

Making Space for Rosa Parks: Democratic Authorship as Political Autonomy

Paul Gowder – pgowder@stanford.edu

Draft: please do not cite or quote.

The great men in history are those whose own particular aims contain the substantial will that is the will of the World Spirit. The[y] can be called heroes, because they have drawn their aim and their vocation not merely from the calm and orderly system that is the sanctified course of things, but rather from a source whose content is hidden and has not yet matured into present existence.

– Hegel, *Introduction to the Philosophy of History* (Leo Rauch, trans.)

Thank you Miss Rosa

You are the spark

Started our freedom movement

Thank you Sister Rosa Parks.

– The Neville Brothers, *Sister Rosa*.

I. Introduction

This paper defends the following two claims, which, together, I call *the leadership principle*: 1) leadership, by which I mean *an individual citizen's exercise of practical reason, in the public sphere, and in pursuit of the transformative authorship of collective values and norms that are rooted, in good faith, in potentially collectively shared ends*, is a virtue of citizens in a democracy, and 2) states ought to be structured so as to permit this sort of leadership.¹ The two claims obviously go together, and, largely, the argument for 2) shall depend on 1), since I take it as fairly uncontroversial that a state that permits its citizens to exercise civic virtue will, *ceteris paribus*, be better than one that does not do so, although I will say some few words about why this is toward the end.

The object is to defend the idea of active citizen authorship and advocacy as essential to the idea of democracy – as part of what we mean when we talk about a citizen's "political autonomy," and as an essential ingredient in a healthy state. The need for such a defense comes from two directions: first, to help meet the challenges to active citizen participation by the long line of skeptics whose

¹ This is not meant to describe political leadership in the sense of running for electoral office or becoming a member of a political party, although a citizen may do those things in the course of the sort of leadership under discussion here, which is more about advocacy and authorship than about the adoption of any particular social/political role.

inspiration comes from Plato and Schumpeter, but which continues today.² Second, such a principle can be applied to give us traction on several major debates in contemporary democratic theory. I will sketch its application to the questions of judicial review, direct democracy, and Rawls's idea of public reason. Detailed elaborations of these applications will not be pursued. The idea is not to offer any arguments on those vexed questions, but to show that the leadership principle helps us think about them. I mean that to be a point in favor of the leadership principle: one reason we might accept an idea in democratic theory is that it allows us to advance our understanding of some of the major issues of democratic organization that we actually face.

The general argument can be laid out syllogistically as follows:

1) Citizens of a democracy are praiseworthy (virtuous), *as citizens*, *ceteris paribus*, to the extent they contribute to their state's instantiating the things that are virtuous for states to instantiate (*public virtues*), such as justice, respect for liberty, social equality, the common welfare, and a democratic decision-making process.

2) Citizen leadership on the domain of values, rather than along the domain of instruments, can, reliably and in the aggregate, contribute to the public virtues.

Therefore, 3) leadership on the domain of values is a virtue of citizens of a democracy.

4) A state that permits its citizens to exercise civic virtue will, *ceteris paribus*, be better than one that does not do so.

Therefore, 5) democratic states ought to permit their citizens to exercise leadership on the domain of values.

Premiss 2) is obviously the boldest claim in this paper, and the one to which most of the effort will be developed, though I do not pretend that I can address all the objections to it in this paper. The argument for it relies on the formal or structural features of citizen leadership in a democratic society,

² Relevantly, for the virtual location of this symposium, consider the paper Jason Brennan recently mentioned here, and which he has forthcoming in *The Australasian Journal of Philosophy*: "Polluting the Polls: When Citizens Should Not Vote."

rather than on the actual content of the norms citizens defend or whether that content is correct or incorrect, and it draws heavily on social choice theory and the theory of history. The claim is that citizen leadership has a positive general tendency, with respect to the virtues of states, in virtue of three features: 1) its ability to promote electoral competition, checking corrupt rulers; 2) its ability to disclose information about citizen preferences (in the broad sense, including beliefs about moral truths), counteracting the power of agenda setters to manipulate outcomes; and 3) its ability to create beneficial instability in incompletely virtuous societies, leading to progress, over time, to better societies. These three benefits, I argue, are sufficient to establish that citizen leadership as I have specified it will work overall toward the betterment of democratic societies, in light of the special features of the sort of leadership I've identified.

The leadership principle is meant to be a freestanding political claim of the sort that can be incorporated into the political doctrine of a society following Rawls's political liberalism -- nothing in the argument below depends on, for example, perfectionist claims about the best human life incorporating political involvement. Rather, the argument is about what constitutes goodness (virtue) for a citizen, *as* a citizen, and under democratic traditions that depend on no comprehensive doctrine. This strategy excludes a variety of arguments that are more traditionally brought to bear in support of an active role of citizens in a democracy; excluded are, for example, neo-Aristotlean claims about humans as political animals,³ the sorts of claims Rawls identifies [ci: *Political Liberalism* pg. 206] with civic humanism, and Arendt's claims about the value of action. Such a freestanding claim serves two useful purposes. First, it will strengthen the case for citizen leadership. That's simply a function of its compatibility with a wider set of comprehensive doctrines -- a politically liberal account of citizen leadership will not be subject to objection by (for example) comprehensive liberals on the grounds that it demands people refrain from pursuing their private ends, nor by those who reject the naturalistic premises of an Aristotlean

³ As in Josiah Ober, 2007, "Natural Capacities and Democracy as a Good-in-Itself," *Philosophical Studies* 132:59-73.

argument about human essences. Second, it will permit the account to be used as an internal critique of elements of Rawlsian theory. In particular, the account I develop in this paper can be used as a critique of Rawls's idea of public reason, and the constraints on political justification that it contains. (As will be apparent upon perusal of the argument, there is a strong Rawlsian strain to this paper: at numerous critical points, claims are established by what is ultimately a straight appeal to Rawls; a practice that is justified primarily by the object of providing such an internal critique).

Also, the argument is squarely in the tradition of non-ideal theory, and indeed depends on so being: it is necessary to many of the claims below that we think of ourselves as in a society that is, and will for the foreseeable future be, imperfect with respect to the virtues of states (imperfectly just, imperfectly democratic, imperfectly liberty-respecting, imperfectly welfare-maximizing, etc.), such that we can coherently understand the idea that progress toward greater state virtue, in a democratic fashion, is, if not an obligation of all citizens, then at least something that is praiseworthy for a citizen. As a society becomes more virtuous, the need for leadership may decrease (and in a perfectly virtuous society, instability-inducing citizen leadership might even be a danger), but we are not in that world.

I also, by way of introduction, want to put up front what I mean by the rather complex idea of leadership I have articulated. There are two main elements. First is the idea of *transformative authorship*. This is meant to mean leadership on the big questions – on the fundamental moral values and social relations that make up a political community – not on the little questions, like the particular way in which a given value is to be implemented.⁴ The borders of this concept are ineluctably fuzzy, but the broad idea should be clear enough to proceed.

Second is the idea of the good faith pursuit of *potentially shared ends*. This is meant to be a constraint on citizen leadership to ensure that it remains minimally democratic, which we will see requires it be rooted in values to which the prospective leader genuinely and reasonably (or not grossly

⁴ Doubtless a case could be made for the value of (competent) citizen leadership on the little questions too. That is not the aim of this paper.

unreasonably) believes all other citizens could conceivably endorse after an ideal process of persuasion. Importantly, this is a weaker constraint (in ways to be discussed toward the end of this paper) than those implied by Rawls's idea of public reason, which can be seen as serving a similar function.⁵ It is meant to be a very weak standard indeed, meant to exclude from civic virtue those who advocate gross offenses against the public like the reinstatement of racial apartheid, the end of women's suffrage, and the like. In addition to a democracy constraint, it is also a substantive justice constraint: by limiting citizen leadership to potentially shared ends and excluding those things that are beyond the pale in the ordinary social world, it eliminates a possible way in which citizen leadership could lead to societies becoming less rather than more virtuous over time.

II. The Central Argument: Leadership as a Virtue of Citizens

II.A. Leadership: Superogatory, not (necessarily) Obligatory

Premiss 1: Citizens of a democracy are praiseworthy (virtuous), as citizens, ceteris paribus, to the extent they contribute to their state's instantiating the things that are virtuous for states to instantiate (public virtues), such as justice, respect for liberty, social equality, the common welfare, and a democratic decision-making process.

I take it that the first premiss that I identified in the introduction will be uncontroversial. It is a minimal premise, meant to capture the weakest plausible claim about the relationship between a citizen in a democracy and the state of which she is a citizen: that citizen is praiseworthy to the extent she contributes to the instantiation of state virtues. In its weakest form, this premise is beyond dispute: if a citizen contributes (by permissible means) toward making her nation more just, or more democratic, that citizen has done something good. No more discussion is required.

What does require discussion, however, is the importance of the claim's in fact being weak. Particularly, I want to insist that my argument only requires that this contribution be superogatory,

⁵ So suggests Joshua Cohen in *Procedure and Substance in Deliberative Democracy* [cite], where he argues that to fail to appeal to public reasons amounts to excluding one's fellow-citizens from the *polis*

rather than obligatory. The argument is compatible with stronger claims, and those who hold certain positions in normative ethics might require stronger claims. Act-utilitarians, for example, ought to be committed to the nonexistence of superogatory duties. For an act-utilitarian, a citizen's contributing to the goodness of the state is morally obligatory, at least so long as she doesn't sacrifice any value of equal import to do so. This is fine.

But *this* argument cannot rely on any stronger version of the claim, on pain of running afoul of (comprehensive) liberalism. A major objection to participatory democracy, which I find compelling, is the idea that participatory democracy requires "too many meetings" – that it requires citizens to sacrifice too many of their private interests in order to pursue the public good. Similar objections could be made to any defense of democratic leadership that required the strong (obligatory) version of premiss 1 rather than the weak (superogatory) version: perhaps citizens ought not to be required to spend all their waking hours trying to figure out how to make their state slightly more virtuous.

Of course, this objection might not destroy the strong version of premiss 1: the appropriate answer might just be to abandon comprehensive liberalism. But this seems far too strong to casually do, when the argument can go through perfectly well on the basis of a superogatory virtue. The aim here is, after all, to articulate a form of democratic virtue that is compatible with a Rawlsian politically liberal society, which would presumably include reasonable comprehensive liberals. This argument thus relies only on the weak version of premiss 1.

II.B. Citizen Leadership and Social Progress

Premiss 2: Citizen leadership on the domain of values, rather than along the domain of instruments, can, reliably and in the aggregate, contribute to the public virtues.

Imagine a naive democrat. Such a democrat, heartened by arguments about a "marketplace of ideas" and structural features of the epistemic environment like the Condorcet jury theorem, might casually assert premiss 2 because he is under the misimpressions that a) citizens can reliably be

expected, over time, to come to consensus on shared values and b) even to the extent that citizens don't come to consensus, things like the Condorcet jury theorem will guarantee that citizen involvement is positively associated with better outcomes. Our job is more complicated, in virtue of two features of the political world:

The Fact of Reasonable Pluralism (per Rawls): Even in the best practicable society, there is likely to be deep and abiding disagreement on the moral and political right, as well as the more basic questions of how society is to be structured to achieve any such moral and political right. Moreover, this disagreement is *reasonable*, in that it can be traced to the structure of human reasoning on values and questions of ultimate truth, such that no citizen can reasonably demand that others adopt his views on those matters.⁶

The Fact of Citizen Incompetence: Citizens are not reliably good at getting the right answers to political questions, particularly complicated questions of policy formation and implementation. There is no reason to believe that citizens ordinarily have even the better-than-even odds of getting many political questions right that the Condorcet jury theorem demands. This is a particular problem for domains such as economics and science policy, which require specialist knowledge unavailable to most citizens.

How can we preserve premiss 2 in light of the two facts just identified, which cast doubt on any arguments relying on the tendency of citizens to converge to similar beliefs or collectively bring about the right answer to individual social questions by ordinary electoral processes? Before getting into the specifics, I will highlight two general features of the approach: a) the limitation to questions of value and b) the focus on structural, rather than substantive, features of citizen leadership.

⁶ This paper takes on this claim *arguendo* only. I am not certain that there really is such a thing as reasonable pluralism.

Leadership in values

One reason to limit premiss 2 to citizen leadership in the domain of *values* rather than *instruments* is to avoid some of the bite of the fact of citizen incompetence. By “instruments,” I mean concrete policy choices directed at achieving the values that those who constitute the state have decided are to be pursued. For example, the decision to pursue economic prosperity is a value, while the decision to raise or lower interest rates in order to do so is an instrument.

The argument for citizens’ deferring to specialist expertise is much stronger with respect to instruments than to values, because instrumental expertise unequivocally exists: there are people who simply know much more about what causes unemployment, global warming, and terrorism than does an average citizen. On the other hand, there is serious dispute about whether *moral* expertise exists in the same sense as does policy expertise [citations from moral expertise literature go here].

This paper takes no position on the general question about moral expertise. Political morality in a democratic society, on the other hand, is a special case, and for that case, we have reason to assert the absence of expertise. As this paper is addressed primarily to those who have some of the commitments of Rawlsian political liberals, the main point to be made in favor of what I will call the no-expertise thesis is that it is entailed by the fact of reasonable pluralism. Essential to the idea of reasonable pluralism is that one is not blameworthy for moral disagreement with one’s co-citizens or for refusing to defer to the claimed moral expertise of (say) a Kantian philosopher or a Catholic priest. But one essential feature of expertise is that those who hold it can reasonably demand that those who do not hold it change their views to conform. At least, one is not publically blameworthy in such a society for this disagreement: one’s views cannot be discounted because one fails to accept the moral views that one’s peers accept, supported though they may be by an army of experts.

By contrast, I am epistemically blameworthy, and perhaps morally blameworthy (to the extent my epistemic failings have harmful consequences to others) if I ignore the advice of a climatologist in

casting my vote on global warming policy, conditional on my having the same values as that climatologist with respect to whether the consequences of global warming are undesirable or not. This difference shows us that whatever the relationship is between citizens of a (politically liberal) democracy and moral specialists, it is not a relationship of expert-layperson, entitling the former to deference.

Of course, the absence of moral expertise does not entail the conclusion that we will collectively get moral questions right. Perhaps we're all just really bad at moral questions. If that's true, it represents a challenge to democracy in general – indeed, to all systems of government – and one that I don't purport to meet here.⁷ What the limitation to values, in conjunction with the no-expertise thesis, achieves is the elimination of a standard argument for *deference*, by limiting the domain of the virtue of citizen leadership to those questions for which there is nobody to whom to defer.

The focus on structural features of citizen leadership

The ideal form of an argument for premiss 2 would show that social progress is promoted as a consequence of the mere fact of citizen leadership, rather than anything about the content of the ideas expressed in that leadership or the ends for which that leadership is exercised. The canonical example of such an argument is Hegel's theory of history: spirit's development in history is advanced regardless of the individual aims of the world-historical individual who represents the next stage in that development; Napoleon's private desires drive matters, but it does not matter what those private desires are – the "cunning of reason" brings it about that they align with the teleology of history.

⁷ At one level, if humans are generally terribly bad at making moral judgments, *and* there are not even moral experts (by the no-expertise thesis), then we would be forced into a pessimistic kind of anarchism: if all I can know is that everyone's judgment of the right is not to be trusted, I certainly have no reason to subject myself to law made by any of my fellow humans; although I can do no morally better myself, at least in anarchy I won't be subject to the forcible imposition of the bad moral judgments of others. At a higher level, the combination of the belief that the masses make bad moral judgments and acceptance of the no-expertise thesis strips one of the normative resources to judge any governments at all, for they entail in the holder of both beliefs the conclusion that one's own normative judgment is faulty as well, and some level of trust in one's own normative judgment is necessary to critique any state.

Something like this (although not nearly as pure or well-developed) is the aim of this argument – an account of how citizen leadership promotes the virtues of states that does not depend on the content of the norms promoted by citizen leaders, or their correctness.

II.B.1. Agenda-Setting and Corruption

One major vice of a democratic society is *concentration of power*, and that this is a vice that causes two problems: first, it risks the thwarting of the legitimate political will of the citizenry, and thus threatens the virtue of popular sovereignty. Second, it permits the use of public power for private gain – that is, *corruption* – and the consequent loss of public welfare. In the next two subsections, I argue that citizen leadership can reliably ameliorate those two consequences of the vice of concentration of power.

II.B.1.a. Leadership and Schumpeterian competition

The exercise of practical reason by citizens in the public sphere bears some relationship to citizens' participating in electoral competition: citizens may choose to run for office as a way to express their considered judgments about the direction in which their society is going, or other citizens may be inspired to run for office by the arguments of moral critics, in hope of making a change. Yet even the most skeptical of commentators on democracy will accept that electoral competition is a good thing for a democratic society, because it serves as a check on corruption and mismanagement.

The standard Schumpeterian line on elections is that they serve as an escape valve against bad rulers: if the rulers get too bad, the citizens can always throw them out and select new ones. But this isn't true if electoral competition is not meaningful, either because there are no viable competitors to the current rulers or because all of the competitors are just as corrupt or incompetent as is the lot in power. But if citizens are running for office based on strong moral critiques of the existing system, then

this increases the probability of finding a viable candidate to run against the crooks and fools in power. This is a good for any democratic political society.

II.B.1.b. Social choice theory and leadership as an information-forcer about preferences

Recall that the sort of leadership I propose to defend is an exercise of practical reason, in the public sphere, about moral norms. The previous subsection might seem somewhat beside the point, for surely not all who exercise that sort of practical reason compete for political office, and, equally, not all who compete for political office exercise that sort of practical reason. However, there is another, broader, way that citizen leadership promotes the virtue of day-to-day politics: by providing information about public preferences, which makes it harder for dominant agenda-setters to manipulate outcomes.⁸

One of the pathologies of democratic government is undoubtedly the power of those who control agenda – who determine which issues are presented for choice, in what order, and in what bundles – to manipulate electoral outcomes.⁹ We know from Arrow's theorem, the Gibbard-Satterthwaite theorem, and their various extensions that it is impossible to construct a collective choice rule, otherwise satisfying minimal rationality conditions, which guarantees both path independence and independence of irrelevant alternatives – that is, which is not subject to potential agenda manipulation by an actor with knowledge of the underlying preferences and the power to alter either the order in which options are considered or the set of options that are considered. There is no need to go into the details of the social choice theory in this paper – it is sufficient to appeal to Riker for the explanation: in many

⁸ It is important to note that by preferences, here, I don't mean the narrow sense of preference to which the work of economists is directed, but rather a more capacious notion that tracks the will of a given voter with respect to public policy – a voter's all-things-considered judgment about what electoral choice is best, given a set of moral beliefs and a disposition to bring state policy in accord with them.

⁹ The propositions in this subsection about the power of agenda control are taken almost entirely from chapter 7 of Riker, *Liberalism Against Populism*. [cite]

circumstances, even if there is a Condorcet winner,¹⁰ or even what Riker calls a “strong equilibrium,”¹¹ an agenda setter can manipulate the order of voting or add irrelevant alternatives in order to bring it about that some other option is selected. This is a result that is often seen to pose serious danger for the ability of democracy to meet any reasonable normative ideals. What, after all, good is even the liberal notion of popular sovereignty as a check on the powerful if those powerful, who are better-informed than the public, can achieve any electoral results they want merely by tweaking the agenda?

But the crux of that point is in the “better-informed.” In order to effectively manipulate the agenda, of course, agenda-setters need to have some idea of the preferences of a given electorate. As they have incentives to do so, it’s not so easy to see how such information could be concealed from them (or if it would even be a good idea). But the electorate also needs to *lack* information about its own preferences. As Riker also points out, strategic voting can, in many practical circumstances, defeat the power of agenda setters, if the electorate is informed about the preferences of its members.

The role of leadership as I’ve described it ought to be obvious in this context. By exercising practical reason in the public sphere about moral norms, citizens express their own preferences and those of the constituencies for whom they stand, as well as spark debates that may motivate other citizens to do the same. Consequently, they make it more likely that choices that are strongly preferred (such as Condorcet winners) and that happen to exist will become known. This can only weaken the power of dominant agenda-setters: surely a service to democracy.¹²

¹⁰ That is, an option within a set of alternatives that can beat all the other alternatives in the set in pairwise comparisons, though it may not win in a simultaneous comparison of all alternatives.

¹¹ An alternative most preferred by an absolute majority of the electorate.

¹² This point need not be expressed only in social choice-theoretic terms. Intuitively, it is better for democracy that citizens know where their fellows stand. Surely, for example, one of the great achievements of the civil rights and anti-war movements in the 60’s was that they let those who were opposed to racism and the war know that they were not standing alone against the state, and this doubtless inspired further activism and drove the change achieved by those citizens.

II.B.2. What the owl said to Riker: A connection between the theory of democracy and the theory of history

In this (lengthy) subsection, I will argue that citizen leadership can promote progress toward societies that have more of the civic virtues, particularly justice, by creating instabilities that threaten worse societies more than they do better ones. In effect, my object is to turn Riker on his head: Riker famously argued that the populist model of democracy fails in part because of the instabilities caused by policy entrepreneurs in light of the impossibility results of social choice theory. In this subsection, I argue that this is *a good thing* – that this instability promotes social progress, and that citizen leadership can contribute to this beneficial instability.¹³

On the Notion of “Stability.”

In this subsection, I appeal extensively to the idea of a society’s “stability.” This is, unfortunately, a somewhat undertheorized concept, and one that is put to a variety of potentially incompatible uses through political philosophy as well as political science. Just consider, for example, the roles stability plays in Rawls’s concept of a well-ordered society and Riker’s social choice critique of populist theories of democracy (examples chosen advisedly, as they are both deployed below). Can they be reduced to one idea?¹⁴

Clarifying the literature on stability is a different and very difficult project. But I’m obliged to say at least something tentative about the concept, as it plays a major role in my argument. Perhaps the common core of “stability” can be expressed as follows. Various facts about a state indicate that state’s instability at various levels of severity, which we can imagine laid out on an ordered continuum. The

¹³ We can see Riker’s own primary example, in chapter 9 of *Liberalism Against Populism*, that of the civil war, as an excellent data point in support of this proposition. If we take all of Riker’s claims at face value, the preferences of citizens in the antebellum period, as well as the way in which democracy was structured, led to dangerous instability that were exploited by policy entrepreneurs to break up political coalitions, leading to the unsettling of the prior constitutional equilibrium and bringing about the civil war... *which caused the end of slavery in the United States.*

¹⁴ I thank Ruth Kricheli for pressing me to make myself much clearer on this point.

most severe sort of instability is one in which the state's form of government is frequently changing, as from democracy to dictatorship and back, or where those who hold power in the society are violently replaced. Less unstable, but still very unstable, is a society in which its fundamental orientations on major political divides changes – a society that changes from capitalist to socialist and back, for example, or from religious to secular. Another, still lesser sort is experienced by a society whose major policies or presuppositions of social life are under serious and sustained challenge – consider conditions in the U.S. in the 1960's. I think this is the sort of instability Rawls was worried about with respect to the two principles of justice: a stable society in the Rawlsian sense is one in which the public has common knowledge that the principles of justice are endorsed by all. Still less unstable, but unstable, is a society whose governing party coalitions are in flux, or where legislation is regularly changed and re-changed. The weakest sort of instability is the sort discussed by social choice theorists – call it “social choice instability” which is instantiated in the relationship between the preferences of those in the society entitled to make decisions on official questions (the electorate, and/or legislature) and the function used by that society to convert those preferences to decisions (e.g., simple majority vote, borda count, etc.). Social choice instability exists to the extent that the preferences on some issues, given the choice function in place, lead to cycles, or may lead (as in the previous subsection) to potential manipulation by agenda-setters.

There are a few fairly uncontroversial propositions about stability, which I cannot defend here, but which seem true. First, instability can be caused by shocks that are exogenous to the ordinary political life of a society (wars, economic upheaval, diseases, famines, etc.). Instabilities can also be caused by internal shocks or the actions of citizens within the ordinary process. Particularly, although not exclusively, instabilities at lower levels of the continuum can be causal factors in instabilities at

higher levels. Social choice instabilities such as preference cycles¹⁵, for example, can contribute to instabilities in legislation or in governing coalitions, when there are dissatisfied people in society to make use of the cyclic preferences. Riker gives us an account of the use of cyclic preferences by policy entrepreneurs to bring about legislative and coalitional change in *Liberalism Against Populism*.

Intuitively, those can cause still higher-level instabilities: consider (again from Riker) how instabilities in the governing coalitions of the United States and policy toward slavery and other issues led to the U.S. Civil War.¹⁶ Generally, higher-level instabilities incorporate lower-level instabilities: a revolution, for example, ordinarily brings with it rather a lot of legislative change. Finally, it is uncontroversial that instabilities, particularly at the higher levels, can lead to major social change, in either direction, along the dimensions of the virtues of states under discussion in this paper. The classic examples are (on the negative side) the instabilities in Weimar Germany that led to the rise of the Nazis, and, on the positive side, the instabilities that led to the abolition of slavery in the United States, and, later, the victories of the civil rights movement.

This is the conception of stability with which the paper will work. It should be contextually obvious which level of stability is under discussion at any point, and, where it is not obvious, the reader may assume that all levels are being discussed as a single phenomenon. Doing so is permissible because of the cross-causation between levels of instability discussed in the previous paragraph: a fact that causes one level of instability thereby raises the probability of other levels of instability.

¹⁵ A situation where, under the voting rule in effect, something like the following happens: option a is preferred to option b, which is preferred to option c, which is preferred to option a.

¹⁶ This is the proposition that connects Riker's social choice instability to Rawlsian instability.

Two Claims About History

If we look at the major works in philosophy of history starting with Hegel and Marx, we can see a central idea that seems to animate them at a high level of generality, which we can state in fairly weak form as follows:

The Generalized Historical Assertion: Better states tend, on average, to be more stable than worse states, at least temporarily, and the instability of bad states leads to their progress to better forms.

Yet this idea seems too strong, and its defense depends on a number of contestable empirical claims about the relationship between the virtues of states and occurrences within those states.¹⁷

However, it is possible to state a weaker, and much more easily defended, thesis that can give us much of what we might want out of a teleological theory of history, which I will call *the Stability-of-Virtue*

Thesis:

More virtuous states are, *ceteris paribus*, more robust against the sorts of shocks that tend to create social instability than are less virtuous states.

Before considering the plausibility of the stability-of-virtue thesis, note how it is weaker than the generalized historical assertion. Unlike the GHA, the SVT does not require that “evil contain the seeds of its own destruction.” It might not do so: lots of wicked societies have persisted for a long time.¹⁸ The SVT does not require that unjust states inevitably perish, for they may exist in an environment that does not generate shocks sufficient to destabilize them. Put differently, GHA suggests that just and unjust states create their instability-generating shocks endogenously, and respond in accordance with their

¹⁷ The strongest defense of which I’m aware of the generalized historical assertion, at least as applied to the specific case of the demise of slavery, is Joshua Cohen’s “The Arc of the Moral Universe” [cite]. While Cohen is convincing in that respect, it’s not clear to what extent his argument extends beyond slavery.

¹⁸ In game-theoretic terms, it might be that a given set of environmental conditions (parameter values) supports multiple stable social equilibria, some more unjust than others.

levels of justice. SVT is merely a claim about what happens to just and unjust states when subjected to exogenous (or endogenous) instability-generating shocks.

Of course, “weaker” is not the same as “does not require defense.” SVT is still highly controversial as well as a difficult proposition to defend, and I cannot begin to defend it in anything like a satisfactory manner here. Readers who do not accept it may, as a consequence, reject this part of the argument, and, in doing so, perhaps reject the argument as a whole, as the argument for premiss 2 may be fatally weakened without it. I will, however, offer some considerations that might move one to at least consider SVT somewhat plausible.

More importantly, however, notice that SVT represents something like a minimal position of optimism about the course of human affairs – using Cohen’s (1997, pg. 96) words from the context of the discussion of the stronger GHA, it serves as a test as of whether it is “reasonable, from a moral point of view, to hate the world.” As a consequence, SVT is something to which many of us are committed (those of us with a shred of optimism about where the world is going), whether or not we can demonstrate its truth. Even if one does not accept my case for SVT, one can profitably use this section as an exploration of the consequences of an optimistic proposition that many of us believe, plus some less difficult claims, for our theory of democracy.

I will defend SVT on two levels: conceptual and empirical.

II.B.2.a. The Conceptual/Intuitive Relationship Between State Virtue and Stability

SVT is, I think, true in large part because we don’t even say a society is virtuous unless it is stable – that is, one necessary condition of the social virtues is their stability. I appeal here primarily to an important element of Rawls’s argument for theory of justice: on his argument, one of the things that makes a society just is just that the principles governing the basic structure constitute a stable social

order on which people can rely, and that they can endorse. A just society is a well-ordered society, and a well-ordered society is stable at least on the level of its fundamental principles.

This is an idea that is *constitutive* of Rawls's theory of justice, not merely an implication of it, and it is tied to two important features of his argument that ought to be compelling even to those who do not think that justice as fairness expresses the best conception of justice. First, one of the regulative principles that is important in setting the notion of a just democratic society is the Kantian idea that such a society's principles of justice are public – and thus that it is common knowledge that all accept those principles. An unstable society fails to satisfy the publicity condition, since the political agreements that make up its principles of justice are not commonly known to be accepted, because not commonly accepted, and (for present purposes) we can know that the publicity condition is not satisfied because people are working, or would like to work, to undermine the principles of justice on which the society is based (to have a revolution, in the extreme case). Second, it is also an implication of the contractualist method of finding the principles of justice with a conception of an original position made up of parties with a sense of justice: the parties to the original position are constrained to only choose stable social orders because, as people with a sense of justice, “they cannot enter into agreements that may have consequences they cannot accept.” (Rawls 1999, p. 153.)¹⁹ For that reason, they will choose principles acceptable to all, and, thus, stable.

The upshot of this is that a conception of justice is worse, *ceteris paribus*, to the extent that a society that follows it does so less stably, because, for example, it is not acceptable to all reasonable citizens.²⁰ Since conceptions of justice match social conditions (for each conception of justice, there is a

¹⁹ Rawls offers an alternate account of the stability of a liberal conception of justice in *Justice as Fairness, a Restatement*, p. 185-6, in terms of public reason rather than the publicity requirement. Nothing in this argument depends on that version of the account.

²⁰ Rawls puts the point this way:

[A] strong point in favor of a conception of justice is that it generates its own support. When the basic structure of society is publicly known to satisfy its principles for an extended period of time, those subject to these arrangements tend to develop a desire to act in accordance with these principles and to do their

possible society that instantiates it), it follows, if we accept the Rawlsian approach, that a society that is less stable is, for that reason, less close to instantiating the best conception of justice – that is, less just.²¹

Nor is justice the only virtue of states for which this sort of argument can be made. Consider democracy. One aspect of democracy – one property in virtue of which we can call a state more or less democratic – is the amount of effective popular sovereignty such a state has, that is, the extent to which the will of the people (to the extent there is one) is actually translated into government policy. But this is a property that is thwarted in the face of instability: citizens have a narrower range of effective policy instruments, and their will can be expected to be less efficacious and over less time to the extent that they cannot rely on the persistence of their government or the social order they put into place with a given legislative act.

Thus, we can assert SVT through the back door: we need not rely on an empirical connection between state virtues and stability, for stability is required before we attribute two of the most important state virtues, justice and democracy, to any given state. While a non-virtuous state may be stable or unstable, a virtuous state *must* be stable.²²

part in institutions which exemplify them. A conception of justice is stable when the public recognition of its realization by the social system tends to bring about the corresponding sense of justice. (Rawls 1999, pg. 154). The point is that this is not an empirical claim about just societies, it's a *conceptual* claim about what it is for us to say that a society is just: given two conceptions of justice, we are to prefer the one that leads to a stable society, where "prefer" here means "judge as a more accurate understanding of what justice is." Much of Rawls's defense of his theory of justice amounts to an attempt to show that it leads to a more stable society, and is thus better at capturing that element of what justice *is* than alternative theories, such as utilitarianism.

²¹ Rawls is, regrettably, not consistent in taking this position. In section 76 of *Theory*, he suggests that it is possible for a just society to be unstable (Rawls 1999 pg. 434-5). But this is inconsistent with the rest of his argument: if it is possible for a just society to be unstable, then a theory of justice could be correct – could more closely approximate correct principles of justice – *even though* it would not lead to stable social arrangements. Were that the case, the superior stability-generating properties of justice as fairness would not constitute an advantage with respect to utilitarianism (not, that is, unless both were equally just, and thus stability constituted a tiebreaker for theories of justice – a claim Rawls nowhere makes), as Rawls repeatedly claims.

²² This Rawlsian notion of a well-ordered society is intuitively plausible even without the Kantian and contractualist basis on which it rests in Rawls's argument. An unstable society is a society which people cannot depend on over time, and thus one in which they will be unable to order their affairs while trusting that they can predict the

II.B.2.b. The Empirical Connection Between State Virtue and Stability

Intuitively, we might imagine some obvious connections between state virtue and stability. Many of the important ways in which a state is good are also factors that directly relate to that state's persistence in its current form. Thus, one of the ways that a state might be good is if it is peaceful and on good terms with its neighbors – and such a state is obviously going to be less vulnerable to external threats. Internally, another prominent way that a state can be good is if it meets the legitimate interests of its members in a fair way. Such a state is more likely to have the support of its members, obviously, than an unjust state – its members are less likely to rebel against it or attempt to change its laws by legal or extra-legal means, they are more likely to comply with its laws, and they are more likely to support it against threats posed by others.²³ A state that ensures the poor are well cared-for need have less fear of revolutions by the poor, for example. Likewise with democratic virtues: laws that are enacted with the consent of the masses are less likely to be resisted by those masses. Moreover, people have senses of justice and of the other social virtues, and tend to care about the virtues of their society. It is a truism that even those who are not directly victims of a society's injustice will see that injustice and will oppose it, or, failing to actively oppose it, will not defend it as vigorously as those who live in a just society.

Happily, these intuitions find support in the social science literature. Several features of a society that are often considered virtuous by political philosophers have strong connections with stability.

behavior of their fellow citizens – at an extreme, an unstable society is one in which the rule of law falls by the wayside. Such a society is also one in which people have fewer incentives to avoid doing harm to their fellows: a society where the rules are not consistent, or consistently enforced, is one where people may be tempted by potential impunity to act unjustly toward others.

²³ “The Moral Law causes the people to be in complete accord with their ruler, so that they will follow him regardless of their lives, undismayed by any danger.” -- Sun Tzu, *The Art of War*, Lionel Giles, trans. (1910). (Relatedly, the imperial Chinese used to talk of the “mandate of heaven” as the force that preserved just rulers in power.)

Consider democracy. Research on the “democratic peace” shows that democracies tend to not wage war between themselves and tend to win wars with non-democratic enemies,²⁴ features that obviously tend to preserving stability in the face of shocks. Similarly, political scientists have seen positive effects from expansion of the franchise: Acemoglu & Robinson strongly suggest that expanding the electoral franchise to include the lower classes in Western democracies was necessary to promote social stability by ensuring that the lower classes could enforce wealth transfers, and thus reducing social unrest and the threat of revolution.²⁵ The idea can be generalized: democracy, to the extent it allows leaders to credibly commit to social policies in the interest of powerful groups (e.g. because it creates a signal for coordinated sanction²⁶), allows for peaceful rather than violent resolutions to social conflict, because it raises the expected utility from remaining in society for groups that would otherwise be revolutionary. Third, Sen has repeatedly and persuasively argued that democracy prevents widespread famine,²⁷ and obviously, countries where there is famine or the risk of famine will be less stable than countries where there is no such risk, both because they are more vulnerable to external threats and because they are more vulnerable to internal unrest.

Second, consider a wealth-transfer policy that distributes resources to the poor. Acemoglu, Robinson & Johnson offer a model demonstrating that, in general, even non-welfare-maximizing kinds of redistributive policies can increase social stability (in the form of keeping the current leaders in power).²⁸ The strategic idea here is that redistribution can increase the power of the bloc to whom redistribution is given, thus ensuring that they have the power to enforce future commitments from leaders (once again increasing the payoff from the non-revolutionary option). Less strategically elegant,

²⁴ Bruce Bueno de Mesquita, James D. Morrow, Randolph Siverson & Alastair Smith, 1999, "An Institutional Explanation of the Democratic Peace," *American Political Science Review* 93:791.

²⁵ Daron Acemoglu & James Robinson, 2000, "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective," *Quarterly Journal of Economics* 115:1167-1199.

²⁶ As in James Fearon, 2006, "Self-Enforcing Democracy," paper presented at the 2006 Annual Meetings of the American Political Science Association, Philadelphia, PA.

²⁷ e.g., Amartya Sen, 1999, "Democracy as a Universal Value," *Journal of Democracy* 10.3:3-17.

²⁸ Daron Acemoglu, James Robinson & Simon Johnson, 2001, "Inefficient Redistribution," *American Political Science Review* 95:649-661.

but also somewhat compelling, analyses of ways that redistributive politics increases social stability abound in the literature, and tend to hinge on the idea that redistribution is enacted when the lower classes seriously threaten the social order.²⁹

II.B.2.c. Leadership and the Evolution of States

The results of social choice theory have shown us how active democracies can promote instability. The main argument is given by Riker in *Liberalism Against Populism* (1982), and is too complicated to do more than summarize the chief conclusions here. The basic idea is as follows: at any given time, there are latent social choice instabilities in democratic populations – there is, for example, the possibility of cyclic preferences under various configuration of the agenda.³⁰ In many cases, these instabilities are kept latent by features of the structural environment, such as restrictions on agenda-setting power, as well as on legislative practices like vote trading and coalition forming.³¹ Instabilities may also be kept latent by happenstance: it might be that a policy dimension is stable when considered in one way (as, for example, a simple binary choice between raising taxes a given amount and keeping them the same), but becomes unstable when combined with other dimensions (how shall the money be spent?) or when more points in the set of possible outcomes along that policy dimension (such as lowering taxes, or raising them a different amount) are introduced.

Dissatisfied citizens can add new issues and re-bundle old issues, such that old allegiances are changed, and perhaps policy cycles that were previously latent have an effect on policy. Such policy

²⁹ Possibly the most classic example is Frances Piven & Richard Cloward, 1971, *Regulating the Poor: The Functions of Public Welfare*, New York: Pantheon Books. A different version of the argument is made in J. Falkinger, 1999, "Social Stability and Redistribution of Income," *European Journal of Political Economy* 15:35-51. This argument goes at least back to Aristotle (Politics, bk. 2, pt. 5-6, bk. 5, pt. 1). "Everywhere inequality is a cause of revolution, but an inequality in which there is no proportion- for instance, a perpetual monarchy among equals; and always it is the desire of equality which rises in rebellion."

³⁰ McKelvey's chaos theorem shows that cyclic preferences are pervasive under majority rule. See McKelvey, 1979, "General Conditions for Voting Intransitivities in Formal Voting Models," *Econometrica* 47:1086-1112.

³¹ A classic discussion is Shepsle, 1979, "Institutional Arrangements and Equilibrium in Multidimensional Voting Models," *American Journal of Political Science* 23:27-59.

entrepreneurship upsets coalitions and leads to unstable decisions – to legislative actions that change from term-to-term, to shifting factional and regional alliances, and, in extreme cases, to the fall of governments. In chapter 9 of *Liberalism Against Populism*, Riker gives an example of how the actions of such policy entrepreneurs, attempting to build winning coalitions in an environment where the underlying preference structure was (on his argument) cyclic, lead to the fracturing of a political coalition that was holding the various regional factions together in the antebellum United States, contributing to the civil war.

Citizen leadership, in the sense articulated in this paper, is an important aspect of this process. Citizens who articulate new moral norms for their community are acting as policy entrepreneurs – without the abolitionists who articulated a strong moral critique of slavery, the instabilities that, on Riker's account, led to the civil war could not have happened. Engaging in a strong moral critique of the status quo, and offering new alternatives to the policies that are in place, rooted in such a moral critique, changes the agenda on the legislative plate. In terms of contemporary politics, consider the potential interaction between the efforts of radical Christians, such as adherents of the liberation theology movement, who are attempting to deploy Christian moral doctrine in favor of economic egalitarianism, and the current political alliance between U.S. Christians and economic conservatives within the Republican party: by strongly articulating a moral critique of conservative economic policy, those activists could undermine the existing coalitions in the United States and reconstitute the legislative agenda in just the sense that Riker discussed.

Having seen how citizen leadership can promote instability, it remains to be shown how that instability is a good, in that it promotes the virtues of states. But this, of course, is the role of SVT. If states that are less virtuous are, for that reason, more vulnerable to the instability-generating function of citizen leadership, then increasing the amount of citizen leadership should, on the whole, lead to

more change in non-virtuous states than in virtuous states.³² Of course, we cannot know which direction this change will lead – a non-virtuous state could easily be destabilized toward a less virtuous condition rather than a more virtuous one. But, unless there is an implausible tendency for destabilized states to go drastically toward the worse, the trend over time should be positive, because of SVT – that is, because states in just and democratic conditions will tend to last longer than states in unjust and undemocratic conditions. To see this, consider the way that natural selection optimizes a species. There is no guarantee that a given mutation will be optimal in the environment – mutations are created by a random process. But mutations that are worse for the species die off, while mutations that are good for the species persist and reproduce. *Mutatis mutandis* for social mutations brought about by instability: SVT is a principle of *the evolution of states*.

II.B.3. The risks of leadership?

The previous subsections made out the basic case for the proposition that citizen leadership can, regardless of the content of the normative principles advocated, bring about positive changes in imperfect societies. It is now time to address a triad of objections that might be raised.

First, one might worry that the argument I've offered in this section does not extend to citizens in societies that are already virtuous. In a very just democracy, we might think that there is an adequate level of electoral competition, that public preferences are already known, and that there is little or no room for historical progress to positively shape the ethical life of the community. In such a society, citizen leadership might not promote the virtues I've ascribed to it.

This worry need not detain us long, for we may simply answer "we'll cross that bridge when we come to it." This paper is consciously a work of non-ideal theory, meant to specify the virtues of citizens

³² One of the obvious ways that SVT works in the social choice context is that a just state is likely to have citizens whose preferences over public policy are similar, at least on the most important issues that might threaten the stability of their core institutions, because (returning to Rawls) they will all endorse the same principles regulating basic structure of the society, and know that they all endorse it.

living in democracies that are imperfectly just and imperfectly democratic. The virtues of citizens might be different in ideal societies: the idea that one's moral duties and virtues are different in ideal and non-ideal situations is no stranger to moral theory.

Second, one might worry that individual instances of citizen leadership might be so bad that they outweigh the good consequences. Suppose, for example, a citizen gets believes that African-Americans are morally inferior and should be treated as second-class citizens, and successfully wages a political campaign in favor of racist state policy. It is at least conceivable that citizen leadership could give rise to enough cases like that that it would have a generally harmful tendency for political communities, notwithstanding the formally positive features I've identified.

That second worry can never be completely eliminated, but there are some things that might be said against it. First, recall that the sort of citizen leadership under consideration is citizen leadership in good-faith pursuit of potentially shared ends. The shared ends proviso is designed to exclude many of those sorts of problems: in American society, a citizen cannot seriously and in good faith think that instituting racial apartheid would be a potentially shared end. At bottom my argument is limited to the proposition that *sane* exercises of political-moral practical reason are virtuous.

Still, we might think that public moral incompetence is severe enough that we ought not to create institutions to promote citizen leadership in general, because the bad kinds of citizen leadership (promoting injustice, say) might be so prevalent as to make matters worse off, all things considered, even after taking into account the good things that might come from citizen leadership.

The problem with that objection is that the risk of society moving toward injustice exists whether or not there is a lot of citizen leadership, and there is no reason to believe that an increase in citizen leadership would make it worse. That is, the state must make many decisions in order to function *as* a state: who counts as a citizen, how it should manage its affairs with other states, how its property rights are to be constructed, what is to be done with the poor, how its officials should be

chosen, what actions are to be criminalized, etc. *Someone* will make those decisions, if only by default, either with or without lots of citizen leadership. But by the no-expertise thesis, we cannot claim that that someone is any more likely to be better at getting the moral questions right than the general public. So even in a world with no citizen leaders (rule by technocrats or guardians), those who are making the decisions are just as likely to, e.g., reinstitute racial segregation than citizen leaders – or, at least, those who think that ordinary citizens are particularly prone to drive great injustices owes us an argument for this proposition, which has not yet been presented.

Related is the third worry. Even though the citizen leadership I defend here is limited to citizen leadership on *moral* questions, it might be that moral leadership leads to instrumental errors. For example, citizens might make moral appeals for more redistribution that will then cause the economy to collapse because of their instrumental failure to understand that redistribution leads (arguendo) to bad consequences. If the citizens had refrained from raising their moral problems with the economy as it was, the claim might go, everyone would have been better off, and had citizens been aware of the economic consequences of their moral claims, they would have chosen to accept the moral loss.

Let us distinguish two worries that the previous paragraph may be thought to be raising. It may be that the state of affairs after the disastrous economic redistribution is morally better, even though the citizens would have preferred it not come about had they known what they were doing. In that case, the answer is a simple appeal to the priority of the moral. “So much the worse for the preferences of the public,” we can just say, “citizen leadership has brought us to a morally preferred state of affairs, and so its consequences are, for that reason, acceptable.”

The sharper case is where the bad consequences, of which instrumentally incompetent citizens are unaware, are themselves morally relevant. Suppose that the poverty brought about by the redistribution is so bad that the state of affairs afterward is morally worse than before (people are starving in the streets, there are violent battles over clean water, etc.). And suppose that, were the

economists running things, this disaster would have been averted.³³ This is a story of good intentions run amok, and not one that the no-expertise thesis can address, since there is, we have acknowledged, instrumental expertise, and there are instrumental facts that are relevant to moral realities.

That poses a problem that requires me to make a small revision to the claim at issue. Citizen leadership on moral questions is virtuous, we should say, *only so long as they consult and listen to the experts available to them on the instrumental questions*. This really reflects an obvious truth that is buried in the notion of leadership I've given: it's a virtue, I'm claiming, for citizens to exercise practical reason on certain questions and in the public sphere. There are standards for the exercise of practical reason, and, at a minimum, good practical reasoning includes taking an appropriate amount of theoretical care. And because citizen leadership is superogatory rather than mandatory, this standard may be imposed on citizens without imposing unreasonably burdensome mandatory duties on citizens, even if consulting experts is costly. Citizens may either decline to lead, or, leading, undertake to lead responsibly.

Let us take stock of the argument so far. In section II.A, I defended the proposition that citizens are virtuous, qua citizens, to the extent they contribute to things that are virtuous for their states. In section II.B, I defended the claim that citizen leadership has a general tendency to bring about virtues of states, whether or not citizen-leaders are correct in the moral principles they advocate, by constituting a check on concentrated power (II.B.1) and by destabilizing vicious societies, leading to historical-ethical progress (II.B.2). Leadership, in my sense, may not be the only thing that brings about those happy results, but it is a thing that does so. It follows that citizen leadership is a virtue of citizens: the road has been long, but we have made it to claim (3) of the syllogism I presented in the introduction. In the remainder of the paper, I will discuss the implications of this claim for the conduct of *states*.

³³ If that's not the case, then we are again in a situation where the no-expertise thesis applies: citizens aren't doing any worse than the alternative.

III. States, Leaders, and Political Autonomy

III.A. Making the translation from individual virtue to civic virtue

As I said in the introduction, there is very little to actually do in order to complete the argument. It is, one might think, fairly obvious that if a behavior is a virtue of a citizen, then a state ought to allow citizens to do it. We would think less of a state that, for example, forbade its citizens to give to charity, because if a state ought to do anything, it ought to permit its citizens to live good lives, characterized by the virtues.

However, matters are not quite so simple. A politically liberal state might not be permitted to encourage its citizens to live virtuous lives generally, where those virtuous lives are specified by a controversial comprehensive doctrine. A politically liberal state, that is, may not take a position on what exactly the virtues are, that individuals should instantiate. But, we might think, if that's true, then we ought not to baldly state the principle that a politically liberal state should even *permit* its citizens to follow virtuous lives, for deciding what is, and is not, a virtuous life is no province of the state, and we cannot impose a principle on the state ("permit people to live virtuous lives") which requires, for its implementation, a judgment that politically liberal states cannot make. I'm not sure how serious this problem is – we might simply be able to say that politically liberal states must permit the *uncontroversial* virtues. But we need not make this sort of move.

Instead, recall that citizen leadership is a *civic* virtue, relating to citizens as such, in their role as citizens, and in virtue of the fact that civic leadership promotes the virtues of states. Whatever else we can say of states, even politically liberal states, surely we can (and must) say that states are obligated to promote those things that are their own virtues – states are obligated to promote their own justice, their own democracy, their own peacefulness, etc. They may not be permitted to promote those virtues by *coercion* (if, for example, the liberal constraints of permitting their citizens to pursue their own ends

make it impermissible for states to force citizens to spend all their time promoting state virtues). But they are certainly obligated to permit them, and to pursue them by non-coercive means.

If, then, as I've argued, citizen leadership promotes state virtues, it follows that states ought to permit, and non-coercively promote, it. Even if citizen leadership is not the *best* way to achieve state virtues, such that the state cannot justifiably expend limited resources promoting it rather than achieving greater justice, democracy, peacefulness, etc. by some more efficient means, it does not seem controversial to say that a state must *at least permit* its citizens to pursue even a second-best method of achieving its own virtue.

And with that very short step, we have reached the end of the main body of the argument. The remainder is application: I contend that the model of citizen leadership can ground some majoritarian intuitions about various objects of public policy, but without suffering from some of the objections to which majoritarian accounts are traditionally subject.

II.B. Applications of the leadership principle

As a citizen, I have the responsibility and the privilege to become an *author* of the ends of the community, if I so choose, and if I can convince my fellow citizens to agree. Such has been the topic of the previous sections. We can call this idea a *principle of political autonomy*, in the following sense: the ability to act as a citizen-leader is a necessary condition for citizens to have the powers that they ought to have in a democracy. A citizen is not fully a citizen – does not have the ability to manifest her will in the public sphere – to the extent she has been denied a democratic power that the state is obligated to give her, and, as has been argued above, the state is obligated to give its citizens the power to act as leaders.

Let us take a further step. Principles of political autonomy do not only constrain the state, they constrain our normative theorizing. A normative theory of the duties of citizens is false to the extent it counsels citizens to refrain from doing something that they ought to be encouraged to do (as a

superogatory duty), or to the extent it counsels citizens to do something that they ought not to do. This is trivial, but it is important to keep at the front of consciousness for the application of the leadership principle to accounts, like Rawls's idea of public reason, of the obligations of citizens in a democracy which may not be instantiated in law.³⁴

I now turn to precisely that application.

*II.B.1. Leadership and public reason*³⁵

Rawls's idea of public reason specifies the rules under which Rawls thinks citizens, living in a politically liberal society, ought to offer reasons to each other in order to justify the most important kinds of state policy (which Rawls specifies as "constitutional essentials and questions of basic justice"). In principle, it is separable from the other core ideas in political liberalism; for example, an overlapping consensus on a political conception of justice could exist without requiring citizens to follow the constraints of public reason by justifying constitutional essentials in terms of that conception.

The fundamental idea is that citizens are only to offer reasons to one another that they reasonably believe the others can reasonably accept. This excludes certain kinds of *controversial* reasons: citizens cannot offer the whole truth as they see it, as represented by their comprehensive doctrines (or even by controversial scientific claims), in public justification, because those claims are not reasonably acceptable to those who hold other reasonable comprehensive doctrines. The underlying idea here appears to be one of non-contradiction: because citizens recognize that other comprehensive doctrines are reasonable, they recognize that they can't reasonably demand that those who hold other comprehensive doctrines accept anything that is contrary to those doctrines.

³⁴ Rawls has repeatedly pointed out that the constraints of public reason are to be understood as moral constraints on citizen behavior, not as legal restrictions (e.g. *Political Liberalism*, pg. 213), presumably because implementing them as legal restrictions would violate the political liberties.

³⁵ This section is drawn from a work in progress that is tentatively entitled *Political Liberalism Without Public Reason*, which contains a full, and vastly more complex, critique of public reason. The instant paper was inspired by the need to build a normative superstructure to capture my intuitions about democratic participation and leadership that came out in the course of preparing that paper. The aim of this section is only to show how the principle of leadership might be deployed in such a critique, not, in any way whatsoever, to elaborate such a critique.

One way to criticize public reason, and an idea that it's easy to find intuitively compelling, is to criticize it as undemocratic. The whole point of the idea, after all, is that citizens, in a democracy, are not to attempt to bring it about that the state does what they think is right. That can't be democratically legitimate. On a sort of naive Rousseauianism, the idea of public reason stands directly in the way of the a citizens' manifesting their collective will through the state.

But the question is not so simple as all that. It turns out to be quite difficult to fill out the ideal of democracy to which the "public reason is antidemocratic" intuition appeals. For some defenders of public reason claim that it is an essential *part* of a good conception of democracy.³⁶

One approach is to embrace a majoritarian conception of democracy, one in which the ideal of democracy consists in the majority getting its way. But majoritarian conceptions, whether naïve (such as the notion that what matters is aggregating the brute preferences of the majority) or sophisticated, are subject to some serious criticisms. Common among those criticisms at least since the *Federalist Papers* is the worry that majoritarianism sanctions vicious majorities, who violate other democratic principles (such as those of inclusion), for example, by instituting racial segregation. This, of course, is a classic debate in the literature, one that highlights still unresolved tensions between liberalism and democracy as it is ordinarily conceived. The standard majoritarian line is to insist that democratic principles can give us warrant to reject some of the more outrageous possibilities for vicious majorities (like denying suffrage to their opponents), but to simply bite the bullet and acknowledge that democracy has costs, and that we might well have to accept some of the consequences of vicious majorities. Perhaps the strongest defense of the majority principle against this majority tyranny objection is a *tu quoque* offered by Dahl: if a system other than majority rule is in place, that system will empower some minority to inflict harm on the majority, and we might end up with *minority tyranny*

³⁶ I am again thinking here in particular Joshua Cohen's *Procedure and Substance in Deliberative Democracy*, which contrasts a deliberative conception of democracy, including the idea of public reason, with the more conventional majoritarian conception of the sort articulated by Dahl, Waldron, etc.

instead of majority tyranny.³⁷ But this isn't a very satisfying point: one thing we ought to demand out of our theories of democracy is that it avoids *all kinds* of tyranny.

The leadership principle shows us how we can do better. Because it only sanctions citizen leadership directed toward potentially shared ends, it avoids the most serious objection directed at the majoritarian conception: treating the authorship of potentially shared ends as the content of political autonomy, rather than the brute instantiation of the will of the majority, does not require us to accept outrageous but conceivable majority results like the reinstatement of racial segregation.

Yet the leadership principle still gives us some ground on which to begin a critique of public reason. First, note the difference between "shared ends," as I've described it, and public reason. The notion of a potentially shared end appeals to an ideal epistemic situation: something is a potentially shared if and only if a citizen can reasonably believe that, after an ideal persuasive process (something like Habermas's ideal speech situation), other citizens would agree with it. This is a much more lax standard than public reason, since public reason forbids citizens from contradicting the reasonable comprehensive doctrines that other citizens actually have, regardless of what they would believe under ideal epistemic circumstances.³⁸

To fill out this idea, suppose a citizen wishes to have the Ten Commandments taught in the public schools, and thinks the only sound argument for that position relies on the truth or falsity of particular religions. She further genuinely believes that the aforementioned truth or falsity is accessible to human reason: after an ideal persuasive process, her political opponents would agree.³⁹ Exercising practical reason in the public sphere in support of this end would, on my account, be an act of political

³⁷ Robert Dahl, *Democracy and its Critics*, pg. 155-6 (1989).

³⁸ To be more accurate, it forbids citizens from contradicting the reasonable comprehensive doctrines that could be expected to exist in a well-ordered society. But the two ideas probably amount to the same thing.

³⁹ Note that this is not inconsistent with a Rawlsian belief in the fact of reasonable pluralism. Reasonable pluralism may be true because an ideal persuasive process is impossible. Thus, a citizen may believe that reasonable pluralism is true now, but that an ideal persuasive process would lead to convergence of beliefs on moral questions.

autonomy.⁴⁰ Not so under public reason, since no public reasons are (ex hypothesi) available with which to justify such legislation.⁴¹ Thus, if we accept the shared ends conception of citizen leadership as giving content to the principle of political autonomy at the heart of democracy, we have somewhere to stand in offering a critique of the idea of public reason as antidemocratic. We can say, that is, that our account of what it means to be a democracy – one that is not subject to the objections that pose such a threat to majoritarianism – require that citizens be allowed to do this thing that public reason tells them they must not do. We do not have a full critique of public reason, for Rawls has argued that voters’ satisfying the constraints of public reason is necessary to serve a variety of other important goods of a political society.⁴² But we have a starting point, and one that the extant competing conceptions of democracy cannot provide.

II.B.II. Direct democracy as an instance of leadership

There is a long tradition of criticizing direct democracy (a term that I take to include institutions such as ballot initiatives and referenda, as well as “thicker” forms of direct governance as were practiced in the Athenian *polis*) on various of the familiar grounds used to criticize majoritarian institutions in general – that such institutions promote majority tyranny, that they pose a risk to the public good due to citizen incompetence, etc.⁴³ These are serious concerns: at the time of this writing, the voters of California are considering a ballot initiative, proposition 8, which is intended to strip gay and lesbian citizens of the state of marriage rights recently granted them by the state Supreme Court. The battle is

⁴⁰ This is not to say that such legislation would ultimately be permissible. The opponents of that legislation might rightly respond that the state should be neutral between religions, and, hopefully, they will prevail. But the *advocacy* of such legislation, in the public sphere, seems like a basic act of democratic agency.

⁴¹ I leave aside the fact that Rawls would have public reason limited to “constitutional essentials and questions of basic justice.” This is unimportant to the basic contrast, as equivalent questions could be found in that domain – consider, for example, objections to assisted suicide rooted in controversial moral doctrines. Judging from Rawls’s participation in the Philosophers’ Brief in *Washington v. Glucksberg*, I imagine that Rawls would count that as a constitutional essential or question of basic justice.

⁴² Rawls would also claim that public reason would permit the most important instances of leadership defended in this paper. I don’t think that’s true, but a discussion will have to be elsewhere.

⁴³ One particularly good example of this sort of critique is Barbara .S. Gamble, “Putting Civil Rights to a Popular Vote,” *American Journal of Political Science* 41:245-269 (1997), who offers empirical support for the worry about majority tyranny, as well as a useful literature review on the subject.

hotly contested, and the outcome is far from certain, though, as I write these sentences, it is feared the measure will pass⁴⁴ – a result that many who think seriously about these issues, at least since Hart defeated Devlin on the question of the legal enforcement of morality, find extremely worrisome. Yet majoritarians seem to be committed to supporting direct democratic institutions, and, once again, biting the bullet when it comes to measures like proposition 8, which do not appear to directly threaten democratic principles, but which are offensive to the liberal notions of individual rights and a robust private sphere of autonomy.⁴⁵

Can the idea of democratic leadership worked out here shed any light on this debate? I think it can, when we recognize the beneficial effects of measures like proposition 8. Even if it passes, the democracy benefits from the fact that citizens have been exposed to a vigorous debate, in the public sphere, on the issue – a debate that may, over time, lead to a more secure foundation for gay rights, as the defenders of those rights organize and engage in moral advocacy and appeals to the empathy of their fellow citizens. Moreover, the mere fact that citizens have this method of influencing public acts available to them may encourage more citizens to invest their time and resources in democratic leadership, as they may rightly expect more of a payoff from this investment, in the form of public policies reflecting their moral beliefs. Direct democracy *in general* seems to help us get the benefits of leadership.

These words will be cold comfort to those who would get married, but for the narrow-mindedness of the California electorate. But I have some warmer comfort: while the leadership principle sanctions direct democratic institutions in general, it does not sanction uses of direct democratic institutions that violate the proviso noted above about shared ends, and it is in just such

⁴⁴ By the time of presentation, this fear has been realized.

⁴⁵ Majoritarians do not speak with one voice on this question – some, such as Sherman J. Clark, “A Populist Critique of Direct Democracy,” *Harvard Law Review* 112:434-482 (1998), argue that direct democratic institutions are poor ways to capture what the majority will actually is. Modulo such technical critiques, however, majoritarians seem to me to be obligated to support institutions that can accurately convert popular demands into public policy.

cases where fundamental liberal principles are cast aside that the shared ends proviso is actually violated. Regardless of one's religious beliefs, one cannot, in good faith, believe that all citizens would agree – even after an ideal epistemic process – that marriage rights should be denied to gay citizens. We can say that cases like proposition 8 are situations where an institution that in general promotes a form of democratic goodness (citizen leadership) does not do so in the specific situation presented to it. Unlike majoritarianism, in other words, the leadership principle has the resources to criticize this sort of majority tyranny.

It is important to compare the Ten Commandments example from the previous section with the gay rights example of this one, as this will allow us to see the difference between the shared ends proviso and stronger principles like Rawlsian public reason, and yet how the shared ends proviso allows us to condemn actions of majority tyranny. It might appear that the two examples are identical: in each case, religiously motivated citizens, who believe that, if only the rest of the public would get their thinking right, they would see the truth of those religious beliefs, organize in order to have those beliefs codified as law. How can the shared ends proviso sanction one and condemn the other?

The idea at play here is one that has driven the notion of majority tyranny since its inception: the dangers of picking on discrete and insular minorities.⁴⁶ In the proposition 8 case, its advocates are not just trying to get the rest of the public to accept their religious principles as a fit subject for law, they are applying those religious principles to directly make life worse for people to whom they object. It might be reasonable to believe that, at the end of an ideal epistemic situation, all citizens could see the truth of the Ten Commandments and their suitability for enactment into law and teaching in the schools. It is much less reasonable to believe, in good faith, that, even after an ideal epistemic situation, gay citizens would agree that they are to be condemned to second-class status. The latter fight is *personal*.

⁴⁶ The term “discrete and insular minorities” is, of course, taken from Justice Stone's famous opinion in *United States v. Carolene Products Company*, 304 U.S. 144 (1938).

This can be a general principle: *laws targeting discrete and insular minorities can rarely or never meet the shared ends proviso*. And if that's true, the most important cases of majority tyranny are excluded from democratic sanction by the leadership principle, without thereby endorsing the full constraints of public reason.

II.B.III. The perennial problem of judicial review

Another focal point of the debate between majoritarians and other theorists of democracy is in the legitimacy of judicial review. Majoritarians often find themselves pressed to reject judicial review, in whole or in part, for reasons much like Waldron's.⁴⁷ Here, unlike in the previous two cases, the leadership principle does not support the majoritarian position. Judicial review does not threaten citizen leadership, and, in fact, judicial review might be a way to promote the ends of citizen leadership.

On the first of those claims: nothing in the argument for citizen leadership requires that this leadership lead, at the end of the entire political process, to actual law. Leadership that leads to legislation that is eventually ruled unconstitutional has already had the structural beneficial effects that I've noted above: elites have already been checked, agendas have already been expanded, and political coalitions have already been destabilized by the time some court has the chance to strike it down, at least in a system like the one in place in the United States, where the courts refuse to rule on "advisory questions." It is true that the prospect of judicial review might limit the incentives of citizens to invest in political activity that could be nullified by a court, but the day-to-day experience in the United States has not given us much reason to believe that political actors have any particular reluctance to promote unconstitutional legislation. And political activity does not stop when the courts speak: if the court's judgment does not match the ethical convictions of the community, politics continues, with the court decision representing another voice with which activists interact. *Roe v. Wade* is a chief example of this

⁴⁷ Jeremy Waldron, 2006, "The Core of the Case Against Judicial Review," *Yale Law Journal* 115:1346.

phenomenon: the fact that the court resolved the question of the legality of abortion has not kept abortion's opponents from organizing politically to change matters, nor its proponents from organizing in its defense.

On the second claim, it is notable that very prominent Supreme Court cases in the United States seem to become symbols of the development of our collective ethical life. Consider *Brown v. Board of Education*. The Supreme Court, at the behest of civil rights activists, participated in a social upheaval that fundamentally reshaped the political community. This is exactly the sort of thing that I've argued that citizen leadership can promote, and *Brown* shows us that one way for citizen leadership to do so is via the courts – citizen leadership in the form of litigation brought *Brown* to the Supreme Court, and then the decision came to represent, and to grant political authority and prestige to, the reforms of the civil rights era. Thus proceeds social progress.

IV. Conclusion

I have argued that leadership – morally rooted social advocacy – is a virtue of citizens, to be encouraged by the state and by our normative theory, and that this gives us some traction on some of the major debates in democratic theory. I would like to close with some intuition-pumping.

We have strong intuitions about the greatness of certain kinds of democratic leaders – from Pericles to Martin Luther King, Jr. – that are not tied to beliefs about the human good generally. Even a doctrinaire liberal (political or comprehensive), who rejects the use of ideals of human goodness as a ground for political judgment, can acknowledge the greatness of Rosa Parks, *qua citizen*. I might be wrong about this: I can merely appeal to an intuition that I believe will be shared by my readers, who were raised and educated as was I, in a democratic political culture. But this paper is in part an attempt to tap into our shared *endoxa* about our citizen leaders through history. Consider the following names:

Pericles, Demosthenes, Cicero, Cato the Elder, Martin Luther King Jr., Rosa Parks, Susan B. Anthony, Frederick Douglass, Mahatma Gandhi, Nelson Mandela, Cesar Chavez, and Thomas Paine.

If I may impose again on the *endoxa*, I think that we can see two things that all those leaders share. First, each had a transformative impact on the political community of which he or she was a member. Nothing was the same after each passed through.

Second, each was democratic, in the strong sense that each pursued ends that could be attributed to the political community as a whole, or were defensible to that political community. There is a difference between a leader who happens to be active in a democratically organized state and a democratic leader. Leaders like Savonarola and Robespierre, for example, can be excluded from the latter category because they turned on their own political communities. All of the people I've listed so represent something arising out of their political communities, something that a fellow-citizen could endorse just because he or she was a fellow citizen. These were not always uncontroversial ends – the civil rights movement, for example, was of course subject to bitter opposition from the political community – but they were always ends that the leader could ask, in good faith, the community as a whole to endorse – that is, they were not mere factional interests. In some cases (like the case of the civil rights activists) the community was *morally obliged* to endorse the ends pursued by the leader – and, in most cases, the community came to endorse those ends.

A corollary of the democraticness of these leaders is that their leadership was instantiated in the public use of practical reason. A democratic leader must lead by persuasion rather than force, and the transformative work of the leaders with whom we are concerned was carried out with words, and with other communicative interactions such as civil disobedience.

Both of those notions can be expressed in something like (a democratic version of) Hegelian terms. Like the Hegelian world-historical individual, they can be understood as advancing the ethical life of a political community by transforming its most fundamental assumptions and way of life. The leaders

of the American civil rights movement perhaps best exemplify this. The culture of American life was very different before and after the civil rights movement. Before, segregation was taken for granted, an unsurprising aspect of daily life. Today, it has become anathema. Our self-understanding as a political community has changed in fundamental ways as a result of the civil rights movement.

Moreover, Hegel tells us that the world-historical individual's aims arise out of the movement of spirit – his passions are not his own, but they give being to the development of spirit in history. There is an analogous feature of our democratic leaders: their aims arise not out of their own idiosyncratic or self-interested desires, but out of the development of the political culture of their communities – Rosa Parks appeared at a time when entire communities were looking for a lever to end segregation, and when the tensions between American ideals of equality and its practices of racial discrimination were at a high point – when the cognitive material existed in the political culture for the project of the civil rights movement. With Hegel, we can say that many of our leaders emerged from the tensions in the ethical life of the community.

I write, then, in praise of what we might call *the world-historical citizen*, and its exemplar is not Napoleon Bonaparte but Rosa Parks.