RMM Vol. 4, 2013, 178–190
Special Topic: Can the Social Contract Be Signed by an Invisible Hand?
http://www.rmm-journal.de/

Danny Frederick

Social Contract Theory Should Be Abandoned

Abstract:

I argue that social-contract theory cannot succeed because reasonable people may always disagree, and that social-contract theory is irrelevant to the problem of the legitimacy of a form of government or of a system of moral rules. I note the weakness of the appeal to implicit agreement, the conflation of legitimacy with stability, the undesirability of 'public justification' and the apparent blindness to the evolutionary critical-rationalist approach of Hayek and Popper. I employ that approach to sketch answers to the theoretical, historical and practical questions about the legitimacy of government or of systems of moral rules.

Keywords: critical rationality, disagreement, evolution, legitimacy, social contract.

1. Introduction

Social-contract theories of the legitimacy of government, or of a society-wide system of moral rules, are enduringly popular. That is a puzzling fact because

- social-contract theories cannot solve the problem they are intended to solve,
- that problem can be solved in a more straightforward manner,
- it is surprising that social-contract theorists cannot see both those things.

In *section* 2, I explain why social-contract theory cannot succeed. In *section 3*, I explain why social contract theory is redundant. In *section 4*, I comment on some infelicities in some contributions by social-contract theorists to the current debate. In *section 5*, I outline an alternative approach which draws on the evolutionary and critical-rationalist approaches of Friedrich Hayek and Karl Popper (whose outlooks are very similar), but I do not discuss Hayek's own penchant for social-contract theory, which is irreconcilable with his evolutionary view. In *section 6*, I conclude.

Henceforward, I usually leave the qualification 'society-wide' implicit when I speak of systems of moral rules. Throughout I assume that moral theories are objectively true or false and that there are objective facts about what specific agents ought to do in particular circumstances, though I deny that we can ever know for certain which moral propositions are true. I leave it to the moral

subjectivist, non-cognitivist or anti-realist to reformulate my criticisms of social-contract theory in her own terms.

2. Social-Contract Theory Cannot Succeed

The point of social-contract theory is to show that a form of government or a broader system of moral rules and institutions is legitimate and therefore ought to be obeyed or complied with. If a set of free, adult individuals made a contract with each other to accept a form of government or a system of moral rules, then, at least barring unusual circumstances, that form of government or set of rules would be legitimate for those people, who would be morally obliged to honour their agreement. However, that is not the way that governments or society-wide systems of moral rules come into being. The social-contract theorist realises that. He proposes that a form of government or system of moral rules is legitimate for a group of people *if and only if*, under favourable conditions, everyone in the group would (provided he is reasonable) agree to be subject to that form of government or that system of moral rules.

Thus, social-contract theories require that a proposition of the following form is non-vacuously true:

(SC) necessarily, all reasonable persons under condition C will agree that p.

The condition represented by 'C' is some favoured condition, such as being behind a veil of ignorance or being impartial. The proposition represented by 'p' is a complex proposition affirming that everyone in the society should comply with a specific form of government or a particular system of moral rules (or, perhaps, any one of a set of such forms or systems, but I ignore that complication here, since it makes no difference to my argument). The qualification 'necessarily' indicates that the universally-quantified conditional that follows it holds counterfactually. Without that modal qualification, (SC) would be vacuously true due to the falsity of its antecedent, if no reasonable person happened, in fact, to be under condition C. I offer no account here of the kind of necessity involved. A proposition of the form (SC) is vacuously true if and only if the description for which 'C' does duty specifies either an impossible condition or the fact that all reasonable people agree that p, where 'p' represents the same proposition as does 'p' in (SC).

However, it seems clear that no proposition of the form (SC) is non-vacuously true. The reason is that reasonable persons are distinguished by their *critical rationality*, which means that it is, in principle, rationally open to them to question, criticise or deny *any* proposition. What makes critical rationality reasonable is that persons are thoroughly fallible: for all they can know, the beliefs, views or theories they have so far acquired may be mistaken (Popper, 1972[1957], 49–52; see also my 2013, *section 3*). These include not only all

empirical propositions but even one-time supposed self-evident truths of mathematics and logic. For instance, Gottlob Frege thought he could derive all of arithmetic from self-evident axioms of logic, but Bertrand Russell showed that one of Frege's axioms was self-contradictory (Russell 1959, 58–59; Whitehead and Russell 1927, 59); and in contemporary logic all the traditional logical laws have been impugned, as a means to resolving serious logical problems, by competent logicians at leading research institutions (Priest and Thomason 2007, 96–98). If even the law of non-contradiction can be doubted or denied by reasonable people under the best available conditions, it seems that there can be no proposition which could be represented by '*p*' to render (SC) non-vacuously true.

3. Social-Contract Theory Is Redundant

We should not be troubled by the fact that social-contract theory cannot succeed, because social-contract theory is redundant.

Why should the social-contract theorist think that the legitimacy of a form of government or of a system of moral rules depends upon the hypothetical agreement of those who are subject to them? An answer seems to be disclosed in the following thought of David Hume (1888[1739–40], 556): "Nothing is more advantageous to society than such an invention [i.e., government]; and this interest is sufficient to make us embrace it with ardour and alacrity." Ignoring the complication that several forms of government, or systems of rules, may meet the interests of individuals equally well, we might express Hume's thought as follows:

(a) the form of government or system of moral rules in question best serves the interests of the individuals who make up the community,

therefore,

(b) necessarily, every reasonable member of the community will (under favourable conditions) agree to be subject to it.

It should be observed that (b) is a proposition of the form of (SC). It should be clear that (b) does not follow from (a), since reasonable individuals are such that

- they can easily be mistaken about
  - their interests (it may take a lifetime of trial and error to discover what one's interests are)[1]
  - the interests of others
  - what social arrangements will best serve those interests (which is the subject of ongoing social-scientific research),
- some of them (for example, some kinds of environmentalists) might not wish to act in the best interests of themselves or of the individuals who make up the (human) community.

[1] See my forthcoming, *section 4*.

However, what is curious is why the fallacious transition to (b) is made at all. For, surely, if (a) is true, then the form of government or system of moral rules in question is legitimate and the members of the community ought to comply with it. Thus, not only is (b) false and not a consequence of (a); it is entirely unnecessary for the argument for the legitimacy of government or of a system of moral rules. All that is required for that purpose is an argument for (a). Appeal to a hypothetical social contract is otiose as well as cack-handed.

Propositions of the form (a) are contingent. Whether or not a particular form of government or system of moral rules best serves the interests of the individuals who make up a given community depends upon the peculiarities of the environment they inhabit. Further, keeping the environment constant, differing forms of government or systems of moral rules may be appropriate for communities containing differently constituted persons (humans, Martians or whatever). However, it seems a moral truth, and plausibly a necessary one, that:

(c) the form of government or system of moral rules which best serves the interests of the individuals who make up a given community in a given environment is legitimate for that community in that environment.

Some theorists may be reluctant to leave behind (b) in favour of (c) because of fears about authoritarianism or paternalism (see, for example, Sugden 2013, 61–62). If we accept that a form of government or system of rules can be legitimate for a community even though some (perhaps many) people in that community reject it, it may seem that we will be led to accept that all manner of schemes may be imposed on the community, in the face of opposition, in the name of the general good or the dissenters' own good. However, that is not so—at least, in the case of human persons—because the separateness of human persons, in conjunction with their fallibility and capacity for critical rationality, means that their interests are best served by having a right to direct their own lives (see my forthcoming, section 4), which must therefore be recognised in any legitimate form of government or system of moral rules.

4. Social-Contract Infelicities

Robert Sugden seems to be offering an alternative to the usual hypothetical agreement of social-contract theory when, following Hume, he suggests that political institutions and moral norms that obtain in a society may be legitimised through *implicit* rather than explicit agreement. Such implicit agreement

> "rests on a general sense of common interest; each member of society has this sense and is aware that the others have it too [. . .] the idea of mutual advantage is an essential part [. . .] a convention is a mutually beneficial practice that is generally followed in a society; each individual is motivated to follow it both by interest and by a sense of

> justice; but both motivations are conditional on the expectation that others will follow the practice too." (2103, 65–66)

However, if an 'implicit agreement' does not mean the same as a *hypothetical* agreement, as in (SC), it seems it must mean an agreement (albeit implicit) between all people in the society in *actual* circumstances, which is even less plausible than propositions of the form (SC). It is not the case that everyone actually agrees that government is better than anarchy; and even if we pretend that they do, we should acknowledge that they disagree as to whether the *existing* form of government, or the *prevailing* system of moral rules, is legitimate. Similar objections apply to the attempt of Peter Vanderschraaf (2013) to understand the relationship between ruler and ruled as an implicit contractual relationship.

A more useful notion of 'implicit agreement', which will be the one employed henceforward, is illustrated by an example given by Anthony de Jasay:

> "Consider the case of two men Will and John who both want to cross a river. Will has a boat and John has a pair of oars. I believe they would without further ado get into the boat and row across without previous agreement about taking turns at rowing and the fee John would pay Will for using the boat and Will would pay John for using the oars." (2013, 56)

Such implicit agreements, depending on unspoken mutual acknowledgement, are possible only for a small number of people in proximity to each other. They therefore seem unsuitable for the social-contract theorist, who needs a society-wide agreement to moral norms. It seems true that a moral norm implicitly agreed in a small group may spread throughout a society by means of chains of such implicit agreements between members of different small groups. However, unfortunately for the social-contract theorist, the society-wide norm thereby produced would not be one that is implicitly agreed by all. For example, if an implicit agreement to behave in some reciprocal fashion has been established within one small group, some of the members of that group may establish a similar agreement in their dealings with some members of another small group. Those members of that other small group may then go on to extend the norm by implicit agreement within their group; and then other members of that group may spread the norm to another group by implicit agreement with the members of that group with whom they deal; and so on. Through such chains of unspoken mutual acknowledgement a common moral norm could be established throughout a great society. However, it would not be the case that all the members of the society have agreed with each other, implicitly, that the moral norm should obtain in the society. No one has made an implicit agreement with anyone but a small number of people. Indeed, it would be possible for some people, even for everyone in the society, to insist that the norm should not govern the whole of society, even while they are keen to retain it for their interactions with the small number of people with whom they deal. For, each may hold that, while the norm is appropriate for her type of person (white, heterosexual, noble, or

whatever), it is quite inappropriate for some other types who inhabit the same society. The norm may thus be recognised by all as an existing society-wide moral norm established by implicit agreements, even though many in the society do not agree that it is a *legitimate* society-wide moral norm. A chain of implicit agreements to the same moral norm, even where that chain spreads the norm throughout a society, does not amount to a society-wide agreement to the norm holding society-wide.

Gerald Gaus defends the transition from (a) to (b) as follows:

> "[I]t is not enough to show that such [common social] norms are in the interests of all; a person must not only conform to a moral norm, but internalize it, feel guilt, experience the reactive attitudes of resentment and indignation, and be prepared to sanction those who violate it." (2013, 76)

That appears to confuse different questions. What makes a form of government or a system of moral rules legitimate is that there is no better alternative possible way of securing the interests of all. A separate question, to which matters such as conformity, internalisation and sanctions are pertinent, is what it takes to make a system of moral rules (including those pertaining to government) stable over time. However, even with regard to the latter question, (b) is not required. It is sufficient if *enough* of the populace think that the form of government or system of moral rules is legitimate. After all, we can expect every society to have its radicals and anarchists.

Indeed, it would be a bad thing if everyone did internalise the prevailing system of moral rules, at least in a developed society in which rules are codified and legislated. Such a society, if it is to be capable of adapting its form of government or system of moral rules to changed circumstances, or in the light of new knowledge about unchanged circumstances, will have, from time to time, some members who propose novel changes to the existing forms, thereby disagreeing with those of their fellows who still endorse the current arrangements. In such a society, disagreement over the current form of government or system of moral rules is a condition of adaptation and of progress. Thus, it is not only highly unlikely, but also highly undesirable, that we should ever achieve what Gaus calls "public justification", that is, "the public knowledge that a stable practice is morally acceptable to all the participants" (2013, 80).

Gaus eschews the view that he labels "The Moral Autonomy Conviction" (2013, 73), or "the MAC", for short. On the MAC:

- "[e]ach competent moral agent [. . .] properly arrives at her own judgment as to what MORALITY requires [. . .] and this judgment does not necessarily refer to any collective determination" (Gaus 2013, 73);
- the agent has "no need to have her judgments confirmed by others; if she concludes that she has deliberated as well as she can, and that she is justified in believing that IT [i.e., MORALITY] requires φ, then that is what she proclaims, and so, in the name of IT, she concludes that all must φ" (Gaus 2013, 74).

The problem with the MAC, says Gaus, is that it sees morality as a personal norm, though one that purports to direct all. But "a society in which most have this, and only this, conception of morality, will find it exceedingly difficult to adopt common social norms, for they require a collective determination of what rules are to be internalized, when I can expect others (or myself) to express guilt and remorse, and when I expect others (or myself) to be sanctioned" (2013, 76). Gaus, mistakenly, wants to spell out this 'collective determination' in terms of a social contract.

There are at least two problems with the MAC. First, the MAC seems to claim that the competent moral agent accepts any moral proposition only as a consequence of her own rational evaluation. That claim seems to be absurd because a rational evaluation would involve considering reasons pro and con; and such reasons would themselves be propositions which the agent has accepted; but if the agent's acceptance of these propositions in turn must be a consequence of the agent's own rational evaluation, then the agent is involved in a vicious infinite regress. The fact is that most of the propositions we accept (including the moral ones) are not propositions we have accepted as a consequence of rational evaluation but are rather inherited biologically or culturally and are, *in that sense*, products of 'collective determination'. However, there is no requirement that, if we are reasonable, we will all agree on these inherited views. Our critical rationality enables us to stand back from any of our inherited views, criticise them and, sometimes, replace them with better ones, even though, in this process of rational appraisal, we inescapably make use of other inherited views which are the result of biological or cultural evolution (Hayek 1967[1963], 60–63; 1978[1970]; Popper 1972[1949], 131–32; 1994b, 134–39).

Second, and somewhat tangentially, while one who recognises that reasonable persons are critically rational need not accept the mistaken notions that judgements or beliefs can be "confirmed" or "justified" (Popper 1983, 18–34), she will insist that knowledge is social, not only because it depends on theories which are products of species and societal evolution, but also because the part of that knowledge that results from rational evaluation depends upon intersubjective criticism and testing (Popper 1945, chapter 23). An agent's rational evaluation of a proposition must survive the severest criticism of her intellectual adversaries before it can count as knowledge; though the process of intersubjective criticism may involve the agent responding to criticism by criticising it, or by modifying her view in ways which increase its explanatory value, rather than by simply giving up her view (Popper 1959, sections 19-20). So, whereas the MAC may encourage "each person to claim the status of the truth giver [. . .] [and thus] to invite civil conflict" (Gaus 2013, 74), the recognition of our capacity for critical rationality encourages each person to co-operate in learning from others while retaining her independent initiative.

Jan Narveson appears to claim that, in order to show that government, or a moral rule, is legitimate, "[w]e 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" (2013, 89). That claim seems false. It would be sufficient for us to show that the rule *would* benefit each individual better than any alternative rule, whether or not each individual can see that we have shown it. We might have to convince each individual of the truth of what we have shown *if* we had to obtain the agreement of each individual to the rule before she complied with it. That would be the situation of separate individuals coming together to form an *association* (a club, for instance). However, any such individuals are already members of a *society*, through membership of which each has been inducted in a system of moral rules without ever having agreed to that system of moral rules. Of course, *any* of those rules is, in principle, open to question by the critically rational individual; but in the meantime that individual will for the most part comply with those rules out of habit.

5. Sketch of an Alternative

The problem of the legitimacy of government or of a system of moral rules has a theoretical, a historical and a practical aspect. I consider these in turn.

The theoretical question is:

> (q1) what, if any, form of government, or system of moral rules, is legitimate for a community of individuals of type I in an environment of type E?

Given (c), the answer depends upon which form or system "best serves the interests of the individuals who make up the community". The quoted phrase needs spelling out, and different theorists will spell it out in different ways. Given any sufficiently detailed spelling out, the answer to (q1) becomes a matter for the social sciences, and thus revisable in the light of new social-scientific knowledge. The results of the social sciences may also be brought to bear on some of the theoretical disputes between alternative detailed specifications of the notion, *best serves the interests of the individuals who make up the community*, since it is, at least in part, a factual matter what is in the best interests of a person of a particular type. It should therefore be possible to evaluate rival answers to (q1) in terms of empirical and explanatory adequacy, as well as in other ways.

Given specified types I and E, the legitimate form of government, or system of moral rules, will be that identified by the true answer to (q1) for those types. However, each of the rival answers to (q1) which we produce will be a theory which, given our thoroughgoing fallibility, we cannot know for certain to be true. Even if, for specified types I and E, we manage to obtain a theory which answers (q1) and which is far superior to its rival theories, it is always possible that, in future, an even better theory will be proposed; and, indeed, theorists should be encouraged to find better alternatives even to our best theories. In consequence, we will *never* be in a position to say for sure which form of government, or system of moral rules, is legitimate. The best we can do is to keep improving on our

current theories. We start with the answers that we have; we try to improve them; and within that context we try to find better ones. If we ever find ourselves in a position where everyone not only agrees which theory is currently best, but also accepts that theory as true (instead of searching for a better one), our epistemic task will be to stimulate the disagreement, and thus the search for alternatives, on which the growth of knowledge depends (Popper 1945, chapter 23; 1983, 67–71; 1994a).

The historical question is:

> (q2) how could it come about that the majority, perhaps the vast majority, of the members of a society view the form of government or the system of moral rules under which they live as legitimate?

We know that forms of government and systems of moral rules have not been introduced by explicit agreement of all who are subject to them. However, we noted in *section 4* that small-scale norms introduced by implicit agreement can spread to become society-wide norms that no one agreed to (either explicitly or implicitly). There is also a tradition in economics of explaining how society-wide norms can evolve as the unintended by-products of actions which had other aims, in Ferguson's famous phrase, "the result of human action but not the execution of any human design" (1782, 205). Such evolutionary processes can result in society-wide norms, such as the use of gold as a common medium of exchange, even though none of the barter exchanges from which the regularity evolved involved any implicit agreement on a common medium of exchange (Menger 2007[1871], 257–62).[2] However, evolved society-wide norms, as we noted in *section 4*, need not be acceptable as norms to all, or even to most. What explains the fact that mere regularities of social behaviour can come to be seen by most members of the society as legitimate moral norms is that people are not only rule-following creatures, but are also *rule-seeking* creatures; and, specifically, they are inherently on the lookout for moral norms. Just as our inborn tendency to search for laws of nature helps us to satisfy our need to make sense of the natural world (Popper 1972[1957], 42–48), so our inborn tendency to search for moral laws helps us to satisfy our need to make sense of the social world, and it predisposes us to hypothesise that existing social regularities are legitimate moral norms (though we hold such hypotheses open to criticism and testing insofar as we are critically rational).

The practical question is:

> (q3) how do we bring it about that we live under a legitimate form of government or a legitimate system of moral rules?

[2] Hillel Steiner (2013) criticises accounts of the greater efficiency of money payments over barter which assume contemporary decision theory; but, contra his final sentence, his criticism does not seem to impugn Menger's account of how money evolves out of barter transactions, since that account can be propounded independently of the assumptions of contemporary decision theory.

The general approach to answering (q3) is similar to that taken in answering (q1): we start with what we have, we try to improve it, and within that context we try to think up new ways of doing things that are better. Our theories giving answers to (q1) will be our guide in identifying improvements to our current systems of government and morals. However, the debate over the practical question, (q3), will be of interest not only, or even primarily, to theorists. All sorts of 'practical people' whose interest in theory, or in epistemic virtue, is minimal or non-existent, will be keenly interested in bringing about or blocking changes to current systems of government and morals; and they may indeed be reasonable in pursuing their non-epistemic ends. So, *even if* all the theorists agree on which answer to (q1) is currently best, and *even if* they all accept that answer as true (instead of seeking alternatives to it), they can expect to meet disagreement of diverse sorts from many non-theorists. However, not only is unanimous agreement among reasonable people with regard to (q3) not to be expected; it is not needed in order to bring about substantial change in existing moral norms and institutions. The best we can do to move in the direction of a form of government or system of moral rules that is legitimate, is to make use of existing institutional arrangements, sympathetic organised interests and available techniques of persuasion to bring about changes, in accord with our favoured answer to (q1), that a significant proportion of the population (perhaps a majority) do not agree with. That is doubtless trite; but who has a better alternative to offer? Not the social-contract theorist, whose offering seems irrelevant to (q3), as well as to (q2) and (q1).

6. Conclusion

Social-contract theory is intended as a solution to the problem of what makes a government or a system of moral rules legitimate. It cannot solve the problem because reasonable people are critically rational: they are always rationally entitled to disagree. Further, the problem can be solved without social-contract theory by invoking the interests of the members of society. Social-contract theory should therefore be abandoned.

An account of the social contract in terms of implicit agreements cannot work because even a society-wide chain of small-scale implicit agreements to the same norm does not add up to a society-wide agreement that the norm is legitimate society-wide. Some social-contract theorists seem to mix up the question of legitimacy with the question of stability; but even stability does not require the agreement of all reasonable people. An approach which I have suggested may solve that problem is to combine an account of the evolution of social regularities with the human propensity to look for moral norms in their attempt to make sense of social reality. However, it would not be a good thing if *everyone* viewed the existing form of government or system of moral rules as legitimate, because

progress in morals and politics, as in other areas of knowledge, depends upon disagreement.

When it comes to the matter of how we improve our institutions or rules to get them closer to those we think legitimate, social-contract theory has nothing useful to say.

References

De Jasay, A. (2013), "Conduct and Contract", *Rationality, Markets and Morals* 4, 53–60.

Ferguson, A. (1782), *An Essay on the History of Civil Society*, fifth edition, London: T. Cadell.

Frederick, D. (2013), "Popper, Rationality and the Possibility of Social Science", *THEORIA* 28(1), 61–75.

— (forthcoming), "Voluntary Slavery", in: *Las Torres de Lucca*.

Gaus, G. (2013), "Why the Conventionalist Needs the Social Contract (and Vice Versa)", *Rationality, Markets and Morals* 4, 71–87.

Hayek, F. (1967[1963]), "Rules, Perception and Intelligibility", in: Hayek, F., *Studies in Philosophy, Politics and Economics*, London: Routledge and Kegan Paul, 43–65.

— (1978[1970]), "The Errors of Constructivism", in: Hayek, F., *New Studies in Philosophy, Politics, Economics and the History of Ideas*, London: Routledge and Kegan Paul, 3–22.

Hume, D. (1888[1739–40]), *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, Oxford: Clarendon Press.

Menger, C. (2007[1871]), *The Principles of Economics*, Auburn: Ludwig von Mises Institute.

Narveson, J. (2013), "Social Contract: The Last Word in Moral Theories", *Rationality, Markets and Morals* 4, 88–107.

Popper, K. (1945), *The Open Society and its Enemies*, London: Routledge and Kegan Paul.

— (1959), *The Logic of Scientific Discovery*, London: Hutchinson.

— (1972[1949]), "Towards a Rational Theory of Tradition", in: Popper, K., *Conjectures and Refutations*, fourth (revised) edition, London: Routledge and Kegan Paul, 120–135.

— (1972[1957]), "Science: Conjectures and Refutations", in: Popper, K., *Conjectures and Refutations*, fourth (revised) edition, London: Routledge and Kegan Paul, 33–65.

— (1983), *Realism and the Aim of Science*, London: Routledge.

— (1994a), "The Myth of the Framework", in: Popper, K., *The Myth of the Framework*, London: Routledge, 33–64.

— (1994b), *Knowledge and the Body-Mind Problem*, London: Routledge.

Priest, G. and N. Thomason (2007), "60% Proof", *Australasian Journal of Logic* 5, 89–100.

Russell, B. (1959), *My Philosophical Development*, London: George Allen & Unwin.

Steiner, H. (2013), "Invisible Hand Processes and the Theory of Money", *Rationality, Markets and Morals* 4, 44–52.

Sugden, R. (2013), "Contractarianism as a Broad Church", *Rationality, Markets and Morals* 4, 61–66.

Vanderschraaf, P. (2013), "A Governing Convention", *Rationality, Markets and Morals* 4, 131–56.

Whitehead, A. and B. Russell (1927), *Principia Mathematica*, Volume 1, second edition, Cambridge: Cambridge University Press.