Justification, Choice, and Promise:
Three Devices of the Consent Tradition in a Diverse Society

Gerald Gaus

1. CONSENT IN THE SOCIAL CONTRACT TRADITION

The twin ideas at the heart of the social contract tradition are (*i*) that persons are naturally free and equal, and (*ii*) that genuine political obligations must in some way be based on the consent of those obligated. In the *Second Treatise* Locke rejects the claim of Filmer — and the entire conception of the Great Chain of Being on which the medieval order was founded — that some people have natural authority over others, and thus those over whom authority is exercised have a natural duty to obey. Filmer explicitly attacked the "dangerous" view of the "natural liberty of mankind."[1] "Every man that is born," wrote Filmer, "is so far from being free-born, that by his very birth he becomes a subject to him that begets him. Under which subjection he is always to live, unless by immediate appointment from God, or by grant or death of his father, he became possessed of that power to which he was subject."[2] Locke, in reply, insists that "Men being...by Nature, all free and equal and independent, no one can be put out of this estate and subjected to the Political Power of another, without his own *Consent*."[3] Almost right from the beginning the ideas that persons are free and equal, and that political authority must be based on the consent of the governed, went hand-in-hand.

Over the last thirty years A. John Simmons has brought home a simple but fundamental truth: while consent can indeed reconcile our status as free and equal with submission to political authority, the undeniable fact is that few actually consent to — in the sense of agreeing to, choosing, or promising to obey — political authority. To Simmons the "legitimacy" of the state is "its exclusive right to impose new duties on subjects by initiating legally binding directives, to have those directives obeyed," which generates corresponding obligations on its citizens.[4] If

---

[1] Robert Filmer, *Patriarcha* in *Patriarcha and Other Writings*, edited by Johann P. Sommerville (Cambridge: Cambridge University Press, 1991), p. 1.

[2] Filmer, *Observations Upon Aristotle's Politiques Touching Forms of Government Together with Directions for Obedience to Governours in Dangerous or Doubtfull Times*, in ibid., p. 282.

[3] John Locke, *The Second Treatise of Government* in *Two Treatises of Government*, edited by Peter Laslett (Cambridge: Cambridge University Press, 1960), §95. Emphasis in original.

[4] A. John Simmons, *Justification and Legitimacy: Essays on Rights and Obligations* (Cambridge: Cambridge University Press, 2001), p. 137.

such legitimacy requires that people have actually consented to such a state, it turns out that there are no states that possess general legitimacy *vis a vis* their citizens.[5] This gives rise to a certain sort of "philosophical anarchist" position according to which, while it is conceptually possible that there be such a thing as a genuinely authoritative state which imposes obligations on its citizens, as a matter of fact no present state meets the necessary condition — general consent has not been obtained — and so all states are, at least with respect to the great majority of their citizens, illegitimate.[6] Such philosophical anarchists maintain that citizens should "act morally" toward each other but not accept distinct moral obligations to obey the state.[7]

Simmons acknowledges that Kantian-inspired views do not lead to this result. The Kantian stresses that since each person has good reason to endorse the existence of states insofar as they are "necessary for the realization of freedom and rights and justice," each person has an obligation to "enter civil society and accept the duties society imposes."[8] Thus, for a Kantian such as Thomas Nagel, "[t]he task of discovering the conditions of legitimacy is traditionally conceived as that of finding a way to justify a political system to everyone required to live under it."[9] There is a type of "rational consent" here, since the justification is owed to each person, appealing to her distinct ends and values. On a typical contemporary Kantian view, justification requires the hypothetical consent of each idealized person (persons who are aware of their reasons and act on them); each endorses principles and institutions given her own reasons. Simmons, in reply, insists that the Kantian conflates the justification of the state (are states necessary for justice? Do I have reasons to do as the state directs?) with their legitimacy (are we obligated to obey the laws of our own state because they have been legislated by our state?). Even if each could be shown that there are good reasons to endorse the existence of states, and even the existence of one's own particular state, it would not follow that one is bound to acknowledge that one's state has a moral power over one to create new obligations. Simmons concludes that "[a]ppeals to hypothetical choice, acceptability, or reasonable nonrejectability have a very different moral basis and force than do appeals to actual choice….Even appeals to what ought to be chosen in the light of the

---

[5] Ibid., pp. 155-6. On the difficulties of obtaining general consent, see Simmons' *Moral Principles and Political Obligations* (Princeton: Princeton University Press, 1979), chaps. 3 and 4.

[6] Simmons, *Justification and Legitimacy*, chap. 6.

[7] Ibid., p. 153.

[8] Ibid., p. 140.

[9] Thomas Nagel, *Equality and Partiality* (New York: Oxford University Press, 1991), p. 33.

individual's own interests and values are quite different from appeals to that individual's actual choices."[10] Justification via the consent of idealized agents who respond only to their reasons to endorse or reject is very different from legitimacy based on real people's actual choices. The Lockean and the Kantian views, Simmons insists, have "quite different" moral foundations.

In this essay I seek to bring the Kantian and Lockean contract traditions together. Kantian rational justification and actual choice are complementary devices through which our freedom and equality can be reconciled with moral and political authority. We should not think that there is simply one way by which relations of moral and political authority can be reconciled with our status as free and equal. I defend here three distinct devices through which freedom and authority may be reconciled: justification to others, actual social choice, and promise. All three are aspects of the "consent tradition" broadly construed. As all-too-often occurs, philosophers have succumbed to endless disputes about *which one* must ground authority and its power to obligate when, we shall see, all three are important ways in which a free and equal person can live under the authority of others.

2. MORAL AUTHORITY AMONG FREE AND EQUAL PERSONS: THE ROLE OF JUSTIFICATION

2.1 The necessity of authority: the instrumental and constitutive interpretations

As Simmons noted above, according to Kant the source of political obligation is our duty to enter civil society.

> Although experience teaches us that men live in violence and are prone to fight one another before the advent of external compulsive legislation, it is not experience that makes public lawful coercion necessary. The necessity of public lawful coercion does not rest on a fact, but on an a priori Idea of reason, for, even if men to be ever so good natured and righteous before a public lawful state of society is established, individual men, nations and states can never be certain they are secure against violence from one another because each will have the right to do what *seems just and good to him*, entirely independently of the opinion of others.[11]

Kant goes on to insist that justice is absent in the state of nature because each relies on his own judgment, and thus "when there is a controversy concerning rights (*jus controversum*), no competent judge can be found to render a decision having the force of law."[12] Thus Kant argues that individuals are obligated to abandon the "state of nature" and enter into a "juridical state of affairs."[13]

---

[10] Simmons, *Justification and Legitimacy*, p. 147.
[11] Kant, *The Metaphysical Elements of Justice*, 2nd edition, edited and translated by John Ladd (Indianapolis: Hackett, 1999), p.116 [§43]. Emphasis added.
[12] Ibid.
[13] Ibid., pp. 114-15.

The idea that there cannot be justice among free and equal persons when each has the right to act on his own view of justice is open to two interpretations: the instrumental and constitutive. Locke clearly presents an *instrumentalist* view of the injustice that arises in the state of nature when each acts on his own private judgment about justice. In the state of nature, Locke maintains,

> [t]here wants an *establish'd*, settled, known *Law*, received and allowed by common consent to be the Standard of Right and Wrong, and the common measure to decide all Controversies between them. For though the Law of Nature be plain and intelligible to all rational Creatures; yet Men being biassed by their Interest, as well as ignorant for want of study of it, are not apt to allow of it as a Law binding to them in the application of it to their particular Cases.[14]

Self-interest, passion, and desire for revenge lead people to misapply the law of nature, and over-punish perceived violations against them.[15] Consequently, to secure a condition in which justice obtains, "all private judgment of every particular Member" must be

> excluded, [and] the Community comes to be Umpire, by settled standing Rules, indifferent, and the same to all Parties; and by Men having Authority from the Community, for the execution of those Rules, decides all the differences that may happen between any Members of that Society concerning any matter of right; and punishes those Offences which any member hath committed against the society, with such penalties as the Law has established…."[16]

A system of political authority, we might say, provides an impartial system of justice, and so overall a more just order. If this is how the authority of the state provides for the justice lacking in the state of nature, Simmons' sharp distinction between justifying the state, and submitting to the authority of the state, makes perfect sense. To justify the state is to realize that it performs the task of providing an impartial umpire, and so an effective system of justice. Realizing that, one can think the state which does its job is a good thing and be delighted when we have such a state. It would not follow, however, that one should submit to its authority in the sense of acknowledging an obligation to do as it instructs (just because it instructs it). One can, in the spirit of the philosophic anarchist, say "it is well and good we have the state, but I shall act towards my fellows on morality as I see it." The state is instrumentally effective is securing a more just society, but this instrumental effectiveness does not entail that you must submit to its authority.

---

[14]   Locke, *Second Treatise*, §124.
[15]   Ibid., §§125-26.
[16]   Ibid., §87.

Kant points to a much deeper sense in which justice is absent from the state of nature (an "apriori idea"), one which leads us back to the problem of authority among free and equal moral persons. Stephen Darwall has recently stressed the way in which interpersonal moral claims involve "authority relations that an addresser takes to hold between him and his addressee."[17] When I make a moral claim on you not to φ, I am not making a request that you refrain from φ, or calling attention to my opinion that φ is immoral: I am issuing an imperative that you must not φ.[18] If I decide that your action falls under the purview of social morality, I see it as my business that you refrain from φ-ing.[19] I claim authority over you in the sense of claiming a *standing* to direct your action. And I suppose that you are under obligation to comply. As Margaret Gilbert observes:

> To say that someone has the standing to do something means simply that he is in a position to do it. If someone lacks standing to do it, the question whether he is justified in doing it does not arise. For he *cannot* do it. One who lacks the standing to make a certain demand or issue a rebuke can, of course, utter a purported rebuke or make a purported demand. He can speak in a rebuking or demanding tone. His target, meanwhile, may have little interest in this if it is possible to question his standing to rebuke or demand. His target may well respond in some such words as these: "It's none of your business, so… forget it!'.[20]

And here we confront the core problem: how can one have the standing to command the performances of others while treating them as free and equal moral persons? You are claiming that *they* are obligated to obey *you*. Recall that for Locke, fundamental to our natural freedom and equality is that "the *Natural Liberty* of Man is to be free from any Superior Power on Earth, and not to be under the Will or Legislative Authority of Man, but to have only the Law of Nature for his Rule."[21] To be morally free is to have *only* the law of nature as one's rule. This implies, though, that the private judgment of others about the demands of morality does not rule one. As a morally free person one employs one's reason to understand the requirements of the law of nature, and one submits to rule by that law, which is the rule of reason. Moral freedom is not freedom from morality, but freedom to directly employ one's (private) reason to determine what morality requires. Such freedom appears

---

[17] Stephen Darwall, *The Second-person Standpoint: Morality, Respect and Accountability* (Cambridge, MA: Harvard University Press, 2006), p. 4.

[18] Ibid., pp. 10-11, 76.

[19] Kurt Baier, *The Moral Point of View*, abridged edn. (New York: Random House, 1965), xviii-xix.

[20] Margaret Gilbert, *A Theory of Political Obligation* (Oxford: Clarendon Press, 2006), p. 147. See also pp. 103ff, 147ff, 245ff. Gilbert stresses the close relations between the concepts of standing, authority, command and obligation (p. 46).

[21] Locke, *Second Treatise of Government*, §22.

straightforwardly at odds with the claims to authority over others that constitute the practice of interpersonal moral authority.

Unless there is an authority endorsed by public reason — the reason of all — a moral claim is simply the attempt of one person's private judgment to rule over others. For the constitutive view, unless there is an authority that expresses a public reason — a judgment that you and I both share — the practice of interpersonal morality is inconsistent with our fundamental status as free and equal. Or to put it more personally: unless one's moral claims are endorsed by public reason, they manifest disrespect for the freedom and equality of others, and unless one submits to the authority of public reason one cannot have respectful moral relations with others. Notice that on the constitutive view, Simmons' philosophical anarchism leads to either moral authoritarianism or moral nihilism. If you say that public authority is well and fine, but you see no reason why you should submit your judgment to public reason, you are no longer able to proceed with moral demands that treat others equally. You must either forgo moral claims entirely (nihilism) or you are thrown back to claims of moral authority over others based on your private reason (authoritarianism). Because the philosophical anarchist does not appreciate that all moral claims are exercises of authority over others, his effort to reject submission to authority while keeping interpersonal morality is ultimately incoherent.

2.2 The public authority of justified social norms

The social contract tradition makes a very quick move from the conflict of private judgment about morality to the need for an umpire with political authority to judge all disputes about right according to an authorized public reason, to which all submit.[22] However, as T.H. Green pointed out in his review of the social contract tradition, Hobbes, Locke, and Rousseau overlooked the fact that between individual judgments about individual rights, and the determinations of public political authority, lies the social authority of practices, norms, and conventions.[23] We might think of this as taking seriously Kant's idea that the solution to the conflict of private judgment is to enter into *civil society* — while we need to avoid a condition in which each does what "seems just and good to him, entirely independently of the opinion of others," this need not immediately drive us into *political* society. When a moral

---

[22] I have argued this more fully in "The Property Equilibrium in a Liberal Social Order," forthcoming in *Social Philosophy & Policy*.

[23] Green, *Lectures on the Principles of Political Obligation* in *Lectures on the Principles of Political Obligation and Other Writings*, edited by Paul Harris and John Morrow (Cambridge: Cambridge University Press, 1986), §113.

rule *R* is a social norm it is commonly recognized as a rule followed within the group and in fact there is general conformity to it.[24] It is common knowledge that it is the rule of our group: each knows that it is our rule, and knows that others know it to be.

Now when a moral rule that is accepted as a social norm is also endorsed by the private reason of each free and equal moral person — each moral person's deliberative rationality — moral authority is reconciled with our freedom and equality. Suppose rule *R* is endorsed by the members of a social group *G* in the sense that *G** — somewhat idealized members of *G* who deliberate solely on their relevant values, ends, and personal moral intuitions — all have sufficient reason to endorse *R* over all the alternatives. It is true that what is endorsed by *G** is not the same as what is actually endorsed by, or would if asked be actually endorsed by, *G*. Different theories of public reason relate *G* and *G** different ways, but in all plausible versions a member of *G** deliberates on the reasons to endorse a moral norm that her counterpart in *G* has, but to which her actual counterpart does not always pay attention. This is, I think, a far less controversial idea than Lockeans such as Simmons would have us believe. We are a complex combination of selfish and moral creatures: the moral system, we might say, has developed on top of an earlier selfish set of motivations.[25] In less psychological terms, we are often tempted to put aside our normative commitments and cheat, even when we accept that this violates a norm we have good reason to endorse and internalize. Morality is required just because our will does not always align with our reasons to restrain our will and abide by social rules. As a rational moral agent, to know whether a rule is one to which my private reason accords moral authority, a test which consults my present will ("Do you now choose the rule?") is inherently flawed, for it supposes that our moral commitments and reasons will be dominant in our thinking (rather than, say, narrowly self-interested reasons that induce us to bargain and bluff).

If *R* is a social rule endorsed by the relevant reasons of all, when I demand that you conform to *R* I am not merely employing my private reason to instruct what you must do: I am appealing to a commonly recognized rule that our joint reasons endorse as a rule to live by. In familiar Kantian terms, the rule expresses legislation

24 I provide a detailed analysis of social rules in *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded Word* (Cambridge: Cambridge University Press, 2011), Chap. III.

25 See Peter J. Richerson and Robert Boyd, "The Evolution of Free Enterprise Values," in *Moral Markets: The Critical Role of Values in the Economy*, edited by Paul Zak (Princeton: Princeton University Press, 2008), p. 114; Daniel Freidman, *Morals and Markets*: *An Evolutionary Account of the Modern World* (New York: Palgrave Macmillan, 2008), chap. 1.

in the commonwealth of ends, in which the law comes from all and applies to all. In addition, when *R* is a social rule, it is common knowledge among us that *R* is the rule to which all submit. Because the authority of the rule is self-imposed, the claim to authority over others is reconciled with the status of all as free and equal.

3. MORAL AUTHORITY AMONG FREE AND EQUAL PERSONS: THE ROLE OF CHOICE

3.1 Two shortcomings of reconciliation through justification

Unfortunately for Kantians, justification as a device of reconciling moral authority with recognition of others as free and equal falls short of a fully adequate reconciliation, and this for at least two reasons. (*i*) In a pluralistic social world, in which individuals have a plethora of values, ends, and moral ideals, it is most unlikely that any candidate social rule will be endorsed by all members of *G** as the best. On any plausible account of the deliberations of members of *G**, we can imagine that they will endorse very different rankings of candidate moral rules. The only way to induce unanimity is to greatly, indeed radically, idealize the members of *G** so that they essentially evaluate candidate moral rules in the same way, as does Rawls' argument from the original position. The veil of ignorance first excludes "knowledge of those contingencies which set men apart…."[26] Rawls then attributes to the parties a concern with primary goods that provides a basis for their common deliberation. Insofar as we consider ourselves as agents devoted to some ends, they are what we need. When abstracted to the common status of agents devoted to their own (unknown) evaluative standards (values, comprehensive conceptions of the good and so on), because "everyone is equally rational and similarly situated, each is convinced by the same arguments."[27] So although the original position begins by posing a problem of collective choice, the problem is reduced to choice by one person, and so there is no disagreement.[28] This, though, invites the Lockeans' charge that the subjects of justification are so far removed from actual people and their reasons as to be irrelevant what real agents have reason to do. Although, as I have argued, it is appropriate to employ a degree of idealization that allows us to model a person's reasons relating to her relevant values, ends, and moral ideals, an extreme idealization that reduces the deliberations of members of a pluralistic society to that of a representative, single, abstract, person ignores our real problem: the justification of authority under conditions of wide disagreement.

---

[26]   John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1999), p. 17.

[27]   Ibid., p. 120.

[28]   Ibid., pp. 120-121.

(*ii*) Suppose, however, that a deliberative model was able to show that all members of G^* would concur on some rule R. The problem of disagreement in judgment still would not be solved, for people — members of G^* — would continue to disagree, invoking different interpretations of R. Hobbes stressed this problem. "All laws, written and unwritten," Hobbes tells us, "have need of interpretation."[29] Hobbes insists that the exercise of our rationality is always fallible: "no one man's reason, nor the reason of any one number of men, makes the certainty."[30] Rational people aim at what Hobbes calls "right reason" — true rationality, which reveals the truth. However, because everyone's exercise of rationality is fallible, we often disagree about what is right reason; the private use of reason leads to disagreement and, thought Hobbes, conflict. It was because he saw all substantive rules as open to different interpretations by different people's use of reason that Hobbes insisted that adjudication of moral disputes required a sovereign.[31] So long as we disagree about the interpretation of the rules, our basic problem remains. On the basis of your interpretation of R, you demand that I not ϕ, I reply that as I see it, R allows me to ϕ. For you to continue to demand that I refrain from ϕ is to assume authority over me on the basis of your merely private reason. Or at least so it seems.

3.2 *The eligible set*

The two shortcomings of reconciliation through justification have the same form: when people seeking moral relations with others through authoritative common rules consider their relevant reasons, either to endorse a rule or to interpret a rule, they disagree. In a pluralistic society, self-legislation would appear to lead to disagreement. If we take this problem seriously, we must revise the typical Kantian-inspired justification model. Instead of constructing a deliberative justificatory model in which all members of G^* concur on the same rule (or the same interpretation of the rule), we should suppose that the outcome of any plausible deliberative model among members of G^* will yield a set of proposals.

Let us assume that a member of G^*, G_i^*, proposes the rule R_i, regulating some area of social life that, given her values, ends, and moral ideals, is optimal. If she were the moral dictator — if social morality conformed perfectly to her private reason (as many moral philosophers seem to suppose it should) — this would be the

[29] Hobbes, *Leviathan*, edited by Edwin Curley (Indianapolis, IN: Hackett, 1994), p. 180 (chap. 26, ¶20).

[30] Ibid., p. 23 (chap. 5, ¶3).

[31] I consider Hobbes' position in detail in "Hobbes' Challenge to Public Reason Liberalism" in *Hobbes Today*, edited by S.A. Lloyd (Cambridge: Cambridge University Press, forthcoming).

rule that all would endorse as authoritative. The problem is that under any remotely realistic level of idealization, other members of G^* are almost certain to disagree that this is the optimal rule given *their* values, ends, and moral ideals. Other members of G^* must consider how well R_i stacks up against other proposals. To simplify, let us suppose that each member of G^* is able to order all proposals based on her relevant values, ends, and moral ideals. Having constructed such an ordering she has evaluated the proposals from best to worst in relation to how well they advance or express her values, ends, and moral ideals.

   Not all the proposals, however, will qualify as eligible moral rules. *First*, following Kurt Baier, R. M. Hare, Rawls and many others, we must suppose that there are some formal constraints on what can constitute a *bona fide* moral rule.[32] An eligible rule must have a certain generality, the rule must be public at least in the sense that its content can be taught to new members of the group, it must be determinate enough to resolve conflicts within the group about what is the correct action, it must generally be accepted that its requirements override personal aims and inclinations, at a minimum the rule must not be viewed as hostile to the basic good or well-being of any member of the group, and a person must be able to endorse the rule whatever role she occupies under it (reversibility).[33] *Secondly*, it is important to recall that to accept a rule as authoritative entails that others can demand that you comply regardless of your personal aims and inclinations, appropriately rebuke and punish you for not complying, and that you appropriately feel guilt for failing to comply. For almost all members of G^* these will be real costs of moralization: one allows that one's actions become the business of the public, and that you are to conform your actions to public norms regardless of your own aims, and are answerable to others for failing to do so. Now a member of G^* will not only deem ineligible any rule that fails to conform to the formal constraints on moral rules, but as she reviews the proposed rules in terms of her reasons to endorse them, she must take into account the costs of moralizing an activity. At some point she will almost certainly decide that, given the extent to which some R_i advances her overall values, ends and moral ideals, these benefits are exceeded by the costs of moralization. She will thus deem any such rule ineligible as a genuinely morally authoritative rule that she can endorse.[34]

---

[32]   Another way to model this idea is that only proposals that meet these constraints are advanced.

[33]   I explain and justify such traditional constraints, and compare them to those advanced by Baier, Hare and Rawls, in *The Order of Public Reason*, §15.2.

[34]   Though it could still be adopted as a useful social convention. We should not think that to reject moral regulation of some area of life entails that it will not be socially regulated at all. In

Each member of *G** will, then, divide her ranking of proposals into an eligible and an ineligible set. Let us call *S* the socially eligible set — the set of all proposed rules that are not in the ineligible set of any member of *G**. If *S* is null there is no eligible rule on this matter: that would be simply to say that no rule R_i conforms to the formal principles of right in the eyes of every member of G*, and/or there is no moral rule that, for each member of *G**, the reasons to endorse it outweigh the costs of moralization. Of course any Kantian theory must worry that this will generally be the upshot of attempts at moral justification — that would entail the dismal result that there is no such thing as a publicly justifiable moral rule.[35] Kantian theory has, in my view, plausibly argued against this dismal result (where all we are left with is the choice between moral authoritarianism and nihilism), but then has pushed on to a much more controversial claim: *S* is a singleton. Given evaluative pluralism this is most unlikely. A plausible Kantian view under conditions of wide-ranging pluralism will be left with a socially eligible set *S* that is neither null nor a singleton. No social order has been able to exist without moral rules covering, for example, basic rights of the person, harm to others, and property rights. Effective moral rules on these matters are a great good, and this would be recognized by an even moderately idealized *G**. But there is no reason to think that there is only one acceptable moral rule to regulate each of these matters.

3.3 *Moral equilibrium*

Can a group of free and equal people come to accept a common rule, even though plausible justificatory models are indeterminate? Suppose that members of *G** take the results of the justificatory argument, and assume that it is the limit of the morality dictated by impartial reason. Moral justification, we might say, is indeterminate, specifying only a set *S*. Insofar as justification has had an outcome, it is *S*: what is common to *S* defines the set of actions that are certainly morally prohibited, required, or permitted. Beyond that, each is free to act on the moral rule within *S* as she chooses. Now each person has two concerns: (*i*) to induce the group to come to adopt his preferred rule and (*ii*) to come to act on a common moral rule with others. The first concern is obvious; a moral agent has good reason to induce others to adopt the rule that, given his private reason, he thinks best. In addition,

---

the last hundred years many rules of sexual conduct have become "de-moralized," though in many areas there are still significant social conventions.

[35] It would show that Filmer was ultimately correct: "We do but flatter ourselves, if we hope ever to be governed without an arbitrary power. No, we mistake. The question is now, whether there shall be an arbitrary power, but the only point is who shall have that arbitrary power…." *The Anarchy of Limited or Mixed Monarchy* in *Patriarchia and Other Writings*, p. 131.

though, moral persons are concerned with converging with others on a common moral rule. Until there is a shared common rule, known by all to be the rule of the group, there cannot be a real moral life that respects all as free and equal. Should all come to accept the same eligible rule, then, finally, moral authority is reconciled with freedom and equality.

Two members of G* can be modeled as playing an impure coordination game. Display 1 gives the game in its simplest form, a simple 2x2 game (numbers indicate ordinal utility, highest being most preferred).[36]

	Betty	
	R_A in S	R_B in S
R_A in S	1 2	0 0
R_B in S	0 0	2 1

Alf (label at left for rows)

Display 1

Should Alf and Betty find themselves coordinating on R_A neither would have reason to change his or her action. For each, given his or her overall evaluative standards, each has the most reason to act on R_A. Should they instead find themselves at R_B, each will then have most reason (given his or her evaluative standards) to act on R_B. Note that in neither case is either induced by some external consideration to conform to a rule that is not, from his or her perspective, optimal: *consulting simply his or her own standards for evaluating proposed rules, each has decisive reason to freely endorse whichever eligible moral rule on which they have coordinated.* When coordinating on R_A Alf can demand that Betty conform and, consulting only her own values, ends, and moral ideals, she will have decisive reason to conform; and if they have coordinated on R_B Betty can demand that Alf conform, and he will have decisive reason to act on it, considering only what he thinks is important — his own "utility function." And this even though, from the initial deliberative perspective, neither had reason to act on the other's preferred moral rule.[37]

A one-shot two-person game can give us some insight, but it is clearly an inadequate way to model the selection of a moral rule from the eligible set. The relevant coordination problem is not a single-play game, but an iterated game. We

---

[36] It is important to stress that "utility" is a mathematical representation of an agent's judgment as to how well a rule satisfies her values, aims, and moral ideals. It is not either itself a unique value (such as welfare), nor it is a goal to be pursued.

[37] Again, we should not be misled by the language of "preference." To prefer R_A to R_B is simply to rank R_A over R_B for purposes of choice; in our terms one's evaluative standards indicate reason to rank R_A over R_B — this is all that is implied by saying one has a preference for R_A over R_B.

have a number of encounters with others, and each can be understood as a play in a series of impure coordination games over many options. Now in an iterated game a person's utility (again, remember this is defined solely in terms of her relevant values, ends, and moral ideals) is a combination of her utility in this play, plus her expectations for utility in future games. Thus a person might sacrifice utility in one play to induce play in future moves that will yield her a more favored result. Moreover, it is certainly the case that in iterated games the play can move from one equilibrium to another. Peter Vanderschraaf and Brian Skyrms have shown how taking turns on each of the two equilibria emerges in iterated two-person impure coordination games.[38] However, in large *N*-person impure coordination games with multiple equilibria such solutions are, I think, practically impossible. In such large iterated games a bandwagon effect takes over. As I have argued elsewhere, such large-person iterated coordination games exhibit a strong increasing returns effect:[39] the more people come to embrace a particular rule, the more reason others have to also embrace it.[40] Even if one ranks a rule near the bottom of *S*, under a wide range of conditions one will have strong reason to embrace it if most others are acting on it. This dynamic is illustrated in Display 2.

Display 2[41]

Time
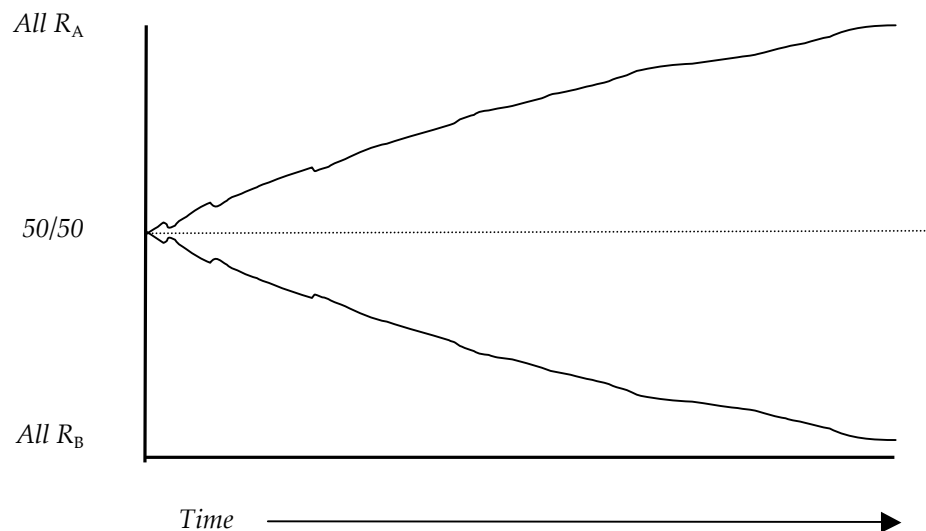
[38] Peter Vanderschraaf and Brian Skyrms, "Learning to Take Turns," *Erkenntis,* vol. 59 (2003): 311-46.

[39] The path-breaking work on increasing returns was done by W. Brian Arthur. See his *Increasing Returns and Path Dependency in the Economy* (Ann Arbor: University of Michigan Press, 1994). Those technically-minded will see the point can be put in terms of positive externalities due to network effects.

[40] See my "The Demands of Impartiality and the Evolution of Morality" in *Partiality and Impartiality,* edited by Brian Feltham and John Cottingham (Oxford: Oxford University Press, 2010), pp. 42-64, and *The Order of Public Reason*, chap. VII.

[41] Based on Arthur, *Increasing Returns and Path Dependency in the Economy*, p. 3.

Starting out with a population evenly split between advocates of R_A and of R_B, individual choices and random events can lead the population to all R_A or all R_B equilibria. Which equilibrium emerges will be path-dependent: at time zero there is no reason why one or the other should emerge as the *unanimously-chosen rule*. But once they have arrived at such a convergence, each member of G^*, consulting only her own values, ends, and moral ideals will freely act on the rule in equilibrium. For our purposes what is crucial is that the contingent way in which large groups can come to coordinate on a common practice is no bar to there being a determinate morality that all can endorse given their evaluative criteria *once it has been arrived at*.

We can weaken our assumptions about the motivations of the members of the group, and still get a similar result. Suppose that we take our actual group G, and divide it into two subgroups which are roughly equal: the Moralists and the Simple Coordinators. The Moralists reliably recognize members of the socially eligible set, S; they act like the members of G^* we have been examining. In contrast the Simple Coordinators either have no special concern with morality, or they simply have a fuzzy idea of the contours of S. They also have only a vague idea of what rules would best satisfy their values (we can think of Simple Coordinators as those who have a harder time tracking what their G^* counterparts would judge). Simple Coordinators do, though, see the advantage of coordination with others: they are aware of the confusion and disorder produced by contradictory moral demands. Simple Coordinators will coordinate on any rule, in or out of S. In contrast, Moralists deem acting on a rule outside S (the socially eligible set) as certainly no better than, and perhaps worse, than no rule at all. It fails the test of being an eligible moral rule, yet to adopt it as moral rule allows that others can insist that one conform to it, rebuke one for failing to conform, and one is to internalize it and feel guilty for not conforming to it. Display 3 presents a simple coordination game between a Moralist and a Simple Coordinator.

Simple Coordinator

		R_A in S	R_Z not in S
	R_A in S	1 / 1	0 / 0
Moralist	R_Z not in S	0 / 0	1 / -1

Display 3

The Moralist will not coordinate on R_Z, and so she will induce the Simple Coordinator to move to R_A. We thus see that a large enough group of Moralists can move the entire group to an equilibrium within the eligible set.

3.4 From justification to choice (and back again)

In the end, an authoritative social morality that treats all as free and equal is to a large extent the result of the choices of its members. It is not *merely* chosen: not just any rule that a group might converge on is a genuinely moral rule. Bad norms, backed by punishment, can be in (a non-moral) equilibrium: each may have sufficient reasons to conform simply in order to avoid punishment. Cristina Biccchieri analyzes such cases:

> Fear should never be discounted, because there are many cases in which one obeys a norm only because neglecting others' expectations and preferences will bring about some form of punishment. We may conform without any intrinsic value to the norm and without finding others' expectations legitimate. Some Arab women may observe Muslim sexual mores, and Corsican men embrace norms of revenge, for fear of being punished if they break the rules. In both cases they may find their community norms oppressive and ill-suited to modern life, but whoever speaks first or rebels first runs the risk of bearing huge costs. Breaking the rules looks like the risky cooperative choice in a social dilemma. Freedom from a bad norm is a public good that is often very difficult to bring about.[42]

As Boyd and Richerson show, punishment can stabilize just about any norm, good or bad.[43] Social equilibrium does not imply moral equilibrium, and the fact that a bad norm is in social equilibrium can make it very difficult to dislodge. However, when a rule in equilibrium is within the socially eligible set, it is both chosen by the members of a society *and* is acceptable from what Baier called "the moral point of view."[44] The rule passes the tests of impartiality, generality, teachability, and so on, and is endorsed by the values, ends, and moral ideals of all. Each, consulting only his own values, ends, and moral ideals has reason to act on it.

A moral rule in equilibrium is not a collective "we" choice; the group, *as a group*, does not choose its moral rules.[45] Nevertheless, a moral rule in equilibrium is a social fact that arises from the interdependent choices of its members. As interdependent individuals, we choose our moral rules — at least, if we are lucky enough to actually possess genuine moral rules (§4). This is a real choice of actual people; not a

---

[42]   Bicchieri, *The Grammar of Society: The Nature and Dynamics of Norms* (Cambridge: Cambridge University Press, 2006), p. 42.

[43]   Robert Boyd and Peter Richerson, *The Origin and Evolution of Culture* (Oxford: Oxford University Press, 2005), chap. 9.

[44]   Baier, *The Moral Point of View*.

[45]   Cf. Gilbert, *A Theory of Political Obligation*, Part II.

hypothetical choice, or one dictated by impartial reason. To be sure, it is the choice of social creatures, who must take into account what others are legitimately doing, and cannot dictate to the group their preferred rule (although, if they are moral philosophers they will probably do their best to). Or, to recall Kant's formulation (§2.1), it is self-legislation that does not amount to each acting on his own judgment of justice *entirely independent of that of others*. Importantly, although the rule is chosen by members of society, taken individually in their interdependent interactions, and unless some rule was so chosen by the society there would be no authoritative morality, the authority of the morality does not depend on each and every person choosing the rule. Once society has chosen a rule, if the rule in equilibrium is also a member of the socially eligible set, *we have created through our actual interdependent choices what impartial reason could not deliver: a uniquely justified rule*. For once the rule is in social equilibrium (and is a recognized social norm), then all have conclusive moral reason to act on this rule rather than any other in *S*. Thus, having created a justified rule though our interdependent choices, we can then insist that all conform to this rule, for all free and equal persons now have conclusive reason to conform to this rule, rather than any other. That genuine moral authority is created by our choices does not entail than one who holds out, refusing to choose our rule, escapes the moral authority of our justified social-moral rule. In this fundamental sense, society consents to the moral rule to which we are all bound.[46]

Filmer long ago pointed out that the "plausible and dangerous" notion that all are free and equal appears to require that each and every person consents to authority, but of course it is impossible to imagine that all do so. And, Filmer insisted, it is not enough to say that the majority, even the overwhelming majority, have consented: "unless it can be proved by some law of nature that the major, or some other part, have power to overrule the rest of the multitude, it must follow that the acts of multitudes not entire are not binding to all but only such as consent unto them."[47] From the outset this has been the Achilles' Heel of consent theory.[48] On the account advanced here, the consent of the majority — their actual choices in selecting a rule from *S* — does have the power to bind the rest. Once society has consented to this rule, it becomes the sole rule that is uniquely rationally justified: it now is the one rule that all have reason to endorse. It is the only rule that all can

---

[46] It is important to distinguish this notion of consent from promissory obligations, which I consider in section 5. Compare Gilbert: "Actual consent theory invokes an agreement as opposed to a contract in law." *A Theory of Political Obligation*, p. 55.

[47] Filmer, *Patriarcha*, p. 21. Cf. Locke's claim that in civil society "the *Majority* have a Right to act and conclude the rest." *Second Treatise*, §95.

[48] See Gilbert, *A Theory of Political Obligation*, chap. 5.

endorse and which is able to provide the basis of shared moral life in which all are treated as free and equal. Each, consulting only her set of relevant values, ends, and moral ideals, has reason to act on that rule over any alternative. What was *ex ante* indeterminate, is *ex post* uniquely justified.[49]

4. POLITICAL AUTHORITY, JUSTIFICATION, AND CHOICE

4.1 The state as the protector, sustainer, and developer of social morality

T. H. Green insisted that "A state presupposes other forms of community, with the rights that arise out of them, and only exists as sustaining, securing, and completing them."[50] For Green, the great error of social contract theory was that it supposes that individuals somehow confront each other simply with their individual rights before the advent of the state, where in fact all states arise out of a pre-existing system of social morality and authority, and at least in the first instance their task is to protect, sustain, and develop this system of morality. Indeed, unless there is a definite pre-existing social morality, there is no reason to think that government can solve our problems. As David Schwab and Elinor Ostrom have stressed, until we can be confident that those with political authority will themselves abide by the norms of social morality, instituting political authority may simply leave us open to new forms of exploitation — those with authority may simple use the institutions to promote their narrow self-interest.[51] Once we understand that interpersonal morality already supposes a system of social authority, the question is not whether free and equal individuals endorse and choose to be subjected to the authority of others, but whether, while endorsing social authority, they pull back and refuse to endorse political authority.

Rex Martin, following Green, argues that "civil rights must be formulated (carefully constructed), maintained, and harmonized — harmonized, that is, if such rights are to form a coherent set and if internal conflict within given rights is to be avoided. Agencies are required, then, to formulate, maintain, and harmonize civil rights, and these agencies must be themselves well coordinated if the job of harmonization is to be accomplished."[52] Although on many matters a society of free and equal individuals can coordinate on a common interpretation of a moral rule

---

49   In game theoretic terms, in the *ex ante* situation there are many possible Nash equilibria (in fact, every member of *S*); once a rule has been selected by society, it is in Nash equilibrium, and so all have reason to act on it.

50   Green, *Lectures on the Principles of Political Obligation*, §134.

51   David Schwab and Elinor Ostrom, "The Vital Role of Norms and Rules in Maintaining Open Public and Private Economies" in *Moral Markets*, pp. 204-227.

52   Rex Martin, *A System of Rights* (Oxford: Clarendon Press, 1993), p. 166.

(and so, a rule specifying their rights), social coordination fails to fully secure a common authoritative social morality among free and equal persons for a number of reasons.

(*i*) Given social and technological changes, new cases may arise for which the social rule has no clear answer. Although we may expect that over a long enough time a new social equilibrium on this matter may arise, this may entail a long period of disagreement and moral disorder.

(*ii*) As Martin notes, the rules of social morality can conflict. Again, although some rules of priority are no doubt part of social morality (we usually know what to do when a promise conflicts with saving a life), many of these conflicts may be unresolved by social morality, and so a more formal system of adjudication may be required.

(*iii*) The informal punishments of social morality may be insufficient to enforce moral rules when the temptations to cheat are high.

(*iv*) Our social order may be stuck at a social equilibrium outside the socially eligible set — that is, we may have an oppressive social morality that some free and equal moral persons do not have sufficient reason to endorse, and to abide by.

Now a person committed to social morality must see that these problems pose severe setbacks to our ability to treat each other and free and equal moral persons while acknowledging a system of moral authority. Such a person must acknowledge that a state is justifiable: an agency that protects, sustains, and develops our social morality is justified by our very devotion to social morality conjoined with the knowledge that problems (*i*)-(*iv*) are endemic features of life — at least in large–scale modern societies.

Recall that Simmons is ready to accept the Kantian claim that states are justified: what he challenges is that this type of justification implies a political obligation to obey (§1). Note, though, that when states are justified as harmonizers and developers of social morality, this justification implies that the judgments of the state authoritatively articulate social morality, which itself is authoritative and obligatory. Insofar as the task of the state is to, as it were, continue on with the development of social morality by political means, when it fulfills these tasks it imposes obligations via the authority of social morality. The social contract tradition was wrong in holding that a political umpire or adjudicator is always a necessary, constitutive, element of justice among individuals (§2.1), but it was correct that in many instances it is so. And this necessity is not simply an instrumental necessity: the state's actions

as developer of social morality is constitutive of many of our moral relations with others (§2.1). Social morality both constrains the state (where we have a social equilibrium on a member of the eligible set, the demands of morality are clear) and yet it is also completed by the state, when it provides the means for adopting a member of the eligible set. This latter role of the state is the important insight in Jeremy Waldron's argument that law is a response to disagreement about what is best. The law, Waldron argues, can help us to coordinate on a common solution to common problems: it renders one coordination point more salient by attaching sanctions, and so makes it less likely that people will hold out for their favorite outcome. "But before it can do that, the society must have decided which of the coordinative strategies to select as the one to be bolstered in this way. That itself is no mean achievement — and I want to say that it is by embodying that achievement that law commands our respect."[53] The real achievement of the law is to select which coordination point should be sought. Thus, in a way similar to the analysis of section 3.3, Waldron believes that law and legal authority can be modeled on an impure coordination game. As Waldron understands politics, we debate and discuss the merits and demerits of each of the possible coordination points; since it is an impure coordination game, Alf prefers a different coordination point (R_A) than does Betty (R_B), and so they have something to argue about. However, they each prefer any coordination point (in our terms, any member of a rule in S) to lack of coordination. In essence, then, Waldron argues that we need to coordinate on some single reasonable law, even if it is not the one that each of us sees as most reasonable. "*A piece of legislation* deserves respect because of the achievement it represents in the circumstances of politics: action-in-concert in the face of disagreement."[54]

4.2 The role of social consent

The state is morally justified if (*a*) it is required to protect, sustain, and develop social morality. We also should add that for any particular state to be justified it must (*b*) be within the socially eligible set of all institutional structures that fulfill the role specified in (*a*). For a state T to have moral authority qua sustainer and developer of social morality, both conditions must be met. Again, we have arrived at an eligible set that is not a singleton. Members of G^* would almost certainly be unable to settle on a uniquely best regime. Some form of democracy, understood as a representative system with rights of participation and political pluralism, is, I think, abstractly

---

[53]   Jeremy Waldron, *Law and Disagreement* (Oxford: Oxford University Press, 1999), p. 104.
[54]   Ibid., p. 108. Emphasis added.

justified, but members of G* will not concur on which system is best. Democracies differ in fundamental ways. Arend Lijphart famously contrasted basic differences in governance between majoritarian and consensual democracies;[55] democracies also differ whether they are unicameral or bicameral, have proportional representation, are parliamentary or presidential, have a unitary or federal structure, possess a written or an unwritten constitution, have a strong or weak system of judicial review — just to name some of the more obvious factors. There are also reasonable disputes about differences in the effects of electoral laws,[56] differences in the likely outputs of different systems (bicameral regimes, for example, are said to make legislation more difficult), as well as disagreements about the intrinsic value of some regimes (some hold, for example, that basic democratic values incline towards proportional representation). Whether we take these issues as a bundle (members of G* choose regimes types), or divide them up into a series of choices, members of G* will order the options differently.

As with social morality itself, different societies arrive at different equilibria on these matters, largely on the basis of a path-dependent political history. Just as the justification of the rules of social morality from the impartial deliberative perspective of the commonwealth of ends (G*) is insufficient to yield a determinate publicly justified social morality, so too is the abstract justification of democracy and its attendant political rights insufficient to yield the justification of a system of governance. Even the United States, which can trace its constitutional structure to an explicit convention, has accrued two hundred years of development, due often to unpredictable social and political events, and the unpredictable responses to them. Public reason does not mandate a specific democratic regime (how could it?). To be sure, those committed to a democracy that secures publicly justified outcomes may well advance proposals for institutional design that, in their view, do a better job than other arrangements, but members of G* will, of course, disagree about the merits of these proposals. Political authority too relies on informal social authority — an evolution of a political culture leading to the selection of one of a wide range of acceptable political systems.

Understood as an institutional system that is required to protect, sustain, and develop social morality (and its rights), the authority of the state, and the corresponding obligation of its citizens to obey, is indeed based on the consent of the

[55] See Arend Lijphart, *Democracies: Patterns of Majoritarian and Consensus Government in Twenty-One Countries* (New Haven: Yale University Press, 1984).

[56] For a classic study see Douglas Rae, *The Political Consequences of Electoral Laws* (New Haven: Yale University Press, 1971). Of course contemporary political science has much to say about these matters, though it is typically open to reasonable dispute, even among the experts.

governed. For the state thus understood is itself a development of social morality, and so like all social morality it is the result of numerous individual choices over a long process that leads to a specific social equilibrium. Again, this is not a "we choice" that all make together, or a one-time social contract or constitutional convention. It is an ongoing social choice, arrived at by a path-dependent history, and continually reaffirmed by the choices of its members. Filmer was correct: "Mankind is like a sea, ever ebbing or flowing, every minute one is born and another dies. Those that are the people this minute, are not the people the next minute. In every instant and point of time there is variation."[57] But this does not preclude the important idea that political authority rests on the ongoing consent of the governed.

Three caveats must be stressed. *First*, as I argued with relation to social morality, while our social morality is based on the consent of society, it need not be the choice of each and every member of society, nor does the fact that some withhold their consent mean that they are not obligated. *Second*, only if the state structure is within the socially eligible set do its laws have genuine authority. Thus, while we can see how a general obligation to conform to the laws of one's just state can be established, there is no general obligation in Gilbert's more sweeping sense — a general obligation of all those living in almost all political societies, even deeply flawed and unjust ones, to obey the law.[58] *Lastly*, I have only explored the obligation to obey one's just state insofar as it is understood as a protector, sustainer and developer of social morality. Green noted that once states are instituted to perform these tasks, they grow to take on others (for example, securing public goods or achieving collective goals). We should not suppose that the same account of the authority of those laws which are needed to articulate the requirements of justice also applies to policies that, say, seek efficient provision of collective goods. That is another issue.

5. THE ARBITRARY WILL AND THE POWER OF PROMISING

It is often thought that the social contract's reliance on consent invokes consent qua promising: one is bound to obey the terms of the contract (to obey the law) because one has promised to do so. I have been at pains to show that consent need not be in the form of a promise. The members of the socially eligible set are justified authoritative rules or institutions; a society exercises real choice in deciding on which of these they shall equilibrate. No promise, explicit or tacit, is involved. But

---

[57] Filmer, *The Anarchy of Limited or Mixed Monarchy* in *Patriarchia and Other Writings*, p. 142.

[58] Gilbert, *A Theory of Political Obligation*, p. 214, 234ff. But see p. 67, where Gilbert seems to have the idea of an eligible set.

what of promises? Can they truly bind, and could they form the basis of political obligation?

Some doubt that actual consent, even if it could be obtained, would yield true political obligations.[59] David Estlund has recently criticized the claim that actual consent to a state's authority is sufficient to render it legitimate.[60] The actual consent view seems to allow that we could all be bound to obey an institution just because no one has yet thought of the decisive problem with it; once the problem occurs to someone and she withdraws her consent, then the institution immediately is de-legitimized. It would seem that actual consent cannot provide the basis for assured continuing obligations — there is, we might say, justificatory instability. Now relying on consent qua promising avoids this problem: to promise is not to just to agree for now, but to bind oneself. Even if it later occurs to you after the promise that perhaps it was not such a great idea to promise, you are, generally, nevertheless bound. On Hobbes' view of promising and contract, unless a crucial new piece of information becomes available between the time of promising and the time of performance, one is bound to perform, regardless of second thoughts.[61]

This response to the problem of justificatory instability seems simply to lead to a deeper problem, stressed by Hegel in his criticism of Rousseau (and Fichte):

> ...he takes the will only in determinate form as the individual will, and he regards the universal will not as the absolutely rational element in the will, but only as a "general" will which proceeds out of the individual will as out of a conscious will. The result is that he reduces the union of individuals in the state to a contract and therefore as something based on their arbitrary wills, their opinion, and their capriciously given consent....[62]

Hegel makes two crucial points. First, he is entirely right that a fundamental aspect of promissory obligations is our moral power to bind our future action to our present "arbitrary wills." Many find this power mysterious, and seek to base it on more fundamental, and non-arbitrary, moral principles.[63] I believe this is a mistake. One of the most basic features of moral agency, especially in a diverse society in

---

[59] For an excellent analysis, and criticism, of this view, see Gilbert, *A Theory of Moral Obligation*, pp. 75-83.

[60] See his *Democratic Authority* (Princeton: Princeton University Press, 2008), p. 49-50. Estlund develops these criticisms further in his forthcoming "Reply to Critics" in *Ethics*.

[61] "The cause of fear, which maketh such a covenant invalid, must be always something arising after the covenant made, (as some new fact or other sign of the will not to perform), else it cannot make the covenant void. For that which could not hinder a man from promising ought not to be admitted as a hindrance of performing." Hobbes, *Leviathan*, p. 85 (chap, 14, ¶ 20).

[62] G.W.F Hegel, *The Philosophy of Right*, translated by T.M. Knox (Oxford: Oxford University Press, 1952), p. 157.

[63] See, for example, T.M. Scanlon, *What We Owe Each Other* (Cambridge, MA: Belknap Press of Harvard University Press, 1998), chap. 7.

which people fundamentally disagree about values, ends, and moral ideals, is the ability of individuals to, as it were, extend the bounds of morality through their individual choices and agreements. A framework of common moral rules is necessary for ordered moral relations, but insufficient. We not only need to invoke the authority of social morality in our relations with strangers, but we need specialized moral claims with respect to those with whom we share ends, ideals, and projects. One way — though not the only way[64] — to extend moral obligations is through our distinctive moral power as agents to create moral obligations through promising. Just as a distinctive legal power of a legislature is to create obligations, a distinctive self-legislative power of a free and equal moral person is to create obligations on herself. This power follows from our very status of self-legislative agents. In terms of the categories I have employed, here it is not our legislative power qua members of G^* (in Hegel's terms, our rational wills), but our legislative power simply as members of G (our actual, arbitrary, wills) that legislates morality via promising.

An effective social morality needs to respect the self-legislation of arbitrary, actual, wills. We can say some things about what members of G^* would endorse, but this is, I have argued, surprisingly modest. Often it is simply unclear what a person has reasons to endorse — what his G^* counterpart would agree to. A person's reasons are complex, and his deliberations are often inconclusive. As an actual agent in the world, he must often act on his own, often incomplete and imperfect, deliberations about what he has reason to do. Fundamental to being an actual agent is to have the authority to decide in these sorts of cases what one has reason to do. If I had deliberated on your beliefs and values, I would have decided that ϕ is the thing to do, but you deliberate and decide that γ is. This is not simply a disagreement, in which we both have competing claims. That you have decided γ rather than ϕ is normally decisive: to be an agent is to be guided by your deliberations about what you should do, even if from an impartial epistemic point of view, my decision is superior. In Hegelian terms, actual agents are typically guided by their actual, arbitrary, wills. Because a purely rational morality is impossible (or, at least, we have no idea what it would be), real moralities must respect our arbitrary wills and, indeed, must allow that our moral self-legislation is often based on them. Our power to bind by promising is perhaps the most important way this is accomplished in our social morality.

---

[64] Though I believe that she over-extends her analysis, Gilbert perceptively shows how common commitments can generate specialized obligations.

The problem, then, is not that arbitrary wills cannot morally bind themselves. The real difficulties for the promissory view of political obligations are two. The first is recognized by all: it is simply implausible to think that many have bound themselves in this way to any state, so obligation through promise is never a plausible ground for a case that there is an obligation of almost all citizens to obey the law, just because it is the law. However, even if that obvious — and as far as I can see insurmountable — difficulty could be overcome, an account of political obligation based solely on promissory obligations would remain objectionable. In their different ways, and from their very different perspectives, Estlund and Hegel are correct: to base obligations *simply* on the arbitrary wills of a group of people is insufficient to assure us that our basic framework is not unjust and irrational. If all happen agree we may leave well enough alone, but the worry lingers: is the entire framework based on errors, superstitions, and the effects of social indoctrination? Unless we are convinced that the foundations of the framework are set in good reasons that all can endorse, the worry gnaws that our moral and political lives are grounded in nothing but arbitrary, perhaps oppressive, wills.

6. CONCLUSION: CONSENT THEORY AS DEVICES OF RECONCILIATION

I have stressed in this essay how the authority of social morality and political obligation is the result of a complex interplay of our rational and actual wills, or what we might think of as hypothetical and actual consent. Arguments from the hypothetical consent focus on an idealized group G^*, seeking to show what authoritative rules and institutions all members of G have reasons to endorse. As Kant stressed, until we have common rules that reflect public reason there simply is *no* public justice: we have only different people employing their private reason, each asserting what he thinks justice is and demanding that others comply. But, as I have stressed, to respect others as free and equal is inconsistent with claiming such authority of one's private reason over them. Only rules endorsed by public reason can reconcile authority with our status as free and equal. However the reasoning from this perspective (that of G^*) is indeterminate: in a pluralistic and diverse society no single moral rule is endorsed by public reason. Here, I have argued social morality employs a second device of reconciliation: a social choice of one member of the eligible set on which we coordinate. This is a real choice. In Hegelian terms, we might say that it is based partly on our rational, and partly on our joint arbitrary wills. Insofar as we are rational, we restrict our choice to the socially eligible set; but within that set a society chooses via the interaction of its members' arbitrary wills, producing a social equilibrium. The itinerary from the rational to the actual will is completed in promising, a device that grants each free and equal person the moral

power to employ her arbitrary will to grant others moral authority over her. It allows self-legislation by actual free and equal moral persons, based on their full set of ideals, values, and ends.

The consent tradition and liberalism developed together. Liberalism's devotion to the moral freedom and equality of all can only be reconciled with moral and political authority through forms of consent — all of which are exercises of self-legislation. The problem has been the insistence that one form of consent is the universal device of reconciliation. Different aspects of the consent tradition perform some tasks of reconciliation well, but all are implausible when pressed as the sole method of reconciliation. Once we appreciate that "consent theory" is not a specific device, but an array of devices for reconciling our status as free and equal with relations of authority, we can stop battling about which is correct, and investigate what tasks each can perform.

Philosophy
University of Arizona