Is Justice a Fixed Point?*

Abstract

Following the work of John Rawls, political theorists have fixated on the comparative stability of different equilibrium states of justice. This article identifies a crucial gap in this literature, namely, the lack of attention paid to non-equilibrating systems. Drawing on Kakutani's Theorem, I present plausible cases in which society will fail to exhibit any equilibrium states of justice. An important implication for political theory is that, rather than focusing exclusively on stable equilibria, theorists should examine dynamic processes of justice which need not exhibit equilibrating forces. To this end, a more useful concept than stability is that of *robustness*, or the ability of a system to maintain general desiderata in the face of an evolving conception of justice.

Wordcount (with references): 7,480

*ACKNOWLEDGEMENTS: [Redacted for anonymity]

Is Justice a Fixed Point?

> *An equilibrium centered view is essentially static and provides little insight into the transient behavior of systems that are not near the equilibrium. Natural, undisturbed systems are likely to be continually in a transient state; they will be equally so under the influence of man.*
>
> — C.S. Holling, "Resilience and Stability of Ecological Systems"

I. Introduction

"It is evident," John Rawls tells us, "that stability is a desirable feature of moral conceptions... However attractive a conception of justice might be on other grounds, it is seriously defective if the principles of moral psychology are such that it fails to engender in human beings the requisite desire to act upon it" (1999, 398). Rawls has inspired legions of political theorists to fixate on the comparative stability of different conceptions of justice (Chung 2020; Thrasher and Vallier 2018, 2015; Kogelmann and Stich 2015; Weithman 2010). Within this discussion, theorists typically examine different *justice equilibria*, that is, states in which a particular conception of justice "generate[s] its own support" over time (Rawls

1999, 138). To test stability, they search for subversive tendencies or imagine subjecting society to diverse shocks of various sizes (Chung 2019; Vallier 2017).

This perspective on justice, however, makes nontrivial assumptions about the nature of social-moral systems. To examine the properties of an equilibrium state, such as stability, one must first justify the assumption of an equilibrium state.[1] Before discussing *stability*, the theorist must discuss *existence*.

Showing that an equilibrium exists in a complex system like society can be difficult. Some such systems simply do not exhibit equilibrating tendencies. In fact, there is good reason to suppose that the dynamism and flux of contemporary society are the result of endogenous disequilibrium. This has been forcefully argued in the case of economic systems (Arthur 2015; Schumpeter 2008), as well as other social systems (Miller and Page 2007). It is reasonable, therefore, to suspect that social-moral systems may also resist equilibration. If this is the case, then to accurately characterize the nature of justice, political theory may need to refocus its gaze. As Holling explains in the epigraph, applying equilibrium analysis to nonequilibrium phenomena can blind the theorist to important features of the phenomena being examined.

In this paper I argue that philosophers have granted undue attention to just *states* of affairs, thus ignoring the *process* by which such states arise and are swept away. This paper begins by presenting a challenge for the *static approach*, i.e. the approach to political philosophy that examines the features of just *states*, rather than just *processes*. The problem is that equilibrium states, also known as *fixed*

[1] In this paper, the term "equilibrium" is used to refer to fixed point equilibria. It does not include cycles or any other type of pattern that some might refer to as an equilibrium.

points, often don't exist in complex, dynamic systems, such as society. The standard way of demonstrating the existence of an equilibrium state, Kakutani's theorem, fails to convincingly demonstrate the existence of justice equilibria. Instead, it points us towards plausible counterexamples in which equilibria do not exist. Static theory therefore faces an *existence problem*: the focus of analysis may be a phantom, something that never (or rarely) occurs in real social systems. In exploring ways to meet this challenge, section III suggests that, rather than attempting to prove the existence of an equilibrium state, *dynamic justice theorizing* should instead focus on nonequilibrium processes. In this approach, the traditional concern with *stability* is set aside and replaced by the more promising concept of *robustness*.

II. The Problem of Existence

Karl Popper influentially argued that, as a complex open system, society will not equilibrate. "Society," he claims, "is changing, developing. Its development is not, in the main, a repetitive one..... Conditions are changing, and situations arise (for example, in consequence of new scientific discoveries) which are very different from anything that ever happened before" (Popper 1963: 457). Many historians share Popper's belief, eschewing the search for general laws of history, and, as we will see, some have even extended this stance to conceptions of justice. How do we know that any conception of justice, including justice as fairness, can support an equilibrium in the form of a well-ordered society? In other words, what does it take for an equilibrium to exist, and under what conditions might we suspect that it does not?

The standard approach in the social sciences is to apply *Kakutani's Theorem*.[2] From game theory to general equilibrium theory, social scientists rely on Kakutani's Theorem when demonstrating the existence of equilibrium states in a model system. This theorem specifies a set of conditions, which, when satisfied, guarantee that there exists at least one fixed point as a mathematical fact. Understanding how these conditions might fail to be satisfied will be instructive in assessing the plausibility of the static theorist's assumption that there exists some fixed point of justice.[3]

For our purposes, we can provide an interpretation of Kakutani's Theorem that translates it into a theorem about justice:[4]

-Let A be a set of points in $\mathbf{R}^n$ that contains all possible "justice states" of society, with n being the number of justice-relevant parameters. Hence, every state of society is described by a vector of length n, which is why A is a subset of $\mathbf{R}^n$.

-f is a function that maps one justice state, any point in A, to another point contained within A.[5]

-There are four conditions required for Kakutani's Theorem to apply:

[2] This theorem is used in two of the most celebrated advances in the history of economic theory: Nash's (1950) proof that all finite games have equilibrium solutions, and the Arrow-Debreu (1954) proof of the existence of general equilibrium for a competitive economy. Rawls was heavily influenced by contemporary economics, and his technical terms, such as "general reflective equilibrium," are likely drawn from his exchanges with economists (Wolff 1977).

[3] As discussed below, these conditions are not necessary for the existence of a fixed point, but relaxing any one of them allows for us to produce counterexamples, some of which are plausible accounts of how society actually works.

[4] A more general, formal statement of Kakutani's theorem can be found in Osborne and Rubinstein (1994, 20).

[5] It could also map to subsets, but we use points for present expository purposes. See Osborne and Rubinstein (1994, 20).

C1) Suppose the set of possible justice states, A, is *compact*: it contains its own boundaries. That is, the extreme values that any of the n dimensions might take can be realized (A is closed), and these n dimensions do not extend outward to infinity (A is bounded).

C2) Suppose that the set of possible justice states A is *convex*: it includes all intermediate states between any two of its possible states.

C3) Suppose f has a *closed graph*: if f can map arbitrarily close to some justice state x^*, then there is some justice state x that will map directly onto x^*.

C4) Suppose $f(x)$ is *non-empty* for all x in A. That is, a justice state x always maps to one (or more) other justice states via f, even if this means that the state x maps to itself.

Relying on Rawls's framework (1999, 400-30), a "justice state," x refers to some publicly held conception of justice. The transition function f refers to the process that takes society from its public conception, through the institutional manifestation of that public conception, and back to a new public conception of justice. When (C1) - (C4) are satisfied, a fixed point *must* exist: there will be some well-ordered society, under some public conception of justice, that is in equilibrium, meaning that it maps onto itself.

Crucially, following Rawls, this model focuses on the public conception of justice, not its institutional manifestation per se. The justice state x is not an institutional state, but a state in which a particular conception of justice publicly prevails. This will, of course, have ramifications for the institutional structure. But Rawls, along with many others, recognizes that the same conception of

justice might call for different institutional specifications under different circumstances (1999, 400-1). The question this model investigates is whether the public conception of justice itself will equilibrate, regardless of any institutional variations. That is, will citizens living in a well-ordered society organized around a conception of justice *x* continue to publicly endorse *x* over time?

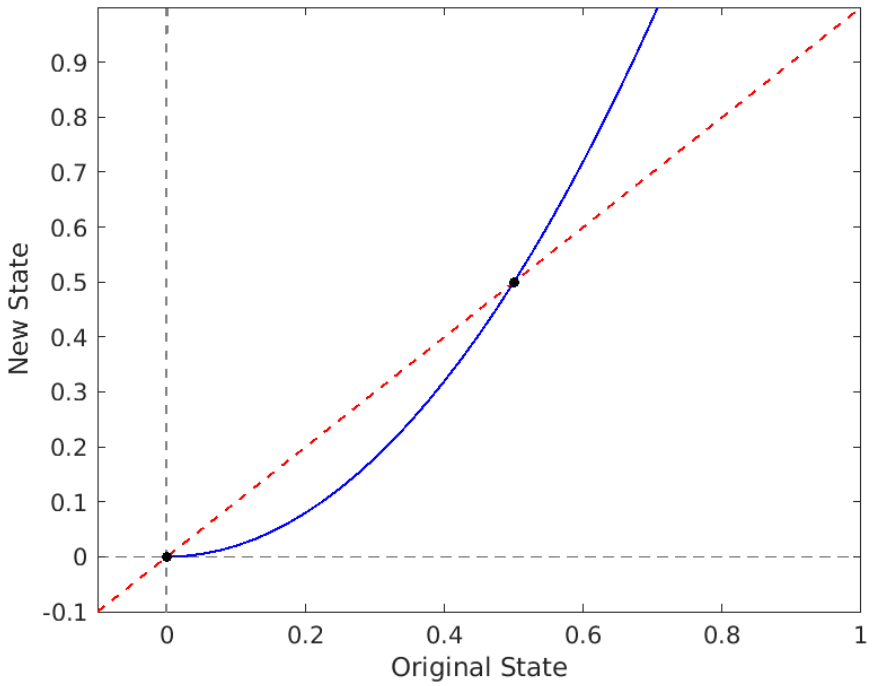

Figure 1: Two fixed points exist, one at (0, 0) and one at (.5, .5).

Graphically, a fixed point will be any point that falls along the 45 degree line represented by $g(x) = x$ (the dashed diagonal line in figure 1). If a function, representing a system, can be drawn that relates all original states to new states *without* intersecting the dashed line $g(x) = x$, then the system has

no fixed points. One way of understanding Kakutani's Theorem is as a set of conditions which ensure that $f(x)$ will, in at least one place, intersect the line $g(x) = x$.

Note that, while these conditions are *sufficient* for a fixed point to exist, they are by no means *necessary*. However, relaxing any one of these conditions allows for counterexamples: we can find some system in which no equilibrium exists. If these counterexamples, once interpreted in concrete terms, sound like plausible descriptions of our own social-moral systems, then we have reason to doubt that there exist any fixed points of justice. The argument, therefore, aims to shift the burden of proof. If there are plausible construals of society in which no justice equilibrium exists, then the theorist may not simply assume that these fixed points of justice exist. The theorist must *prove this*, or at least provide some reason for us to believe it.

Let's start with the simplest possible graphical representation of a fixed point by assuming that $n = 1$. That is, there is only one dimension that is relevant to justice. Moreover, we can measure this justice-relevant feature on a scale of 0 to 1. Thus, let $A = [0,1]$, and let f: $[0,1] \rightarrow [0,1]$ be defined so that $f(x) = 2x^2$. It is easy to check that this system, represented in figure 1, will satisfy all of the conditions of Kakutani's Theorem.

One possible interpretation takes x to be the level of some dimension of justice, say, social equality. In this case, 0 represents the limit of extreme inequality, and 1 represents full equality. On the x-axis are arrayed all possible levels of social equality that a society might endorse as its public conception of justice. The y-axis represents all of the new states that the original states might map onto. The solid line represents the laws of transition, that is, the function f shows how an original state will map onto a new state. Again, the dashed line, $g(x) = x$, is simply there for reference: any point falling on

the dashed line will, by definition, be a fixed point, since on this line the original state and the new state are equivalent. Our particular function f says that when society is unequal, i.e. $x < .5$, social equality tends to decrease until society is entirely despotical. However, f also shows that equality is self-reinforcing in the sense that a society with a high measure of equality will tend to increase in social equality. From Kakutani's Theorem, we know that our graph must possess at least one fixed point. Intuitively, this fixed point should occur when the forces pulling us towards less social equality perfectly balance with those pushing us towards greater social equality. As we can see, this fixed point occurs at $x = ½$. In fact, there exist two fixed points, since the equation: $f(x) = 2x^2 = x$ has two solutions: $x = 0$ and $x = ½$. At both of these points, f maps x onto x. The well-ordered society is in equilibrium when there is either an intermediate level of social equality or when society is entirely unequal.

This example is helpful for visualizing a fixed point, but it is rather unrealistic. What happens, for example, above $x = ½$? According to the graph, equality will continue to increase indefinitely, since the "new state" is above the dashed line, i.e., the new state is more equal than the original state for all $x > ½$. But then, the original state $x = 1$ will map to some new justice state, x, that's greater than 1. Given that equality is measured on a scale of [0,1], this model hardly makes sense. Let's consider another example. For instance, suppose $f(x) = x^2$ instead of $2x^2$. Then we get the following:

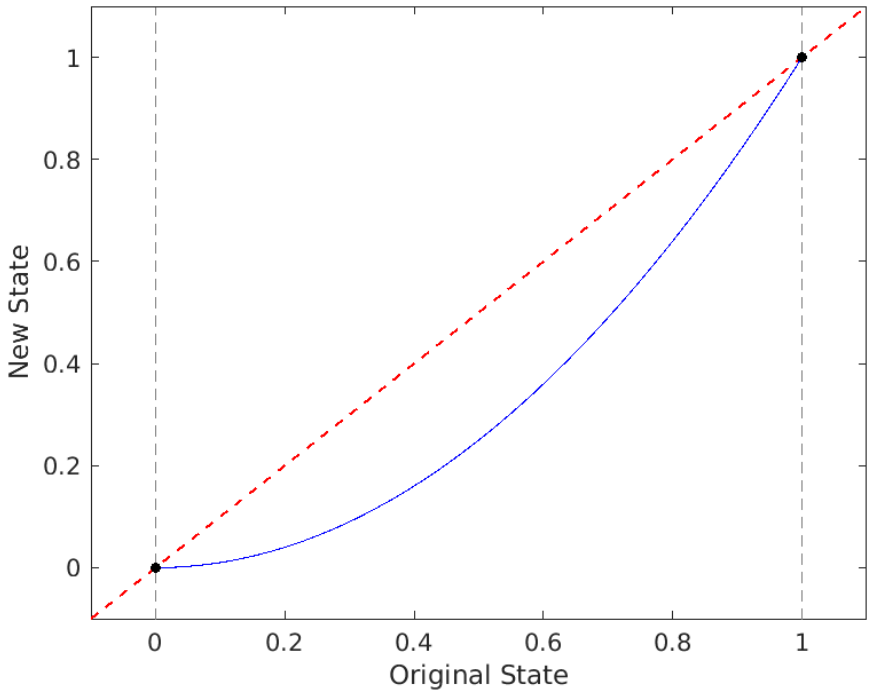

Figure 2: Two fixed points exist, one at (0, 0) and one at (1, 1).

In this new system, the only two fixed points are total equality or total despotism.

Now we can begin to cash in on all of this formalization. Conditions (C1) and (C2) concern the nature of the set of possible justice states, that is, the shape and properties of the set of public conceptions that a society may come to endorse. Conditions (C3) and (C4), on the other hand, concern the nature of the transition function *f*, that maps from one justice state to another. Rather than considering a counterexample for each condition, we will examine three potential cases, one where the *set* of justice states fails to satisfy the conditions of Kakutani's theorem and two where the *transition function* fails to do so.

Consider what happens if our society fails to satisfy (C1). Recall that (C1) stipulates that our space of possible justice states, previously [0,1], must be *compact*, which means, roughly speaking, that

it contains its (finite) boundaries. To violate (C1), let's replace [0,1] with the very similar, though not compact, space (0,1). In terms of our example, this is tantamount to supposing that there is no way for society to occupy a state of equality with a score of 0, i.e. total despotism, or a score of 1, i.e. perfect equality. We can get extremely close — arbitrarily close even — but, due to the confluence of myriad social and natural facts, the universe simply will not permit a justice state to realize perfect (in)equality. The result is figure 3:

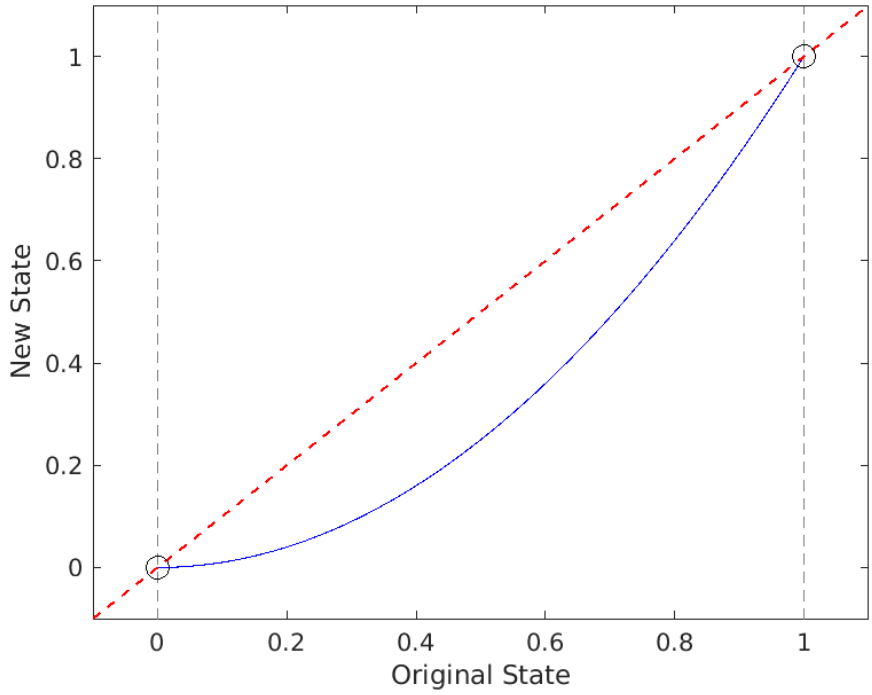

Figure 3: By violating (C1), we have eliminated the fixed points at (0, 0) and at (1, 1).

Notice that both of our fixed points have vanished: the points (0,0) and (1,1) have been replaced with empty circles to show that these are no longer valid inputs or outputs. With an extremely minor

change in the nature of the set of possible justice states — we reduced this set by an *immeasurably* small amount — we have eliminated the existence of any fixed point of justice. Moreover, the supposition employed to do so is not unrealistic. It's not difficult to imagine that we can approach full equality or full despotism without ever completely realizing either.

Perhaps the counterexample depicted in figure 3 seems to be of mostly theoretical interest, since the system will asymptotically approach the point (0,0), even if it never arrives. This seems, on the one hand, unrealistic, and on the other, pretty similar to a fixed point. But this is a weak interpretation. Perhaps the story accompanying the graph is that, however close we get to the ideal of total equality, society is constantly pulled closer and closer to total despotism, and dissolves upon reaching this state.[6] Under this plausible interpretation, the dynamics depicted in figure 3 accord with highly influential political theories that posit a degeneration into tyranny. For instance, one might read Book VIII of Plato's *Republic* in just this manner. Another example is Marx's theory that the capitalist class will control an ever greater share of wealth, while imposing a capitalist ideology on the populace, ultimately resulting in a revolutionary regime change. In these theories, society is drawn towards tyranny, but total tyranny is not a fixed point, since society will not subsist (or will be radically altered) upon reaching this state.

A very similar counterexample serves to demonstrate that violating (C2) also permits systems lacking fixed points of justice. However, different and more interesting counterexamples arise when we focus on conditions (C3) and (C4), which concern the transition rule f. If this transition rule is

[6] Technically, society will never reach 0, but it will get so close that even infinitesimally small error or noise will push it there.

undefined for certain states or discontinuous at certain points, a system may not exhibit any fixed points. What would these abstract properties look like in a real social-moral system? Gustav Schmoller offers one possible interpretation:

> ...the same theory which proposes a demand of justice as its consequence often is made by individuals, but repudiated by public opinion; and then suddenly with irresistible elementary force it takes hold of the masses, leads them on new paths, radically influences legislation and puts a changed stamp on whole epochs. (Schmoller 1894, 712)

A public conception of justice might change slowly and continuously, like a stick gradually bending into an arc, until it reaches a critical point where the stick snaps, disrupting the continuous trend that preceded this new state. The snapping of the stick may correspond to the dissolution of anything recognizable as a "public conception of justice," or it may correspond to an entirely new trajectory along which the public conception of justice will evolve. There is no way to rule out the possibility of such social-moral dynamics *a priori*, and indeed, a cursory glance at social movements reveals fits and starts that belie the assumption of smooth continuity.[7]

This idea of building pressure and sudden change corresponds to a common violation of Kakutani's theorem, namely (C3), which demands that the graph of f be *closed*, i.e., it should contain all of its endpoints. In particular, one way to violate this condition is to introduce a discontinuity in

[7] Many historical examples of discontinuous transitions are examined by Kalvyas (2006).

society's conception of justice. Perhaps, to take a simplistic example, society is committed to equality even if equality is declining for some lengthy period of time. There may come a critical point, however, call it the "threshold point," when society undergoes a major transition, and finds itself pulled toward a completely different basin of attraction.[8]

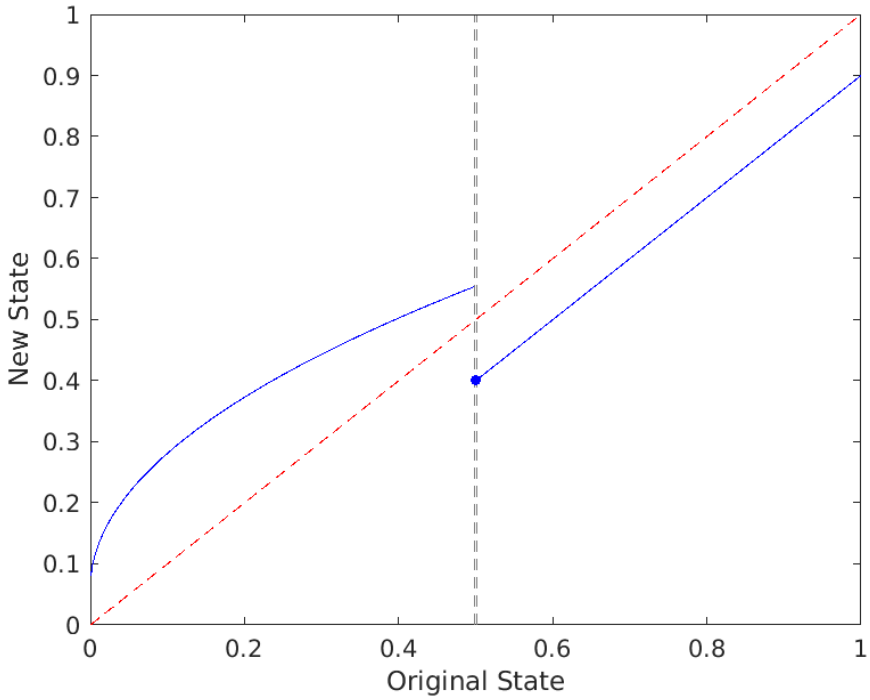

Figure 4: By violating (C3), a discontinuous transition function can exist that does not exhibit any fixed points.

In figure 4, there are no fixed points. A discontinuity occurs at $x = .5$, where society finds its public conception of justice to be radically transformed. Supposing the .5 threshold is crossed, a different

[8] Aligica and Tarko (2014) discuss "slippery slopes" that lead to discontinuities, or "catastrophic thresholds."

mindset takes over and citizens become more prone to rejecting equality. At no point does the system equilibrate. There is no justice state that maps onto itself.

The introduction of discontinuity provides a graphical representation of Schmoller's description of sudden shifts in public conceptions of justice. Again, our one-dimensional analysis of justice is highly stylized and rules out more interesting dynamics. As depicted, a society approaching $x = .5$ from the right would leap to a new state, located on the other curve, and then forever flutter about without settling down. Although more interesting dynamics would be possible with a higher-dimensional model,[9] the representation allows us to visualize one of the ways in which a public conception might fail to equilibrate. Systems with discontinuous dynamics may lack equilibria.

Schmoller's view that society exhibits discontinuous changes finds support in contemporary social science. Following Schumpeterian lines of reasoning, complexity economists have argued that endogenous novelty prevents equilibrium states from arising in the economy. "[N]onequilibrium," Brain Arthur assures us, "is the natural state of the economy" (Arthur 2015, 5).[10] Just as the economy might be approaching a fixed point, a new technology, product, or organizational form appears, and "a gale of creative destruction" radically alters the economic situation (Schumpeter 2008, 84). Analogously, we might imagine cultural or "ideological entrepreneurs" who produce "discontinuous change" in our public conception of justice (Storr 2009, 105).

[9] A multidimensional model is developed below.

[10] One source of disequilibrating novelty is simply the recombination of past ideas (Gaus forthcoming, sec. 15.4; Arthur 2015, 141; Beinhocker 2007, 249-53; Axelrod and Cohen 2000, 42-3).

There is one more counterexample worth considering. Focusing on (C4), we might relax the assumption that, for any justice state x in the domain [0,1], $f(x)$ is defined. That is, some justice states will fail to map onto another justice state. Violating this condition might be interpreted in many ways. Some states may lead to the disappearance of any and all consensus on justice, so that a society in such a state simply ceases to possess any recognizable public conception of justice. Or perhaps a conception of justice is so corrosive that its wide adoption among the populace leads to the dissolution of society altogether. Alternatively, remaining within a Rawlsian framework, an undefined area in the domain could indicate that one of the subjective or objective circumstances of justice fails to hold (Rawls 1999: 109-10). According to Rawls, a conception of justice would then have no role to play, and citizens would not require (or be able to instantiate) principles of justice in their institutions. There are other possible interpretations, as well, but this suffices to show that condition (C4) is not guaranteed to hold for all social systems, even within Rawls's conceptual schema.

The last two counterexamples were confined to a one-dimensional justice space, greatly limiting the behavior of the system. Adding more dimensions would allow for a much larger variety of non-equilibrium behaviors. In a scenario with two justice-relevant dimensions, for example, a violation of (C4) could involve a scenario in which the justice states continually oscillate around some point, never settling down into a static equilibrium state. Although it will unfortunately complicate the simple model, a representation of the two-dimensional scenario will also provide substantial insight.

An illuminating graph of the new dynamical system calls for a different kind of plane, known as the "phase space." Rather than placing the current state on the x-axis and the new state on the y-axis, as above, the new plane places one justice-relevant parameter on the x-axis and a different

justice-relevant parameter on the *y*-axis. Lines on the graph, accordingly, represent trajectories of two-dimensional states as they evolve over time. The point of this new plane is to represent the relationship between the two variables. For this illustration, let us select two justice-relevant variables to which Rawls grants paramount importance. Let *F* represent the extent to which citizens endorse and strive to implement Rawls's first principle of justice and *E* represent the extent that they do so for his second principle of justice. For purposes of exposition, let *F* be referred to as *freedom* and *E* as *equality*. To keep the equations simple, we will measure both *F* and *E* on a plane centered at the origin. Suppose, as is likely the case, that the extent to which freedom and equality shape society's institutions affects how citizens view and prioritize these two values. The simplest relationship of this kind would result in a linear system of differential equations:

$$\frac{dE}{dt} = aE + bF \qquad (*)$$

$$\frac{dF}{dt} = cE + dF$$

where *a, b, c, d* are real parameters, and the left hand sides of each equation represent the change over time of equality (*E*) and freedom (*F*), respectively. In this scenario, the transition function *f* will take two inputs, one value for *E* and one for *F,* and will map to a new point in the next period. The change induced by *f* will no longer be represented by the difference between the *x* and *y* values, but by arrows indicating the direction of a trajectory.

For different parameter specifications, the system (✶) will behave in very different ways. For instance, if we fix $a = d = 0$, and if c and b have opposite signs, then the system is guaranteed to exhibit cycles. The origin will always be a fixed point, but given that we have relaxed assumption (C4), f need not be defined for all input pairs. Consequently, if f is undefined for any region containing the origin, this system will not exhibit any fixed points, only cycles. There are, therefore, infinitely many possible systems exhibiting cycles and lacking equilibria (figure 5).

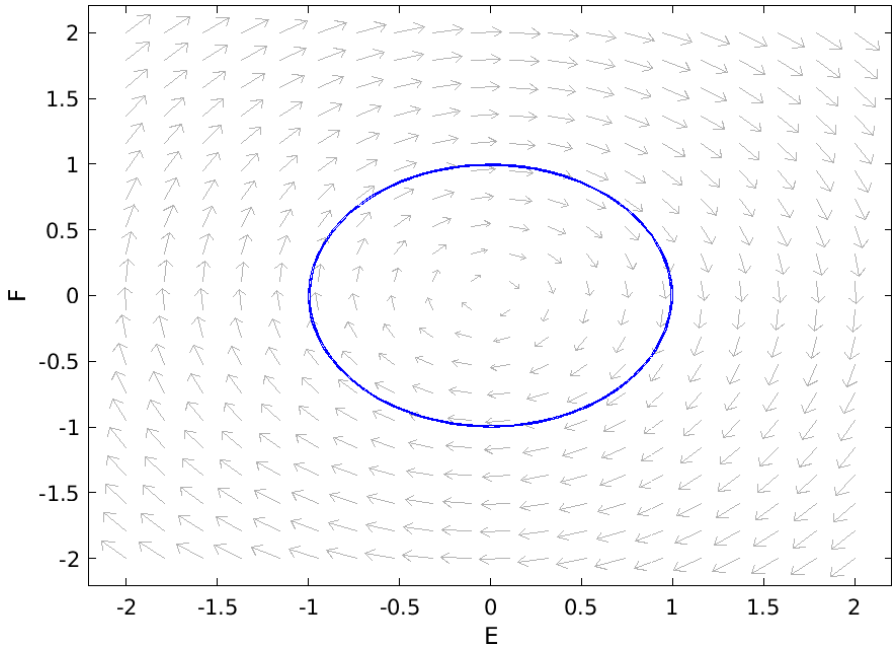

Figure 5: A multivariable system can exhibit cycles, one of which is indicated by the dark circle. By violating (C4), we can depict this system as lacking fixed points.

In broad outline, figure 5 conveys a story in which a greater preference for freedom produces a greater preference for equality. However, a greater preference for equality diminishes citizens' preference for freedom. Recalling the Rawlsian framework, the details of the story depicted in figure 5

will involve an bidirectional interaction between the public conception of justice and its institutional manifestation. Perhaps freedom is instantiated as economic liberalization, which creates socio-economic inequality, thus producing a reaction in thought and feeling that causes a greater desire and higher esteem for egalitarian values. And perhaps equality is instantiated in a more communitarian social organization that suppresses the desire for individual liberty. The exact details need not occupy us here. Instead, the key point is that the possibility of cyclical systems lacking equilibria calls upon static theorists to provide a more direct defense of the assumption that social-moral systems will contain fixed points.

Even this more complex model is, of course, highly simplified. Rawls's theory, as well as other static theories, include many intricate features not represented here. For example, both Rawls's liberty and equality principles are multi-part and are regulated by certain priority rules (1999: 266). Complications — for example additional variables or non-autonomous terms — could be introduced for the sake of realism or for textual fidelity. Most promisingly, the introduction of nonlinear terms would permit the emergence of more complex behavior, such as limit cycles, which aptly model a wide variety of cyclical phenomena (Strogatz 2015: 198). Though worthy of closer study, these more realistic models are unnecessary for the present argument. The point of this formalization is not to definitively prove that Rawls's conception of justice — or any other for that matter — will never equilibrate. Instead, the point is to indicate that static theorists need to justify the assumption of a fixed point. There are many scenarios in which a fixed point will not exist, some of which involve perpetual cycles.

Although this model is highly stylized, these stylized scenarios are analogous to processes described more richly by political theorists and social scientists. Political theorists should not dismiss

the cyclical view of society as empty formalism, for such views hold a place of prominence in the history of political thought.[11] In fact, such cyclical oscillations provide a dynamic representation of one of the most influential results in political theory: Arrow's impossibility theorem (Arrow 1970). If a society continually cycles between several justice states, then the social choice function necessarily violates transitivity. A more intricate version of the model in figure 5 could thus depict the scenario in which society, preserving Arrow's four conditions, violates transitivity in determining its public conception of justice. As many have noted, the prospect of intransitive social choice raises important normative questions.

Kakutani's Theorem has guided us in imagining simple scenarios where the sort of system envisioned by Rawls will fail to equilibrate. As argued above, one cannot begin to discuss the stability of an equilibrium until one has established that the equilibrium actually exists. Static theorists of justice thus face a challenge: they must justify the assumption of equilibrium. The most common way of demonstrating the existence of an equilibrium, viz. invoking Kakutani's Theorem, leads to difficulties. As this section has argued, the conditions of Kakutani's Theorem plausibly fail to hold in many scenarios, allowing for perpetual disequilibrium in justice systems. Perhaps there is some way to meet this challenge and, thus, to justify static theory. However, the next section argues that there is a promising alternative. Examining justice as a non-equilibrium process opens up a new avenue of research with the potential to yield novel insights into the nature of justice.

[11] See, for instance, Ibn Kalduhn (2015) and Peter Turchin (2003).

III. Dynamic Theorizing

In contrast to views that examine the stability of different conceptions of justice, the present paper has asked a different question: will the well-ordered society, if realized, be in equilibrium? In many situations, there will be no fixed point of justice. Society will flout the conditions of Kakutani's Theorem, opening up the possibility that no public conception of justice can be iteratively reaffirmed, even in the absence of exogenous shocks. Social systems, as many theorists have argued, may exhibit constant, endogenous disequilibrium. As John Miller and Scott Page put it, "Stock markets soar and crash... Political parties rise and topple..." — and conceptions of justice, I would add, surge and recede — "While the notion of social equilibria is an important one... we may need to go beyond equilibria to truly understand the social world" (Miller and Page 2007, 222).

Given the prevalence of endogenous disequilibrium in social systems, how do we theorize about justice? For those committed to focusing exclusively on static justice equilibria, there are two possible responses. One approach is to drastically narrow the focus of political theory. By restricting attention to cases in which justice equilibria actually exist, and can be shown to exist, justice theorizing may proceed as normal, with the qualification that it focuses only on these special cases. If continual autocatalytic change characterizes contemporary societies, however, then these cases will be the exception rather than the rule (Beinhocker 2006: 70-4, 99). The major drawback of this approach is, first, the difficulty of demonstrating the existence of equilibria, and second the highly contextual nature of any conclusions or prescriptions that the theorist offers. In being so constrained, static models will lack what scientists call "external validity." That is, even if a static theory identifies the

correct principles of justice for a particular time or place, changes in the justice-relevant variables imply that such theories will not generalize to different contexts.

Another possible response is to impose conditions on society that make it more amenable to static theorizing. This may be what Rawls had in mind when he entertains the notion of a steady state economy ([1993] 2001: 107, n.33). Commentators on Rawls have noticed the inconsistency between the notion of stable equilibrium and the dynamism of our contemporary, globalized economy (Schrepfer 2019: 159). Thus, one possible response to the endogenous change exhibited by today's complex, globalized social systems is to suppress social complexity, as many states historically have (Scott 1999; Hall 1985). By reducing the diversity and interdependence of contemporary societies, we might foster the existence and stability of static justice equilibria. The major drawback of this approach is that such simplifications will likely come at severe cost, not just to our material well-being, but also to our political freedom. Indeed, later Rawls concedes that a diversity of perspectives spontaneously arises as the natural result of free institutions and that the suppression of this process could come about only through the "oppressive use of state power" (2005: 37). The cost of facilitating stable fixed points, therefore, may be prohibitive.

Assuming that broad applicability is a desirable feature of a political theory and that the radical simplification of contemporary society is undesirable (or unattainable), then we must ask if there is an alternative approach to theorizing that adequately responds to the problem of existence. Indeed, there is: we may explore methods of theorizing that do not assume the existence of equilibria. The approach considered and criticized in section III, which I have called *static theory*, studies conceptions of justice insofar as they exhibit equilibrating properties. By contrast, the alternative approach, *dynamic theory*,

will examine social-moral systems exhibiting continual flux. These two approaches differ in their objects of study and, consequently, in their preferred methods.

With respect to the objects of study, static theories of justice consider the array of social values to be either exogenous or self-reinforcing and, thus, changeless. These values are held by citizens who must formulate just terms of engagement to facilitate cooperation. Rawls's version of static theory strives for a "reflective equilibrium," in which the principles underlying society's basic institutions match our coordinated, considered judgments (Rawls 1999: 17-8).[12] The static theorist thus focuses on a justice *state*, seeking to understand the institutions that will best instantiate and support the hypothesized justice state. By contrast, dynamic theory accepts that values are endogenous and evolving. The dynamic theorist thus focuses on how these constantly changing values can be accommodated and adjusted to maintain cooperation over time. That is, the dynamic theorist focuses on a justice *process*, seeking to understand how institutions will accommodate coordination amidst a network of agents whose considered judgments are reflexively adapting to the political and social structures they help to produce. Dynamic theory examines systems that permit constantly evolving rules in response to constantly evolving values, thus eschewing the assumption of fixed point equilibria of justice.

Given these different objects of study, the two types of theory employ very different methods. Static theorists tend to be utopian, because they take their task to involve characterizing a stable social equilibrium that instantiates the given values of its members. Again, taking Rawls as the exemplar, the

[12] See, also, Rawls's discussion of "general" reflective equilibrium in *Political Liberalism* (2006: 384, fn.16), which emphasizes the interpersonal, coordinated aspect of reflective equilibrium.

assumption of full compliance is introduced, since a fully just society will reinforce its own underlying values, constituting a fixed point of justice. The project that Rawls undertakes is thus "realistically utopian," since it aims to determine how "a democratic regime can attain complete realization of its appropriate political values" (2001: 13). For the dynamic theorist, there is no utopian social order. Instead, the basic fact of politics is disagreement and disputation: as values evolve and change over time, even a (previously) ideal social order will prove to be inadequate. Rather than a state without errors, dynamic theorists seek a process of error correction.[13]

Furthermore, since stability as a concept presupposes equilibrium, it is unavailable to the dynamic theorist. Rather than *stability,* therefore, the dynamic theorist should focus on the analogous concept of *robustness*. Thrasher and Vallier (2018) have recently offered an important exploration of this concept. For Thrasher and Vallier, robustness is not a feature of a system in equilibrium, but of a process of equilibration: a robust system "tends toward equilibria" (406). This is a major advance over prevailing static theories, but the reference to equilibria is insufficiently radical, since, as argued above (sec. III), social systems may not exhibit any equilibrating tendencies. Luckily, referencing equilibria is neither necessary nor standard in definitions of robustness. In general systems theory, robustness often refers to a system's ability to "maintain functionality" in the face of shocks (Page 2010: 8). Put differently, system robustness refers to "feature persistence" (Jen 2005: 10). A system is robust, in other words, when, despite exogenous and endogenous shocks, it maintains some of its qualitative features, even as it shifts between equilibria or remains in a non-equilibrium state. Accordingly, the task of the

[13] In some construals of the debate, this means that while static theorists skew *ideal* in their methodology, dynamic theorists skew *nonideal*. See Valentini (2012).

dynamic justice theorist is to identify desirable features that a process of justice might possess and to determine whether these features will be robust as society shifts between distinct states of justice.

As an example, consider the multivariable model depicted in figure 5. Here, society exhibits a constantly evolving conception of justice, one which goes through periodic cycles. A dynamic theorist, without assuming any fixed point conception of justice, may consider the flexibility and revisionary nature of this social-moral system to be a normatively desirable property. In fact, in the context of discussing the intransitivity of democratic procedures, James Buchanan defended the normative appeal of cyclical social choice. By allowing diverse perspectives to prevail at different points in time, the formal inconsistency of such collective decisions "provides one of the most important safeguards against abuse" (Buchanan [1954] 1999: 96).[14] Moreover, policies induce changes in beliefs and preferences, and "[i]individual values are... constantly changing; so a postdecision ordering may be different from a predecision ordering" (Buchanan [1954]1999: 99). If, following Buchanan, the dynamic theorist holds that cycles, or intransitivity in general, is a desirable feature of social choice, then the next step is to ask whether or not the social-moral system is *robust*: will this desirable property persist as the public conception of justice evolves? For the system depicted in figure 5, the answer is "yes." Although an exogenous shock will shift the system to a new trajectory, it will remain cyclical, thus preserving the intransitive nature of social choice. If the answer were "no," the dynamic theorist would consider which institutional reforms could secure greater robustness.

This example reinforces the point expressed by Holling in the epigraph about the advantage of a dynamic approach. Whichever (defined) point in the phase space the theorist selects as most just,

[14] Cf. Buchanan and Congleton (1998: 28).

society will resist any attempt to pin it there. To the static theorist, therefore, the fluctuating and inconsistent system could only be viewed as a messy failure to instantiate or maintain a state of justice. By contrast, the dynamic theorist observes features of the process itself. This adds an additional dimension to the analysis; what appears irrational or immoral to the static theorist manifests an underlying order and value to the dynamic theorist. The static theorist sees only the failure to remain at the correct state, while the dynamic theorist sees a process of error correction in response to evolving values.

Although static theory, *on its own*, is deficient, this does not mean that there is no value to static theory or that the static theorist and the dynamic theorist are necessarily antagonists. To the contrary, these two approaches can complement one another. The important distinction between processes and states parallels a distinction between *process desiderata* and *state desiderata.* Some, though not all, state desiderata are *sectarian* in that they are internal to a particular conception of justice. For example, in a strict egalitarian conception, a society is more just to the extent that it prohibits unequal material holdings. On the other hand, some state desiderata are *general* to all conceptions. For example, reducing violent conflict is a desirable feature of a conception of justice whether one is an egalitarian, a libertarian, a conservative, a socialist, or whatever. The static theorist can complement the dynamic theorist by identifying such general desiderata, and determining the institutions that enable them to persist as stable equilibria. While it is unlikely that whole conceptions

of justice will constitute stable equilibria, certain values seem to transcend time and place, even in cases when our institutions sadly fail to instantiate these values.[15]

Sometimes these transcendent values are process desiderata, in which case they fall within the purview of the dynamic theorist. Sometimes, on the other hand, these transcendent values are state desiderata, in which case the static theorist can make an important contribution. On this analysis, static theory can make an important contribution, but historically it has made a twofold error: (1) an error of *commission* in thinking that whole conceptions of justice, rather than a few general values, will likely remain in equilibrium, as well as (2) an error of *omission* in neglecting the nonequilibrium processes that, I have argued, are crucial to understanding justice.

Post-Rawlsian political philosophy has seen a burgeoning interest in dynamic theorizing about justice. A central figure within this new program, the late Gerald Gaus, explores the implications of complexity, diversity, and dynamism in several works, most notably *The Tyranny of the Ideal* (2017). Gaus argues that in the process of pursuing an ideal, we learn more about our institutions and our values. As a result, individuals must reevaluate the features of that ideal, as well as their criteria of evaluation. What was once an ideal no longer is; the "fitness landscape" has shifted, and the search must go on.

According to Fred D'Agostino, Gaus, along with several others,[16] forms part of a "new program," which "more or less abandons the project of providing an end-state description of a justified

[15] For instance, about half of the ten commandments seem to be cross-cultural.

[16] Especially notable here are Ryan Muldoon (2016) and Julian Müller (2019), who have both contributed book-length explorations of dynamic theory.

social order" (D'Agostino 2016: 31). Instead, these theorists focus on processes that enable stakeholders to instantiate, legitimate, and continuously reevaluate their own social and political arrangements. Future work on dynamic theories of justice promises to be one of the most fruitful and exciting developments in post-Rawlsian political thought.

IV. Conclusion

This paper has argued that static justice theorizing faces severe limitations in a dynamic system characterized by constant disequilibrium. Theorists such as (early) Rawls have insisted that the requirements of justice are static and unchanging, even if their institutional manifestations may shift over time. However, drawing on Kakutani's Theorem, it is easy to imagine cases where the justice equilibrium, far from being stable, will not even exist. How can the justice theorist accommodate the possibility of nonequilibrium?

I have argued that the key is to focus on robustness, rather than stability, where robustness is understood as maintaining certain *process desiderata* in the face of changing institutions and conceptions of justice. This requires the justice theorist to identify these desiderata and to study diverse institutions and their respective abilities to satisfy these desiderata. Doing so will allow her to illuminate an otherwise obscure aspect of any system of justice. In particular, the process of disputation, adjudication, and resolution comes to the fore.

References

Aligica, Paul D., and Vlad Tarko. 2014. "Institutional Resilience and Economic Systems: Lessons from Elinor Ostrom's Work." *Comparative Economic Studies* 56:52-76.

Arrow, Kenneth. 1970. *Social Choice and Individual Values*. New Haven, CT: Yale University Press.

Arrow, Kenneth J., and Gérard Debreu. 1954. "Existence of an equilibrium for a competitive economy." *Econometrica* 22 (3): 265-290. https://doi.org/10.2307%2F1907353.

Arthur, Brian. 2015. *Complexity and the Economy*. New York: Oxford University Press.

Axelrod, Robert, and Michael D. Cohen. 2000. *Harnessing Complexity*. New York: Basic Books.

Beinhocker, Eric. 2006. *The Origin of Wealth: The Radical Remaking of Economics and What it Means for Business and Society*. Cambridge, MA: Harvard Business Review Press.

Buchanan, James. 1954/1999. "Social Choice, Democracy, and Free Markets." In eds. Geoffrey Brennan, Hartmut Kliemt, and Robert D. Tollison, *The Logical Foundations of Constitutional Liberty*. Indianapolis: Liberty Fund.

Buchanan, James and Roger D. Congleton. 1998/2003. *Politics by Principle, Not Interest*. Indianapolis: Liberty Fund.

Chung, Hun. 2019. “The Instability of John Rawls’s ‘Stability for the Right Reasons’.” *Episteme* 16, no. 1 (March): 1-17.

Chung, Hun. 2020. “The Well-Ordered Society Under Crisis: A Formal Analysis of Public Reason vs. Convergence Discourse.” *American Journal of Political Science* 64, no. 1 (January): 82-101.

D'Agostino, Fred. 2009. “From the Organization to the Division of Cognitive Labor.” *Politics, Philosophy & Economics* 8, no. 1 (February): 101-129. https://doi.org/10.1177/1470594X08098873.

D'Agostino, Fred. 2018. “How Can We Do Political Philosophy.” *Cosmos + Taxis* 5, no. 2: 29-37.

Gaus, Gerald. 2013. “The Turn to a Political Liberalism.” In *A Companion to Rawls*, 233-250. New York: Wiley.

Gaus, Gerald. 2017. *The Tyranny of the Ideal*. Princeton NJ: Princeton University Press.

Gaus, Gerald and Chad van Schoelandt. 2017. “Consensus on What? Convergence for What? Four Models of Political Liberalism.” *Ethics* 128, no. 1 (October): 145-172

Hall, John A. 1985. “Capstones and Organisms: Political Forms and the Triumph of Capitalism.” *Sociology*, 19(2), pp.173-192.

Jen, Erica. 2005. “Stable or Robust? What's the Difference?” Edited by Erica Jen. In *Robust Design*, 1-20. New York: Oxford University Press.

Kalduhn, Ibn. 2015. *The Muqaddimah*. Translated by Franz Rosenthal. Princeton, NJ: Princeton University Press.

Klosko, George. 1988. “The Nocturnal Council in Plato’s Laws.” *Political Studies* 36 (1): 74-88.

Kogelmann, Brian. 2017. "Justice, Diversity, and the Well-Ordered Society." *The Philosophical Quarterly* 67 (269): 663-685. https://doi.org/10.1093/pq/pqw082.

Kogelmann, Brian, and Stephen Stich. 2015. "When Public Reason Fails Us: Convergence Discourse as Blood Oath." *American Political Science Review* 110 (4): 717–730.

Miller, John H., and Scott E. Page. 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton: Princeton University Press.

Muldoon, Ryan. 2016. *Social Contract Theory for a Diverse World*. New York: Routledge.

Müler, Julian. 2019. *Political Pluralism, Disagreement and Justice: The Case for Polycentric Democracy.* New York: Routledge.

Nash, John. 1950. "Equilibrium Points in N-Person Games." *Proceedings of the National Academy of Sciences of the United States of America* 36 (1): 48-9. 10.1073/pnas.36.1.48.

Osborne, Martin J., and Ariel Rubinstein. 1994. *A Course in Game Theory*. Cambridge, MA: MIT Press.

Page, Scott. 2010. *Diversity and Complexity*. Princeton, NJ: Princeton University Press.

Popper, Karl. 1963. *Conjectures and Refutations: The Logic of Scientific Discovery*. New York: Routledge.

Rawls, John. 1996. *Political Liberalism*. New York: Columbia University Press.

Rawls, John. 1974/1999. "A Reply to Alexander and Musgrave." In ed. Samuel Freeman, *Collected Papers*. Cambridge, MA: Harvard University Press.

Rawls, John. 1999. *A Theory of Justice: Revised Edition*. Cambridge, MA: Belknap Press.

Rawls, John. 1993/2001. *The Law of Peoples*. Cambridge, MA: Belknap Press.

Rawls, John. 2001. “Kantian Constructivism in Moral Theory.” Edited by Samuel Freeman. In *Collected Papers*, 303-58. Cambridge, MA: Harvard University Press.

Schmoller, Gustav. 1894. “The Idea of Justice in Political Economy.” *The Annals of the American Academy of Political and Social Science* 4 (March): 1-41.

Schrepfer, Matthew. 2019. "A Rawlsian Case for Economic Nationalism: Globalisation and Distributional Autonomy in the Law of Peoples." *Journal of Applied Philosophy* 36.1: 155-163.

Schumpeter, Joseph. 2008. *Capitalism, Socialism, and Democracy*. New York: Harper Perennial Modern Classics.

Scott, James. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CT: Yale University Press.

Storr, Virgil. 2009. “North’s Underdeveloped Ideological Entrepreneur.” In *The Annual Proceedings of the Wealth and Well-Being of Nations*. Vol. 1. Beloit, WI: Beloit College Press.

Strogatz, Steven H. 1994. *Nonlinear Dynamics and Chaos*. New York: Perseus Books.

Thrasher, John, and Kevin Vallier. 2015. “The Fragility of Consensus.” *European Journal of Philosophy* 23 (4): 933-54.

Thrasher, John, and Kevin Vallier. 2018. “Political Stability in the Open Society.” *American Journal of Political Science* 62, no. 2 (April): 398-409.

Turchin, Peter. 2003. *Historical Dynamics: Why States Rise and Fall.* Princeton, NJ: Princeton University Press.

Valentini, Laura. 2012. "Ideal vs. Non-ideal theory: A Conceptual Map." *Philosophy Compass* 7.9: 654-664.

Vallier, Kevin. 2017. "Three Concepts of Political Stability." *Social Philosophy & Policy* 34 (1): 232-259.

Weithman, Paul. 2010. *Why Political Liberalism?: On John Rawls's Political Turn*. New York.: Oxford University Press.

Wolff, Robert P. 1977. *Understanding Rawls*. Princeton, NJ: Princeton University Press.