

Jan Narveson

Social Contract: The Last Word in Moral Theories

Abstract:

Most meta-ethical theories fail either for lack of real content or because they fail to make needed distinctions, or to give sufficient account of what a moral theory is about. Positing that values are intuited is useless or worse, since the very problem that gives rise to the need for morals is that people's values vary, greatly, from one to another, thus leading to conflict, and 'intuition' is no basis for interpersonal agreement if we have initial disagreement—as we surely seem to do. Conflicts can, of course, be steamrollered when the philosopher proposes that everybody should do things *his* way. But the trouble is that people may not (almost certainly will not) have motivation to conform to what the philosopher proposes. Yet motivation is of the essence. A moral theory that everyone is free to recognize with no discernible effects on his behavior is useless. The 'Social Contract' idea is, very simply, to account for morals by starting with the actual motivations and deliberations of individuals, then considering the effect of placing such individuals in a society of other, especially differing ones. We play our cards right if, in some relevant and useful sense, everyone can expect to do better by embracing moral constraints than not. Then the motivation problem is squarely faced from the start. The idea is that it will transfer to morals by virtue of the relations we can expect to have in society, given our various interests. This essay explores these issues, explaining why there is simply no alternative to the social contract idea.

Social contract theory in the past several centuries has been focused especially on government. The idea is that government is just if, and because, we *consent* to it. Or at least, that we would if we thought about it more. Well—or, at *very* least, that we would if we thought about it under certain specified idealized conditions. What gives this what plausibility it has would be the thought that we do better by accepting government—*provided that* everyone else do so as well. Thus the 'social contract'. Consent is driven by the putative favorability to each of the terms, provided that all others go along. The catch, that motivates most of the zillion or so words of discussion in the literature, lies in that little proviso.

Who are all those 'others'? The temptation is to make them all clones of the theorist, thus guaranteeing in advance the claimed agreement. They're behind, perhaps, a Veil of Ignorance, or they are perfectly informed or utterly impartial, or any number of other things that we all know perfectly well people are not.

But if our story is to have any plausibility whatever, our characterization of ‘all those others’ has to be true to life. If we impose any conditions, they have to be ones that people can actually meet, and would be motivated to meet.

To start with, we should give up on this classic idea of social contract for specifically political institutions. It is, I am inclined to think, hopeless, for reasons substantially like this one, from Kliemt (as quoted in Sugden 2009, 12):

“[T]here is absolutely nothing that can be done by government that would fulfill the norm of unanimous agreement in real fact.”

Government, says Sugden (2009, 16) himself, is “fundamentally a matter of mutual advantage”. Assuming he means by this the sort of government that could possibly get this kind of unanimous contractual acceptance, he’s right, of course—that’s what will be argued here. And he goes on, “[w]hat we should not do (but what the rhetoric of soft paternalism invites us to do) is to imagine that we are privileged advisors to a benevolent government, designing a set of interventions that will steer people in general towards the decisions that we think are best for them” (Sugden, 2009, 21). This too must be agreed, and is far too often not observed. Our fundamental starting point has to be people with *their* values *as they are*, not somebody else’s.

In other words, we start with fully liberal premisses. As Sugden goes on to say, “The conception of politics as voluntary exchange [...] imposes the discipline that policy proposals must be addressed to citizens as individuals, and must be justified to each of them as the means by which he or she can achieve what he or she wants. I share Buchanan’s conviction that this is the proper role of normative economics in a free society.” (22)

Other ‘foundations’ have the problem of other-worldliness. There is no point predicting the actions of people not found in reality. The question is only whether this conception is really compatible with any kind of government at all, including democracy. I rather suspect not, in fact. If we seek to engage the reason of all, then we can’t tailor our audience to our own specifications. ‘All’ takes in way too much for that. We need to show each person that *that person* could expect to gain, in the sense of expecting an improvement as measured by *that person’s* values, again with the caveat that this is providing all others subscribe to the same rule.

And Jasay says this about contract: “Linguistic cognateness is no excuse for ignoring the incompatibility in what two closely related terms signify. In convention, including its most stringent form, the contract, individual choice prevails throughout. The social contract, on the contrary, is an hypothetical agreement by all members of a society (the ‘all’ invoking an idealized unanimity, to be treated with Wicksellian tolerance) to adopt a meta-rule, a rule of rule-making or constitution, by which non-unanimous decisions can be reached that are binding for all. The social contract legitimizes collective choices to which all individuals have committed themselves to submit.” (Jasay 2010, 400) That is indeed the argument of those who favor government. But on both the Kliemt view and, really, Jasay’s, the prospects for government look very slim indeed.

Kliemt's is an important claim. On the face of it, it's enormously plausible. One could argue that it's really more than just plausible. After all, governments' actions are all devoted to imposing on someone or other, and *prima facie*, this virtually by definition means that those persons did not or do not consent. If the Social Contract is to look promising, it certainly needs to be made consistent with that fact. So how do we do that?

Sidetracking Government

The first thing to do is to follow Gauthier (1986) in unhooking social contract theory from *government* in particular. Instead, as its best contemporary exponents see it, what we agree to (if we do) are *principles*, and in fact moral principles—"morals by agreement" in the words of Gauthier's title. Moral principles are principles for the ordering or regulation of society generally. They are not necessarily or essentially promulgated nor administered, and certainly not legislated, by a central agency—which is what government claims to do. Morality is 'administered' not by one person or a committee, but by persons, all and sundry: A criticizes B on the basis of A's ideas about those rules—and A's might differ from C's, for example. Morality is notional, and actual only in the sense that (a) moral ideas both concern and often really do affect people's actions, and (b) every now and then, someone criticizes someone's behavior (his own or someone else's) on what purport to be moral grounds, and often that criticism in turn does have some effect. How much effect, and how great the probability of effect, varies hugely, as we know. (More will be said about that below.) Meanwhile, the idea of attempting to formulate a general program of restrictions to be placed on our and everyone's behavior, in principle by everyone, while difficult, is not fantastic, but useful. The idea of social contract, I suggest, is what makes morality useful—it explains why and how it is so.

What Makes Social Contract Tick

What is it about social contract that makes it so promising? Here's a partial answer, though not conclusive. When two people make a particular agreement, such as a legal contract, what they agree to is that they will do x, *provided that* the other party does y, x and y having been identified in the negotiation leading up to the agreement. If the first person acts while the other party has not already acted, and thus this other party is in the position of being able to benefit by defaulting on his own part of the bargain, there is an obvious problem, for it now looks as though default is the later-actor's best move. And sometimes default does happen, as we all know. But, as we will have frequent occasion to recall in this essay, we are discussing the foundations of *normative* principles.

It is characteristic of 'oughts' that they do not necessarily, and in general do not always, eventuate in 'is'. It can still be true of Jones that he ought to have done *x* even though he did not. What can't be true is that there is a good reply to the claim of 'ought' that 'not necessarily. . .' The latter, if made out, genuinely defeats 'oughtness'; nonperformance as such does not.

Disagreement

So there are two areas of actual or potential disagreement of a sort, despite the agreement having apparently been made. In the first place, the two do not agree *in* their specific values—there we observe, and must reckon with, great variety, and in fact this sort of disparity is actually a prerequisite to the usefulness of contracts—for if we already had perfect agreement of that kind, we would not need to make them. And in the second, unfortunately, people sometimes cheat, violate, defect on their agreements.

Disagreement.1: Preferences

To agree to the sort of terms that define contracts (not restricted to legal agreements), one party must rank the situation of her having *y* rather than *x* higher than that of her having *x*, which she currently does have, rather than *y*, which she does not; while the other party must have the reverse preference schedule: he prefers *x*-and-not-*y* to *y*-and-not-*x*, the latter being what *he* already has. Thus contracts presuppose partial disparities in preference rankings—they would be impossible without it. (Think of the apocryphal story of the two old prospectors who exchange five dollar bills at Christmastime.) So we would want to be careful about Kliemt's observation: 'disagreement' in specific preferences is not only to be expected, but it is a precondition for contract taking place at all.¹

Disagreement 2: Defection

Despite—indeed because of—these differences, contract unites: if we both act as agreed, we both do better than we would if no agreement had been made. A ends up with what she prefers more, B ends up with what he prefers more. And what it is all contingent on is the confidence that the other party will act as agreed. That brings us to the other sort of 'disagreement': defect. The trickier

¹ Of course, if our preferences are expressed in unqualified value-terms: not, I prefer *x* to *y*, but '*x* is better than *y*', then this difference of preference would become a disagreement, at least formally. But if we use the subjective rendering, then to be sure, 'disagreement' means only that their preference orderings are opposite. Jones doesn't disagree that Smith *has* different preferences, nor need they come to blows over the disparity—though they often do anyway, alas. Indeed, they each prefer that the other person's preference be realized. So the disparity in itself, rather than precluding possible contract, is just what calls for it. Indeed, it is the essential part of what explains its rationality, so long as it works—that is, that people actually perform as they claim to be committing themselves to performing in signing up.

problem is the defector, Hobbes's Foole. Now, the beginning of wisdom here is to point out that, on the face of it, the defector has said what is not actually true: that he *agrees* to do the thing in question on the specified condition. For the condition obtains—the first party has acted—and yet he doesn't do what he said he would. We are justified in saying that in that case, he didn't '*really*' agree. Why? Because when we say, 'I'll do x', we make a *prediction*: viz., that I will in fact do x (Ardal 1968, 225–237). Nonperformance falsifies that. Such predictions could, of course, fail in other ways, ways that we don't regard as impugning the contract—e.g., party B dies before he can act, or it in some other way becomes impossible for him to do it, and so on. But those are loose ends—footnotes, and not to worry about. The big problem, as so widely recognized in the literature, is Hobbes's Foole, for whom agreements are “but words and of no strength to secure a man” (Hobbes 1950[1651], 139).

The question is whether the Foole acts rationally. Hobbes thought not (Hobbes 1950[1651], 119–121). But the irrationality does not lie in failing to pursue his advantage in the specific case. It lies, if at all, in its effect, of opting out of the social system in which people can rely on each other to do what they've agreed—the very thing that makes agreement worth bothering with; but the Foole's behavior negates the value of agreements, and invites the sort of retaliation that would more than negate it for *him*—at least, if the retaliators play their cards (e.g., their policing institutions) right.

Genuinely undermining that social system would indeed be irrational for anyone with future interests whose pursuit depends on further agreement. The trouble is, though, that in a particular case it may not seem as though one *does* 'undermine' that system. There's the chance that it will go on much as before, enabling the con man, the thief, and others to thrive. How good that chance is depends, of course, on the efficacy of the control system in that social setting: how likely is it that someone will intervene, and how effectively, to undo the expected utility gains of defection? But there is also a very important collateral question: what has society done to try to bring it about that our potential defector does not become an actual one? If he has a built-in reticence to do this, we can all be way ahead. If not, the likelihood that enforcement by threat will be effective is far lower.

I think it clear that no amount of tinkering will come up with a plausible case for *government* from this source. Rather, as has been emphasized by the best modern thinkers (Gauthier especially), what we can hope for is 'agreement'—in the relevant sense—on certain general principles. Hobbes thinks that morality is really summed up in just *one* general and very fundamental principle, from which all the rest follows. That one principle, which he tags the “first Law of Nature” is set out by Hobbes in the language of peace-keeping: we are to “seek peace” unless we can't get it, in which case we are entitled to “all the helps and advantages of war” (Hobbes 1950[1651], 107). I think he is right. In turn, that agreement implies the 'agreement to keep our agreements': meaning, that we all recognize that certain forms of words, when employed in recognized ways, entail obligations, because that's what they are designed to do; and that when prop-

erly used, against the background of peace, those obligations are to be upheld by members of society, both in their own cases and, when feasible, reinforced regarding others. But our question here, of course, is: what if the recognized forms of words are *not* used? How then can we impute obligation? The social contract, after all, is precisely the case in which those words are *not* used; how, then, do we extract nevertheless the same sort of obligation as is generated by literally employed words?

The obligation to keep formal promises and articulated agreements (set forth in Hobbes's Third Law of Nature) is, in Hobbes's elegant model, an entailment of the First Law. To break a promise is to violate peace. More will be said shortly about the argument for the First Law, but here I note that the language of war and peace will sound a bit overdramatic for many cases. Still, what Hobbes can reasonably be thought to have had in mind is not only outright war, with intended death and bodily harm, but rather more nearly what is captured in Gauthier's beautifully generalized formulation, which he calls "the Lockean Proviso": that no one is to seek to better his own situation *by* worsening that of someone else, unless one's own situation will be made still worse by one's adherence to peace (Gauthier 1986, 203). I take that to explain the Hobbesian idea: to 'make war' on someone is to seek to worsen his or her situation, in order to better one's own. And since our measure of worsening and bettering is the values of the persons themselves, rather than the values of some imposed Platonic idea from elsewhere, the connection with consent and volition is obvious: no one consents to be made worse off, and so to prohibit worsening is to require consent, at least in the form of acquiescence, to the Hobbesian First Law.²

Words are not Fundamental

Therefore, to defect on agreements (specifically, what Hobbes calls "covenants")³ is to take advantage of one's interactee by accepting an agreed benefit from him, yet failing to come through with the benefit to him that motivated the whole transaction from his point of view, thus leaving him worse off. Thus it is, contrary to the First Law, to "invade for gain": gaining at his expense—making oneself better off by making the other party worse off. Thus the obligation due to articulate agreements is deeper. We have moral reason to keep our agreements only if, and because, we have reason to engage in peaceful rather than warlike relations with our fellows.

² Other construals and discussions of the Lockean Proviso, as it has come to be called, in Locke's own misleading words about it, abound. See, e.g., Nozick 1974, 174–182; Schmidtz 1991; Narveson 1999.

³ Covenants: the subset of agreements such that one party acts significantly prior to the other, thus creating for the other the opportunity to collect his part of the bargain without providing the other's benefit to him. Where both parties act simultaneously, we don't have this problem, though conceptually we have others.

This last requires one important understanding: it is assumed that we may be presumed to consent not only to actual improvements in our situations, but also to *not being made better off* so long as we aren't thereby made *worse off*. The consent is passive: it simply consists in not dissenting. The effect is that we are to proceed by what I call the 'Gospel according to St. Pareto': our interactions with others are always to have net non-negative impact on anyone else concerned, insofar as that is possible at all. ('Net' being necessary because, of course, we very frequently do what would make us worse off were it not for someone else's supplying something, as an exchange, that makes us better off overall, as with successful visits to the dentist. 'Net' is to be understood as a matter of it being on the whole preferable, or at least no worse, from the point of view of each interactee to get whatever he or she gets from the exchange on balance.) This bloodless consent *is* a rational requirement. If something doesn't matter *to me*, then so far as I'm concerned, it *doesn't matter* and I have no reason to disagree. Given that I do have reason to refrain from war, it means that the peace called for by Hobbes's First Law of Nature is automatically presumptively agreed to by all nonparticipants in any transaction. Peace among transactees plus peace among non-transactees is peace, pure and simple. Thus we disagree with Shaw, who says "[t]he worst sin towards our fellow creatures is not to hate them, but to be indifferent to them: that's the essence of inhumanity".⁴ On the contrary: there is no injustice in indifference, whatever else there might be. Our fellow creatures are much better off from our total lack of attention than from the malevolent intentions in which Hobbesian 'war' consists.

The Problem with Government

Government, however, is inevitably coercive in the sense of violating the Pareto requirement: it inevitably makes some better off by making others worse off—through taxation, or through imposition of unnecessary restrictions. If the social contract is to lead to politics, this last has to be somehow responded to. Hobbes provides the classic argument: he thought that we *needed* government in order to make agreements effective, and for that matter, more generally, to make effective the agreement on the First Law of Nature—which in his (and my) view implies all the other such Laws, as he calls them. The idea is that we need government to keep the peace. So imperative is this device, he thinks, then if Hobbes is right, we would all be better off under *any* government, however bad, than in what he takes to be the only alternative—the State of Nature with its alleged war of all against all. We don't get to pick and choose: the government we've got, whatever it is, is the best we can do, since civil war, he thought, also gets us back to the dread War of All against All.

⁴ George Bernard Shaw (1856–1950), *The Devil's Disciple*, Act II (1901).

We may surmise that when Hobbes declares that in a Commonwealth, the Word of the Sovereign *is* the voice of the Law of Nature, he should be envisaging not, as he seems to imply in his discussions, real-world governments with their innumerable violations of that Law, but rather, the ideal Sovereign who makes no mistakes ('God' being a suitable instance...) The trouble is, there are no such. Our reason for any obedience we may give, aside from sheer fear in view of government's ham-handed powers, is contingent on his enforcing that Law and nothing else. Beyond that, Hobbes doesn't have a case.

Thus Hobbes's authoritarianism is surely a mistake. It isn't just that Government cannot be guaranteed to make us all better off than we would be in the State of Nature—though indeed it can't. It's that Morality *could* do that, if we could all be counted on to adopt it, and therefore we would not need to take the further step to government. Since there is an additional cost involved in the move to government, given its unreliable wielding of extreme power, it would not be rational to choose it if one could have reliable morality instead. Indeed, since governments are sure to be very imperfect, a not completely reliable morality arguably could and probably would be preferable to a proportionately imperfect government. Obviously the question is how reliable the anarchic condition with morality would be. We needn't agree with Hobbes that it just couldn't work.

Hobbes's successor Locke understood the point: his depiction of the 'State of Nature' is of a predominately peaceable society, and his depiction of (good) government is of tolerance and democracy. Because the State of Nature is better than really awful government, we don't have to put up with the latter, and we can hope, and try to work, for limited government. Thus Locke wins on both fronts—moral and political.

The politically ambitious are out to persuade us that he's wrong on the first front. Anarchy, they hold, is impossible—it will degenerate into the intolerable "war of all against all". If so, then the Hobbesian argument for government kicks in. Yet the Lockean modification which plumps for some choice on our part also holds: despotism *non grata*—the politically ambitious work within the constraints of citizen control via election. So, Locke is headed for all-out democracy, though perhaps he didn't quite realize it. As we have come to see, though, government, however democratic, inevitably and inherently crosses the line—the line of individual rights as proclaimed by both Hobbes and Locke. Thus Locke proclaims that "[t]he *Supreme Power cannot take* from any Man any part of his Property without his own consent" (Locke 1980, sec. 138). So—no government!

Because of roughly Lockean considerations, which we may identify with the social fact of the indefinite iteration of dilemma games that change defect from the dominant to the overwhelmingly dominated strategy, we can expect that morality will do pretty well. There will be some who step out of line, but we all have the right to deal with those. Of course, Locke himself thought that in order to do so effectively, we again need government, to that extent siding with Hobbes: government, he thought, would have to take over both the policing and the adjudicative functions. But he thought at least that we could improve things a lot relative to Hobbes's absolutism, even if the cost in the worst cases is civil

war. We don't need to resolve that issue here—happily—but it's enough to see that the philosophical case for accepting the man-made laws of governments as definitive guides to behavior is far from irrefutable.

Government by 'Acquisition'

Hobbes distinguished between government by "institution" and government by "acquisition". The former is the ideal model, where we all get together and elect our monarch. The latter is where some powerful person and his henchmen simply walk in and take over, brooking no opposition, and overcoming such opposition as there is. Yet, having made the distinction, Hobbes goes on to say, interestingly, that it doesn't really matter much in the end which we happen to have. Most of us nowadays don't buy that. But if we don't, the problem is the one advanced previously: there would appear to be no such thing as *legitimate* 'government by institution'—*all* governments are by acquisition. If we think some legitimate, it is purely a matter of comparing better ones with worse ones. Few would deny that some governments are worse—some a *lot* worse—than others, after all, and if we have our choice, we'll take one of the 'good' ones. But just as some thieves are kinder than others, so some governments are, even though they are fundamentally not legitimate in the best sense of the term—which is, that everyone genuinely does come out as well as possible, given the presence of others.

Governments, on Hume's better view, are originally founded on "usurpation and rebellion"—in short, by one or another kind of violence. When we think they arose by better means, it's always by the actions of an elite, such as the propertied and influential men who wrote up the American constitution. On the view of the philosophical anarchist, our grudging acquiescence in government is due partly to the fact that we are wimps, but considerably also because we suppose that any realistic alternatives will be as bad as or worse than what we've got. Should we be ready to dignify that with the blessing of Social Contract? I don't think so. It is, in Rawls's terms (1990, 147), a *modus vivendi* rather than a real commitment.

It may be asked whether Rawls's distinction isn't really a distinction without a difference. Isn't morality, as well as government, basically a 'way of life'—an accommodation to the exigencies of social living? Well—yes, it is, to be sure. Nevertheless, there is a distinction (of degree, at least) between being so habituated to the requirements of morality that we do not seriously pause to consider the options we regard as prohibited, and so unhabituated to them that we must work our way through the argument every time we face an immoral choice, opting for the morally preferred one only because it wins in an elaborate calculation. We do, after all, know people who exemplify the latter—we read often enough in the papers of people who make the choice *against* morality, though we think that they haven't done their homework well. And we, most of us, probably exem-

plify the ‘committed’ most often in our daily lives, when we do not even consider the option of helping ourselves to the candy bars while the checkout person’s attention is diverted, or when we return lost property to its owner without even considering the option of keeping it. Just the day before this was written, I think of a case where I began to walk off with my purchase, and was grateful when another customer dashed over with my wallet, which I had left on the counter. I am sure there was not discussion of the option either by the checkout person himself nor by the chap who brought it over to me. If we were *all* like that, life would be better—and so many of us are like that, that it is indeed pretty good. The Social Contractarian holds that this is what the ‘contract’ should strive for, and that in ‘favorable circumstances’ it very often succeeds.

This gets us back to morality, which we here understand as the set of general rules for the social informal regulation of the conduct of everyone—rules that are to be internalized by all. We may take the ‘rules’ in question, at least for the present, to be just the Hobbes’s First Law, or Gauthieran ‘Lockean proviso’ described above, together of course with its implications. If we can’t even get *that*, then social contract is, I would agree, truly hopeless.

The trouble with government by social contract, by contrast, is that it’s a matter of ‘consent’ to a Chicken game; the government is in considerable part acting as a protection *racket* rather than, or in addition to being, a protection *agency*. This is no way to go, if we have our choice. Whether we do have that choice is, to be sure, a serious question. So, more on this next.

The War of Some Against Some

Hobbes, as I say, thinks that the war of all against all is the *only* alternative to accepting his First Law of Nature, which prescribes peace whenever it can be had—the universal ‘peace of all against all’ instead of the war against all. But what about war against *some*? That is to say: why may we not use violence to achieve *some* of our ends against *some* others, even though we agree not to use it *whenever* we could gain an advantage by doing so? When Pizarro conquered the Incas, he did not think that force was ruled out against *those* people, though he felt morally bound to treat his fellow Spaniards quite well. And any number of tribes, throughout history, have regarded the *other* tribes as largely beyond the perimeter within which moral protections were operative.

At this point it is well to remind ourselves once more of the project we pursue here: viz., to find a *universal* moral principle or principles to apply generally in all situations. That is a matter of definition: we seek a principle or principles that is or are to have the ‘force of reason’ *for all*. Wars of some against some, it seems to me, cannot meet this requirement. Thus, for example, the condition, say, of slavery-for-most is no doubt better than that of war-for-all, and there are an unlimited number of other arrangements in between. If we construe the

reasoning for peace as Paretian, isn't this, then, sanctioned, given the Pareto principle? Why do we need to strive for universality?

The answer to this, I think, is that limited 'moralities' *are* sanctioned in the purely relative sense that in each case, what is proposed is better than what it would replace. But as an open-ended, general rule, that won't meet the requirement of reason. A war of some against some is *not* agreed to by all if the social situation improved on is itself worse for some than it could be if some others were not violating the peace against *them*, so that *their* acceptance of that situation is not considered. Rationality may indeed tell us to put up with evils we can do nothing about. But the question isn't whether the wimps (most of us) have a better strategy than the active rebel. It is, rather, whether there is not a rational basis for regarding evils *as* evils. Aristotle regarded slavery, of the right kind, as a good rather than an evil. He thought he had a rational basis for not asking the slave's opinion in the matter. In this, as we all today think, he was wrong. Rationally speaking, the opinion of anybody who has an opinion and can articulate it counts. Thus, from the universal point of view, there is always an alternative or alternatives to any partial or restricted set of rules that are better. That is what grounds the expectation of universality as the indicated ideal. A rational morality is one that has the support of all. Anything less invites retaliation from those left out.

And thus, in support of Hobbes, one alternative trumps all others, namely Hobbes's First Law. Why? Consider any world in which almost all are required to be at peace, *except* some subset. Bear in mind that the Peace Principle, the First Law of Nature, is not pacifist. We are entitled to use force against those who use it against *us*, or against any innocent for that matter. But defensive actions are not contrary to the Law, and so are excluded from consideration here. What we are considering here, rather, are principles calling not merely for the use of force against violators, but for the legitimacy of force for *other purposes*—principles by which some can gain an advantage against some others, *even though those others are innocent* by the terms of the principle—such as the slaves in our example. Or to take a closer-to-home example, take the 'welfare state' which on the face of it allows the poor to make war against the middle-and-upper earners ('the better off', as we'll call them). Yet these latter are not enemies of the poor—they have done nothing to *make* them poor. (Innumerable contemporary writers claim to deny the latter, but the denial is always question-begging. If some work successfully and enrich themselves in the process, they need not, and I would think typically have not done so by *taking* anything from the very poor.⁵ And that is the situation of most taxpayers in modern societies, after all.) Thus if governments, at the behest of voters, compel the better-off to contribute to the welfare of the erstwhile poor, then the situation is that the poor, and their confederates, make (a fairly limited) war on the classes above them, on the face of it. And on that same face of it, the situation in which those better-off are not thus forced, leaving the poor wherever they were prior to the onset of the

⁵ For a discussion of one who argues, as many have, that they *do* "take something from the poor", see Narveson 2013.

welfare-state initiative, is Paretian preferred to the situation in which they are taxed to someone else's benefit. A welfare state is preferred by some sizable fraction, no doubt—but not by everyone, some of whom outrightly prefer not to have it. No one *likes* to be poor, of course, but that's irrelevant in the social contract. What matters is how the poor got that way—in particular, whether it was a result of aggressive activities by their fellows. If not, then he's no worse off when someone else becomes wealthy than he would be if that someone had remained poor; but realistically, he's better off, for the wealthy spend their money, either on productive investments or on consumption items that someone makes and sells to them, to the benefit of the makers. Thus the relatively unrestricted enterprise which has enabled a half-billion of contemporary Chinese to rise above the dire poverty of their predecessors, even though another half-billion remain (this is written in 2013) in the poverty of the countryside, gets the approval of the social contract, even though it would, of course, be still better if the lot of the rest were also improved. (That they likely ere long will be is a further point, of course, but not essential to the approval in question.) Introducing enforced transfers and restrictions only impedes the process of wealth—as depressingly well exemplified by the previous regime of Maoist communism. Social Contract simply says that *everyone* is entitled to peace, which may be broken only to prevent attacks.

For that reason, modern writers need to justify their initiatives of an apparently unfair kind by arguing that the better off are really better off for being in a welfare state, despite the taxes and despite appearances. That's an uphill battle, but perhaps not impossible—we don't get into that here. It would require a lot more information than we have merely from the raw descriptions given so far.⁶

But again, why universality? Why should we agree with the previous assessments instead of going along with the political majorities of the present and past? The only answer permissible within the Hobbesian framework is Hobbes's: that "Nature has made men so equal, in the faculties of body, and mind; as that though there be found one man sometimes manifestly stronger in body, or of quicker mind than another; yet when all is reckoned together, the difference between man, and man, is not so considerable, as that one man can thereupon claim to himself any benefit, to which another may not pretend, as well as he." (Hobbes 1950[1651], 101) On the face of it, this seems quite false. Pizarro was manifestly superior in military technology to his rivals; the strong are stronger than the weak, and often gain advantages because of it. But while that is surely true, it does not follow that we will—even the strongest among us—embrace a *morality* of force, whose rule is that the strongest are to prevail. This is a nonsensical rule, since morality calls upon us *not* to use all means at our disposal for advancing our ends, whereas not using them to defend against force is exactly what reason calls for. It is also in the longer run stupid: no one is able to

⁶ For an interesting example of how one *might* argue, in limited circumstances at least, see Putnam 2013.

maintain his exceptionality over the rest for long. Everyone, on the other hand, can gain in peace.

The advantage of universality, if available, is precisely that *everyone* is someone to fear, and if rules can be found that allay such fear, we all benefit, so far as that goes.

Hobbes observed that those who prosper at the expense of the rest can do so only by indirection: the rest would not permit it if they knew what was going on. They “cannot be received into any society, that unite themselves for Peace and Defence, but by the error of them that receive him” (Hobbes 1950[1651], 121). Perhaps that error can be maintained, but, again, the point here is that no publicly admissible principle can allow it.

Social Contract and The Paper It’s Not Printed on

We move now to the question of the Invisible Hand and its ‘signature’ on the Social Contract. A familiar criticism of social contract is that it’s ‘not worth the paper it isn’t printed on’. I want to argue, however, that it *is* worth that paper and a good deal more. To be sure, as we all know, we do not ever all explicitly say to our neighbors and everyone, ‘I hereby promise to abide by . . .’ OK; but, do we do something that is *good enough*?

The first point to make about this is Hume’s. Hume deftly disposes of the idea that political or moral obligation arises from a promise. Suppose that we actually did all go through those verbal motions. Then the question is: well, how does that actually bring about the obligation intended by those words? As he says, “we are not surely bound to keep our word because we have given our word to keep it” (Hume 1957[1752], 122). Here again, Hobbes’s point looms before us: promises are ‘just words’ unless—well, what? Why is it that those who *say*, ‘I agree . . .’ and then act contrary to their agreement, are understood to be acting ‘*contrary*’—badly—and why does that matter to us? Hobbes thought that the ‘something’ was—could only be—the sword of the sovereign. But after all, swords can be used for many purposes, good and bad. How do we know that the sword is descending on the right heads? The principle emerging from the ‘social contract’ is to *answer* that question: the sword should be falling on the guilty, not the innocent, and the guilty are those who undo the peace.

The Humean story is the modern one: morality is a device that works a terrific trick for us members of society, demonstrating the superiority of a certain sort of obligation to an unconstrained condition. With contracts, we aren’t obligated *unless* we agree, and we *are* obligated if we *do* agree. In both cases our having *said* ‘I agree’ is taken to be a sufficient sign that we have indeed agreed and may be so taken and treated accordingly. There is a visible, voluntarily controlled sign to the public indicating agreement-thus-obligation. What makes this such a useful system is that it preserves the voluntariness of the actors; they know, with this social/verbal system in effect, that people can’t pin duties

on them that they haven't spoken up for, and thus that the ones they do speak up for are guaranteed, if reciprocity is forthcoming, to make them better off, apart from bad luck.

In what sense, then, do we 'agree' on the Peace Principle, since, almost always, we surely haven't 'spoken up' for it? There's a good answer to this: we aren't idiots! "Two men, who pull the oars of a boat, do it by an agreement or convention, tho' they have never given promises to each other. [...] In like manner are languages gradually establish'd by human conventions without any promise." (Hume 1957[1752], 123) What we perceive is the benefits to be had from cooperation—real cooperation, not enforced cooperation. The Peace Principle gives rise to 'iffy' obligation: I do this if—*provided*—that you will do that, and for many important thises and thats, while we both gain if we do our parts, there is nevertheless a potential for one party to gain at the expense of the other by not doing his part when the other *has* done his. We both know this. And we know the consequences of thwarting this principle. Life is better for all of us if we both accept and abide by it. Could it be *still* better—not for all, but for me—if I were the one to take advantage of the other? Surely—but we know that *each* can say that, and yet if each acts accordingly, none of the benefits envisaged will in fact obtain. It is the *sense* of all this that constitutes the Social Contract at work. We agree in our perception of the situation, and we therefore are correctly taken to agree *to* the requirements imposed on us by these workable social arrangements. The presumptive agreements mean that our fellows can, if we default, exclude us from further cooperation, including by putting us in jail or the like, with approval from all for those exclusionary activities. Social contract works both because of that and because we are smart enough to see in advance that the way to go is for us all to internalize the disposition to keep the peace, and thus our agreements, without needing oversight by the police.

The case for regarding promises as *obligatory*, then, is clear: knowing that you can gain from defection, yet that we'll both be worse off if we can't make agreements at all, we also understand that the defected-upon party has been wronged, will rationally retaliate if he can, and if he can't by himself, will rationally seek the support of his fellows, all of whom have the same interest—as indeed does the violator himself, in almost all other moments of his life. Violators, then, are in no position to complain.

And that's what it *is* to be in the wrong, morally speaking: to have no publicly sustainable case for one's actions. Critics seem to think that the fact that it can be in one's interest to cheat is a fatal criticism of social contract as foundation for morals. On the contrary: any theory according to which we all always do the right thing should be laughed out of conceptual court. We can't prove what isn't so. And if morality were rational in the same way that, say, arithmetic is, then we would all be moral all the time, apart from sheer confusion or mistake. The Social Contract makes it clear both why we are indeed obligated by these self-imposed arrangements, and also why morality, as a social institution, will not be perfect in its control of actual behavior. That is how it should be, rather than constituting a relevant, let alone a definitive, criticism.

Social Contract and Morality

So we are back to hoping for light on morality by exploring the social contract idea. And as soon as we do that, we need to make some distinctions, especially these of Jasay:

0. "If I meet a friendly stranger and he tells me that it is a lovely day, I will agree that it is indeed lovely. "The agreement commits to nothing and stirs to no action."

1. "If I see all drivers driving on the right, I conjecture that they find such coordination a good thing and I agree with them, so much so that I, too, will do as they do. My agreement stirs me to emulation. But it is not an agreement with them."

2. "I conjecture that all my neighbours would find it a good thing if their neighbours stole less, or not at all, from them, and they would rather steal less frequently, or indeed not at all, if the neighbours responded to such restraint by like restraint. I agree, and will join them in the corresponding convention by stealing less often or (if the most efficient of the alternative equilibria gets selected) not at all. Note that Agreement 2 is confined to a recognition of identical interests and an identity of responses to be expected from one another. The agreement does not entail a binding undertaking to perform an action or forbear from one. [...] Mutual advantage results from the convention, but that is not the reason why I adhere to it."

Then we come to

3, which "has two variants, the unilateral or unrequited promise and the reciprocal promise or contract. Hume pays little attention to the former. The latter, rendering exchanges of non-simultaneous performances feasible, expands the scope for generating mutual advantage, beyond the constraint of the cash-on-the-barrelhead type and is a gift of incalculable value to civilization. By establishing the binding undertaking, Agreement 3 lays what looks like a footbridge leading from non-cooperative to co-operative game theory, and also to social contract theory. This terminates the crescendo of agreements." (Jasay 2010, 399)

These are useful distinctions. Regarding '0', we should note that though it commits us to nothing, yet agreement on factual matters, and especially on the relevant ways both of collecting more information about the world around us, and of how to resolve disagreements, by resort to more careful or more extensive observation and calculation, clearly is broadly presupposed at all other levels. If we don't 'live in the same world' there is no hope for rational agreement of any other sort, not to mention commitments. And unfortunately, it is only a fact that we

do live in the same world—which doesn't keep some people from living at least partially in a dream world, impervious to scientific reason, yet whose irrational ideas are extremely effective in determining their behavior.

Agreement at level 1 is also of fundamental importance, as is easily overlooked. That you and I both perceive something to be to our advantage, even though it doesn't require cooperation, strictly speaking, is again requisite to moving to higher levels. No expected advantage, and there's an end of the matter. Coordination is often important. That it is often obvious what kind of behavior in a given situation is called for by coordination, and because there is no real conflict of interest among those, doesn't mean that we may not reasonably talk of the 'rule' of the road, and many other such.

We may say much the same about level 2: I need to be able to appreciate your situation, and your expectable pattern of responses to it, before we could be ready to deal at level 3. And his example, of course, is over toward the edge—people *do* steal, after all, and when there's a prospect of someone violating the non-theft custom, we move to a higher level. We will now, unhesitatingly, invoke obligation.

This brings us to the fundamental level for social contract: agreement *with all* that really does commit—where I act only if I think you will act, and vice versa. With a view to making these important conditions likely, indeed reliable, we bring in the language of commitment, of rights and duties.

Morals: Two Roles

We need here to invoke, and lean rather heavily on, a distinction I have emphasized throughout my work on morals (Narveson 2009, 222–223). This is the distinction between 'embracing' or 'accepting' morals in the sense of having a disposition to *do* what morality asks, on the one hand; and on the other, of embracing it in the sense of being disposed to *advertise* it, and to help administer it—to recommend, to advise, to swear to uphold, and also, and especially, perhaps, to help in its reinforcement by visiting your fellows with appropriate positive verbal (and other) feedback when they do what is right, and with the appropriate negative level—criticism, condemnation, and possibly reinforcing action, when they do what is wrong. The commitment at these verbal levels is important, for two reasons. First, once things are public, the cost of coming out in public as an enemy of the rest is high. And second, once one verbally supports some principle, the cost of failing to act in accordance with it oneself also becomes high. We are vulnerable when we have gone on record as proclaiming that the thing to do in situation S is x, and yet don't do x when S eventuates.

Part of the importance of this distinction is that it makes it easy to see how people can be hypocritical. For, I want to argue, whenever mutual advantage by way of agreement at level 3 is possible, but the hoped-for advantage is dependent on the commitment of others, as Hobbes observed, then we may quarrel with

his assertion that promises are ‘but words and of no strength to secure a man’. Indeed, for most of us, most of the time, they *are* of ‘strength to secure . . .’ There are people we know not to trust, and there are bargains that we know it would be very hard to keep, but that doesn’t mean we’re lost. On the contrary, whenever your commitment is fairly ‘cheap’ (there is little in the way of what Rawls calls “strains of commitment”), and the obtainable benefit real and worth having, we can rely quite well on most people.

What to do about the others, though? This is where Hobbes thought that government was necessary. The government commands a big club—bigger than any other set of persons in the polity in question, as Hobbes insists—and it will, he supposes, use it against the bad guys, who make contracts but then take the money and run, and against murderers, rapists, and robbers who aren’t, when they do so, breaking promises made to specific people. But Hobbes thinks that they have nevertheless violated, as it were, a commitment they *should have made*, which the rest of us can and do expect them to make. How much does that help?

War

To see how much, let’s contemplate the hypothetical condition of society in which nobody makes any sort of commitment at all, let alone to the Hobbesian Law of Nature. These are the people who will kill, steal, maim, and rape whenever they feel so impelled, and morality be damned—and, Hobbes thinks, the people who will so behave are, in the absence of constraints, *everybody*.

Focussing our scrutiny from the vantage point of present day social life back to a Hobbesian, premoral ‘state’ is not too easy. On the one hand, there is the anthropological evidence about primitive man, who, as has recently been argued, was not a nice guy.⁷ What to make of such reports, even if well supported by evidence, as we will suppose this one is, is a bit of a puzzle, for it is also plausible to think that primitive communities, being after all communities, must have had institutions of cooperation to a considerable extent. Moreover, they could not have had killer instincts of a truly pervasive kind, for if they did, we would not be here. (Is there *anyone* who *might* not have benefited from someone else’s murder, if he could get away with it and was, as most of us are not, inclined to do such things?)

But this need not be an anthropological issue. Our question is merely, how would it be *if* people behaved in the envisaged way? We know that some few do so now—violent behavior is obviously not impossible. To understand what we

⁷ Lawrence H. Keeley (1996) estimates casualty rates “of up to 60%, compared to 1% of the combatants as is typical in modern warfare. Despite the undeniable carnage and effectiveness of modern warfare, the evidence shows that tribal warfare is on average 20 times more deadly than 20th century warfare [. . .].” (See summary at en.wikipedia.org/wiki/War_Before_Civilization, retrieved September 18, 2013.)

oppose when we get into morality is as much a work of imagination as of observation. It is at that level that we can get a grip on the reasons for subscribing to the inhibitions of morality, whatever may be our own tendencies to do without them.

Paul Viminiz (2009) has set out a general theory of war, of which the essence is that we have war when people regard other people as more useful dead or in chains than alive and free; we have peace when they have the opposite evaluation. Then there is the transition from something approaching the Hobbesian condition to the one we are in today—roughly, a pretty high level of acceptance of basic moral constraints, such that few of us even think in terms of advantage from the proscribed acts. If we are rational, moreover, the general usefulness of our fellow men is obvious on consideration, to the point where we may think that there is something rationally defective about the ‘warlike’ among us. In any case, we also agree—almost all of us—that we should, and certainly are entitled to, deal with the ‘bad guys’ much differently than with the rest of our fellows. They deserve, we think, to be jailed or perhaps executed, etc., in contrast to what any innocent person deserves. In short, we accept Hobbes’s Law: we seek peaceable ways of relating to our fellows, except only when they intentionally attack us.

Now, if we all come to adopt civilized values, we will have no war to worry about, and will go about our business to the huge benefit of all concerned. But—*will* we ‘come to adopt them’? If so, why and how? What is needed is commitment in the sense that people agree that they really do have obligations to conform to the Hobbesian law. In practice, this will issue in persons who refuse to violate the law even when they could do so with impunity. Some will think them foolish; but all will think them moral. The Social contract idea is that everyone is to embrace those principles in the conditional way just explained. That is a tall order. But it’s not out of reach.

Rationality and Morals—One Last Time

An important question here, however, is whether and just how this is all a matter of *rationality* as Hobbes supposed. Our assumed view of practical reason is that a person acts in such a way as to promote the satisfaction of his desires or values. We may think we are capable of evaluating our desires, so that our *values* are in some sense improvements on those desires. But nevertheless, our considered desires are still desires, and the consideration we give might leave us with some pretty anti-social desires. Some people, it seems, will be pretty unpeaceable despite, so to say, having read their Hobbes. And lots of us will be sympathetic to the Aristotelian dictum that when it comes to morals, we have to start them young. If someone isn’t moral by the time he’s fifteen or so, moral philosophers likely aren’t going to be able to do anything with them.

Thus consider the recidivist with his short time discount. He doesn't reckon beyond the morrow, and either (low class version) isn't capable of thinking further, or (high class version) doesn't see why he should. If TV programs are to be believed, there are quite a few of those. Our problem here, if it is one, is that we seem to have non-signatories of the Social Contract: people who refuse to accept the general injunction toward peace. What to do?

This is a philosophical investigation, and not a practical document for constitution-making. Philosophically, the point to be made about these supposed counterexamples is that they *have* signed a sort of 'contract' with all: for we will credit them with a sort of general declaration of war, 'against all'. So understood, we are all, without being asked, 'signatories' to *that*: if a man is trying to kill me, then I may kill him. We might say, 'May the best man win!'—but that's misleading. For what we really want to say is that this man, who insists on making war against those who are perfectly ready to have peace if only he'll do his share in providing it, is an idiot. That is what we *say* to him, even as we oil up our weapons, with which the practical problem presented by his presence on the scene will be solved if we can. His persistence will then impel us to pull the trigger. The war of all against all will turn into a war of a whole lot of us against some relatively few, and we have decent hope that we will win. Since at that point there is no alternative, that will have to be enough.

And after all, aren't we, simply, *right*? To be sure, the dice as I have depicted the dialectical situation are loaded. That peace is 'better than' war is a simple matter of Paretian rationality: with war, as compared with peace, life is dangerous and probably short. With peace, it is, with any luck, a lot more comfortable, a lot safer, and a lot longer. Why isn't that *enough*?

There is, to be sure, no guarantee that peace will win. There is only a guarantee that no social *normative* principle—which is what we are talking about here, after all—can plausibly beat it.

References

- Ardal, P. (1986), "And That's a Promise", *Philosophical Quarterly* 18, 225–237.
- Gauthier, D. (1986), *Morals by Agreement*, New York: Oxford University Press.
- Hume, D. (1957[1752]), *An Inquiry Concerning the Principles of Morals*, Indianapolis: Bobbs-Merrill, Library of Liberal Arts.
- Jasay, A. de (2010), "Ordered Anarchy and Contractarianism", *Philosophy* 85, 399–403, doi:10.1017/S0031819110000240.
- (2013), "Conduct and Contract", *RMM* 4, 39–46, URL: http://www.rmm-journal.de/downloads/Article_Jasay.pdf.
- Keeley, L. H. (1996), *War Before Civilization: The Myth of the Peaceful Savage*, Oxford: Oxford University Press.
- Locke (1980), *Second Treatise of Government*, ed. by C. B. Macpherson, Cambridge–Indianapolis: Hackett Publishing.

- Narveson, J. (1999), "Property Rights: Original Acquisition and Lockean Provisos", *Public Affairs Quarterly* 13, 205–227 (reprinted in Narveson, J. (2002), *Respecting Persons in Theory and Practice*, Lanham: Rowman & Littlefield, 111–130).
- (2009), *This is Ethical Theory*, Chicago–LaSalle: Open Court.
- (2013), "Liberty, Property, and Welfare Rights: Brettschneider's Argument", *Libertarian Papers* 5, URL: <http://libertarianpapers.org/2013/9-narveson-liberty-property-welfare-rights/>.
- Nozick, R. (1974), *Anarchy, State and Utopia*, New York: Basic Books.
- Putnam, R. B. (2013), "Crumbling American Dreams", *New York Times*, Sunday August 4, Sunday Review.
- Rawls, J. (1990), *Political Liberalism*, New York: Columbia University Press.
- Schmidtz, D. (1991), "Property", in: *Limits of Government*, Boulder: Westview Press, 15–32.
- Sugden, R. (2009), "Can a Humean Be a Contractarian?", *RMM* 0, Perspectives in Moral Science, ed. by M. Baumann & B. Lahno, 11–23, URL: http://www.rmm-journal.de/downloads/002_sugden.pdf.
- Viminiz, P. (2009), *A Theory of War*, unpublished manuscript