

# Political Stability in the Open Society

**John Thrasher** Monash University  
**Kevin Vallier** Bowling Green State University

**Abstract:** *We argue that the Rawlsian description of a just liberal society, the well-ordered society, fails to accommodate deep disagreement and is insufficiently dynamic. In response, we formulate an alternative model that we call the open society, organized around a new account of dynamic stability. In the open society, constitutional rules must be stable enough to preserve social conditions that foster experimentation, while leaving room in legal and institutional rules for innovation and change. Systemic robustness and dynamic stability become important for the open society in a way that they are not in the well-ordered society. This model of the open society and the corresponding model of stability have interesting implications for thinking about the goals, norms, and institutions of liberal political systems.*

One of the most important parts of a theory of justice is the description of the society that a conception of justice is supposed to govern. The most common description is John Rawls's notion of a *well-ordered society*. A well-ordered society is one governed by a conception of justice and whose citizens comply with the conception of justice based on their sense of justice. That is, they see the conception of justice as worth endorsing and as worth complying with for its own sake.

We argue that Rawls's model of a well-ordered society, as an account of a realistic utopia, is defective for two main reasons. First, the societal model most appropriate for assessing the justification and stability of a conception or conceptions of justice must allow for deep disagreement and diversity in general, and not just for conceptions of the good (Dryzek and Niemeyer 2006; Gaus 2016). For this reason, the societal model must allow for more pluralism than is seen in the well-ordered society (Muldoon 2017). Second, focusing on disagreement may lead us to care more about *discovery* in the liberal society than in the well-ordered society (Anderson 2006; Hayek 2011; Mill 1978).

A liberal society is one in which people are free to experiment with different approaches to the good life. That liberal ideal will lead us to want to experiment in order to discover the conceptions of justice most appropriate for governing our shared social order. If we are therefore committed to recognizing deep diversity and the need for social discovery, we must modify the

idea of a well-ordered society and the ideas most closely associated with it. In particular, a new notion of *stability for the right reasons* is required—a more dynamic notion appropriate for what we will call the *open society*.

Developing an idea of stability for a society where conceptions of justice may be controversial and partially unknown poses several new challenges. Most importantly, stability for the right reasons may now seem to have a *cost*. To the extent that certain social rules are stable, i.e., in equilibrium it will be difficult to discard those rules or principles in order to adopt better ones. And yet, we still want many of the social rules that comprise our order to remain in equilibrium so that they can form a sturdy basis for social cooperation.

Toward this end, we distinguish between types of social rules and the forms of stability most appropriate for each type. We distinguish general constitutional rules, the “rules of the game” that specify how we settle upon first-order legal and institutional rules that apply directly to behavior, from first-order legal rules themselves. In an open society, we want our constitutional rules to be stable insofar as they preserve the social conditions that foster experimentation.<sup>1</sup> But we do not want a similar stability of the first-order legal and institutional rules, given that they can be improved upon. Instead, we want legal and institutional rules to be able to achieve a new equilibrium after being challenged by social innovators, a notion of stability that we can think of as *robustness*.

---

John Thrasher is a Lecturer in Department of Philosophy, Monash University, 6th Floor Menzies Building, 20 Chancellors Walk, Monash University, 3800 Victoria, Australia (John.Thrasher@Monash.edu). Kevin Vallier is Associate Professor at Department of Philosophy, Bowling Green State University, 1025 Clark St. Bowling Green, OH 43402 USA (Kevinvallier@gmail.com).

<sup>1</sup>As well as, of course, providing for more basic needs like reducing civil strife and criminal behavior.

So our thesis is that stability in the open society involves second-order rules that are stable for the right reasons—based on our varied conceptions of the good and justice—and first-order rules that are *robust* for the right reasons. When challenged, the first-order rules can change, but they tend to return to a new equilibrium on a new set of rules.

## Beyond the Well-Ordered Society

Rawls's (1996, 35) most mature account of a well-ordered society describes a realistic utopia where (a) the basic structure of society is regulated by principles of justice, (b) citizens are aware of this fact, and (c) citizens all have an effective sense of justice so as to comply with the directives of the basic structure, which they regard as just. The well-ordered society is important because “the notion of a well-ordered society articulates a formal and abstract conception of the general structure of a just society” (Rawls 1999b, 235). The well-ordered society is a model of a complete conception of justice that allows us to evaluate that conception of justice and to rationally evaluate a world that realizes that conception. Although Rawls changes his analysis of the well-ordered society considerably over the course of this work (Kogelmann 2017), Rawls's mature view of the well-ordered society is the conjunction of three conditions:

- (1) Assurance—everyone accepts, and knows that everyone else accepts, some member of a limited set of reasonable political conceptions of justice, which establish shared points of view from which citizens' claims on society can be adjudicated (Rawls 1996, xlvii–xlvi).
- (2) Publicity—its basic structure—the main political and social institutions and how they fit together as one system of cooperation—is publicly acknowledged or with good reason believed to satisfy these principles (or some mix of them).
- (3) Stability—its citizens have a normally effective sense of justice and so generally comply with society's basic institutions, which they regard as just (if not fully just).

It is the sense of justice that drives people to internalize the rules of their society, and so develop a stable disposition to comply with those rules. This preparedness to comply with the rules of their society based on their sense of justice renders the well-ordered society stable *for the right reasons*. A society is stable for the right reasons when its stability is based on its citizens' acting on *moral* reasons rather than merely penal sanctions. Thus, a well-ordered

society must usually be able “to generate its own support” rather than have it imposed from without (Rawls 1996, 119). Importantly, compliance based on one's sense of justice alone is not always enough to ensure that the well-ordered society generates support; there is a role for the enforcement of the law in generating stability, for as Rawls (1999a, 211) says, “the coercive powers of government are to some degree necessary” to ensure stability. However, since citizens' acting on their sense of justice is the predominant factor securing compliance, we will focus on this predominant element of stability in this essay.

The move from the well-ordered society to the open society is driven by two objections to the above conditions. First, Rawls claims that a well-ordered society is constructed around a single conception of justice or a small family of such conceptions. Everyone must accept and know that everyone else accepts some member of that set. But one of the hallmarks of a political liberal approach to justice is the recognition of reasonable pluralism, where free and equal persons deeply disagree about the nature of the good.

There is no reason, though, why disagreement and diversity only generate pluralism about the good. Reasonable pluralism also extends to disagreements about many other issues, including, plausibly, the right. Pluralism about the right is likely to run *just as deep* as pluralism about the good (Gaus 2016; Kogelmann 2017). In any case, there is no way to limit the impact of evaluative and doxastic diversity. Even if a liberal society did have considerable agreement on a small set of principles, the diversity that liberal pluralism allows and encourages would eventually introduce new and competing visions of justice into the society (Bruner 2015; Sparrow 2016).

A citizen's conception of the good is an organized combination of that citizen's beliefs and values, as is a citizen's conception of justice, though the two conceptions will typically range over different domains. This suggests what we call the *symmetry thesis*.

**Symmetry thesis:** Reasons to expect pluralism about the good in liberal societies are also reasons to expect equally deep pluralism about other central features of liberal societies.

This expands to all domains in society, but here we are primarily interested in pluralism about the right. Those who reject the symmetry thesis must show that beliefs and values about the domain in question are somehow insulated from the diversity about beliefs and values that lead to the same pluralism about the good.

Some political liberals object to the symmetry thesis. Jonathan Quong argues, for instance, that disagreements about the good and about justice have fundamentally

different characters. Disagreements about the good are often “foundational” in that the parties to the disagreement do not share enough common premises to ever hope to agree about the content of the good or what the good life consists in (Quong 2011, 204–8). Disagreements about justice, by contrast, are not fundamental because all parties share common principles that allow for resolutions of their disagreements, at least in principle. This argument, however, is importantly limited to those who are already liberals, if not already *political* liberals. If we are interested in evaluating liberalism in relation to its various competitors, however, we should not assume that all reasonable members are already committed liberals. Further, as Kogelmann argues, there are certain forces within political liberalism that naturally lead to pluralism about justice within a free society once we accept the reality of reasonable pluralism and the burdens of judgment (Kogelmann 2017, 5–7).

Our second deviation from Rawls’s conception of the well-ordered society is more radical. Given reasonable disagreement about politics, we should also be prepared to acknowledge that there are at least some features of principles of justice and their application to society that are currently *unknown*. It may be that the true conception of justice is, at present, beyond our ken. If we do not know what the true theory of justice is, it must eventually be *discovered* over time. Given that this is how we think of science, it seems natural that we would apply a similar standard to the most fundamental principles and rules that govern our society as well.

There is also a deeper analogy. Virtually all past scientific theories have been ruled false, substantially changed, or abandoned as we have developed new techniques and learned more about the world. The germ theory replaced the miasma theory that, in turn, replaced the humoral theory of disease. Scientific views once held with supreme confidence were deeply flawed. The same is true of political theories. The “humane” liberalism of the nineteenth century in Britain—with its penchant for imperialism, hypermoralism, and paternalism—looks extremely unappealing to today’s liberals. There is good reason to think that our near-universal history of being wrong about the nature of the world has implications for how we should think about the future (Laudan 1981). We should take seriously the possibility that we are overconfident and parochial in many of our views. In the future, many of the eminently *reasonable* views of today will probably seem as strange as the humoral theory of disease. Consequently, we should be humble about how many of the “truths” that we think we know are final truths. So we should not construct a model of a society to be regulated by principles of justice based on members of that society being convinced

that a particular conception of justice, or a small set of conceptions, is true.

However, even if members of a well-ordered society agreed that one of the presently recognized conceptions of justice is the right one (even if they disagree about which is right), they would still disagree about how best to apply it to structure social institutions. This is because a conception of justice significantly *underdetermines* how to institutionalize it. Call this the *underdetermination thesis*.

**Underdetermination thesis:** For any conception of justice, there are many ways of implementing that theory. No conception even mostly specifies all the important details of its institutional implementation.

The underdetermination thesis introduces a new dimension of disagreement. While Rawls did think that reasonable people could disagree *somewhat* about how to institutionalize justice as fairness, this was limited, ultimately, to whether to embrace property-owning democracy or liberal socialism (Rawls 2001, 138–39). The underdetermination thesis implies a much broader range of reasonable disagreement about how to institutionalize a conception of justice. Rawls would reject the underdetermination thesis because he believed institutional variation in implementation would be minor. Given the realities of reasonable disagreement, there is little reason to think he is correct on this point, as we argue below.

In addition, we also disagree about the best way to move to a more just social world. A wide range of normative and non-normative claims bears on how far away our society is from fully realizing justice, and about how best to transition from our present circumstances to a just society given all the inevitable complications in transition (Gaus 2016; Sen 2009).

The symmetry and underdetermination theses show Rawls’s model of a well-ordered society is defective in a society characterized by genuine, reasonable pluralism. Disagreement about justice should be just as deep as disagreement about the good in a pluralistic society. But, even if there were complete agreement about justice, there would still be disagreement about how to implement those principles institutionally since principles of justice strongly underdetermine their institutionalization. Disagreement about justice in the well-ordered society is likely to be so great that the first condition (assurance) is not satisfied. This, in turn, can undermine the common knowledge that is necessary for the second condition (publicity), which can ultimately undermine the third condition (stability; Kogelmann and Stich 2016; Thrasher and Vallier 2015). We must now search for an alternative to the well-ordered society that is consistent with pluralism.

## The Open Society

In light of the problems with Rawls's conception of the well-ordered society, we develop a more dynamic model of a diverse society, what we call, following Popper (2013) and Gaus (2016), *the open society*. The open society has the following features:

- (A) Everyone accepts, and knows that everyone else accepts, a political *limiting principle* that is a general basis for a political and social constitution.
- (B) It is publicly acknowledged, or with good reason believed, that there is general agreement on this constitution and its dispute resolution mechanisms.
- (C) Citizens recognize the legitimacy of the limiting principle, the constitution, and general dispute resolution mechanisms.

These new conditions accommodate additional pluralism and are compatible with the symmetry thesis. Citizens can now disagree about what justice requires within much broader limits than in the well-ordered society. This is because the open society only requires convergence on the limiting principle. This principle can take several different forms and may include some conception of basic human rights or, more minimally, a rule of recognition. The limiting principle acts to certify constitutional possibilities or conceptions of justice by excluding those possibilities or conceptions that are incompatible with that principle. This contrasts with Rawls's requirement that we converge on a single conception of justice or one of a small set of conceptions, as it only specifies broad *limits* on acceptable conceptions that allow for a quite large set. We will discuss the content of the limiting principle further below, but for now, we can understand it as specifying membership in the set of acceptable conceptions of justice.

Condition A also specifies that the limiting principle is both widely accepted and widely recognized as accepted by all, which establishes an important principle of something like common knowledge or publicity (Lewis 1969). Unless each person has a credible belief that everyone else accepts the limiting principle, she will lack assurance that compliance with the limiting principle is practically rational for her. As Rawls repeatedly stressed, justice does not require unconditional compliance, but rather specifies norms that we are morally bound to follow so long as others do likewise (Rawls 1999a, 436–37). This means that citizens are prepared to comply with the limiting principle so long as they believe others will do likewise; they are practically rational in complying with the

limiting principle, but in cases where common knowledge breaks down, citizens may see themselves as permitted to disobey the principle, though they need not explicitly think as much (Lister 2017).

Condition A also allows that citizens have different points of view from which they attempt to resolve their competing claims. This is because they adopt different conceptions of justice and because conceptions of justice just *are* perspectives on how to resolve society's competing claims (Rawls 1999a, 112–18). Condition A may thereby raise the possibility of conflict between conceptions of justice; different conceptions of justice held by different members of the public will lead them to try to resolve competing claims differently. However, conflicts between citizens' points of view about justice will not necessarily be problematic for several reasons. First, citizens might share enough of a perspective about justice to converge on mechanisms of dispute resolution about other aspects of justice. Second, citizens might nonetheless come to agree about how to resolve competing claims even if they have different perspectives on how that resolution should occur. Citizens can converge from different perspectives on solutions. Finally, citizens may be able to organize society such that they have different social domains in which they can resolve competing claims differently. In a federalist (Levy 2007) or polycentric order (Aligić and Tarko 2013; Kogelmann 2017), for instance, different groups may be able to use their diverse perspectives to resolve competing claims within their own jurisdictions.

Condition B is relatively straightforward in light of Condition A, since it requires that the basic structure only be governed in accord with the limiting principle and associated constitutional dispute resolution mechanisms. Since we wish to allow different parts of a basic structure to be governed by different features and perhaps at different levels (as in a federalist or polycentric order), we do not want to commit the open society to holding that all of a society's institutions somehow "fit together" into one system of cooperation governed by a conception of justice. Thus, convergence on the limiting principle need not involve social convergence on a single reasonable conception of justice, even one chosen from a relatively capacious set of reasonable conceptions of justice.

## Limits, Constitutions, and Institutions

### The Limiting Principle

We can begin explaining the idea of a limiting principle by appealing to Rawls's own limiting principles from



the introduction to the paperback edition of *Political Liberalism* (1996). There, Rawls claims that all reasonable liberal political conceptions of justice must meet three conditions:

- (1) They must all identify a set of basic liberties.
- (2) Assign them a special, though not lexical, priority.
- (3) Ensure adequate all-purpose means to exercise those liberties.

These conditions are limiting principles because they specify which conceptions of justice lie within an acceptable range and define the boundaries of that range. To be more specific, the limiting principle tells us which conceptions of justice are compatible with the sense of justice of each person. This does not mean that each person must endorse any member of the set as *best*, but rather that the members of the set are good enough that a person's sense of justice is not offended or violated.

We can distinguish, then, between unacceptable, even repugnant, conceptions of justice and conceptions of justice that are suboptimal from the perspective of each person's perspective on justice, but are nonetheless candidate moral conceptions on which social life might nonetheless be based. In this way, the conceptions of justice within the set can be in some sense *legitimate* in that they are acceptable, though far from optimal, ways of organizing social life. Broadening the range of reasonable pluralism, the political liberal must grapple with allowing for an enormous number of persons and ideals to form an overlapping consensus, including deep religious and moral disagreements. And our assumption of diversity is sufficiently broad that people can engage in a wide range of experiments, and so they avoid discharging the great cognitive burden of expressly articulating their sectarian commitments.

It is important to note that the limiting principle is not an overlapping consensus in the Rawlsian sense. Members of the open society need not endorse the limiting principle from the point of view of their distinct conceptions of the good, as in the well-ordered society. Rather, there is convergent agreement on the limiting principle for various reasons, such as purely conventional reasons (Hardin 2003; Hart 1961; Sabl 2012), instrumental reasons (Moehler 2014; Thrasher 2016), or some combination of conventional and moral reasons (as in the "Kantian coordination game" in Gaus 2011). So, the limiting principle limits both the acceptable conceptions of justice in a society and permissible constitutions, but it is not necessarily accepted purely for moral reasons.

## Constitutional Stability

The open society needs a different conception of stability than the well-ordered society. First, there will be less agreement about justice in the open society than in the well-ordered society, which means that stability cannot be founded on extensive agreement about justice. There will be universal agreement on the limiting principle, but the limiting principle underdetermines the justification of any particular conception of justice or constitution. Second, since one of the chief benefits of the open society is its dynamic nature, too much stability of the sort we find in the well-ordered society would not be beneficial, as it would hinder the very change that makes the open society dynamic. In the open society, stability of the traditional sort can prevent change that can lead to improvements. On the other hand, with insufficient stability, the rules that govern social life may vary too much. If these rules are unpredictable or erratic, they will not be able to maintain an environment where productive social experimentation can occur.

Recognizing the trade-off between different types and amounts of stability in the open society, we can distinguish between two types of social rule: first-order *institutional* rules and higher-order *constitutional* rules, with different forms of stability appropriate for each. In so doing, two distinct notions of stability emerge that can be combined to produce the type of stable social environment appropriate for the open society.

We can define a constitutional rule as a rule that ratifies, reforms, or repeals institutional rules that themselves directly govern action. Constitutional rules are the highest-order rules, the final court of appeal in determining and altering all other social rules. The analogy here is to Hart's (1961) distinctions between primary and secondary order legal rules. The important point is that the highest-order rules do not directly govern the behavior of citizens, but rather govern the rules that govern behavior. Constitutional rules may also govern the behavior of *political officials* who have the capacity to alter the legal environment within which citizens of the open society interact.

Constitutional rules are the rules most directly responsible for both the process of institutionalizing justice and the process of preserving the capacity of citizens to engage in social experiments. Imagine, for instance, a federalist constitutional order that protects basic rights for all of its subunits but that also allows for the law to vary considerably across its subunits. We can imagine that citizens of the open society agree upon certain basic rights as setting the boundaries of reasonable disagreements about justice, and that they also recognize how much of the

content and application of justice depends upon social experimentation with different legal regimes. For this reason, citizens believe in preserving the federal system and move to different states or provinces based on the social experiments active in those states or provinces. Of course, as Levy (2007, 2014) has argued, there is a tension between a universal political order that secures a certain amount of freedom and equality and the pluralism and disharmony that federalism allows. The dynamism of the open society, we argue, largely derives from this disharmony, something that the Rawlsian well-ordered society does not capture.

Constitutional rules that respect justice pluralism should be stable for the right reasons in the open society. These rules will be in equilibrium based upon the free compliance of political officials. But we can distinguish between two ways in which constitutional rules can be stable—based either on the rule's *meaning* or its *function*. *Semantic* stability means that the content of a constitutional rule remains the same over a given period of time. *Functional* stability means that the constitutional rule performs more or less the same function over time, despite changes in the social environment. The rule resists shocks or deviations of various types. In principle, these two types of stability can differ; the meaning of a constitutional rule might change while its function stays the same, whereas the function of a rule might change even if its meaning stays the same. In practice, semantic and functional change should largely track one another in an open society. Changes in meaning lead to associated changes in function, and changes in function will tend to be recognized as such and so manifest themselves in semantic change. On the other hand, these two types of stability might vary independently of one another.

The rule must also be compatible with the sense of justice of each citizen, as bound by the limiting principle, such that people want to comply with it so long as others do. The constitutional rule should not only produce benefits that citizens enjoy—it should also be normatively attractive to those who follow it. The fact that it has normative force means that an individual who does not comply with the rule realizes that she will face some kind of personal or social consequence, such as feeling guilt or being shamed by others. In this way, constitutional rules are social norms in Cristina Bicchieri's (2006) sense that they are backed by normative as well as empirical expectations. Consider Robert's Rules of Order as an analogy, where unilateral deviation from the rules will not merely bring some form of censure, but generally be an ineffective way to get what one wants (Robert and Robert 1990).

Stable constitutional rules also have a conditional aspect. Constitutional rules must be somewhat effective;

otherwise, the constitutional rule is not a constitutional rule. In this way, constitutional rules are only binding for each person insofar as she believes that others will comply with the rule. Furthermore, in light of the previous paragraph, conditional compliance must involve some normative aspect. Each person must conditionally comply because the rule is compatible with her sense of justice, as well as because she believes that others will do likewise.

Constitutional stability can be undermined due to a breakdown in semantic or functional stability. The common understanding of the constitution might shift due to ideological changes in society, or certain provisions of the constitution might fail to be observed, such that the provision becomes an effective dead letter. To illustrate a breakdown in functional stability, we can imagine that self-interested officials decide to break a constitutional rule in order to secure benefits for themselves or those they represent. These officials may be able to avoid sanction, such that violations become a normal state of affairs. We can also imagine that defection begins to spread, as other officials who would otherwise cooperate observe that the system of sanctions has broken down. Increasingly, then, the effective provision of the constitution fades, and the stability of the constitutional rule breaks down. The breakdown of stability does not mean that we will face social chaos, but rather that unconstitutional norms will form that lack the benefits associated with the once-effective constitutional rule. We also cannot be confident that the new practices will be consistent with the limiting principle.

A constitutional rule is an equilibrium of a mixed-motive cooperative game—there are temptations to defect from the rule, but each person values coordinating on the rule as well. So, each person has reasons to follow the rule, because the rule is seen as normatively binding, but also has practical reasons to follow the rule in order to benefit from coordination. As Bicchieri (2006) argues in the context of social norms more generally, norms that act as stable equilibria in mixed-motive games will involve a combination of empirical and normative expectations. For a constitutional rule to be an equilibrium, and so stable, it must be backed by some degree of assurance, such as when the individuals involved have common knowledge or common belief that the equilibrium point will remain where it is (Lewis 1969). Based on the foregoing, we can represent a version of this game as follows (Table 1).

We can imagine that both constitutional rules A and B are consistent with the limiting principle and are, hence, acceptable to both parties. Nevertheless, Official 1 in our example prefers Rule A to Rule B and Official 2 prefers

**TABLE 1 Constitutional Coordination Game**

		<i>Official 2</i>	
		<b>Rule A</b>	<b>Rule B</b>
<i>Official 1</i>	<b>Rule A</b>	3, 1	0, 0
	<b>Rule B</b>	0, 0	1, 3

the opposite. They would rather coordinate on either rule, however, than not coordinate. Over time, there will be a temptation, a legitimate temptation, to want to move the coordination point toward one's preferred equilibrium point. Although this process can undermine the stability of the rule, it need not. It will all depend on how and why the equilibrium shifts.

Widespread compliance with constitutional rules should tend to generate the requisite assurance. However, rule compliance alone is unlikely to provide adequate assurance without the possibility of external sanction for violations. Assurance based purely on the observed behavior of others will tend to be fragile on its own (Thrasher and Vallier 2015). But even external sanctions must occur from within the system, from the goodwill or conviction of public officials. These rules must be seen as legitimate and normatively binding by the officials themselves to be effective. Thus, constitutional rules are always somewhat vulnerable to defection, given that there are fewer tools available to discourage deviation; external sanctions are only effective within a system of constitutional rules that officials see as providing them with reasons to comply with those rules.

### **Institutional Stability**

Institutional rules include all rules governed by some higher-order rule. They include, for instance, procedural norms that govern law enforcement and legal norms produced by the legislative process as well as judicial and regulatory rules. Indeed, most legislation falls into this category. In the open society, *all institutional rules are subject to change*. In order to reap the benefits of open institutions, though, this institutional change must occur in the context of stable constitutional rules and widespread agreement on the limiting principle.

The main benefit of open institutions is that they allow for social experiments that can lead to improvements. Social experimentation can include experiments with new kinds of norms, as well as economic, cultural, and religious experiments. Many of these experiments will fail by generating poor consequences or improving the status quo. The cost of failure is potentially high. But

the key point is that persons can observe the results of the experiments and then transmit the results across society. Even failures, then, can provide social benefits. We therefore set up a social process of trial and error that draws benefits even from costly failed experiments. Society gets data from successes as well as from failures (Hayek 2011, chap. 2; Mill 1978).

What reason do we have to think that the experimental process will lead the social order to improve over time? At a minimum, we think the alternative of insisting that a single just arrangement be imposed on the populace will be less likely to generate progress. Without allowing for sufficiently many social experiments, we are likely to congregate around an unacceptable conception of justice, or at least an inferior institutionalization of the conception of justice and institutions we have chosen.

But with sufficiently diverse perspectives in society, we have at least some formal reason to think that social discovery will be a generally positive process. These results can be lumped together under the name "diversity trumps ability" (Anderson 2006). The core idea comes from work by Lu Hong and Scott Page (2004). Although their model is mathematically complex, the basic idea is that a group of problem solvers who see the problem in question differently (but not too differently) will tend to overperform and, in some special circumstances, outperform more knowledgeable or talented but less diverse groups. Several other thinkers have adapted this work to political decision making. H el ene Landemore (2012) uses the "diversity trumps ability" results of Hong and Page to defend an inclusive form of deliberative democracy. Similarly, Ryan Muldoon (2017) incorporates these insights into his larger theory of justice, and Elizabeth Anderson (2006) uses Hong and Page to argue in favor of greater inclusiveness and integration of society as a whole.

We can now distinguish between stability and *robustness*. Stability obtains when a system returns to the *same* equilibrium point when disrupted, whereas robustness obtains when a system returns to *some* equilibrium point, but not necessarily the original equilibrium point. If institutional rules are robust, then they can move from one equilibrium point to another. Note the difference with stability in constitutional rules, where it is important that a particular equilibrium be maintained. So, while there will be important transition points between equilibriums, we still move to an equilibrium, and so the change process has a terminus. The system of institutional rules should not collapse under change, but rather adapt to change by driving the system to an equilibrium point.

Institutional equilibria have properties of social dilemmas and coordination problems. There are incentives to defect, as in social dilemmas, but the rational

**TABLE 2 Social Dilemma and Coordination Game**

	A	B	C	D	E	F
A	4, 4	1, 8	0, 0	0, 0	0, 0	0, 0
B	8, 1	2, 2	0, 0	0, 0	0, 0	0, 0
C	0, 0	0, 0	5, 5	3, 6	0, 0	0, 0
D	0, 0	0, 0	6, 3	4, 4	0, 0	0, 0
E	0, 0	0, 0	0, 0	0, 0	6, 6	4, 10
F	0, 0	0, 0	0, 0	0, 0	10, 4	5, 5

**TABLE 3 Mixed-Motive Multi-Equilibria Coordination Game**

	A	B	C	D	E	F
A	4, 2	0, 0	0, 0	0, 0	0, 0	0, 0
B	0, 0	2, 4	0, 0	0, 0	0, 0	0, 0
C	0, 0	0, 0	5, 4	0, 0	0, 0	0, 0
D	0, 0	0, 0	0, 0	4, 5	0, 0	0, 0
E	0, 0	0, 0	0, 0	0, 0	6, 5	0, 0
F	0, 0	0, 0	0, 0	0, 0	0, 0	5, 6

equilibrium may also have coordinative elements. This is represented in Table 2 in an example taken from Andrew Schotter (2008, 25).

In this game, there are three embedded social dilemma subgames (shaded). The overall best coordination result for both players is (E, E), which yields a payoff of 6 for each player. However, in this social dilemma game, only (B, B), (D, D), and (F, F) are equilibria. Since there are multiple mutually beneficial equilibria, players are now in a meta-social dilemma coordination game. Of those coordination equilibria, (F, F), which gives each player a payoff of 5, is Pareto optimal.

In institutional settings with substantial opportunities for defection, stability can be maintained if there are benefits from coordination. Within that institutional framework, the coordination point need not stay the same so long as coordination moves between the various equilibria. In Table 2, there is no reason to move between coordination points, but we can imagine a social dilemma that has been transformed into a mixed-motive coordination game along the lines suggested in Table 1. Here, the stabilization of norms reduces the incentive for deviation (Bicchieri 2006, 2010, 2016).

In this case, the institutional game will lose its social dilemma aspect and transform into a mixed-motive coordination game, such as the one represented in Table 3.

This model of institutional stability as robustness—that is, the process of changing institutional rules, but

in a way that changes are made from one equilibrium to another—avoids chaos and instability so that social expectations can be set and social experimentation may take place. So, institutional rules must tend toward some equilibrium state, if not the equilibrium state from which they began. Interestingly, the development of institutional rules is a path-dependent process that may be impossible to predict even at an extremely general level (North 2010, 19–22).

We imagine several types of deviations that ground social robustness rather than mere stability. First, agents might attempt to use constitutional rules to change institutional rules. In those cases, the law is used to move society from one equilibrium to another, or even sometimes create a new equilibrium. Either way, the transition process is relatively quick and semipermanent, but it may also involve a great deal of coercion at times. Another way in which an equilibrium can break down in favor of a new one is through social entrepreneurship. An agent will start to flout the norms and legal rules present in her order because she thinks there is a better equilibrium in the offing. Consequently, the social entrepreneur may pay significant costs to carry out her experiment, but if she gains enough adherents, then over time the practice she opposes can wither and the practice she supports can take hold.

To illustrate both kinds of challenges to institutional rules, we can look at two important recent social changes in the United States. The civil rights movement changed a variety of institutional rules, but it sought a permanent change in the law from state laws that required segregation to federal laws that forbid racial discrimination. Civil rights protestors not only demonstrated within the bounds of the law, but they also often broke laws that they considered unjust; nevertheless, their goals were partly centered around legal change.

In contrast, the gay rights movement, despite seeking legal recognition in the form of gay marriage, has focused largely on making the cultural case that homosexual relationships are neither unhealthy, unnatural, nor immoral. Social entrepreneurs have violated old institutional rules that forbade homosexual activity and homosexual relationships and have been able to abolish these moral prohibitions over the span of a few decades. Legal changes have largely followed suit.

There are as many ways to experiment with social change as there are diverse perspectives in society. We have only illustrated two kinds of challenges to institutional rules, but we can imagine still more kinds. For instance, we can imagine social change created by experiments in one part of a polity that are exported to other parts. Colorado's legalization of marijuana thus far seems to be a success



and will likely be exported to other states. But this social experiment is different from attempts to change the law or to change social norms. Instead, the experiment was legally warranted by a statewide ballot referendum. Other states have engaged in social experiments with school vouchers in much the same way. But again, our point is not to enumerate all the ways in which experiments might occur, but rather to illustrate that there are many possible choices.

We grant that experimentation will come at a cost to some citizens. Not everyone benefits from social change. Those heavily wedded to the status quo or those who would preserve traditional ways of life may lose in any system that allows for substantial experimentation. This point highlights a substantial difference between the open society and the well-ordered society. In the language of Levy (2014), the well-ordered society is a rationalist society based on a universal conception of justice, whereas the open society is largely pluralistic, lacking a shared conception of justice, or at least lacking a shared interpretation of that conception. While the limiting principle and the constitution will presumably protect basic rights, there is no guarantee that a particular conception of justice will be met in every jurisdiction. In many ways, this will make the open society more *political* than the well-ordered society since the social and democratic process of experimentation and contestation is never fully completed. Even so, experiments in living need not be experiments for everyone. In an open society, there are ample opportunities to develop distinct communities dedicated to pursuing different forms of life in the ways suggested by radical decentralists like Nozick (1974) and Kukathas (2003), but governed by a stable social superstructure.

Appealing to the benefits of experimentation and the idea of robustness allows for a fruitful comparison between the benefits of an open society *in comparison* to the well-ordered society. The best-case scenario for the defender of the well-ordered society is if citizens are coordinating around the best available conception of justice or an excellent but perhaps not optimal conception. Further, reaching another excellent conception of justice may be extremely costly and likely to yield less justice in transition. The well-ordered society does well when status quo bias is not so bad. But in other conditions, where justice is an object of dispute and much of the content and application of justice are unknown, we can expect the open society to perform better than the well-ordered society from the perspective of most of the members of the well-ordered society (Gaus 2016, 230–40). Basically, anyone who has a conception of justice other than the one on which the well-ordered society is based (if it is based on any one conception) has reason to favor more open

institutions that allow for experimentation with new ideas, such as new conceptions of justice. For this reason, we think that the open society is a *good bet* for most people. This means that we should change our conception of stability of institutional rules to the idea of robustness, as robustness both allows for change but also generates temporary stability for new sets of rules. Again, a robust order tends toward equilibria, but not toward any one equilibrium.

We must now address the sorts of compliance problems we expect to find at the level of institutional rules. Institutional rules represent an interesting balance between compliance and much noncompliance. Some noncompliance is necessary in order for social entrepreneurship and experimentation to work. Too much noncompliance, however, will render the rule unusable and undermine stable expectations. The optimal level of noncompliance then is not zero. Consequently, we must welcome some kinds of error. To economize on error and to leave non-experimenters with rules they like means allowing experimentation at the margins of society, protecting social entrepreneurs and giving them space to engage in new experiments in living, and then structuring a portion of their polity around a new set of laws. In a way, then, we want to allow for experimentation, but we also want to confine experimentation to certain limited jurisdictions.

Free societies have developed two main mechanisms to allow for experimentation—freedom of association and federalism. Freedom of association allows social entrepreneurs to create groups within which experimentation can take place without forcing nonmembers to engage in the experiment (Alicicá and Tarko 2013). For instance, people form new religions all the time in liberal societies, confining their experiments to their small groups of adherents. From time to time, a new religion arises from these experiments; most recently, we have seen the rise of the Bahai and Mormonism. Federalism divides legal jurisdictions into subunits within a broader order such that states or provinces can try out legal experiments before exporting them to the nation as a whole (Levy 2007). In recent decades, we have seen experiments in drug policy, education policy, and healthcare policy in the United States. Healthcare policy experiments in Massachusetts led to the formation of the Affordable Care Act, and experiments in marijuana legalization and school vouchers are gradually spreading to more and more states. People will also need the freedom to communicate their experimental results and to assess and sharply scrutinize the social experiments that have been proposed, those that are practiced, and the assessment of results. For this reason, strong rights to freedom of speech must be protected.

## Levels of Stability

So far, we have distinguished stability at the level of constitutional rules and robustness at the level of institutional rules. We must now deal with the question of how the two levels of rules interact. The main threats to harmony between these forms of stability are as follows: (a) stability at the level of constitutional rules somehow hampers or undermines robustness among institutional rules, or (b) robustness among institutional rules somehow hampers or undermines stability at the level of constitutional rules.

We can understand the two threats with an analogy. Imagine a farm that is cordoned off into large segments, and each segment is devoted to growing a different kind of crop. Weeds, disease, or adverse weather conditions will threaten some crops, but not others. We envision two types of problems that could threaten the farm. First, the boundaries between crops might be so rigid and controlled that successful crops cannot be expanded into new areas, such as parts of the farm where another kind of crop is failing in some way. Second, problems with one kind of crop might spread to other parts of the farm if the boundaries between crops are insufficiently strong. Perhaps weeds spread from the potato farm to the wheat farm, or a certain kind of blight that afflicts wheat comes to afflict the corn crop as well. The interaction between constitutional stability and institutional robustness is subject to two analogous problems. Again, either the constitutional rules can become so rigid that experiments are undermined and limited in their ability to expand, or the institutional experiments can become so destabilizing that the constitutional rules that respect the limiting principle lose their semantic or functional stability.

This interaction between constitutional and institutional stability may seem complicated or potentially unruly, but this very interactive untidiness, which may seem aesthetically unattractive to our analytical minds, can have unappreciated benefits. Continuing the analogy with the garden, James Scott explains how native agriculture in Guatemala struck European settlers as “overgrown, riotous dump heaps” (Scott 1999, 275). In many places, they replaced these “dump heaps” with more visually organized European gardens and farms, mostly with disastrous effects. Upon closer investigation, what seemed like madness in the native approach was a complex and well-developed method. Farmers plant different types of plants together, and this polyculture farming is a common indigenous technique in the tropics. Polyculture can help prevent thin soil erosion, maximize fertility, and protect

plants from the harsh sun. Its greatest benefit, however, is robustness against threats. One type of crop, arranged in visually appealing rows, is susceptible to catastrophic failure if a new invasive insect or climatic event occurs, but multiple crops are not. The relationship between constitutions and institutions in the open society is complex in a similar way. Change or instability in one area will not necessarily affect other areas, and hence there is little danger that the entire system will be destabilized at any given time.

Generalizing, constitutional rules are more rigid than institutional rules for a variety of reasons. First, stability in the open society requires convergence on the limiting principle. This principle constrains the constitutional possibilities more than the constitution constrains institutional possibilities. Second, constitutional stability makes institutional experimentation possible. Diversity and experimentation at the level of institutions will only be welcomed and maintained insofar as we are confident that a failed experiment at the institutional level will not infect the constitution. Despite this, a constitutional rule can also be so stable that it effectively shuts down experimentation and stops the spread of experimental policies at the institutional level. Too much stability at the constitutional level prevents *good* “bubbling up” where state-level innovations are transmitted to the federal level, and it exerts downward pressure that limits the capacity for experimentation at the institutional level.

These two problems—excessive constitutional rigidity and excessive institutional fluidity—are addressed by recalling the open-ended nature of the limiting principle. An open society requires agreement on the limiting principle, but the limiting principle does not need to have a detailed and substantive content. Perhaps it protects the integrity of the person, guards the rule of law, and specifies certain freedoms of speech and association along with establishing strongly federal arrangements. In that case, a limiting principle is appropriately open-ended such that making it rigid should have more benefits than costs. Because it is open-ended, it should not excessively restrict the institutional level, but because it is rigidly institutionalized, it should not destabilize the institutional level. In sum, the limiting principle is the linchpin of a kind of compound political stability.

## The Open Society as an Ongoing Project

This argument is not merely meant to be an abstract exercise. The open society has tangible benefits for our world

and, we believe, is a compelling vision of social organization. Consequently, we end by sketching the sort of social structures that we think are implied by the limiting principle and that are capable of achieving both constitutional stability and institutional robustness. The sort of society we have in mind is rife with justice pluralism and with social experiments. It is truly open-ended: Little is settled beyond the limiting principle, and the basic structure of society is organized for discovery. To appeal to a well-worn analogy, an open society is like Neurath's boat. It has an architectural structure that persists despite the total replacement of its parts. The architecture is given by the limiting principle, and social experiments allow us to gradually replace old institutional planks with new and better ones.

The limiting principle will specify the bounds of the acceptable. The constitution will include rights to freedom of expression and association, along with the delegation of powers characteristic of a federal order. Members of the open society must endorse these constitutional essentials based on their diverse senses of justice. Not all members will find the limiting principle and its constitutional implications best, but they must all recognize it as more or less just and so legitimately governing their lives. The political order must secure assurance that others are prepared to comply with the limiting principle. There are compelling reasons to be skeptical that norms of public discourse are required (Thunder 2006) or that such norms are sufficient to provide assurance (Thrasher and Vallier 2015). Instead, we suspect that constitutional stability will generate assurance through public acts of compliance with the demands of the constitution, especially by political officials (Kogelmann and Stich 2016; Thrasher 2016).

With proper protections of expression, association, and jurisdictions, we expect institutional rules to differentiate and evolve according to the dynamics between reformers and conservatives in each group or subunit. No encouragement of innovation by the state is necessary. Given the dynamics of diversity in a free society, social change will be the norm. However, legal systems must not be in perpetual flux; so, we may need to establish legal rules at the state or provincial level and additional legislative mechanisms to keep laws in place long enough to run legal experiments. Social change may lead to new equilibria on social practices on its own, but in some cases, we may need a legal order to ensure sufficient robustness at the institutional level. Again, the legal order is not meant to return us to an original institutional equilibrium, but rather to stabilize changes so that social progress can proceed by means of social experimentation. If we are to foster robustness rather than stability, we must err on the

side of allowing more social change given the rigidity of an open society's constitution. While it should be quite difficult to amend the national constitution, the legal order of federal subunits can be much more fluid. We might find, for instance, that an open society has supermajoritarian voting rules at the national level to bias the order in favor of its stable status quo, but that it employs simple majority rules at lower levels of social organization.

In an open society with constitutional stability and institutional robustness, citizens must be concerned, as we noted above, about the potential for conflict between these two forms of stability. Again, excessive constitutional stability might degrade institutional experimentation, whereas institutional robustness might bubble up and undermine the semantic or functional stability of the constitutional order. Our hope is that there will be sufficient congruence between constitutional stability and institutional change, with both governed by the limiting principle. Insofar as the limiting principle specifies a set of basic rights that facilitate institutional experimentation, the experiments themselves should tend to respect those rights. The potential for instability will be most serious when constitutions impose legal rules that frustrate experimentation. A federalist order that protects freedom of expression and association should reinforce a barrier between the two levels of social organization we discuss. We admit, though, that there is probably no mechanical or structural feature that will eliminate friction between institutional change and constitutional rigidity; but we have reason to hope that the two kinds of stability will be congruent.

We think that the open society is a great improvement over the well-ordered society familiar to political liberals. The open society allows for justice pluralism and for more social experimentation than the well-ordered society. Further, given increasing disagreement in free societies, we need a social system that facilitates diversity and discovery.

## References

- Aligică, Paul Dragoș, and Vald Tarko. 2013. "Co-Production, Polycentricity, and Value Heterogeneity: The Ostroms' Public Choice Institutionalism Revisited." *American Political Science Review* 107(4): 726–41.
- Anderson, Elizabeth. 2006. "The Epistemology of Democracy." *Episteme* 3(1–2): 8–22.
- Bicchieri, Cristina. 2006. *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press.
- Bicchieri, Cristina. 2010. "Norms, preferences, and conditional behavior." *Politics, philosophy & economics* 9(3): 297–313.

- Bicchieri, Cristina. 2016. *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press.
- Bruner, Justin P. 2015. "Diversity, Tolerance, and the Social Contract." *Politics, Philosophy & Economics* 14(4): 429–448.
- Dryzek, John S., and Simon Niemeyer. 2006. "Reconciling Pluralism and Consensus as Political Ideals." *American Journal of Political Science* 50(3): 634–49.
- Gaus, Gerald. 2011. *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*. Cambridge: Cambridge University Press.
- Gaus, Gerald. 2016. *The Tyranny of the Ideal: Justice in a Diverse Society*. Princeton, NJ: Princeton University Press.
- Hardin, Russell. 2003. *Liberalism, Constitutionalism, and Democracy*. Oxford: Oxford University Press.
- Hart, H. L. A. 1961. *The Concept of Law*. Oxford: Clarendon Press.
- Hayek, F. A. 2011. *The Constitution of Liberty: The Definitive Edition*. Ed. Ronald Hamowy. Chicago: University of Chicago Press.
- Hong, Lu, and Scott E. Page. 2004. "Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers." *Proceedings of the National Academy of Sciences of the United States of America* 101(46): 16385–89.
- Kogelmann, Brian. 2017. "Justice, Diversity, and the Well-Ordered Society." *The Philosophical Quarterly* 67(269): 663–684, <https://doi.org/10.1093/pq/pqw082>
- Kogelmann, Brian, and Stephen Stich. 2016. "When Public Reason Fails Us: Convergence Discourse as Blood Oath." *American Political Science Review* 110(4): 717–730.
- Kukathas, Chandran. 2003. *The Liberal Archipelago*. Oxford: Oxford University Press.
- Landemore, Hélène. 2012. *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton, NJ: Princeton University Press.
- Laudan, Larry. 1981. "A Confutation of Convergent Realism." *Philosophy of Science* 48(1): 19–49.
- Levy, Jacob T. 2007. "Federalism, Liberalism, and the Separation of Loyalties." *American Political Science Review* 101(3): 459–77.
- Levy, Jacob T. 2014. *Rationalism, Pluralism, and Freedom*. Oxford: Oxford University Press.
- Lewis, David. 1969. *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Lister, Andrew. 2017. "Public Reason and Reciprocity." *Journal of Political Philosophy* 25(2): 155–72.
- Mill, John Stuart. 1978. *On Liberty*. 8th ed. Ed. Elizabeth Rapaport. Indianapolis, IN: Hackett.
- Moehler, Michael. 2014. "The Scope of Instrumental Morality." *Philosophical Studies* 167(2): 435–51.
- Muldoon, Ryan. 2017. *Beyond Tolerance: Re-Imagining Social Contract Theory for a Diverse World*. London: Routledge.
- North, Douglass C. 2010. *Understanding the Process of Economic Change*. Princeton, NJ: Princeton University Press.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.
- Popper, Karl R. 2013. *The Open Society and Its Enemies*. Princeton, NJ: Princeton University Press.
- Quong, Jonathan. 2011. *Liberalism without Perfection*. New York: Oxford University Press.
- Rawls, John. 1996. *Political Liberalism*. New York: Columbia University Press.
- Rawls, John. 1999a. *A Theory of Justice*. Revised. Cambridge, MA: Belknap Press.
- Rawls, John. 1999b. "Reply to Alexander and Musgrave." In *Collected Papers*, ed. Samuel Freeman. Cambridge, MA: Harvard University Press, 232–53.
- Rawls, John. 2001. *Justice as Fairness: A Restatement*. Ed. Erin Kelly. Cambridge, MA: Harvard University Press.
- Robert, Henry M., and Sarah C. Robert. 1990. "Scott, Foresman Robert's Rules of order newly revised."
- Sabl, Andrew. 2012. *Hume's Politics: Coordination and Crisis in the "History of England"*. Princeton, NJ: Princeton University Press.
- Schotter, Andrew. 2008. *The Economic Theory of Social Institutions*. Cambridge: Cambridge University Press.
- Scott, James C. 1999. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CT: Yale University Press.
- Sen, Amartya. 2009. *The Idea of Justice*. Cambridge, MA: Harvard University Press.
- Sparrow, Robert. 2016. "If People Were Movies? Free Speech and Free Association." *Journal of Political Philosophy* 24(2): 227–44.
- Thrasher, John. 2016. "The Ethics of Legislative Vote Trading." *Political Studies* 64(3): 614–29.
- Thrasher, John, and Kevin Vallier. 2015. "The Fragility of Consensus." *European Journal of Philosophy* 23(4): 933–54.
- Thunder, David. 2006. "A Rawlsian Argument Against the Duty of Civility." *American Journal of Political Science* 50(3): 676–90.