to submit to certain commands of government the claim that continued residence on a territory attached to a political society and governed by a government to which the members of that political society had entrusted their executive powers, and from which exit was not blocked by force, constituted tacit consent to submit. David Hume accused Locke and his followers of building “a tory consequence of passive obedience, on a whig foundation of the original contract.”⁹⁶ But it is a mistake to assume that advancing a Lockean or Lockean-type argument entails asserting that exit rights *guarantee* the justice of a political arrangement; other conditions may also be necessary. If, for example, you were deliberately to create a dependency of some sort in another party through coercive means (for example by injecting a poison to which only you have the antidote), and only after that had been accomplished were you to unlock the door and offer to allow her to leave, the resulting dependent relationship would not be warranted as just merely because of the now available exit option. Exit options alone do not guarantee justice, but they do, in a Lockean-type argument, satisfy three desirable purposes:

⁹¹ Ibid., p. 41.

⁹² Ibid., p. 277. Rawls does not tell us why the principles adopted would “no doubt” allow for emigration, nor why any “suitable qualifications” would be necessary, unless all of the justificatory work is being done by the term “suitable,” in which case “no qualifications” might be consistent with “allowing” emigration.

⁹³ Ibid., p. 136.

⁹⁴ Ibid., p. 136.

⁹⁵ Ibid., p. 137.

⁹⁶ David Hume, “Of the Original Contract,” in David Hume, *Essays, Moral, Political, and Literary* (1777; Indianapolis: Liberty Classics, 1987), p. 487.

1. they are just in themselves, as recognitions of a fundamental human right;
2. they are an engine of further liberalization and extension of human freedom and prosperity, by constraining political associations and authorities to recognize the rights of subjects or risk losing their human capital to other states and systems that offer more attractive arrangements; and
3. they can warrant the justice of political arrangements that *are* entered into voluntarily and in accordance with just procedures (i.e., they serve, in effect, as a background requirement for “justice preserving transformations” in matters of political order).⁹⁷

To avoid the “tory consequences” of which Hume warned, it is only necessary to specify further conditions relating to how one enters into a political society and what is necessary to exit it.⁹⁸ Eliminating entry and exit options from the choice situation is unwarranted, adds a needless element of *ir*reality, and contradicts a fundamental liberal principle.

The right to leave does not, of course, entail any correspondingly symmetrical right to enter. My right to quit my membership in a chess club need not entail a right to be accepted by any other chess club. In a world governed by cosmopolitan principles, it is possible (but quite unlikely) that no one would be willing to associate with or accept some particular emigrant. In the world of national states, such as largely exists today and as almost all other theorists posit, exit rights actually are restricted in some places and entrance rights are tightly controlled everywhere. For example, to keep out potential cooperators (or recipients of taxpayer funded benefits) the U.S. government apprehends (and sometimes kills) would-be immigrants along the U.S.-Mexico border, despite the willingness of other American citizens (and land owners) to receive them. Although it is logically possible that no current U.S. citizen would accept immigrants on to their land, it does not seem likely, and certainly far, far, far less likely than the stark certainty of encountering the barbed wire, guard dogs, and armed guards currently deployed by nation states.

Natural Assets and Desert

The second motivation for Rawls’s making closure move explicit seems to be to meet David Gauthier’s objection to Rawls’s formulation of the redistributive difference principle on the grounds that it does not meet the criterion of mutual advantage. Gauthier challenges Rawls’s assertion that natural assets are “not deserved,” noting

⁹⁷ On the idea of justice preserving transformations, see Nozick, pp. 150–53.

⁹⁸ For example, territorial monopoly combined with inalienability of land, at least with respect to the political society to which it is attached, may be inconsistent with a properly Lockean or Lockean-type liberalism that seeks to avoid “tory consequences.” One could, for example, exit from the state without having to leave the territory or piece of land on which one currently lives. For such proposals, see: Johann Gottlieb Fichte, *Beitrag zur Berichtigung der Urtheile des Publikums über die Französische Revolution*, Reinhard Strecker (ed.) (1793; Leipzig, 1922); Herbert Spencer, *Social Statics* (1850; New York, 1970), pp. 185–93, and Randy Barnett, *The Structure of Liberty: Justice and the Rule of Law* (Oxford, 1998).

that although “they are not undeserved, they are not contrary to desert.”⁹⁹ Gauthier asks why, if mutual advantage and contractual agreement are to be the bases of the principles of justice, parties would not adopt a difference principle that would guarantee each an improvement over their non-agreement points. Writes Gauthier,

Behind the veil of ignorance, no one knows her natural abilities and talents, and hence no one knows what she would get in the absence of agreement. Yet each knows that she has certain natural abilities and talents, and that people differ in this endowment, so that in the absence of agreement people would secure different levels of well being. It is therefore possible for everyone to take account of the “no agreement point” in their reasoning, even though no particular person knows how it will affect her.¹⁰⁰

In response, Rawls rejects his earlier insistence on mutual benefit, by jettisoning entirely the options of no agreement and of comparison of alternative agreements, and therefore of bargaining. For the “rational” Rawls substitutes the “reasonable,” which is certainly a less exacting and less rigorous concept. In a defense of that move, Samuel Freeman, in a paper cited approvingly by Rawls in *Political Liberalism*, combines the highly questionable assertion of the inescapability of one’s “given” society with an explanation of reasonableness:

What is inescapable ... is one’s being a member of a social group, recognizing the group’s system of norms, and understanding how the norms function as reasons in public argument. Though that requirement does not involve endorsing all the norms of the group, whatever they may be, it does mean that one who is unwilling to cooperate with others on *any* terms except those most conducive to achieving his own particular ends is being *unreasonable*.¹⁰¹

Further, “For Rawls there is no place for comparing the benefits and burdens of cooperation with cooperation-free interaction ... noncooperation is not a viable option for us.”¹⁰²

⁹⁹ David Gauthier, “Justice and Natural Endowment: Toward a Critique of Rawls’s Ideological Framework,” in David Gauthier, *Moral Dealing: Contract, Ethics, and Reason* (Ithaca, N.Y., 1990), p. 161. Robert Nozick makes an objection similar to that of Gauthier, but it is not followed up in such detail and Rawls seems nowhere ever to address it. Nozick questions whether the fruits of cooperation to be distributed on the basis of the difference principle are to include the total product, or the total product minus “what each individual gets acting separately.” If the “noncooperative shares” are to be considered, “we should note that this certainly is not how people entering into cooperation with one another would agree to conceive the problem of dividing up cooperation’s benefits” (Nozick, pp. 184–5).

¹⁰⁰ Ibid., p. 154. It is worth asking why “no agreement,” rather than a range of alternative agreements, is the appropriate alternative to “agreement.” Rawls sees the point in his response to Gauthier and not only insists that we *must* agree, but that we must come to *one* agreement, and one only, without being allowed to consider the possibility of entering into alternative contractual groups with different principles of association.

¹⁰¹ Samuel Freeman, “Reason and Agreement in Social Contract Views,” *Philosophy and Public Affairs*, 19 (Spring 1990): 122–57, p. 132.

¹⁰² Ibid., p. 136.

Rawls associates the political exclusively with territorial monopoly, which is not a timeless feature of political association, whether past or present, and conflates state, nation, and society. As Rawls notes in contrasting the associational and the political, “the government’s authority cannot be evaded except by leaving the territory over which it governs, and not always then.”¹⁰³ (The claim is clearly false. Every day hundreds of millions of people—possibly billions—evade governmental authority without ever leaving home.¹⁰⁴) Rawls concludes from that patently false claim that,

The government’s authority cannot, then, be freely accepted in the sense that the bonds of society and culture, of history and social place of origin, begin so early to shape our life and are normally so strong that the right of emigration (suitably qualified) does not suffice to make accepting its authority free, politically speaking, in the way that liberty of conscience suffices to make accepting ecclesiastical authority free, politically speaking.¹⁰⁵

The social contract is *hypothetical*, rather than actual, because it takes as *already preexisting and unquestionable* an inescapable political unit—the nation state, with its all embracing claim over the full life of the individual—and offers a purely hypothetical justification for its power. No alternative is or can even be considered.

Kantian Constructivism and Inescapable Groups

Immanuel Kant explains the nature of the hypothetical contract as follows:

we need by no means assume that this contract (*contractus originarius* or *pactum sociale*), based on a coalition of the wills of all private individuals in a nation to form a common, public will for the purposes of rightful legislation, actually exists as a *fact*, for it cannot possibly be so.¹⁰⁶

Despite his explicit criticisms of Hobbes, Kant posits a hypothetical social contract establishing an absolute state with an absolute sovereign. He is moved to do so by the specter of anarchy: “universal violence and the distress it produces must eventually make a people decide to submit to the coercion which reason itself prescribes (i.e., the coercion of public law), and to enter into a *civil* constitution.”¹⁰⁷ In the *Metaphysics of Morals*, he claims that

¹⁰³ Rawls, *Political Liberalism*, p. 222.

¹⁰⁴ In extreme cases, they withdraw from the state entirely. See Victor Azarya and Naomi Chazan, “Disengagement from the State in Africa: Reflections on the Experience of Ghana and Guinea,” *Comparative Studies in Society and History*, 29/1 (January 1987): 106–31.

¹⁰⁵ Rawls, *Political Liberalism*, p. 222. I have already dealt with the concern expressed here that exit rights do not by themselves guarantee or warrant the justice of a political arrangement. What is notable here is Rawls’s explicit identification of political order with a territorial monopoly and with a particular nation.

¹⁰⁶ Immanuel Kant, *Political Writings*, Hans Reiss (ed.), H. B. Nisbet (trans.) (Cambridge, 1992), p. 79.

¹⁰⁷ *Ibid.*, p. 90.

Experience teaches us the maxim that human beings act in a violent and malevolent manner, and that they tend to fight among themselves until an external coercive legislation supervenes. But it is not experience or any kind of factual knowledge which makes public legal coercion necessary. On the contrary, even if we imagine men to be as benevolent and law-abiding as we please, the *a priori* rational idea of a non-lawful state will still tell us that before a public and legal state is established, individual men, peoples and states can never be secure against acts of violence against one another, since each will have his own right to do *what seems right and good to him*, independent of the opinion of others.¹⁰⁸

Kant and Rawls are engaged in moral and political constructivism, an exercise in pure “autonomy.” As Rawls remarks,

The parties in the original position do not recognize any principles of justice as true or correct and so as antecedently given; their aim is simply to select the conception most rational for them, given their circumstances. This conception is not regarded as a workable approximation to the moral facts: there are no such moral facts to which the principles adopted could approximate.¹⁰⁹

Kant insists, as Rawls does after him, that the social contract presupposes people who cannot escape one another, that is with whom one “cannot avoid having intercourse”:

[T]he first decision the individual is obliged to make, if he does not wish to renounce all concepts of right, will be to adopt the principle that one must abandon the state of nature in which everyone follows his own desires, and unite with everyone else (with whom he cannot avoid having intercourse) in order to submit to external, public and lawful coercion.¹¹⁰

Thus, the account of social contract for Kant, as for his follower Rawls, presupposes the unavoidability of a natural or inevitable political unit—the nation-state—made up of people “with whom one cannot avoid having intercourse”. The logic of the Rawlsian hypothetical social contract dictates that that is, and must be, a closed group.

The choice to close off exit options in the specification of the choice situation of the social contract is supposed to produce fundamentally liberal conclusions, but it violates at the outset a fundamental liberal principle, the principle of exit rights and freedom of movement, and important elements of the liberal tradition, that is cosmopolitanism and internationalism. It is also starkly at odds with Rawls’s statement that people do have the right to emigrate, a statement offered *en passant* and without justification.¹¹¹ (Is Rawls claiming that when people do decide to emigrate, they should not “consider the attractions of other societies?” But, if so, on what basis would they decide whether to emigrate, and to where?)

¹⁰⁸ Ibid., p. 137.

¹⁰⁹ John Rawls, “Kantian Constructivism in Moral Theory: The Dewey Lectures,” *Journal of Philosophy*, 77 (1980): 515–72, p. 564.

¹¹⁰ Kant (1992), p. 137.

¹¹¹ Rawls, *Political Liberalism*, p. 68.

But why must persons acquiesce in any “given” political arrangement? Why should they not be free to choose those arrangements most conducive to their needs or desires, in voluntary association with others equally free, and to reject those that they find oppressive or unbearable. According to those following the logic of the Rawlsian project, “a coalition that withdraws from society renounces any claim to justice from those who remain.”¹¹² In other words, outside of the nation state, justice is simply unthinkable. Presumably that means that before there were nation states, there was no justice.

“Given” a justified, inevitable, inescapable collectivity (invariably a nation state) within which one comes to be and without which one would be nothing at all, that collectivity has complete ownership of all assets “within” it, none of which may be withdrawn. The problem of allocating the benefits of social cooperation is thereby

¹¹² Gibbard, “Constructing Justice,” p. 272. In pressing the logic of the difference principle toward an extreme enforced egalitarianism, Cohen dismisses trans-communal moral principles applicable between individuals of different “justificatory communities.” He argues that the existence of a “justificatory community” entails “a set of people among whom there prevails a norm (which need not always be satisfied) of comprehensive justification” (Cohen, “Incentives, Inequality, and Community,” p. 282) Further, “an argument for a policy satisfies the requirement of justificatory community, with respect to the people it mentions, only if it passes the interpersonal test. And if all arguments for the policy fail that test, then the policy itself evinces lack of justificatory community, whatever else might nevertheless be said in its favor” (ibid., p. 282). The “interpersonal test” “asks whether the argument could serve as a justification of a mooted policy when uttered by any member of society to any other member” (ibid., p. 280). Some arguments, Cohen asserts (but does not demonstrate) fail to meet this test: “The incentive argument does not serve as a justification of inequality on the lips of the talented rich” (ibid., p. 280). Cohen concludes from this that those who could not justify their behavior in this way fail the interpersonal test and that if they “do not think that they *need* to provide a justification, then they are forswearing community with the rest of us in respect of the policy issue in question. They are asking us to treat them like a set of Martians in the light of whose predictable aggressive, or even benign, behavior it is wise for us to take certain steps, but whom we should not expect to engage in justificatory dialogue” (ibid., p. 282). Cohen makes a significant unjustified move in attempting to press the extreme egalitarian case: he moves from the idea of failure of “comprehensive justificatory community,” defined by the interpersonal test, to failure of *all* “justificatory dialogue,” for which move he offers no argument, although one is clearly needed, as his notion of “comprehensive justification” does not encompass all forms of justification or of justificatory dialogue. Further, his “test” of whether justification fails is not as rigorous as it at first appears, for clearly “could not justify” does not mean “could not utter a claim with meaning,” but rather, something like “would feel queasy or uncomfortable or guilty making the claim if she were to make the claim.” That, however, is a matter of psychology, and intuitions and feelings could differ. Imagine a Cohenite untalented person making the following claim to a talented person (I specify “talented person” and not “talented rich person,” because there would be no rich people, only talented and untalented, in Cohen’s ideal world): “You must work longer and harder at no additional benefit to yourself, entirely so that I may benefit.” At least some untalented people would feel uncomfortable, queasy, or guilty in making such a claim. For an examination of the myriad errors and confusions in Cohen’s argument for radical egalitarian redistribution, see Tom G. Palmer, “G. A. Cohen on Self-Ownership, Property, and Equality,” *Critical Review*, 12/3 (Summer 1998): 225–51.

reduced to a pure bargaining problem. No person and no assets get out; no person and no assets get in (because they would have had to leave another collectivity, which is not allowed).¹¹³ But in order to institute the “fair” distribution, with all “social and economic inequalities . . . arranged so that they are both (a) to the greatest benefit of the least advantaged,”¹¹⁴ recourse must be had to withdrawal of assets the margin, which is excluded from the principles on the basis of which calculations are to be made in the first place.

Slamming shut the exit door creates the problem to which fairness is alleged to be the answer, but it also makes it impossible to put the solution into practice. Not only is the game rigged; it is not even possible to play it by the rules stipulated.

References

- Arneson, Richard, “The Principle of Fairness and Free-Rider Problems,” *Ethics*, 92 (July 1982): 616–33.
- , “Lockean Self-Ownership: Towards a Demolition,” *Political Studies*, 39 (1991): 36–54.
- , “Property Rights in Persons,” *Social Philosophy and Policy*, 9/1 (1992): 201–30.
- Atiyah, P. S., *The Rise and Fall of Freedom of Contract* (Oxford: Oxford University Press, 1979).
- Azarya, Victor and Chazan, Naomi, “Disengagement from the State in Africa: Reflections on the Experience of Ghana and Guinea,” *Comparative Studies in Society and History*, Vol. 29/1 (January 1987): 106–31.
- Barnett, Randy, “A Consent Theory of Contract,” *Columbia Law Review*, Vol. 86, No. 2 (March 1986).
- , “Contract Remedies and Inalienable Rights,” *Social Philosophy and Policy*, Vol. 4, No. 1 (Autumn 1986).
- , *The Structure of Liberty: Justice and the Rule of Law* (Oxford: Clarendon Press, 1998).
- Barry, Brian, *Theories of Justice* (London: Harvester-Wheatsheaf, 1989).
- Bittlingmayer, George, “Property Rights, Progress, and the Aircraft Patent Agreement,” *Journal of Law and Economics*, 76 (1986): 227–48.
- Brubaker, Earl R., “Free Ride, Free Revelation, or Golden Rule?,” *Journal of Law and Economics*; 18 (April 1975): 147–61, reprinted in Tyler Cowen (ed.), *The Theory of Market Failure* (Fairfax, Va.: George Mason University Press, 1988).
- Coase, Ronald, “The Lighthouse in Economics,” *Journal of Law and Economics*, 17

¹¹³ As I argued in section three above, all assets must be employed to their fullest. But then, without prices, no one could know what the “fullest” use of any resource would be. Prices require exchange, which requires property and the liberty to add or withdraw units at the margin, which is precluded by fiat. The enterprise rests for its fulfillment on the very principle—free withdrawal or addition of assets—that is precluded at the outset. See Ludwig von Mises, *Socialism: An Economic and Sociological Analysis* (1922; first English edn, 1936; London, 1972).

¹¹⁴ Rawls, *A Theory of Justice*, p. 83.

- (October 1974): 357–76.
- Cohen, G. A., “Incentives, Inequality, and Community,” in Grethe B. Peterson (ed.), *The Tanner Lectures on Human Values, Volume Thirteen* (Salt Lake City: University of Utah Press, 1992).
- Donagan, Alan, *Morality, Property, and Slavery* (The Lindley Lecture, 1980 [Lawrence, Kans: Philosophy Department of the University of Kansas, 1981]).
- Fichte, Johann Gottlieb, *Beitrag zur Berichtigung der Urteile des Publikums über die Französische Revolution*, Reinhard Strecker (ed.) (1793; Leipzig: Verlag von Felix Meiner, 1922).
- Freeman, Samuel, “Reason and Agreement in Social Contract Views,” *Philosophy and Public Affairs*, 19 (Spring 1990): 122–57.
- Gauthier, David, *Morals by Agreement* (Oxford: Oxford University Press, 1986).
- , “Justice and Natural Endowment: Toward a Critique of Rawls’s Ideological Framework,” in David Gauthier, *Moral Dealing: Contract, Ethics, and Reason* (Ithaca, N.Y.: Cornell University Press, 1990).
- Gibbard, Allan, “Constructing Justice,” *Philosophy and Public Affairs*, 20 (Summer 1991): 264–79.
- Goldin, Kenneth D., “Equal Access vs. Selective Access: A Critique of Public Goods Theory,” *Public Choice*, 29 (Spring 1977): 53–71.
- Grey, Thomas C., “The First Virtue,” *Stanford Law Review*, 25 (1973): 286–327.
- Griffin, James, “Some Problems of Fairness,” *Ethics* 96 (October 1985): 100–118.
- Hampton, Jean, *Hobbes and the Social Contract Tradition* (Cambridge: Cambridge University Press, 1986).
- , “Two Faces of Contractarian Thought,” in Peter Vallentyne (ed.), *Contractarianism and Rational Choice* (Cambridge: Cambridge University Press, 1991).
- Hardin, Russell, *Morality Within the Limits of Reason* (Chicago: University of Chicago Press, 1988).
- Hart, H. L. A., “Are There Any Natural Rights?,” *Philosophical Review*, 64 (1955): 175–91, reprinted in Jeremy Waldron (ed.), *Theories of Rights* (Oxford: Oxford University Press, 1984).
- Hayek, F. A., “Competition as a Discovery Procedure,” in F.A. Hayek, *New Studies in Philosophy, Politics, Economics and the History of Ideas* (Chicago: University of Chicago Press, 1978).
- Hirschleifer, Jack, “The Private and Social Value of Information and the Reward to Inventive Activity,” *American Economic Review*, 61 (1971): 561–73.
- Hirschman, Albert, *Exit, Voice, and Loyalty* (Cambridge, Mass.: Harvard University Press, 1970).
- Hume, David, “Of the Original Contract,” in David Hume, *Essays, Moral, Political, and Literary* (1777; Indianapolis: Liberty Classics, 1987).
- Jasay, Anthony de, *Social Contract, Free Ride: A Study of the Public Goods Problem* (Oxford: Oxford University Press, 1989).
- , *Justice and Its Surroundings* (Indianapolis: Liberty Fund, 2002).
- Kant, Immanuel, *Political Writings*, Hans Reiss (ed.), H. B. Nisbet (trans.) (Cambridge: Cambridge University Press, 1992).
- Klein, Daniel, “Tie-Ins and the Market Provision of Public Goods,” *Harvard Journal*

- of *Law and Public Policy*, 10 (1987): 451–74.
- Klosko, George, “The Principle of Fairness and Political Obligation,” *Ethics* 97 (January 1987): 353–62.
- Locke, John, *Two Treatises of Government*, Peter Laslett (ed.) (Cambridge: Cambridge University Press, 1988).
- Mises, Ludwig von, *Socialism: An Economic and Sociological Analysis* (1922; first English edn, 1936; London: Jonathan Cape, 1972).
- Moon, Parker T., *Imperialism and World Politics* (New York: The MacMillan Company, 1926).
- Musgrave, Richard A. and Peacock, Alan T. (eds), *Classics in the Theory of Public Finance* (2nd edn, New York: St. Martin’s Press, 1994).
- Narveson, Jan, “A Puzzle about Economic Justice in Rawls’s Theory,” *Social Theory and Practice*, 4 (1976): 1–27.
- , “Rawls on Equal Distribution of Wealth,” *Philosophia*, 17 (1977): 281–92.
- Nozick, Robert, *Anarchy, State, and Utopia* (New York: Basic Books, 1974).
- Palmer, Tom G., “Intellectual Property: A Non-Posnerian Law and Economics Approach,” *Hamline Law Review*, 12 (Spring 1989): 261–304, reprinted in Adam Moore (ed.), *Intellectual Property: Moral, Legal, and International Dilemmas* (New York: Rowman and Littlefield, 1998).
- , “G. A. Cohen on Self-Ownership, Property, and Equality,” *Critical Review*, 12/3 (Summer 1998): 225–51.
- Posner, Richard, *Economic Analysis of Law* (Boston: Little, Brown and Company, 1972).
- Pufendorf, Samuel, *The Political Writings of Samuel Pufendorf*, Craig L. Carr (ed.), Michael J. Seidler (trans.) (Oxford: Oxford University Press, 1994).
- Rawls, John, *A Theory of Justice* (Cambridge, Mass.: The Belknap Press of Harvard University, 1971).
- , “Kantian Constructivism in Moral Theory: The Dewey Lectures,” *Journal of Philosophy*, 77 (1980): 515–72.
- , *Political Liberalism* (New York: Columbia University Press, 1993).
- , “The Law of Peoples,” in Susan Shute and Susan Hurley (eds), *On Human Rights: The Oxford Amnesty Lectures 1993* (New York: Basic Books, 1993).
- Rose, Carol, “The Comedy of the Commons: Custom, Commerce, and Inherently Public Property,” *University of Chicago Law Review*, 53/3 (Summer 1986): 711–81.
- Scanlon, T. M., “Contractualism and Utilitarianism,” in Amartya Sen and Bernard Williams (eds), *Utilitarianism and Beyond* (Cambridge: Cambridge University Press, 1982).
- Schmidtz, David, *The Limits of Government: An Essay on the Public Goods Argument* (Oxford: Westview Press, 1991).
- Simmons, A. John, *The Lockean Theory of Rights* (Princeton: Princeton University Press, 1992).
- , *On the Edge of Anarchy: Locke, Consent, and the Limits of Society* (Princeton: Princeton University Press, 1993).
- Spencer, Herbert, *Social Statics* (1850; New York: The Robert Schalkenbach Foundation, 1970).
- Sterba, James, “From Liberty to Welfare,” *Ethics*, 105/1 (October 1994): 64–98.

- Vallentyne, Peter (ed.), *Contractarianism and Rational Choice* (Cambridge: Cambridge University Press, 1991)
- Wheeler, Samuel C., III, "Natural Property Rights as Body Rights," *Noûs* 14/2 (1980): 171–94.
- Wicksell, Knut, "A New Principle of Justice Taxation," in Richard A. Musgrave and Alan T. Peacock (eds), *Classics in the Theory of Public Finance* (2nd edn; New York: St. Martin's Press, 1994).