ANTHONY TAYLOR

STABILITY, AUTONOMY, AND THE FOUNDATIONS OF POLITICAL LIBERALISM

ABSTRACT. An attractive form of social stability is realized when the members of a well-ordered society give that society's organizing principles their free and reflective endorsement. However, many political philosophers are skeptical that there is any requirement to show that their principles would engender this kind of stability. This skepticism is at the root of a number of objections to political liberalism, since arguments for political liberalism often appeal to its ability to be stable in this way. The aim of this paper is to address skepticism about the stability condition by putting forward a novel defense of it. My defense builds on the claim that stable principles are necessary to secure the full autonomy of those who live under them.

Principles of political morality satisfy what I will call the *stability condition* if citizens of a society that was effectively ordered by them would come to comply with them freely in ideal conditions.[1] The question of whether principles of political morality must satisfy this condition has significant and far-reaching implications for theory development in political philosophy. If our principles must satisfy the stability condition, then those principles that would not be complied with in these ideal conditions – due to, say, motivational incapacities – ought to be rejected. To accept the condition is therefore to give claims about human nature a central place in political philosophy.[2]

The question of whether principles must satisfy the stability condition is also significant because an affirmative answer to it shaped much of the work of John Rawls. In *A Theory of Justice*, he

---

[1] What is needed to satisfy this condition will be elaborated in detail below.

[2] For a rejection of the view that claims about human nature should play this role, see David Estlund, 'Human Nature and the Limits (If Any) of Political Philosophy', *Philosophy & Public Affairs*, 39 (2011): 207–237.

aimed to show that his principles would satisfy the stability condition by arguing that the desire to act justly and the desire to live a good life would be congruent from the point of view of each citizen.[3] And later Rawls's commitment to the stability condition quite clearly guided the development of his view in *Political Liberalism*. Famously, it was his dissatisfaction with the original congruence argument for stability that lead him to make his view sensitive to the fact that any liberal society will be marked by reasonable pluralism.[4] Given this shaping influence that the stability condition had on Rawls's thought, it is difficult to take an overall position on his view, and especially on his now much discussed accounts of liberal legitimacy and public reason, without taking a position on it.

Given the significance of the question of whether principles of political morality must satisfy the stability condition, it is surprisingly difficult to find much by way of sustained argument in defense of it in the literature.[5] This would unsurprising if it were simply obvious that the soundness of our principles depended on them satisfying this condition, but this is not the case. Though the stability condition is consonant with a recent trend in political philosophy pressing for theories to be more realistic, the condition itself is neither obvious nor widely accepted. G.A. Cohen, for example, held that while stability is a condition of wise social choice, treating it as a constraint on principles of political morality would be absurd, for it would render the question of whether a just society would persist over time a conceptually confused one.[6] And Joseph Raz argued that while stability has some value, it is 'not the be-all and end-all of the theory of justice'.[7] More loosely, but perhaps more tellingly, few political philosophers writing today aim to show that the principles they favor could satisfy the stability condition, presumably because they do not think that this is necessary to establishing the soundness of those principles.

The aim of this paper is to offer a novel a defense of the stability condition. I begin by examining the suggestion that principles that

---

[3] John Rawls, *A Theory of Justice: Revised Edition*, (Cambridge, MA: The Belknap Press of Harvard University Press, 1999): 450–505.

[4] Rawls, *Political Liberalism: Expanded Edition*, (New York: Columbia University Press, 2005): xv–xvi.

[5] A noteworthy exception here is Larry Krasnoff, 'Consensus, Stability, and Normativity in Rawls' Political Liberalism', *The Journal of Philosophy* 95 (1998): 269–292.

[6] *Rescuing Justice and Equality*, (Cambridge, MA: Harvard University Press, 2008): 327–328.

[7] 'Facing Diversity: The Case of Epistemic Abstinence', *Philosophy & Public Affairs* 19 (1990): 16.

are stable in this way are necessary for the realization of a valuable form of autonomy: the political value of full autonomy.[8] To enjoy this kind of autonomy, citizens must freely endorse the social and political institutions that have played a significant role, via their upbringing, in shaping their character and self-conception. The stability condition is satisfied when citizens give their free and reflective endorsement to the principles that support the social and political institutions they have grown up under. When this is the case, citizens are autonomous in the sense that they do not experience their sense of justice as an alien imposition.[9]

Examining this thought, I will argue, first, that though the connection between stability and full autonomy is suggestive, it is incapable of delivering a compelling defense of the stability condition. A compelling defense of the stability condition needs to explain why the principles of political morality that would be stable in this way are the principles that we have most reason to accept. An autonomy-based defense of the stability condition could, therefore, only succeed if the principles that allow for the realization of full autonomy are not in conflict with the principles we have most reason to accept. This is, in essence, the claim that two aims we might attribute to political philosophy are in harmony: the aim of identifying the conditions under which we can live autonomously, and the aim of identifying the principles that we have most reason to accept. Though the truth of this harmony claim is far from obvious, I offer a defense of it in the latter part of the paper. Overall, this amounts to a provisional case for holding that the principles that

---

[8] Rawls, *Political Liberalism*, *supra* note 4 at 77. For an early but unelaborated account of the connection between autonomy and stability see Joshua Cohen's 'A More Democratic Liberalism', *University of Michigan Law Review* 92 (1994): 1516–1517, where he states that autonomy provides a reason for holding that stable social cooperation is a basic good. Matthew Clayton sets out the relationship between stable principles and autonomy in more detail in his *Justice and Legitimacy in Upbringing*, (Oxford: Oxford University Press, 2006): 11–19. Further, autonomy is also given a central place by Paul Weithman in his various discussions of political liberalism and the stability condition, though as we will see below his concern is ultimately somewhat different to my own. See his *Why Political Liberalism? On John Rawls's Political Turn* (New York: Oxford University Press, 2010); 'Citizenship, Reflective Endorsement and Political Autonomy', *The Modern Schoolman* 78 (2001): 135–150; 'Convergence and Political Autonomy', *Public Affairs Quarterly* 25 (2011): 327–348; 'Stability and the Original Position from *Theory* to *Political Liberalism*', in *The Original Position*, edited by Timothy Hinton, (Cambridge, UK: Cambridge University Press, 2015): 237–238; 'Autonomy and Disagreement about Justice in *Political Liberalism*', *Ethics* 128 (2017): 95–122; and 'In Defense of a Political Liberalism', *Philosophy & Public Affairs* 45 (2017): 397–412.

[9] I will elaborate on this idea in much more detail in §III below. For ease of exposition, I refer to this value as 'full autonomy', though I always mean autonomy understood as a distinctly political value.

satisfy the stability condition are, in fact, the principles we have most reason to accept.

I think that this defense of the stability condition has an important role to play in the debate surrounding Rawls's political liberalism. Proponents of political liberalism often appeal to the stability condition at crucial junctures in their defenses of the view.[10] But the extent to which this appeal is persuasive is dependent on the defensibility of the condition. If the stability condition can be defended in the manner I suggest here, then this may be a significant victory for defenders of political liberalism. But the argument is also of wider importance. For if the case I offer for the stability condition is sound, then even if proponents of political liberalism are wrong to claim that their principles satisfy the condition, the question of what principles *do* satisfy it is an incredibly significant one.

The paper is structured as follows. Section I gives a sketch of political liberalism, paying particular attention to how the view is motivated by the aim of satisfying the stability condition. Section II then sets out in detail what satisfying the condition requires, and introduces the central challenge that I aim to address: why must principles of political morality be stable in this way? Section III explores the relationship between full autonomy and stability. In section IV, I present my defense of the stability condition that builds on this relationship. Section V concludes.

<div align="center">I.</div>

The best way to introduce the stability condition is through a brief discussion of its role in Rawls's thought. One of Rawls's aims in *Theory* was to show that his conception of justice, justice as fairness, would enjoy a special kind of stability.[11] In order to satisfy this stability condition, he first needed to show that the citizens of a society that was well-ordered according to justice as fairness would acquire a sense of justice informed by its principles. This he aimed to do by offering an account of how the desire to act justly would be

---

[10] Burton Dreben, 'On Rawls and Political Liberalism', in *The Cambridge Companion to Rawls*, edited by Samuel Freeman, (New York, NY: Cambridge University Press, 2003): 316–345; Samuel Freeman, *Justice and the Social Contract: Essays on Rawlsian Political Philosophy*, (New York, NY: Oxford University Press, 2007): 175–213; Jonathan Quong, *Liberalism Without Perfection*, (New York, NY: Oxford University Press, 2011): 158–160; Paul Weithman, *Why Political Liberalism?*, *supra* note 8.

[11] A task taken up in part III of that book.

acquired in accordance with three psychological principles.[12] Having shown that citizens would acquire the desire to be just in this way, he needed to show further that they would give this desire sufficient weight in their practical reasoning that it would not be consistently trumped by their other desires. This he aimed to do by arguing that the desire to be just would be a part of each citizen's good. In the conditions of a well-ordered society, the desire to act justly (the right) and the desire to live a good life (the good) would be congruent from the point of view of each citizen.[13] If this argument had succeeded, it would have shown that justice as fairness satisfies the stability condition: citizens who lived under it would not simply comply with its principles due to the threat of punishment, they would freely affirm those principles and comply with them because doing so is part of their good.

The next step in this story is that the argument Rawls made for the congruence of the right and good is incompatible with the fact of reasonable pluralism: the fact that 'a plurality of reasonable yet incompatible comprehensive doctrines is the normal result of the exercise of human reason within the framework of the free institutions of a constitutional democratic regime'.[14] Given that justice as fairness is a liberal democratic conception of justice, this incompatibility is fatal. The argument for stability had supposed that citizens of a society well-ordered by justice as fairness shared a partially comprehensive doctrine, and that this provided a key component of their basis for seeing their sense of justice as a part of their good. This partially comprehensive doctrine was, broadly speaking, Kantian. Every member of the well-ordered society was supposed to share a desire to express their nature as a free and equal rational being. And to express your nature as a free and equal rational being is, Rawls states 'to act on the principles that would be chosen if this nature were the decisive determining element'.[15] Since the choice from the original position is designed such that our nature as free and equal rational beings is the decisive determining element of the choice of the principles of justice, the fact that the citizens of the

---

[12] *Theory*, *supra* note 3 at 397–449.

[13] Ibid. 450–514. For discussion, see Samuel Freeman's 'Congruence and the Good of Justice', in *The Cambridge Companion to Rawls*, *supra* note 10 at 277–315.

[14] *Political Liberalism*, *supra* note 4 at xvi.

[15] *Theory*, *supra* note 3 at 222.

well-ordered society share this desire to express their nature plays a key role in securing the stability of justice as fairness in the argument of *Theory*.[16] In short, since this desire is a desire to act on the principles that would be chosen in the original position, it thereby amounts to a desire to act on the principles of justice as fairness.[17] In coming to think that the well-ordered society would inevitably be inhabited by members adhering to a plurality of reasonable comprehensive doctrines, Rawls came to doubt that they would all share the desire to express their nature in this way. And to doubt this is to doubt that Rawls's argument for the congruence of the right and the good succeeds. The goal of showing that justice as fairness satisfies the stability condition will not have been achieved after all.

It is at this point that political liberalism enters the picture. The central aim of political liberalism is to show that a well-ordered society that contains a plurality of reasonable comprehensive doctrines could nonetheless satisfy the stability condition.[18] To realize this possibility the conception of justice that orders such a society must, Rawls argues, be a *political* one. Distinct from comprehensive doctrines – which include beliefs and ideals that pertain to all of life – political conceptions of justice are concerned with a more limited range of questions: those concerning constitutional essentials and matters of basic justice.[19] These are, roughly, questions about what rights and liberties should be included in a written constitution, and questions about what principles should regulate major social institutions.[20] A conception of justice that is cast as political rather than comprehensive can avoid conflicting with the plurality of comprehensive doctrines that would be held in a well-ordered liberal society, making it possible for proponents of those comprehensive doctrines to endorse it from their varied perspectives.[21] Given the fact that the well-ordered society will necessarily contain a plurality of reasonable comprehensive doctrines, the stability condition can no longer be satisfied by showing that all of the citizens of that society will have a shared desire to express their nature as free and equal

---

[16] For the part of Rawls's congruence argument that depends on this desire, see Ibid., §86, especially 501–503.

[17] Weithman, *Why Political Liberalism?*, *supra* note 8 at 118.

[18] *Political Liberalism*, *supra* note 4 at xviii.

[19] Ibid. xlviii.

[20] Ibid. 11–12.

[21] Ibid. 482–483.

rational beings. Rawls argues instead that stability can be secured by showing that principles of justice could be the subject of an over-lapping consensus. Such a consensus holds when citizens adhering to different comprehensive doctrines can, despite their other disagree-ments, agree on a political conception of justice for evaluating the major institutions of their society. A distinctly political conception of justice facilitates this form of consensus, as it is acts as a 'module [...] that fits into and can be supported by [the] various reasonable comprehensive doctrines that endure in the society regulated by it'.[22]

Since a political conception is not presented as, or derived from, any particular comprehensive doctrine, it 'must contain its own intrinsic normative and moral ideal'. For political liberalism, this is the criterion of reciprocity:

> Citizens are reasonable when, viewing one another as free and equal in a system of social cooperation over generations, they are prepared to offer one another fair terms of social cooperation [...] and they agree to act on those terms, even at the cost of their own interests in particular situations, provided that others also accept those terms. For these terms to be fair terms, citizens offering them must reasonably think that those citizens to whom such terms are offered might also reasonably accept them [...] not as dominated or manipulated, or under the pressure of an inferior political or social position.[23]

This moral ideal gives rise to the two most discussed features of political liberalism: the liberal principle of legitimacy and the duty of civility. The liberal principle of legitimacy states that the exercise of political power is legitimate only when it is 'exercised in accordance with a constitution the essentials of which all citizens as free and equal may reasonably be expected to endorse'.[24] The duty of civility, which falls on individuals, is a moral duty to 'be able to explain to one another on those fundamental questions how the principles and polices they advocate and vote for can be supported by the political values of public reason'.[25]

A justification for the exercise of political power can only satisfy the criterion of reciprocity, political liberalism holds, if it is drawn from a political conception of justice that is also liberal. Liberal conceptions have three features: they specify 'certain rights, liberties, and opportunities (of a kind familiar from democratic regimes)'; they give 'a special priority' to these liberties over demands to further the

---

[22] Ibid. 12.
[23] Ibid.
[24] Ibid. 137.
[25] Ibid. 217.

general good; and they assure all citizens adequate all-purpose means to make use of their liberties and opportunities.[26] There are thus numerous liberal conceptions of justice, of which justice as fairness is one.

It is these features – its status as political rather than comprehensive and its intrinsic moral ideal – that distinguish political liberalism from rival liberal views and rival conceptions of political morality in general. And they are, of course, controversial. They rule out appeals to conceptions of the good life and metaphysical doctrines as potential justifications for the exercise of political power. They also rule out appeals to variants of utilitarianism, which would not give a guaranteed special priority to a set of rights and liberties over the requirement to promote utility. And they rule out appeals to views, such as Robert Nozick's libertarianism, which would not guarantee all citizens sufficient resources to make use of their liberties.[27]

Why should we accept these restrictions? What reasons do we have to restrict the scope of our conception of justice to political questions, and to eschew appeals to comprehensive doctrines as justifications for the exercise of political power? An answer to this question that is accepted in some form by Rawls and a number of his defenders is this: only a political liberal conception of justice can satisfy the stability condition.[28] Given the fact of reasonable pluralism, no conception of justice based on a comprehensive doctrine could enjoy the special kind of stability that a political liberal conception can, which Rawls calls 'stability for the right reasons'.[29]

There are two kinds of concern we might have about this answer. First, we might wonder why the aim of showing that principles of justice would satisfy the stability condition is such an important one. What is wrong with principles that fail to satisfy this condition? In pressing this concern, we need not hold that stability is entirely devoid of value – instead, we can simply ask why a failure to realize

---

[26] Ibid. 6.

[27] *Anarchy, State, and Utopia*, (New York, NY: Basic Books, 1974).

[28] Dreben, 'On Rawls and Political Liberalism', *supra* note 10; Freeman, *Justice and the Social Contract*, *supra* note 10 at 175–213; Quong, *Liberalism without Perfection*, *supra* note 10 at 158–159; and Weithman, *Why Political Liberalism?*, *supra* note 8. I do not mean to imply that they take an appeal to stability to provide a complete answer to the question of why we ought to accept political liberalism, only that they take it to be a key part of the story.

[29] On this kind of stability, see *Political Liberalism*, *supra* note 4 at 140–144. I examine Rawls's stability condition in detail in the next section.

it counts decisively against a conception of justice. Call this challenge *stability-skepticism*. Second, we might question the claim that a political liberal conception of justice, and only a political liberal one, can enjoy this kind of stability. That is, even if we accept the fact of reasonable pluralism, we may wonder whether the contours of this idea fall where political liberalism takes them to: perhaps the principles that actually satisfy the stability condition are quite different from those Rawls defends. Call this challenge *content-skepticism*.

Since the distinction between these two challenges to the argument from stability to political liberalism is central to this paper, let me expand on it further. We can see the distinction most clearly in two contrasting responses to the following question: is it true that a conception of political morality that is distinctly *political* and *liberal*, in the way just described, would come to be the subject of an overlapping consensus in the well-ordered society? The content-skeptic thinks that Rawls and his followers have offered insufficient grounds for answering this question in the affirmative. Why not hold that a different set of principles would be the subject of this consensus? Or, indeed, that there are no principles that would be? The stability-skeptic, by contrast, has a different attitude. Even if we accept that there would be an overlapping consensus on these principles, they ask why this fact *matters*. Why does this fact speak in favor of political liberal principles? In short, they ask why we should accept the stability condition.

In this paper I will only be concerned with stability-skepticism, and I will set content-skepticism aside. Though the question of whether content-skepticism can also be answered is vital for a full defense of the argument from stability to political liberalism, the question raised by stability-skepticism is foundational. There is little point in investigating whether political liberal principles satisfy the stability condition if we end up rejecting the condition entirely.

II.

In the next section I will consider a response to stability-skepticism that goes via an appeal to a conception of autonomy. Before I can set out this argument, however, I need to say more than I have done so far about the stability condition. In order to see how an appeal to

autonomy might answer stability-skepticism, we first need to see more precisely what such stability consists in.

The first component of the stability condition that requires further elucidation is the idea of a well-ordered society. A well-ordered society is one whose institutions are effectively regulated by a conception of justice and are recognized as just by the citizens that populate it. Every citizen of a well-ordered society accepts the conception of justice that effectively regulates their society and knows that each of her co-citizens also accepts it. A well-ordered society also satisfies a full publicity condition: the entirety of the case for the conception of justice that orders that society – the beliefs and modes of reasoning that support it – are known by every citizen.[30] The idea of a well-ordered society is thus 'plainly a very considerable idealization'.[31]

As I noted above, establishing that a society well-ordered by particular principles would satisfy the stability condition is a two-step process. The first step is to establish that the citizens of that society would acquire the desire to act justly – a sense of justice – as part of their upbringing. Rawls takes himself to establish this by providing a psychological account of the process by which such a desire would be acquired.[32] The second step is to show that the citizens of this well-ordered society would give their sense of justice sufficient weight in their reasoning, such that it would not consistently lose out to their other desires when they are deciding how to act. This second step is necessary because showing that the citizens of a well-ordered society would acquire a sense of justice does not suffice to show that they will be disposed to act justly in the long run. If they find that their sense of justice is often in conflict with the pursuit of their other interests, then they may come to resent the desire to act justly and take steps to rid themselves of it.[33] If this were the case, then the stability established by the first step would, over time, be undone by citizens' reflective attitude toward their desire to act justly.

---

[30] *Political Liberalism*, *supra* note 4 at 66–71.

[31] Rawls, *Justice as Fairness: A Restatement*, (Cambridge, MA: The Belknap Press of Harvard University Press, 2001): 9.

[32] *Theory*, *supra* note 3 at 397–434.

[33] Ibid. 295, 451.

The second step is therefore to show that citizens of the well-ordered society would not have this attitude toward their sense of justice. In order to achieve this, it takes them to be considering the question of what place to give their sense of justice in their life from a particular point of view: a point of view from which they are bracketing their desire to be just for its own sake and considering only their other desires.[34] From this perspective, they are to ask themselves whether they have sufficient reasons to preserve and encourage their sense of justice.

This second step therefore requires a theory of justice to, in Rawls's words, 'supply other descriptions of what the sense of justice is a desire for' and use these to show that a person bracketing their desire to be just for its own sake would still 'confirm this sentiment as regulative of his plan of life'.[35] The reason the question must be considered from this particular bracketed perspective is that, given the first step has been completed, these citizens have already acquired an effective sense of justice. If they were to consider the question of what place to give their sense of justice in their life while examining all of their desires, it would be trivially true that they would choose to affirm it.[36] Given that they have an effective sense of justice, they want to act justly. Therefore, examining all of their desires would lead them to conclude that they ought to preserve and maintain this desire. Their situation is in this way analogous to that of a participant in a loving relationship who desires above all else to act for the benefit of their loved one. Suppose we were to ask such a person whether they ought to take steps to rid themselves of this desire. Considered in one way, this question would trivially be answered 'no' – it is what they desire, after all. But they could also reflect on it considering only their other desires. From this perspective, they might find that this desire is quite bad for them, even ruinous. And if they did find this, then there would be a sense in which it is rational for them to take steps to rid themselves of it: in spite of the pull that it has over them, it is detrimental to their good. Citizens of the well-ordered society are in an analogous situation. Though they want to act justly, when they reflect on their desire from this bracketed perspective they may find that it is detrimental

---

[34] Ibid. 499.

[35] Ibid.

[36] As Rawls notes: Ibid. 498.

to their good. However, if the stability condition's second step is satisfied, they will find that, in fact, they have sufficient reasons to preserve and encourage their sense of justice. When this second step is satisfied, then, they have an effective sense of justice that they *reflectively endorse*: they want to act justly above all else – they can see no good reasons to attempt to rid themselves of this desire.

Beyond the satisfaction of these two steps, the stability condition has one final and important feature. For all I have said so far, the desires needed to complete the second step could come about solely in response to the threat of punishment. The citizens of the well-ordered society could desire to act justly above all else due to the presence of an effective coercive power that threatens to punish them severely for acts of injustice. However, principles for which the second step was satisfied in this way would not satisfy the stability condition. In order for the condition to be satisfied the aforementioned desires must come about *freely*. This means that the desires that stabilize the conception must be come about solely via the educative effects of growing up in a society well-ordered according to its principles.[37] If other forces such as the threat of punishment are required in order to bring about these desires, then the principles do not satisfy the stability condition.[38] In this sense, when principles satisfy the stability condition they are ones that citizens of a well-ordered society want to follow on the basis of the free exercise of their practical reason alone.[39] I will label this feature of stability the *freedom condition*, as it will be a focus of my discussion in Section IV.

We can therefore summarize the stability condition as follows: principles of political morality satisfy the stability condition when the citizens of a society well-ordered by these principles would freely and unanimously give them their reflective endorsement.

---

[37] *Theory*, *supra* note 3 at 401. Weithman provides an illuminating account of this distinction between free and forced stability, *Why Political Liberalism?*, *supra* note 8 at 43–51.

[38] Rawls does, however, think that any well-ordered society will need a penal system to solve its mutual assurance problem. Such a problem occurs because, he supposes, the desire to act justly is conditional on sufficient compliance by others. Therefore, if a citizen lacks confident assurance of her co-citizens' motives, she may not act justly despite her desire to do so. Here 'the existence of effective penal machinery serves as men's security to one another' (*Theory*, *supra* note 3 at 211). What is important to note here is that the penal system is not required to ensure that citizens have the desire to act justly in the first place, it is only required to ensure that citizens who want to act justly in fact do so.

[39] It is this idea of free acceptance that is central to stability for right reasons. See *Political Liberalism*, *supra* note 4 at 142–143.

III.

Now that I have set out in detail what it takes to satisfy the stability condition, it should be clear that it stands in need of defense. After all, it is not self-evident that principles that fail to satisfy this condition ought to be rejected. An alternative response one might have to the conclusion that their favored principles would not satisfy the stability condition is to hold that the fault lies not with those principles, but *with the citizens who reject them*. Perhaps the principles that we have most reason to accept are simply not capable of commanding free and unanimous acceptance in the conditions of a well-ordered society.

The thought that principles of political morality need not be suitable for public knowledge and acceptance is not novel. For one, it undergirds Henry Sidgwick's well-known endorsement of utilitarianism as an esoteric morality. Sidgwick held that since ordinary citizens would be likely to make erroneous utility calculations, they should be taught to follow a simpler set of moral rules than the principle of utility. On his view, utilitarianism itself ought to be kept comparatively secret, with knowledge of it confined to an 'enlightened few'.[40] Utilitarianism is unsuitable for public acceptance on Sidgwick's view, but this fact is not taken to be damning of it – it is only damning of ordinary citizens' ability understand and consistently act on the principle of utility. Plato's view in the *Republic* is similarly that the truth about political morality is unsuitable for public knowledge and acceptance. Ordinary citizens could not be persuaded that philosophers ought to rule because they are philosophers. Plato thus recommends the use of myths that aim to persuade ordinary citizens to accept the rule of philosophers. As Larry Krasnoff writes of Plato's view, since ordinary citizens 'cannot be brought to believe that members of the highest-class ought to rule because they are philosophers, they must be brought to believe that it is the gold in their veins that entitles [them] to rule'.[41] I do not mention these two views in order to endorse them, but rather to make clear that stability-skepticism cannot be dismissed without argument. The claim that principles of political morality ought to be

---

[40] *The Methods of Ethics*, (Indianapolis, IN: Hackett, 1981[1907]): 489.
[41] 'Consensus, Stability, and Normativity', *The Journal of Philosophy* 95 (1998): 269–292.

rejected if they would not satisfy the stability condition is a controversial one that stands in need of defense.[42]

What I will call the argument from full autonomy aims to provide precisely this defense. Here is its basic idea: when principles satisfy the stability condition the citizens who live under them enjoy a valuable form of autonomy. This form of autonomy consists in the fact that the principles that constrain them are not externally imposed constraints on their lives but are rather constraints they would choose to endorse in conditions of freedom. Principles that are stable are thus ones that, in Thomas Nagel's words, 'realize some of the values of voluntary participation, in a system of institutions that is unavoidably compulsory'.[43]

To expand on this basic idea, we must consider the relationship between stability and autonomy in more detail. A person is autonomous, so the well-worn analogy goes, when they are the author of their own life.[44] Though there are numerous conceptions of what is required to live autonomously, it is generally agreed that a person fails to be autonomous when their life is the product of alien or external forces. A pervasive and powerful force that acts upon each of our lives is our upbringing in a particular social and political world. Or, to put it another way, the social and political institutions that we grow up under exert a significant *shaping influence* on us.[45] If as adults we come, after a process of rational reflection, to reject the underlying principles that regulate our social and political institutions, then we will thereby be rejecting a central force that has operated to make us who we are. If we find ourselves in this situation, we can aptly be described as lacking a degree or component of autonomy. What is marked out by the stability condition is a situation in which citizens do not find their lives to have been shaped by alien forces in this way. When principles of political morality are stable, the citizens who grow up under them come to give them their free and reflective endorsement. This means that in addition to

---

[42] For a more recent rejection of the claim that principles of political morality must be suitable for public knowledge and acceptance see Katarzyna De Lazari Radek and Peter Singer, 'Secrecy in Consequentialism: A Defence of Esoteric Morality', *Ratio* 23 (2010): 34–58.

[43] *Equality and Partiality*, (New York, NY: Oxford University Press, 1991): 36.

[44] The source of this metaphor is Joseph Raz, *The Morality of Freedom* (Oxford: Clarendon, 1986): 386.

[45] The idea that social and political institutions play a central role in shaping our motivations and self-conception is an important theme of Rousseau's political philosophy. See, for example, the discussion in Joshua Cohen's 'Reflections on Rousseau: Autonomy and Democracy', *Philosophy & Public Affairs* 15 (1986): 275–297.

accepting these principles, they accept the influences that their social and political institutions have had on them.[46] When the stability condition holds, citizens of the well-ordered society thereby experience a key aspect of themselves – their sense of justice – as a product of their own will rather than of alien forces.[47]

The autonomy that citizens of a stable well-ordered society enjoy is what Rawls calls the political value of full autonomy. Citizens realize this value 'in their recognition and informed application of the principles of justice in their political life […] as their effective sense of justice directs'.[48] In explaining the value of full autonomy, Rawls is characteristically at pains to emphasize that it is to be understood as a political value, not an ethical one; stating it is realized 'in public life by affirming the political principles of justice and enjoying the protections of basic rights and liberties' and also 'by participating in society's public affairs and sharing in its collective self-determination over time'.[49]

What conditions must be satisfied for a citizen to be autonomous in this way? Rawls's discussion suggests the following set of conditions, which draw on various features of his view.[50] The first condition should already be clear: to be autonomous, citizens must accept the principles of political morality that order their society and have thus played a role in shaping them into who they are. The second condition is that the principles that citizens accept and act on must be *fully public*. This means that the full justification for the principles of justice must be publicly available to citizens – the principles cannot be an esoteric morality as in the examples of Plato and Sidgwick above. The third condition is that the principles of justice that citizens accept and act on must have the right *content*: for

---

[46] When the stability condition is satisfied by an overlapping consensus of reasonable comprehensive doctrines some citizens see the political conception of justice as merely 'not in conflict with' their other values (*Political Liberalism, supra* note 4 at 140). But even for these citizens, the shaping influence that the political conception has had on them is not an alien imposition, for, *ex hypothesi,* they do not see themselves as having sufficient reason to reject their sense of justice.

[47] Indeed, when Rawls begins to lay out the stability argument in *Theory* one of his central concerns is with the possibility that citizens will come to reject their sense of justice because of their realization that it is 'largely shaped and accounted for by the contingencies of early childhood' (451) and he writes in response that a stable well-ordered society is one in which 'no one's moral convictions are the result of coercive indoctrination' (452).

[48] *Political Liberalism, supra* note 4 at 77.

[49] Ibid. 77–78.

[50] Here I follow the reconstruction of these conditions in Weithman, 'Autonomy and Disagreement about Justice in *Political Liberalism*', *supra* note 8 at 102–105.

citizens to be autonomous these principles must make it possible for them to enjoy the protection of rights and liberties. As we saw earlier, this is achieved in Rawls's theory by the requirement that liberal conceptions of justice both prioritize a set of rights and liberties and ensure all citizens have adequate all-purpose means to make use of those rights and liberties. The third condition appeals to Rawls's idea of *public reason*. For citizens to be autonomous, when public officials make decisions and advocate for policies that bear on constitutional essentials and matters of basic justice, they must do so based on public reasons. Similarly, when citizens debate or vote on such measures in the public political forum they must be 'prepared to show in due course that [they] can be supported by [public] reasons'.[51] The fourth and final condition is that for citizens to be autonomous the principles of justice that they accept and act on must 'specify the fair terms of cooperation they would give to themselves when fairly represented as free and equal persons'.[52] This condition stipulates that citizens must accept and act on principles that would be chosen in the original position, as that is a choice situation in which they are fairly represented as free equals. Another way to put this is to say that for citizens to be autonomous the principles that they accept and act on must have the right *source*: they must be principles they would give themselves in the original position.

Together this set of conditions ensures that when the citizens of the well-ordered society accept the principles of justice that have shaped them into who they are, they do so not out of ignorance or duress. They enjoy the full protection of a set of rights and liberties, and the full justification for the political decisions that constrain their actions is available to them. Their decision to accept and act on these principles can therefore aptly be described as autonomous.

Confronted with this conception of autonomy, some may respond by asking why all these conditions must be satisfied for citizens to live fully autonomously. Suppose a citizen enjoys the secure protection of rights and liberties and makes choices about how to

---

[51] *Political Liberalism*, *supra* note 4 at 444–445.

[52] Ibid., 77. This is what Weithman calls the 'collective self-legislation condition' ('Autonomy and Disagreement about Justice in *Political Liberalism*', *supra* note 8 at 102).

live her life from an adequate range of options, while free from the influence of coercion and manipulation.[53] Even if this person rejects the principles of justice that regulate her society, could she not nonetheless be described as fully autonomous? This is an important question, but by drawing on the preceding discussion I believe it can be given a powerful response. To see this, consider the follow passage from *Political Liberalism*:

> The government's authority [cannot] be freely accepted in the sense that the bonds of society and culture, of history and social place of origin, begin so early to shape our life and are normally so strong that the right of emigration (suitably qualified) does not suffice to make accepting its authority free, politically speaking, in the way that liberty of conscience suffices to make accepting ecclesiastical authority free, politically speaking. Nevertheless, we may over the course of life come freely to accept, as the outcome of reflective thought and reasoned judgment, the ideals, principles, and standards that specify our basic rights and liberties, and effectively guide and moderate the political power to which we are subject. *This is the outer limit of our freedom.*[54]

When it comes to accepting the authority of an association like the church, the secure protection of rights and liberties, freedom from coercion and manipulation, and an adequate range of options from which to choose may suffice to render this acceptance free. But when it comes to accepting political authority – which, as we have seen, exerts a significant shaping influence on us from early in our lives – these weaker conditions do not suffice to render our acceptance fully free. Those who reject the ideals and principles that are used to argue for and specify political decisions in their society are rejecting a central influence that has acted on them to shape them into who they are. They are not, therefore, at the outer limit of their freedom until they can come to give these ideals and principles their reflective endorsement. It is for this reason that the complete set of conditions set out above is necessary for citizens to be fully autonomous.

With the conditions of full autonomy before us, we are now able to set out the defense of the stability condition that appeals to this idea. To fix ideas, we can begin from the argument for political liberalism that I set out earlier: that we ought to endorse political liberal principles because of their ability to satisfy the stability condition, given the fact of reasonable pluralism. The challenge for this line of reasoning that we are hoping to address is stability-skepticism. Why must we accept principles that could satisfy the stability con-

---

[53] These conditions are central to Raz's conception of personal autonomy in *The Morality of Freedom*, *supra* note 44 at 369–399.

[54] *Political Liberalism*, *supra* note 4 at 222. Emphasis added.

dition? Is it not, after all, a live possibility that the correct principles of political morality are simply not capable of securing this kind of consensus?

The argument from full autonomy supplements the original line of reasoning with two further premises. The first is that principles that satisfy the stability condition are necessary for citizens of the well-ordered society to enjoy full autonomy. The second is that we ought to accept the principles that allow citizens of the well-ordered society to enjoy full autonomy. The argument runs as follows:

(1) Only a political liberalism can satisfy the stability condition, given the fact of reasonable pluralism.
(2) Principles that satisfy the stability condition are necessary for the citizens of the well-ordered society to enjoy full autonomy.
(3) We ought to accept the principles that allow the citizens of the well-ordered society to enjoy full autonomy.
(4) Therefore, we ought to accept a political liberalism.[55]

Before evaluating the premises of this argument, let me first consider a more general worry that some may have about it. Some may suspect that, due to political liberalism's concern with reasonable pluralism, any autonomy-based defense of the view will be a non-starter. They might think, that is, that when Rawls acknowledged that any society well-ordered by liberal principles would be marked by adherence to a plurality of reasonable comprehensive doctrines, he thereby acknowledged that there is no conception of autonomy that all reasonable citizens could be expected to accept. As Jonathan Quong puts the point, when considering the possibility of autonomy-based defense of political liberalism: since 'it would not be unreasonable for some citizens to reject this account of autonomy […] this

---

[55] Clayton sets out an argument from autonomy to political liberalism of essentially this form in *Justice and Legitimacy in Upbringing, supra* note 8 at 11–19. My argument in what follows will go beyond his analysis by addressing an important challenge to premise (3) that he does not consider. Though Weithman has given a central place to autonomy in his writings on political liberalism, he does not endorse this argument. In the paper where he writes that 'Rawls's principles of legitimacy and public reason are grounded on his commitment to political autonomy' he does not appeal to autonomy to address stability-skepticism. Rather, when it comes to defending the stability condition, he argues that terms of social cooperation must play a social role – they must 'provide an enduring public basis for justifying the distribution of the benefits and burdens of social cooperation'. And he suggests 'they can play that role only if they are freely accepted over time by those who live under them'. ('In Defense of a Political Liberalism', *supra* note 8 at 398, 410–411). This is a quite a different route to addressing stability-skepticism, and one that I think is not fully satisfactory. The deepest and most challenging form of that skepticism comes, I believe, from those like Sidgwick who would deny that principles must play this social role.

account of public reason's basis would be unstable in a well-ordered society'.[56] The thought here is that since all conceptions of autonomy are subject to reasonable disagreement, none can be the basis of political liberalism.

This objection depends on the claim that the pluralism that a society well-ordered according to political liberal principles would be marked by rules out the possibility of its citizens having a shared commitment to full autonomy. Even among proponents of political liberalism, however, the question of precisely what could be shared by the citizens of a well-ordered society is contested.[57] Those who doubt that a commitment to full autonomy could be shared are likely motivated by the same general thought that motivated Rawls to develop a political liberalism: that the exercise of human reason under liberal institutions will tend to lead to pluralism. However, if they are not proponents of what I called content-skepticism above, then they must hold that – despite this tendency toward pluralism – political liberal principles can nonetheless satisfy the stability condition. In accepting that political liberal principles can satisfy the stability condition, they are accepting that the socializing influences of growing up in a well-ordered society suffice to induce citizens to share a commitment to a set of substantive values and principles, provided these values and principles are limited in their range of application to the political domain. Given that full autonomy is also a value that is limited in its range of application to the political domain, the general thought that the exercise of human reason under liberal institutions tends to engender pluralism will not suffice to support the claim that it could not be among the shared commitments of citizens, at least for those who are not proponents of content-skepticism. The objector must therefore base their support

---

[56] 'On the Idea of Public Reason', in *The Blackwell Companion to Rawls*, edited by Jon Mandle and David A. Reidy, (Oxford: Wiley-Blackwell, 2013): 271.

[57] Rawls's view was that the well-ordered society would be marked by agreement on the view that political constructivism is an appropriate basis for making objective judgments about political morality (*Political Liberalism*, *supra* note 4 at 110) – a consensus which seems no less demanding than a consensus on the political value of full autonomy. Others, such as Leif Wenar, have argued that the shared commitments of the citizens of the well-ordered society would be much more minimal, 'Political Liberalism: An Internal Critique', *Ethics* 106 (1995): 32–62. See also Quong, who seems to hold that that no more than a commitment to the ideal of society as a fair system of cooperation among free equals could be shared by all reasonable citizens (*Liberalism without Perfection*, *supra* note 10 at 37–39). I have argued elsewhere that the scope of reasonable disagreement depends on our reasons for accepting a political liberal or public reason view in the first place. See Paul Billingham and Anthony Taylor, 'A Framework for Analyzing Public Reason Theories,' *European Journal of Political Theory*, OnlineFirst (2020): 1–21.

for this claim on something other than this general thought about liberalism and pluralism.[58]

To address content-skepticism, any defender of political liberalism will at some point have to address the question of what can be shared by citizens of the well-ordered society. At this point they will need to show that the case they make for the stability condition is one that those citizens could be expected to accept. Importantly, though, this is true regardless of the case that is made for the stability condition: there is no reason to single out the argument from full autonomy as especially vulnerable to this potential problem. Given that there is no reason to single out full autonomy on this score, the concern that the argument is a nonstarter is misplaced. We should, at this stage, treat the question of whether citizens of the well-ordered society could share a commitment to full autonomy as open.[59]

Let us now turn to the premises of the argument. The first premise is questioned by what I called content-skepticism above, which asks why we should accept that the possibilities for consensus in the well-ordered society fall where Rawls and his defenders take them to. Since I am setting that particular skeptical challenge aside here, I will assume that this premise holds. Premise (2) holds because, given the conception of autonomy specified above, satisfaction of the stability condition is necessary for citizens of the well-ordered society to enjoy it. The focus of my critical attention here will therefore be on premise (3).

Though I have argued that full autonomy is a plausible conception of autonomy, this is not sufficient to support premise (3) of the argument, as it does not get to the heart of the challenge that stability-skepticism poses. If we think that the fundamental aim of political philosophy is to identify the conditions under which we can live freely – to uncover, as Rawls puts it, 'the outer limit of our freedom' – then perhaps we will think that enough has been said to support premise (3). But that would be to attribute a controversial aim to our theorizing. Proponents of stability-skepticism need not

[58] On this point, see also Anthony Taylor, 'Rawls's Conception of Autonomy,' in *The Routledge Handbook of Autonomy*, edited by Ben Colburn, Forthcoming.

[59] For a reading of *Political Liberalism* that seems to hold that the political value of full autonomy could be shared by the citizens of the well-ordered society, see Rainer Forst, 'Political Liberalism: A Kantian View', *Ethics* 128 (2017): 123–144, at 140.

accept that the point of political philosophy is to identify the conditions under which citizens can live freely. Instead, they may hold that their aim is to identify the principles of political morality that we have most reason to accept. If the aim of identifying the conditions under which citizens can live autonomously comes into conflict with the aim of identifying the principles that we have most reason to accept, then the latter aim ought to be given priority. Indeed, no sensible advocate of the aim of identifying autonomy-realizing principles would deny this: they would hold that the principles that identify the conditions under which citizens can live autonomously *are* the principles we have most reason to accept. If they did not endorse this claim, it would be unclear why we should accept their principles.

Another way to put this thought is to say that proponents of the view that the point of political philosophy is to identify the conditions under which citizens can live autonomously must accept *Harmony*:

> The principles that allow citizens of the well-ordered society to enjoy full autonomy are the principles that we have most reason to accept.

If Harmony were sound, we would have a powerful defense of premise (3): we ought to accept the principles that allow citizens of the well-ordered society to live autonomously, because these are the principles that we have most reason to accept. However, we cannot assume in advance that Harmony is true. On any plausible interpretation of their views, those who reject the aim of identifying the principles that allow citizens to live autonomously do so because they reject Harmony. If we were simply to assume that these two aims are not in conflict, we would be ruling out such views by fiat. The argument from full autonomy therefore needs a defense of Harmony in order to succeed.

<div align="center">IV.</div>

We now have the argument from full autonomy to political liberalism before us. It has the potential to play a foundational role in justifying political liberalism by allowing us to reject stability-skepticism. However, as I have argued, in order for the argument to play

this role a defense of Harmony is needed. In this penultimate section of the paper, I will put forward a provisional defense of Harmony.

To begin this defense, let us return to the point of view of a citizen of the well-ordered society. She is reflecting on whether to affirm her desire to act justly or to take steps to rid herself of it. When the stability condition is satisfied, she concludes that she has sufficient reason to affirm it. From her point of view, then, the principles that allow her to live autonomously and the principles that she has most reason to accept are identical. The principles she accepts and has grown up under allow her to live autonomously, as the role they have played in shaping her desires and self-conception is one that she endorses. And these principles are also the principles that she has the most reason to accept, as she can see no reason to reject them in favor an alternative. However, this fact does not suffice as a defense of Harmony. For what matters to the opponent of that claim is what principles *we* – actual citizens, here and now – have most reason to accept, not what principles the citizens of a well-ordered society have most reason to accept. Given that our circumstances are quite different from the circumstances of citizens of any well-ordered society, it is not obvious that the principles they have most reason to accept are the principles we have most reason to accept.

In order to defend Harmony, then, we need to show that the principles that the citizens of the well-ordered society have most reason to accept are also the principles that *we* have most reason to accept. I will do this by defending the following two claims.

(H1) The principles that we have most reason to accept are those that we would accept if we were fully informed, procedurally rational, and not influenced by any other factors that distort our judgment.

(H2) The citizens of a society well-ordered according to political liberal principles are fully informed, procedurally rational, and not influenced by any other factors that distort their judgment.

(H1) follows from a widely accepted view about how we ought to work out what principles to accept in moral and political philosophy. According to the method of reflective equilibrium, we should work

back and forth between our considered moral judgments and the principles we think best explain them, making modifications at both levels with the aim of reaching the point at which our principles both cohere with and explain our considered convictions. Accepting this method implies accepting (H1), as what we should be doing when we work back and forth between our considered convictions and the principles that might explain them is: availing ourselves of information that might influence our decision; following uncontroversial norms of procedural rationality; and ensuring that our choice of principles is not influenced by any other factors that might distort our judgment. To see how uncontroversial this is, we need only consider what rejecting it would entail: that we ought to choose what principles to accept by avoiding relevant information, violating norms of procedural rationality, or trying to be swayed by distorting factors.[60]

(H2), on the other hand, is considerably more controversial and will therefore be the focus of our discussion. What is controversial about (H2) is not the claim that citizens of the well-ordered society are fully informed and procedurally rational. These idealizing conditions are contained within the definition of a well-ordered society. That the citizens of the well-ordered society are procedurally rational follows from how they reason about whether to give the principles of justice their reflective endorsement: they consider what would best satisfy all of the other desires that they have. And that the citizens of the well-ordered society are fully informed follows from the fact that their society satisfies a publicity condition: its citizens all know the entirety of the case for the principles of justice that order their society, including the beliefs and modes of reasoning that support them. That is, they are fully informed about the considerations that are relevant to their choice of what principles to accept.

What is controversial about (H2) is the claim that citizens of the well-ordered society are not, in their decision to endorse or reject the principles of justice that order their society, influenced by any dis-

---

[60] Some may want to object to (H1) on the grounds that it amounts to the controversial thesis that the relationship between the principles that we have most reason to accept and our responses in these ideal conditions is constitutive. But (H1) does not imply this controversial claim, nor is such a claim needed to defend Harmony. We can accept (H1) solely on the grounds that the best way to work out what principles we have most reason to accept is to think about what we would accept in these ideal conditions, while rejecting the claim that our responses in these ideal conditions are constitutive of such principles.

torting factors beyond the absence of information or failures of procedural rationality. As we saw above, to satisfy the stability condition, principles must be able to secure citizens' endorsement when only certain influences on their desires are permitted: the public educative effects of growing up in a society well-ordered according to liberal principles. This is what I called the Freedom Condition above, as its role is to ensure that citizens' reflective endorsement of the principles of justice is given freely. Whether or not we should accept (H2) depends on whether we think that the Freedom Condition accurately eliminates the set of factors that might distort our judgment about what principles to accept, beyond the absence of relevant information and failures of procedural rationality. That is, it depends on whether we think that when deciding what principles of political morality to accept we should imagine that we have grown up in a society well-ordered according to political liberal principles, and then consider what we would accept in such circumstances. Of course, it is not at all obvious that we should do this. Why should the principles we have most reason to accept be those that we would accept if our lives had been so radically different?

We can put the challenge posed by this question more precisely by noting that there are two ways the Freedom Condition, and thus (H2), could be dismissed. First, it might be argued that the condition is too strong: that it rules out influences on citizens' choices that should in fact be permitted. Second, it might be argued that it is too weak: that it permits influences on citizens' choices that in fact distort their decision of what principles to accept. I will now consider each of these challenges in turn.

A. Too Strong?

Let us begin with the too strong variant of the objection. The Freedom Condition holds that citizens must endorse the principles of justice in the absence of coercion and the threat of punishment. What this claim means is that the decision to accept or reject the principles of justice occurs when citizens are protected by various liberal rights, such as to freedom of conscience, expression, and association. A proponent of the objection that the freedom condition

is too strong must therefore hold that for this choice to be made free from the influence of distorting factors, some violations of these rights must be permitted.

It is hard to see what would motivate this position. These liberal rights give citizens the freedom to form, examine, and revise their commitments. In this way, they are beneficial, and often necessary, for the proper exercise of our capacity for reasoned choice. This is a strong reason for holding that the choice of whether to endorse or reject the principles should occur when the citizens are protected by these rights.

B. Too Weak?

The too weak variant of the objection is much more plausible. Since the influences on citizens that are permitted by the Freedom Condition are the public educative effects of growing up in a society well-ordered by political liberal principles, proponents of the too weak objection must hold that these influences distort their judgment. The most natural reason to think this is if we suspect that such influences are manipulative or indoctrinating.

To evaluate this objection, we first need to be clear about what exactly public educative effects amount to. Growing up in a society in which particular principles are unanimously followed and publicly acknowledged as the appropriate standard for settling competing claims will undoubtedly have a significant influence on a citizen's beliefs, desires, and self-conception. The Freedom Condition holds this process of socialization is not manipulative or indoctrinating on the condition that we would later come to endorse it in conditions of freedom.

Are these socializing effects nonetheless indoctrinating? Before answering this question, let me first note two initial points. First, it is inevitable that citizens are shaped by their upbringing and the conditions under which this takes place. There is no possible upbringing that is entirely free of shaping influences. This means that if the proponent of this objection wants to avoid asserting the skeptical view that everyone's judgement is always distorted as a result of having had an upbringing of some kind, then they will have to accept that some shaping influences are consistent with us not having been indoctrinated. This does not mean, however, that any shaping influences whatsoever are acceptable.

This leads to the second point: shaping influences would be unacceptable if they diminished or destroyed our capacity to reason. Therefore, if we thought that the effects of growing up under political liberal principles diminished or destroyed our capacity to reason, we would have grounds to hold that the Freedom Condition is too weak. But there is no basis for this concern. Political liberalism's rationale for giving priority to a set of liberal rights and ensuring all citizens have adequate all-purpose means to make use of those liberties is precisely that such conditions are beneficial for the development and exercise of citizens' capacity to reason. These principles nurture and protect that capacity rather than diminishing or destroying it.

With these two points in mind, the question of whether the public educative effects of growing up under political liberal principles amounts to manipulation or indoctrination can be answered as follows. Given that citizens are inevitably influenced by shaping influences of one kind or another, the highest standard we could expect these influences to meet is one whereby they would come freely endorse those influences 'as the outcome of reflective thought and reasoned judgment'.[61] If political liberal principles satisfy the stability condition, then the shaping influences that they have on the citizens of the well-ordered society are the ones that meet this high standard. There are therefore no good grounds to hold that they amount to manipulation or indoctrination.

I think, then, that we can also dismiss the claim that the Freedom Condition, and thus (H2), is too weak. Let us now return to Harmony. I have argued that the Freedom Condition is neither too weak nor too strong to capture the conditions under which our powers of judgment operate free from distorting influences. Since the citizens of the well-ordered society are also procedurally rational and fully informed in the relevant sense, this implies we should accept (H2)

> The citizens of a society well-ordered according to political liberal principles are fully informed, procedurally rational, and not influenced by any other distorting factors.

> I have also argued that we should accept (H1)

---

[61] *Political Liberalism*, *supra* note 4 at 222.

The principles that we have most reason to accept are those that we would accept if we were fully informed, procedurally rational, and not influenced by any other factors that distort our judgment.

Together, I have argued that (H1) and (H2) allow us to defend *Harmony*:

The principles that allow citizens of the well-ordered society to enjoy full autonomy are the principles that we have most reason to accept.

The principles that allow the citizens of the well-ordered society to enjoy full autonomy are those that would satisfy the stability condition: those that they would give their free and reflective endorsement to. If this case for Harmony is sound, these are also the principles that we have most reason to accept. When we are thinking about what principles of political morality to accept we should imagine that we have grown up in a society well-ordered according to political liberal principles, and then consider what we would accept in such circumstances, for these are circumstances in which we are fully informed, procedurally rational, and not influenced by any other factors that might distort our judgment.

<div align="center">V.</div>

By appealing to the case I have made for Harmony, we can defend premise (3) of the argument from full autonomy to political liberalism: the claim that we ought to accept the principles that allow citizens of the well-ordered society to enjoy full autonomy. Since only principles that satisfy the stability condition allow citizens of the well-ordered society to enjoy full autonomy, this implies that we ought to accept the principles that satisfy the stability condition. I therefore conclude that the argument from full autonomy, in conjunction with Harmony, can provide a defense of the stability condition.

An important limit to this argument is that it assumes that content-skepticism can be answered: that political liberal principles satisfy the stability condition. I have assumed but not argued for this claim. If it is false, then we must reject premise (1) of the argument from full autonomy. A full defense of political liberalism that appeals to autonomy and Harmony would therefore need to show that political liberal principles are in fact the ones that the citizens of the well-

ordered society would give their free and reflective endorsement to. However, the defense of the stability condition offered here is nonetheless quite significant. For it gives us a provisional case for holding that the principles that satisfy the condition are the principles we have most reason to accept.

ACKNOWLEDGMENTS

OPEN ACCESS

Nuffield College
University of Oxford, Oxford, UK
E-mail: anthony.taylor@nuffield.ox.ac.uk