

Public Reason and Reciprocity*

ANDREW LISTER

Political Studies, Queen's University

Final Prepublication version: Nov. 8, 2016.

Final Published version: *The Journal of Political Philosophy*, Vol. 25, No. 2, 2017, pp. 155-72.

DOI: <https://doi.org/10.1111/jopp.12111>

Wiley: "This is the peer reviewed version of the following article: "Public Reason and Reciprocity," *The Journal of Political Philosophy*, Vol. 25, No. 2, 2017, pp. 155-72, which has been published in final form at <https://doi.org/10.1111/jopp.12111>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This article may not be enhanced, enriched or otherwise transformed into a derivative work, without express permission from Wiley or by statutory rights under applicable legislation. Copyright notices must not be removed, obscured or modified. The article must be linked to Wiley's version of record on Wiley Online Library and any embedding, framing or otherwise making available the article or pages thereof by third parties from platforms, services and websites other than Wiley Online Library must be prohibited."

I. INTRODUCTION

The general issue that interests me in this article is when, why, and how duties are appropriately held to be conditional on reciprocity—compliance on the part of others. Some duties are presumably not conditional on reciprocity at all, for example, the duty not to torture people, but in other cases conditionality seems more appropriate. Why should I share fairly with you, if you would not share fairly with me? Jiwei Ci has argued that the demand for reciprocity is the

distinctive feature of justice, as compared to unilateral virtues such as benevolence.¹ There are different duties associated with justice, however. It seems more reasonable to insist on assurance of reciprocity as a condition for complying with law (for example, paying one's taxes) than it does as a condition of voting for or otherwise supporting just law (for example, voting for just tax laws).

The focus of this article is on public reason. Philosophers working on public reason have tended to follow Rawls in focusing on the ideal case in which the bulk of the citizenry regularly complies with the demand to make political decisions on the basis of public reasons.² Yet in contemporary liberal democracies, many people either do not know about public reason or do not think it is important. Recognition of this kind of non-compliance raises the question of whether the principle is still binding. Perhaps it is, if as many people have argued public reason is a requirement of political legitimacy rooted in respect for persons as free and equal moral agents.³ Yet it may be too much to expect unilateral restraint with respect to convictions that people believe true and of ultimate importance. How could it be reasonable to ask me not to count a reason I think true and relevant on the basis that you reject it, if you are not likewise willing to exercise restraint with respect to reasons that you think true but which I reject? Kent Greenawalt argued that public reason's requirement of restraint could only apply to speech, not political conduct, because the grounds of people's decisions are not publicly verifiable.⁴ Yet I don't have any greater assurance that others are voting for justice (as opposed to self-interest) than I do that they are voting on the basis of public reasons, as opposed to reasons taken directly from their own religious or philosophical doctrine.

It may be a mistake, however, to think of public reason as a basic principle (whether conditional on reciprocity or not), as opposed to a social norm justified by the consequences of its general observance. Recently, a number of people have argued that public reason is not a

*I would like to thank the participants in the 2015 MANCEPT workshop "Theories of Public Reason," the audience members at the 2016 NASSP panel "Stability and Justice," Richard Healey, Blain Neufeld, and two anonymous referees for the valuable comments they provided on earlier versions of this paper.

¹Jiwei Ci, *The Two Faces of Justice* (Cambridge, Mass.: Harvard University Press, 2006), pp. 19–21.

²John Rawls, *Political Liberalism* (New York: Columbia University Press, 1996), p. xlvii.

³Charles Larmore, "The moral basis of political liberalism," *The Journal of Philosophy*, 96 (1999), 599–625; Gerald F. Gaus, "Liberal neutrality: a compelling and radical principle" in *Perfectionism and Neutrality: Essays in Liberal Theory*, ed. Stephen Wall, and George Klosko (New York: Rowman & Littlefield, Inc., 2003), pp. 137–65; James Boettcher, "The moral status of public reason," *The Journal of Political Philosophy*, 20 (2012), 156–77.

⁴Kent Greenawalt, *Private Consciences and Public Reasons* (New York: Oxford University Press, 1995), p. 132.

requirement of legitimacy, but a mechanism intended to stabilize justice by helping solve “the assurance problem.”⁵ Although motivation by reciprocity stabilizes commitment to social rules when people perceive the compliance of others as a benefit that ought to be returned, it also makes compliance conditional on the expectation that others will comply too. In a society “well-ordered” by a given conception of justice it is common knowledge that everyone accepts the conception, but people will disagree about what policies follow. Such disagreement might lead to doubt about whether others are committed to the same principles, and if people’s commitment to justice is conditional on other people’s commitment, this uncertainty would be destabilizing. Accepting the norm of deliberating only on the basis of generally accepted facts and modes of reasoning allegedly helps provide assurance that others do accept the conception in question. According to this point of view, public reason is not a moral principle that *requires* assurance of reciprocity, but a social norm that *provides* assurance—about other people’s commitment to justice.

Roughly speaking, then, there are three ways we might think of public reason: as a unilaterally binding moral principle, as a moral principle that requires assurance of compliance on the part of others, or as a social norm compliance with which provides assurance about other duties that are conditional on reciprocity. Which is the right approach?

In response to the stability account of public reason, I concede for the sake of argument that a norm constraining public deliberation might provide assurance of commitment to shared principles of justice, but ask why these principles must be part of a merely political conception of justice, that is, one intended to be acceptable to all reasonable points of view. Public reason in this more fundamental sense cannot be a norm intended to stabilize commitment to justice, I argue, but must be conceived of as a moral principle. The question is whether this principle is limited by a reciprocity condition, and if so what form it takes. I maintain that the argument for public reason from respect for persons is best construed as the claim that public reason puts us into a relationship of mutual respect despite moral disagreement. If the rationale for the principle is relational, in this way, it is natural to think that it must be subject to a reciprocity condition. It

⁵Paul Weithman, “Inclusivism, stability, assurance” in *Rawls and Religion*, ed. Tom Bailey, and Valentina Gentile (Columbia University Press, 2015), pp. 75–96, at p. 86. I first came across this kind of position in a 2007 paper by Sharon Lloyd (“Private reasons, public judgments, and the requirements of reciprocity,” unpublished). Similar views have been developed in Paul Weithman, *Why Political Liberalism? On John Rawls’s Political Turn* (Oxford: Oxford University Press, 2010); Gillian K. Hadfield and Stephen Macedo, “Rational reasonableness: toward a positive theory of public reason,” *Law and Ethics of Human Rights*, 6 (2012), 7–46.

need not follow, however, that public reason is binding only if some minimum proportion of citizens accept the principle, for we can build reciprocity into the principle by stipulating that unanimous acceptability is required only with respect to points of view accepting the principle. My contention is that if compliance with law is assured at the requisite level (in part through enforcement), then the duties of public reason associated with authorship of law should only be considered conditional on reciprocity in this “internal” sense, which is not proportional but bilateral.

II. PUBLIC REASON, REASONABLENESS, AND RECIPROCITY

The general idea of public reason is that the exercise of political power ought to be justifiable to all those subject to it, and to all those in whose name it is exercised, despite the reasonable differences of religious and philosophical doctrine one must expect to persist when people have the freedom to disagree. Public reason thus involves a partly idealized unanimity criterion or, as David Estlund says, a “qualified acceptability requirement.”⁶ This requirement can be framed in two main ways: as a constraint on reasons for political decisions, with a default of exclusion, or as a constraint on political action (for example, having a common rule), with a default of inaction.⁷ I will focus on what I take to be the conventional Rawlsian view that public reason is a constraint on reasons for decisions, leaving aside the question of whether the principle applies only to fundamental political decisions (that is, about the basic structure of society), or to all decisions about public policy.⁸

For any view based on a qualified acceptability requirement, the standard of qualification will clearly be a crucial issue. I take it that the Rawlsian standard of reasonableness involves at least three, and possibly four criteria, or levels:

- (1) the ability and willingness to reason with others about what is true and good and

⁶David M. Estlund, *Democratic Authority: A Philosophical Framework* (Princeton, NJ: Princeton University Press, 2008), pp. 4, 40–64.

⁷This distinction is inspired by Christopher Bertram, “Theories of public reason,” *Imprints*, 2 (1997), 72–85, and developed more fully in Andrew Lister, *Public Reason and Political Community* (London: Bloomsbury, 2013), pp. 15–23.

⁸Cf. Jonathan Quong, “The scope of public reason,” *Political Studies*, 52 (2004), 233–50.

right.⁹

(2) recognition of the existence of burdens of judgment (which are obstacles to agreement on the part of people who are reasonable in the minimal, level-(1) sense), and the resultant fact of reasonable pluralism.¹⁰

(3) the willingness to abide by fair terms of cooperation provided others do too.¹¹

(4) (possibly) acceptance of public reason.

Lower levels of reasonableness do not entail higher levels, though one hopes acceptance of one will at least tend to be associated with acceptance of the others.

The third level of reasonableness (conditional commitment to fair cooperation) involves motivation by reciprocity, along the lines of the moral psychology that Rawls made use of in his account of the stability of justice as fairness. Reciprocity in this sense refers to a non-instrumental tendency to respond in kind—to treat well those who treat me well, for example, even if I don't stand to benefit in the future.¹² Reciprocity may be perceived as a duty (that is, if *A* benefits *B*, *B* must benefit *A*), or as a limit on general duties (that is, *A* must benefit *B*, unless *B* would not do likewise were circumstances reversed, *or* unless most others are not complying with the duty when it applies to them). In *Political Liberalism*, Rawls used the term “reciprocity” to refer to the principle of public reason, as well as to the psychology of responding in kind.¹³ Yet public reason is at least partly independent of motivation by reciprocity. To be sure, there is a definitional connection. Public reason involves unanimous acceptability to reasonable views, and reasonableness includes a commitment to reciprocity (at level 3). Yet the conditional

⁹My formulation of this criterion paraphrases Charles Larmore, *The Morals of Modernity* (Cambridge: Cambridge University Press, 1996), p. 168. However, the idea is implicit in Rawls's idea of the burdens of judgment, which are obstacles to agreement apart from self-interested bias and irrationality / stupidity, i.e., “the many hazards involved in the correct (and conscientious) exercise of our powers of reason and judgment in the ordinary course of political life” (Rawls, *Political Liberalism*, pp. 55–6).

¹⁰Rawls, *Political Liberalism*, p. 54. Reasonable pluralism is not the optimistic view that only a narrow range of views not-too-different-than-my own will persist, but the more pessimistic view that even among reasonable people, we must expect deep differences of religious and philosophical doctrine.

¹¹Rawls, *Political Liberalism*, p. 54.

¹²John Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1999), p. 440.

¹³Rawls's “criterion of reciprocity” requires that “the reasons we offer for our political action may reasonably be accepted by other citizens as a justification of those actions” (Rawls, *Political Liberalism*, p. xlv). Compare to “the idea of reciprocity,” which involves acceptance of rules that benefit everyone compared to a fair baseline (Rawls, *Political Liberalism*, p. 16). Samuel Freeman distinguishes three different senses of “reciprocity” used by Rawls: the psychological tendency to respond in kind, social cooperation based on this tendency, and the criterion of reciprocity in justification. See *Rawls* (London: Routledge, 2007), pp. 374–5. I am insisting on the distinction between the first two and the third, following Christie Hartley, “Two conceptions of justice as reciprocity,” *Social Theory and Practice*, 14 (2014), 409–32, at p. 426.

commitment to fair cooperation does not by itself imply acceptance of public reason (level 4). One could have the disposition to return to benefits and the unwillingness to make unilateral sacrifices while believing that the only relevant standard in political decision-making is correctness rather than reasonable acceptability—no need for conceptions of justice to be political-not-metaphysical, no bracketing of disagreement, and no duty of epistemic restraint. It is not obvious whether public reason should require acceptability to these otherwise reasonable views that reject public reason. It is also unclear whether one must abide by the principle when most of one's fellow citizens reject it.

A further issue is whether the principle formulates a standard specifying when laws and institutions are legitimate (a goal or target), or a norm for citizens and legislators to follow in their political deliberations (by which I mean primarily their reasoning, though also their speech to the extent that a requirement of sincerity applies). It is natural to suppose that the norm follows from the target—if citizens don't try to reason publicly it is unlikely that their decisions will end up being publicly justifiable—but it is possible to imagine indirect forms of public reason, where the deliberative norm that best realizes public justifiability makes no reference to the principle.¹⁴ The converse is also possible; citizens and legislators might be expected to reason in public terms but for the sake of a goal other than public justifiability—stability, for example. I think the most plausible and most conventional view is that both aiming at and achieving public justifiability are morally important (that is, aiming is important intrinsically, but also because it makes achieving more likely).

III. PUBLIC REASON AS A STABILIZING MECHANISM

Recently, however, a number of people have argued that public reason is not a moral principle in the standard sense (for example, a requirement of political legitimacy), but a norm intended to generate stability by providing assurance of commitment to principles of justice. According to this interpretation, public reason is what G. A. Cohen would call a rule of regulation—a rule of conduct justified by the consequences of its general adoption—rather than a principle of

¹⁴Gerald F. Gaus, and Kevin Vallier, "The roles of religious conviction in a publicly justified polity: the implications of convergence, asymmetry and political institutions," *Philosophy and Social Criticism*, 35 (2009), 51–76.

conviction.¹⁵ At one point in *Political Liberalism*, Rawls presents public reason as a set of “guidelines” selected by the parties in the original position in order to direct citizens’ interpretation of the principles of justice as fairness, a norm instructing citizens to appeal to commonly recognized modes of reasoning and generally recognized facts.¹⁶ This norm can be portrayed as a means of stabilizing justice as fairness by providing assurance of other people’s commitment to the shared conception of justice. Paul Weithman argues that in order for a conception of justice to be stable, citizens must affirm it on the basis of their own comprehensive doctrines, and they must know that others do so too; “each member of the WOS [well-ordered society] needs assurance that everyone else is committed to treating values and ideals of justice as fairness as authoritative.” My first preference is for everyone including myself to treat the ideals of justice as fairness as authoritative, but if others do not recognize the authority of this conception, then I will not either. Weithman argues that public reason helps makes this overlapping consensus common knowledge.¹⁷

[W]hat citizens know about one another’s commitment to the authority of a conception of justice depends, in part, upon what concepts and methods of reasoning they actually use when they argue about basic political questions. That, I believe, is why Rawls introduces guidelines of public reason—to provide a solution to the assurance problem.¹⁸

¹⁵G. A. Cohen, *Rescuing Justice and Equality* (Cambridge, Mass.: Harvard University Press, 2008), p. 276. Cohen acknowledges that the distinction between “fundamental normative principles” and rules of regulation is not exhaustive, because there exist derivative principles of conviction. If public reason is a principle derived from respect for persons, then it is not fundamental, but it is still not a rule of regulation.

¹⁶Rawls, *Political Liberalism*, p. 225. As Blain Neufeld points out, however, Rawls also presents the demand for public justifiability (in the form of “criterion of reciprocity”) as a basic principle that motivates our use of the original position to think about questions of justice, and so not merely a norm chosen in the original position. See Blain Neufeld, “Review: Paul Weithman, *Why Political Liberalism?*” *Notre Dame Philosophical Reviews*, (2011), available at <<http://ndpr.nd.edu/news/27634-why-political-liberalism-on-john-rawls-s-political-turn/>>; citing Rawls, *Political Liberalism*, pp. xlv, xlvi–xlix, 226–7, 381, 446–7; see also James Boettcher, “Book review: Paul Weithman, *Why Political Liberalism?*” *Public Reason*, 5 (2013), 137–50. Applied to the constitutional structure of society, the criterion of reciprocity yields the “liberal principle of legitimacy,” which states that the exercise of political power is legitimate only in accordance with a constitution citizens can be expected to endorse “in light of principles and ideals acceptable to their common human reason” (Rawls, *Political Liberalism*, p. 137). Simon May and Paul Weithman claim, to the contrary, that the principles of liberal legitimacy and/or public reason are chosen in the original position; Simon Cabulea May, “Religious democracy and the liberal principle of legitimacy,” *Philosophy and Public Affairs*, 37 (2009), 136–70; Weithman, *Why Political Liberalism?*, pp. 313, 351. For criticism of this interpretation, see Blain Neufeld, “Reciprocity and liberal legitimacy: critical comment on May,” discussion note, *Journal of Ethics and Social Philosophy*, (2010).

¹⁷Weithman, *Why Political Liberalism?*, p. 340.

¹⁸Weithman, “Inclusivism, stability, assurance,” p. 86.

The worry about lack of assurance arises because citizens will often disagree about the implications of principles for policy. If the conclusions they reach are very different, they may be led to doubt whether their fellow citizens do in fact accept the principles of justice as fairness. By appealing only to generally recognized facts and modes of reasoning, citizens make it clear how they can plausibly maintain that the policies they support follow from shared principles, despite the fact that other citizens disagree. Public reason provides confidence that all do in fact affirm the same principles, disagreeing only about interpretation.

A number of critics have raised doubts about whether public reason will provide such assurance—for example, because the public display of constrained deliberation is simply cheap talk.¹⁹ I will not take a position on that issue, because I am interested in the more basic questions of what exactly people need assurance of, and whether asking for such assurance is appropriate.

Assurance problems arise for people motivated by reciprocity. On the one hand, reciprocity stabilizes social rules, if people perceive others' compliance as a benefit to be reciprocated.²⁰ On the other hand, motivation by reciprocity generates instability, because it makes compliance conditional on compliance by others.²¹ People not only need to be willing to comply, they need to know that others are complying too. In a prisoners' dilemma, no assurance problem arises because each party knows what the other will do; each person's best reply to the actions of the other party is to confess (defect), leaving both in jail for longer than if they had both kept quiet. Amartya Sen called this problem the "isolation paradox" because an outcome that is worse for everyone results from the fact of individual decision-making.²² The assurance problem only arises when the problem of the dominant non-cooperative strategy has been overcome due to motivation by reciprocity. People return good for good, therefore don't defect if others are cooperating. However, they also refuse to make sacrifices unilaterally. In a large group, people may be unsure about the motives of others. The assurance problem is to provide

¹⁹Gerald F. Gaus, "A tale of two sets: public reason in equilibrium," *Public Affairs Quarterly*, 25 (2011), 305–25, at p. 317; John Thrasher, and Kevin Vallier, "The fragility of consensus: public reason, diversity and stability," *European Journal of Philosophy*, 23 (2013), 1–22; Stephen G. W. Stich, and Brian Kogelmann, "When public reason fails us: convergence discourse as blood oath," *American Political Science Review* (forthcoming).

²⁰Rawls argued that if social inequalities benefit all, people will be more likely to see others' compliance with society's rules as benefits that ought to be reciprocated than if inequalities need only maximize the aggregate good; Rawls, *A Theory of Justice*, pp. 154–5.

²¹Rawls, *A Theory of Justice*, p. 296; see also p. 238.

²²Amartya Sen, "Isolation, assurance and the social rate of discount," *The Quarterly Journal of Economics*, 81 (1967), 112–24.

people confidence that enough others will comply so that our tendency to respond in kind dictates compliance rather than shirking (positive rather than negative reciprocity). Once we remove the temptation to act selfishly by putting in place enforcement mechanisms, we remove uncertainty about the conduct of others, leading to general compliance based on the motive of reciprocity. As Jiwei Ci puts it, the function of legal coercion is “to compel the unjust to follow the norms of justice—even if only reluctantly—so that those with a just disposition can do so willingly, though conditionally.”²³

The stability account of public reason is ambiguous, however, as to what specific duties are conditional on assurance of compliance on the part of others. When we talk about not recognizing the authority of a conception of justice, are we talking about not complying with laws based on those principles, or are we talking about refusing to support enactment of such laws? The distinction between our role as subjects of law and our role as authors of law is blurred when we speak simply of acting on principles of justice. In summarizing Weithman, for example, Gaus says that “[o]ur sense of justice directs us to act on justified principles only if we can be assured that others will do so as well.”²⁴ But “acting on” might mean “complying with rules justified by” or “trying to enact rules justified by.” One might be conditional on reciprocity but not the other. We could accept that the duty to obey approximately just law is conditional on others also complying while maintaining that the ethical duty to seek or enact just law is not conditional on other people also seeking justice. I find it plausible that there is not only a duty to comply with approximately just institutions (so long as most others are doing likewise), but a duty to reform unjust institutions, and to create institutions if they do not exist but are necessary for justice.²⁵ These creative dimensions of the duty of justice are binding regardless of whether others are complying, I take it, but only so long as performance is not too costly.²⁶

The stability account is also ambiguous as to whether public reason simply provides assurance or also requires it. This issue arises because public reason can be construed as an animating ideal of political liberalism, or simply a social norm governing the application of principles of justice assumed to be already on hand. Weithman understands public reason in the latter sense, as a rule requiring exclusion of controversial facts and modes of reasoning in the

²³Ci, *The Two Faces of Justice*, p. 32.

²⁴Gaus, “A tale of two sets,” p. 306.

²⁵Cf. Miriam Ronzoni, “The global order: a case of background injustice? A practice-dependent account,” *Philosophy and Public Affairs*, 37 (2009), 232–42.

²⁶Rawls, *A Theory of Justice*, pp. 294, 99.

application of the principles identified by the original position. He stresses that Rawls never repudiated the “pivotal argument” for the two principles of justice, that is, that they are just because they are the principles that are acceptable in a situation that makes our nature as free and equal citizens the determinative element of our choice.²⁷ Rawls did come to see justice as fairness as a *political* conception of justice, however, meaning one that it is reasonable to expect all reasonable points of view to accept, despite their many differences. The conception that I would otherwise think optimal based on my own doctrine abstracting from the principle of public reason will not generally be the same as the conception that is acceptable to all reasonable views. At this fundamental level, the demand for public justifiability functions as a constraint on the target conception of justice rather than merely as a constraint on deliberation about how to realize that conception. This fundamental constraint will come at some moral cost, judged by one’s own religious or philosophical perspective. The question then arises of whether it is reasonable or realistic to ask people to accept this cost if others are not willing to do likewise. Once we focus on public reason at this basic level, the question arises of whether the principle is itself conditional on reciprocity.

Gillian Hadfield and Stephen Macedo share Weithman’s practical, stability-focused approach to public reason, but they suppose that citizens need assurance of each other’s commitment to exercising this kind of foundational restraint. “How can [citizens] have confidence that others will reciprocate their commitment to supporting governing principles that depart from their own ideal conceptions of truth and value in order to be reasonable to all?” Their assumption is that if (most) others are not willing to support a political conception of justice, but are advocating for what they take to be the whole truth, then it is reasonable (or simply to be expected) that I will do likewise. Deliberating about policy without invoking non-public reasons allegedly helps solve this problem of “mutual assurance”²⁸ by making it common knowledge that citizens are committed to political liberalism, and the idea of adopting a merely political conception of justice as the basis of their political conduct. Hadfield and Macedo assume that public reason involves restraint in the definition of justice, we might say, not just restraint in the application of given principles of justice. What restraint in application provides assurance *about* is the principle of restraint in definition. Public reason in the fundamental, justice-definitional

²⁷Weithman, *Why Political Liberalism?*, pp. 21–3, 28.

²⁸Hadfield and Macedo, “Rational reasonableness,” p. 8.

sense is not a mechanism for stabilizing some other duty that is conditional on reciprocity, but a principle that itself requires assurance of reciprocity.

Putting these two distinctions together (compliance vs. authorship, and application vs. definition), there are at least four duties that are potentially in play (each of which could in principle be unconditional, conditional on reciprocity, or a means to provide assurance of reciprocity with respect to some other duty):

- (1) *Compliance with Law*: The duty to comply with (approximately just) law (even when one could get away with breaking the law, and even to some extent if one thinks the law wrong).
- (2) *Authorship of Law*: The duty to vote for and otherwise support the enactment of just law (instead of voting purely for one's own interests).
- (3) *Restraint in Application*: The duty to deliberate about the implications of given principles of justice employing only generally accepted facts and modes of reasoning (instead of on the basis of all facts and modes of reason one thinks valid).
- (4) *Restraint in Definition*: The duty to define justice in political terms, based on public reasons (instead of based on all reasons one thinks valid).

Which of these duties is subject to a reciprocity condition, and so *requires* assurance of compliance on the part of one's fellow citizens? Which of these duties *provides* assurance about compliance with other duties? Can a duty that is subject to a reciprocity condition itself provide assurance about other duties? Or would that involve simply substituting one assurance problem for another? These are issues that advocates of the stability account have not resolved.

One possible view is that the imperative of restraint in definition (the need for a conception of justice to be political rather than comprehensive) stems from the need for stability, as does the imperative of restraint in application. The idea would be that people will not comply with the demands of justice if the operative conception is too far from what they take to be true justice. They will also not comply if too many others are not complying. The result is that stable conceptions of justice will have to be political. Apart from the lucky few whose conception of true justice happens to coincide with the central tendency in their society's distribution of conceptions, people will have to settle for a conception that is acceptable-if-not-ideal according

to a broad range of conflicting points of view. A comprehensive conception of justice simply will not generate assurance, either with respect to compliance or authorship, because people espousing other doctrines can not be counted upon to support or comply with laws that may be antithetical to their own point of view, and because people motivated by reciprocity will not comply if too many others are not complying.

The main objection to this model is that it will not explain the importance of unanimity on the part of the reasonable. If the problem with a comprehensive definition of justice is that it will not achieve a sufficient level of compliance on the part of others, then the solution is to water down the conception in question to make it more palatable to enough other people—real people, not idealized citizens, and only enough of them, not necessarily all of them. Basing political liberalism on stability alone involves quite a radical revision of the position.²⁹

IV. PUBLIC REASON AS A MORAL PRINCIPLE SUBJECT TO A RECIPROCITY CONDITION

At the opposite end of the spectrum is the view that public reason at the basic level (what I have called “definitional restraint”) follows from our commitment to respect for persons as free and equal moral agents.³⁰ According to Charles Larmore, although respect for persons does not rule out threat or use of force, it does prohibit us from securing compliance with rules by threat of force alone. To secure compliance via threat of force alone is to treat the person threatened as a thing to be moved by causal forces, rather than engaging their distinctive capacity as a person, which is to think and act on the basis of reasons. Threats do make use of the person’s capacity to reason, but they do not engage the threatened person’s reason in the same way our own capacity is engaged. We comply because we think there are reasons for believing in the principle that justifies the rule in question, apart from rewards or penalties. Threats provide incentives for rule-compliance; they do not justify the rule itself.³¹

The argument from respect for persons equates the exercise of political power for reasonably rejectable reasons with the naked threat of force. Yet those who exercise political

²⁹A political conception of justice is not one that strikes a “balance of forces” between existing doctrines, nor that aims to be “near to those doctrines’ centre of gravity” (Rawls, *Political Liberalism*, p. 39).

³⁰Larmore, “The moral basis of political liberalism”; Gaus, “Liberal neutrality”; Boettcher, “The moral status of public reason.”

³¹Larmore, *The Morals of Modernity*, p. 137; Larmore, “The moral basis of political liberalism,” p. 608.

power for nonpublic reasons need not simply threaten. They offer what they take to be good reasons, reasons that really do justify the laws in question, in their view. They engage the rational faculties of other citizens in exactly the same way they engage their own, aiming to show others that the laws or policies in question are morally justified. They do not attempt to deceive or manipulate or otherwise bypass their fellow citizens' ability to engage in justificatory reasoning, treating them simply as objects to be moved by force, or as animals to be motivated by carrot and stick. For this reason, the argument from respect is not persuasive.

The argument from respect can be reformulated, however, as a claim about how citizens can have a relationship of mutual respect despite deep moral disagreement. Membership in a political society is not voluntary, for most people, but it claims final authority over what rights and duties will be recognized and enforced by law, despite the existence of significant pluralism of ethical views even among fully reasonable people. Rawls offered his "criterion of reciprocity" as a description of the appropriate "relation of citizenship" given these conditions.³² He was explicit that "the role of the criterion of reciprocity as expressed in public reason ... is to specify the regime as one of civic friendship."³³ It is possible to interpret this line of thought in communitarian fashion, seeing public reason as a condition of political community in a pluralistic society.³⁴ Whether or not "community" is the right term, the argument from mutual respect is relational, in the sense that the point of the principle is to establish the right kind of relationship between citizens. If the justification of public reason is found in the effect that joint acceptance of the principle has on civic relations, the principle must incorporate a reciprocity condition. It will not make sense to sacrifice the moral truth (as my comprehensive doctrine sees it) for the sake of our relationship if you are not willing to make a similar sacrifice. The question is what form this reciprocity condition takes.

The first possibility is to make the principle binding only when a certain proportion of one's fellow citizens are committed to it. The idea would be that I have a moral duty to define justice in public terms when enough of my fellow citizens do likewise. In this *external* version, the demand for reciprocity affects the conditions under which the principle applies rather than its

³²Rawls, *Political Liberalism*, pp. xliii–iv, 213, 217.

³³Rawls, *Political Liberalism*, p. xlix; cf. Blain Neufeld, "Civic respect, political liberalism and non-liberal societies," *Politics, Philosophy & Economics*, 4 (2005), 275–99.

³⁴Paul Weithman, *Religion and the Obligations of Citizenship* (Cambridge: Cambridge University Press, 2002); Kyla Ebels-Duggan, "The beginning of community: politics in the face of disagreement," *Philosophical Quarterly*, 60 (2010), 50–71; Lister, *Public Reason and Political Community*; R. J. Leland, and Han van Wietmarschen, "Political liberalism and political community," *The Journal of Moral Philosophy*, forthcoming.

content. The second possibility is to make acceptance of public reason one of the criteria of qualification that defines the so-called “justificatory constituency,”³⁵ the set of points of view acceptability-to-which is constitutive of public justifiability. In this *internal* version, the reciprocity condition enters into the content of the principle rather than its conditions of application. The internal reciprocity condition accepts the force of the question “why should I exclude a reason I think true on the grounds that you reject it if you aren’t willing to exercise a similar restraint?” If a point of view rejects the idea of public reason, then the fact that it rejects a reason that I accept does not require me to bracket that reason. This version of the reciprocity condition does not, however, involve any threshold level of acceptance of public reason in the population, below which the principle ceases to be binding.

Which version of the reciprocity condition is correct—or should they both apply at the same time? The answer depends on the reasons for holding duties to be conditional on reciprocity.

V. REASONS FOR CONDITIONALITY

Hobbes famously argued that even my duty not to assault or kill others is conditional on assurance that others will not kill me. Without effective enforcement of non-aggression, I am by right of nature permitted to do whatever I think most conducive to my own survival, even if this means launching an attack on you before you have ever threatened me.³⁶ The form of the reciprocity condition is external, in that the duty is only binding if the general level of compliance is high enough. The reason for the reciprocity condition is that in a low-compliance environment, my performance of the duty threatens (or fails to advance) my own preservation. The cost to me of the performance of the duty is too high, bringing it into conflict with my prior right of self-preservation.

Hume’s example of the society of ruffians focuses on a second dimension, which has to

³⁵I take this term from Jonathan Quong, *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011), p. 181. The idea is normative rather than legal; being excluded from the justificatory constituency does not mean that one is denied the rights of citizenship. Rather, it means that in decisions about these rights, others need not take the fact that you reject a particular reason as grounds for them not to draw upon it.

³⁶*Leviathan*, ch. 14, ¶1; Thomas Hobbes, *The English Works of Thomas Hobbes*, vol. III (London: John Bohn, 1839), p.116.

do with the effects of performance of the duty on others.³⁷ According to Hume, the purpose of having rules of property and contract is to stabilize possessions, making everyone better off by reducing conflict, facilitating trade, investment, specialization, and so forth. In a society of ruffians, however, people are blind to future consequences, and so can't see the long-term benefit to having generally observed rules. In this condition, my respecting the "rules of justice" (that is, property and contract) is of no benefit to myself or anyone else, because it doesn't elicit similar compliance on the part of others. Without general respect for property and contract, the social benefits of these rules are not realized. Again, the form of the reciprocity condition the example suggests is external, and proportional; unless enough others are following the rules of property and contract, the expected social benefits do not arise, so there is no point in my complying.

In these first two cases, whether the duty is binding depends on whether others are complying, but only because their compliance affects the consequences of my performance of the duty (its costs for me or its benefits for others). The nature of this relationship (between compliance on the part of others and the effects of my compliance) is a contingent empirical matter. There could be situations, for example, in which compliance is very low, but the compliance of one or a few will trigger a cascade of compliance on the part of others. The duties in question are not conditional on reciprocity per se, but on the effects of performance, which may depend on what others are doing.

The third reason for duties to be conditional on assurance of compliance on the part of others draws a tighter connection with reciprocity. The idea is that *A*'s duty to *B* (to do some action *X*) may be conditional on *B*'s willingness to do *X* with respect to *A* in similar conditions. The main reason for such a bilateral rather than proportional reciprocity condition would be that mutual recognition that *A* and *B* have this duty to each other (and will generally comply with it) affects the character of the relationship between them. For example, I might think that my duty to aid or share fairly with you is binding only if you would comply with a similar duty to me were

³⁷"Suppose likewise, that it should be a virtuous man's fate to fall into the society of ruffians, remote from the protection of laws and government; what conduct must he embrace in that melancholy situation? He sees such a desperate rapaciousness prevail; such a disregard to equity, such contempt of order, such stupid blindness to future consequences, as must immediately have the most tragical conclusion, and must terminate in destruction to the greater number, and in a total dissolution of society to the rest. He, meanwhile, can have no other expedient than to arm himself, to whomever the sword he seizes, or the buckler, may belong: To make provision of all means of defence and security: And his particular regard to justice being no longer of use to his own safety or that of others, he must consult the dictates of self-preservation alone, without concern for those who no longer merit his care and attention." David Hume, *The Philosophical Works of David Hume*, vol. 4 (Edinburgh: Adam Black and William Tait, 1826), pp. 257–8.

our circumstances reversed. The reason for thinking that such duties might be conditional on reciprocity is that lack of willingness to reciprocate involves a failure of recognition. Refusing to comply with a duty to someone who I know would not reciprocate involves affirming my own dignity, as Stuart White says, and indirectly the dignity of others.³⁸ On this third way of thinking, what is relevant to the question of whether *A*'s duty to *B* is binding is whether *B* would comply with the same duty with respect to *A*, were the relevant conditions to arise. It does not matter whether *other people* are fulfilling their duties to *B*, nor whether other people would fulfill their duties *with respect to A*. Compliance should be conditional on what the beneficiary of the duty would do, not on what others are doing for the beneficiary or what others would do for the benefactor.

The distinction between bilateral and proportional reciprocity conditions gets blurred, however, by uncertainty and the difficulty of targeting conduct to specific beneficiaries. I may be uncertain whether you would reciprocate performance of a specific duty (were conditions to arise in which you owed the duty to me). What rule should I adopt, given that discovery of the truth of the matter may be costly or impossible? On the one hand, I could insist that I will only comply if I am highly confident that you would comply. This policy would minimize unnecessary (that is, unreciprocated) performance, while maximizing wrongful non-performance (non-performance where reciprocation would have been forthcoming). On the other hand, I could adopt the policy of complying unless I am confident that you would not comply. This policy would increase the chances of unnecessary performance, but protect against wrongful non-performance. Of course there are possibilities in between. For example, I might opt to comply with a given duty only if I think most others can be counted upon to comply with it. The underlying rationale might still be relational, but the resulting condition would be proportional because of the fact of uncertainty, and the need to balance two types of moral error.

A similar problem arises in cases in which conduct cannot be targeted to specific beneficiaries. Consider my duty to pay my taxes. The beneficiaries of my tax payments are all those who use the public services to which my taxes contribute, but only some beneficiaries pay the taxes they owe. I do not have the option of targeting my tax payments to services used only by reciprocators. Each additional person who consumes services without being willing to pay

³⁸Stuart Gordon White, *The Civic Minimum: On the Rights and Obligations of Economic Relationship* (Oxford: Oxford University Press, 2003), p. 65.

taxes leaves less for others. If rates of tax compliance are high (for example, due to careful enforcement), then the importance of my duty to the large majority of reciprocators presumably outweighs the moral cost of a few people benefitting from my compliance without themselves being likewise willing to comply. If rates of compliance are low, however, the balance shifts, and it may be reasonable for me to refuse to be taken advantage of by many for the sake of fulfilling my obligations to a few.

The impression that public reason should be conditional on compliance in the proportional sense arises from the fact that the duty applies within political decision-making, and is therefore not “targetable” in its beneficiaries. When I demonstrate against a particular policy, or vote for a politician who promises to put in place some new piece of legislation, the costs and benefits of my conduct are felt by everyone subject to the laws and policies in question. Therefore unilateral compliance with public reason may seem to involve sacrifice on my part for the sake of benefits to non-reciprocators. However, the same logic applies to the duty to vote for justice (as opposed to voting purely on the basis of self-interest). Why do we think *that* duty is not conditional on reciprocity (assuming that is in fact what we think)? One might insist that maintaining good relations with reciprocators is more important than withdrawing cooperation from non-reciprocators. This hard line seems plausible with respect to the ethical duty to support adoption of just law. The fact that Charlie is voting for injustice does not give me the right to do so too, since the laws will apply to everyone, including Betty, who is voting for justice. Yet one might equally say that the fact that Charlie is not paying his taxes does not give me the right not to pay my taxes, because these taxes flow in part to Betty, who is paying or would pay her taxes (if she owed any). If there are lots of Charlies and few Bettys, this reasoning won’t be plausible, in the case of paying taxes. What distinguishes the duty to vote for justice and the duty to pay taxes?

One difference is that the duty to comply with law is enforceable, whereas the duty to vote for just law (as opposed to whatever laws best promote one’s own personal interests) is not. Political liberty protects individuals’ right to make their own decisions about how to vote. A second difference (which results from the first) is that those who fail to reciprocate the creative or constructive aspect of the duty of justice are not fully escaping the requirements of justice. They are constrained by enforcement to abide by the law chosen even if they ignored their duty to support just law and voted solely on the basis of self-interest. A third difference is that it is easier

to verify whether people are complying with laws than whether they are complying with ethical duties that govern participation in political decision-making. With the duty to support and vote for just law, it will rarely be the case that we know that people are violating their duty, because voting is secret, and the relationship between personal interests and political ideology is complex. A fourth difference between complying with a law and complying with an ethical duty governing participation in law-making is that in any large-scale democratic procedure, an individual contribution to the decision is unlikely to be decisive. With the duty to pay taxes, a failure to reciprocate imposes a cost on others; you consume public services that might otherwise have gone to others, but you don't contribute. In contrast, failure to vote for just tax law may have no effect on what law gets chosen. As a result, non-reciprocation of the duty to vote for justice (as opposed to self-interest) need not impose any cost on anyone else. Failure to reciprocate the creative part of the duty of justice is regrettable, but so long as it doesn't affect what law is chosen, this lack of reciprocation doesn't consume scarce benefits that could otherwise better be directed to reciprocators.

For these reasons, the duty to support the adoption of just law should not be seen as being conditional on reciprocity in the proportional sense. One might think that the same holds true for the duty to reason in public terms, since it too is an ethical-but-not-legal duty that applies in the context of political decision-making. However, there are good reasons for thinking that public reason must be conditional on reciprocity in some way. The question of public justifiability arises with respect to reasons that I think valid or sound, and which I think politically relevant, but which are reasonably rejectable. The principle asks me to exclude these reasons from my deliberation despite their being in my view true. This exclusion of reasons comes at some moral cost, judged from my comprehensive point of view. Given what is being asked of me, it is reasonable that I should not be willing to exercise restraint unilaterally.³⁹

It need not follow that the principle is only binding if enough others accept it, however. For in determining which reasons count as unanimously acceptable among qualified points of view, I may recognize as qualified only those perspectives that themselves recognize a duty of restraint. The point of the principle is to make possible a relationship of mutual respect or civic friendship across deep disagreement. This kind of relationship is not possible with those who reject the idea of public justification. It is therefore reasonable to make accepting the principle

³⁹Greenawalt, *Private Consciences and Public Reasons*, p. 128.

one of the criteria of qualification, which is the internal version of the reciprocity condition. In this form, the principle might be binding regardless of how popular it is in a given population.

The main objection to making public reason binding regardless of its popularity is that those who reject public reason seem to benefit from the restraint of others without being willing to pay a similar cost. The idea that they benefit from our restraint is misleading, since what is at stake is not advantages or setbacks to personal interests, but moral gains and losses relative to people's comprehensive doctrines. When Andrew and Ben exercise restraint in their political decision-making but Christine does not, the result may be that the legislative outcome is morally better according to Christine's doctrine than it might otherwise have been, but it need not be the case that Christine is herself better off as a result. Still, if there are lots of Christines and only a few Bens, it may seem that those who abide by public reason are being taken advantage of by those who reject the principle.

It is true that I am giving up something, morally speaking, in agreeing to make political decisions on the basis of public reasons, and that Christine is not doing likewise. Yet I am not making this sacrifice for her sake, but for the sake of my relation with Ben, and I would make this sacrifice anyway, even if Christine were not present. This response to the free riding objection assumes, however, that the moral costs of public reason (relative to Andrew and Ben's doctrines) do not depend on the contents of Christine's doctrine. Suppose, to the contrary, that Christine's doctrine is deemed reasonable despite the fact that it rejects public reason. Suppose also that her doctrine rejects some reason r that both Andrew and Ben accept. In this case Andrew and Ben would be asked to accept the moral cost of excluding r , despite the fact that Christine and other adherents of her doctrine are not willing to pay any similar moral cost (with respect to some reason s that Christine's doctrine accepts, but which Andrew and Ben's doctrines reject). Thus arises our original question: why should I not draw upon a reason I think true and relevant because you reject it, if you are not similarly willing to exercise restraint? The solution is not to make the principle conditional on the proportion of people accepting public reason, however, but to define the principle so that restraint is required only with respect to the ideas of doctrines accepting public reason.

What is at stake here is which views count as reasonable for the purpose of testing whether a consideration counts as a public reason. If we accept the internal version of the reciprocity condition, the objection that Christine "benefits" from Andrew and Ben's costly

restraint without herself being willing to show restraint does not succeed. It may be that that law and policy is not as far from Christine's ideal point than if Andrew and Ben were to reject public reason. But no reason is excluded simply because it is accepted by Christine (since her doctrine rejects the principle of public reason). As a result, the content of Christine's doctrine does not affect the set of public reasons nor the moral costs of public reason to Andrew and Ben. Andrew and Ben are not asked to pay any additional moral cost, beyond that which they are already paying for each other's sake. The reason for their exercising restraint is so that they can have a relation of mutual respect with each other. Since the costs of their restraint are not affected by the existence or contents of Christine's doctrine, on the internal view, her non-compliance should not undermine their duty to each other.

If the principle of public reason is conditional on reciprocity only in this internal sense, then it is binding regardless of whether anyone else accepts it. This result may seem implausible. If the justification of the principle is relational, surely it wouldn't make sense to follow it all by oneself. However, the principle applies in the first place to decisions about the institutions of an ongoing political society, not to particular interactions. Institutional decisions need to take into account our relations with incoming generations, whose ideas are not yet formed, but will be shaped by the institutions in place. In this context, one is never really following the principle alone knowing that no one else will reciprocate. It might be true that when compliance (with public reason) is low, my fulfilling my duty to reason in public terms will not have a big impact, in terms of generating feelings of solidarity, trust, or friendship. However, the account offered here does not have this consequentialist structure, nor does it conceive of the value at stake as being an aggregate measure of average attitudes. The claim is not that citizens should be guided by the principle of public reason because this will tend to promote higher levels of friendly feelings. Rather, the claim is that fulfilling the duty is constitutive, in part, of relations of mutual respect and civic friendship considered one by one.

A final objection is that if public reason is conditional on reciprocity, then it seems to follow, counter-intuitively, that we can coerce those who reject public reason for any reasons we please, including perfectionist reasons taken from our own comprehensive doctrines. The first response to this objection appeals to the rule of law. Power needs to be exercised according to general rules known in advance. Trying to have different rules for different people depending on whether they accept public reason would be quite difficult. The second response appeals to

authorship. Those not coerced by a given law may still have a legitimate public-reason-objection to a law on the grounds that the law coerces in their name and on their behalf.⁴⁰ The final response appeals to indirect coercion. Those not directly targeted by a law are still required to support or not interfere with enforcement of the rule, and required not to attempt to enforce an alternate rule. Among those not targeted by the rule (for example, men, with respect to abortion), some might prefer not to enforce any rule, but they are constrained to support such enforcement, at least via taxes they pay. Others might prefer to enforce a different rule, but they are constrained not to enforce their preferred rule, but to accept the rule decided upon by the community. With decisions about coercion in a non-voluntary association, decisions to enforce rules do not only coerce those directly targeted by the rule.

These are three reasons for thinking that it is not permissible to coerce on the basis of nonpublic reasons those who reject public reason (despite the principle's being conditional on reciprocity). It is not feasible to do so, first of all, due to rule of law considerations. And even if it were, others have standing to object to such treatment because we are coercing in their name, and indirectly coercing them too.

VI. CONCLUSION

In this paper I have tried to explain how these two claims could both be true: (1) the principle of public reason is justified in virtue of its constitutive effect on the relations between citizens, but; (2) it is binding regardless of what proportion of one's fellow citizens accept it.

The answer is that the principle applies to decisions about laws and public institutions compliance with which can be assured by enforcement, and that the principle involves an internal reciprocity condition. The duty to define "justice" in appropriately political terms and to deliberate about political questions on the basis of public reasons is an ethical but not legal duty that applies to one's role as citizen and legislator. As with the general duty to support adoption of just law and institutions, the duties of public reason apply regardless of how many of one's fellow citizens are likewise complying. However, the principle only requires the exclusion (from political decision-making) of reasons rejected by points of view that themselves accept the demand for public justification.

⁴⁰Colin Bird, "Coercion and public justification," *Politics, Philosophy & Economics*, 13 (2014), 189–214.

