Strategic Injustice, Dynamic Network Formation, and Social Movements

1. Introduction

In his book, *Strategic Justice: Conventions and Problems of Balancing Divergent Interests*, Peter Vanderschraaf (2018) defends a theory of justice as convention. This theory rests on the idea that conventions secure cooperation and that they are sustained by mutual expectations and the shared understanding of others' conditional disposition to comply with them. That people also sanction deviant behavior provides an added but necessary incentive for conformity. What Vanderschraaf calls "strategic justice" is an account of justice as convention in which a convention is a system of strategies that solves a particular kind of coordination problem among the members of a population. I will explain the details of this conception of convention and its relation to strategic justice shortly. But, at least for the purpose of this paper, the significance of this approach to conventions and justice is their ability to explain the emergence of a just social order without any explicit agreement or promise and without any centralized authority or sovereign.

My goal in this paper is to make a case for a conventionalist theory of *in*justice. I argue that an oppressive social order can emerge from individuals' rational decisions without a need for any explicit agreement or promise, for instance in the form of coercion or collusion, and without any centralized authority or sovereign that acts as a tyrant. What I call *strategic injustice* is a system of formal and informal conventions that results in oppression and emerges from individuals' strategic choices about their social interactions. To highlight the common roots of this conventionalist approach with strategic justice, I structure this paper as a response to three main questions that Vanderschraaf addresses in his book: What are the necessary circumstances of (in)justice? What explains the emergence of (un)just conventions or system of conventions? and What are the proper responses to injustice?

In response to the first question, I argue that individuals' ability to control some aspects of their interactions with others is a necessary condition for the dominance of unjust/unfair but

cooperative strategies. I provide simulation models of a fair division problem to show that unless they capture such an ability for individuals, the dominance of unjust strategies is an anomaly in the evolutionary models that explain the emergence of a cooperative social order. Thus, for injustice to be a likely outcome of social interactions, as testimonial and empirical evidence suggests, individuals' discretion over their interactions should be added to the list of necessary background conditions that Vanderschraaf provides.

I use the second question to show that strategic injustice is possible and, despite what Vanderschraaf tangentially suggests, its unjust conventions are not a product of coincidence or random background conditions. I argue that a dynamic network formation model can best explain the emergence and resilience of unjust conventions, or even unjust social orders, that satisfy the conditions of strategic injustice. Therefore, strategic injustice is possible and the fact that individuals' strategic choices explain its emergence testifies that it is not a product of coincidence. Finally, to respond to the third question, I rely on the effects of the feedback loop between individuals and their networks to show that any interventions that dismiss its role, like uncoordinated individuals' actions or even purely top-down interventions, are likely to be futile. In fact, maintaining a rough equality that would prevent exploitation and other forms of injustice turns out to be another multi-agent coordination problem. I suggest that something like a social movement that restructures the network of social relations is necessary to solve such a coordination problem.

2. Strategic Justice

Strategic justice is a conventionalist approach because it relies on key features of conventions to answer important questions about justice, such as what the necessary circumstances of justice are, what explains the emergence of a just social order, what makes a system unjust, and how to address injustice. Two main bodies of work create the historical and conceptual foundations of strategic justice: one focuses on Hume's approach to conventions and justice and the other on David Lewis's game theoretic conception of convention. In the following paragraphs, I briefly summarize the relevant aspects of these two bodies of work that are relevant to my argument about strategic injustice.

For Hume, convention is "a sense of common interest; which sense each man feels in his own breast, which he remarks in his fellows, and which carries him, in concurrence with others

into a general plan or system of actions, which tends to public utility" (Hume, 1777/1975, p. 257). Thus, for Hume, conventions have at least four key features: (1) following conventions is mutually beneficial for participants, (2) conventions do not presume any explicit covenant, (3) conventions are useful because participants believe that others will obey them, and (4) conventions are rational strategies to follow when one believes others will obey. The significance of these features is that they can explain the possibility of a just social order without any explicit agreement or promise and without any centralized authority or sovereign.

David Lewis's (1969) reformulation of conventions in game theoretic terms is another important foundation for strategic justice.[1] Like Lewis's approach to convention, strategic justice assumes that conventions are *strategies* that solve *coordination problems*. There is a coordination problem when individuals need to coordinate their actions to complete a task or to make a decision, but neither is prior communication possible nor is there an external force or central authority to impose order. For Lewis, a convention is an arbitrary and self-reinforcing solution to such a problem of coordination. It is arbitrary because there could be some other convention that would work equally well to solve the problem. It is self-reinforcing because the fact that other individuals obey it gives one a rational reason to follow suit without any prior agreement or external enforcement.

Strategic justice departs from Lewis's account of convention on several grounds. It assumes that justice conventions solve a special kind of coordination problem, "where the corresponding equilibria solve problems of conflictual coordination, in which the interest of the agents involved both coincide to some extent and diverge to some extent" (Vanderschraaf, 2018, p. xiii). However, this solution need not be the product of a one-shot coordination game in which players independently arrive at a solution, unlike what Lewis famously assumed. Hawk-Dove[2] is an example of a conflictual coordination game in which two players with diverging interests interact to divide some resource between themselves. The strategies available to the players are limited: They can either fight for the resource (Hawk) or share it (Dove). Thus, if we assume that there is no communication or prior agreement between the players to facilitate their interaction, four

[1] Games in this reformulation represent strategic interactions among multiple actors where for every player the results are dependent on other players' decisions as well as on their own. Players are simply the individuals that participate in the interaction, and strategies are the options available to them. A key feature of these games is that for each combination of strategies individuals choose, their payoffs determine the benefits or utility they enjoy.

[2] This was originally introduced by Maynard Smith (1982).

outcomes are possible. Two of these outcomes involve one party playing Hawk and the other Dove, which ends with Hawk seizing all the resource. Another possible outcome is when both parties play Dove and split the resource available to them. The last possible outcome is when they both play Hawk and split the resource but pay the cost of fighting. Figure 1, summarizes these four possible outcomes for two players. The numbers represent the payoff of each player based on their choice and the choice of the other player. For instance, if player 1 plays Dove and player 2 Hawk, the former gains no share of the available resource, and the latter enjoys both shares.

		Player2	
		Dove	Hawk
Player1	Dove	1, 1	0, 2
	Hawk	2, 0	$^1/_2$, $^1/_2$

Figure 1 The Payoff Matrix for Hawk Dove

Strategic justice departs from Lewis's account of conventions by relying on the assumption that conventions are *correlated equilibrium solutions* to conflictual coordination problems. Correlated equilibrium solutions are distinct from the rational solutions that the structure of the game implies if players arrive at decisions independently. For example, if players only play Hawk-Dove once and independently decide what to do to maximize their payoff, the only rational choice for each is to play Hawk. In fact, the Hawk-Hawk outcome is an equilibrium solution because neither of the two players have an incentive to deviate. If each of them decides to play Dove instead of Hawk knowing that the other is more likely to play Hawk, the Dove player risks reducing their payoff from ½ to 0. However, with some alterations in how often the game is played or who already has dominance over the available resource, other strategies become salient and even dominant. But the dominance of these alternatives is not a result of independent decisions of players. For example, if one of the players already controls the resource, a plausible strategy is to play Hawk if one is the owner and dove if one is not (Skyrms B. , Evolution of the Social Contract. , 1996). Among players who repeatedly play Hawk-Dove, such a strategy can emerge as a convention without any prior communication (Sugden, 1986).

This kind of solution is called *correlated* equilibrium because the actions of the players are correlated and not fully independent.

Seeing conventions as correlated equilibrium solutions explains the emergence of cooperation in repeated games even when cooperation seems irrational. Take, for instance, the prisoners' dilemma in which two prisoners are being interrogated separately and must decide whether to cooperate with each other by remaining silent or to confess and thereby defect. Again, there are four possible outcomes, two of which involve one confessing and the other cooperating. The other two options are either both cooperating with each other by remaining silent or both defecting. Figure 2 represents the payoff matrix for this game and shows that there is no strategy that both prisoners can choose for an equilibrium solution. This payoff matrix also shows that cooperation is irrational because it exposes one to the possibility of a higher loss than does defecting. However, over repeated games, a potential strategy that emerges is to cooperate as long as one's opponent cooperates and then to defect for some number of rounds so as to punish the defector before cooperating again. This strategy is commonly referred to as "TIT FOR TAT" and I will explain its significance shortly.

	Cooperate	Defect
Cooperate	2, 2	0, 3
Defect	3, 0	0, 0

Figure 2 The Payoff Matrix for The Prisoners' Dilemma

3. Strategic Injustice

Strategic justice explains the possibility of a just social order that emerges out of individuals' rational decisions and interactions without any explicit agreement or promise and without any centralized authority or sovereign. My goal in this paper is to show that an oppressive social order can also emerge from individuals' rational actions without the need for any explicit agreement or promise in the form of coercion or collusion and without any centralized authority or sovereign that acts as a tyrant. Of course, this claim is not new and has a long history in theories of oppression and injustice. Perhaps Foucault (1977) is the most well-known of those who suggest such a possibility. But I follow the lead of scholars like Iris Young (1988), who argue that instead of a cruel tyrant with bad intentions, a system of conventions and norms that

emerges out of disparity in power and influence or in extreme cases from exclusion or marginalization, can be responsible for oppression.

> But oppression also refers to systemic constraints on groups that are not necessarily the result of the intentions of a tyrant. Oppression in this sense is structural, rather than the result of a few people's choices or policies…In this extended structural sense oppression refers to the vast and deep injustices some groups suffer as a consequence of often unconscious assumptions and reactions of well-meaning people in ordinary interactions, media and cultural stereotypes, and structural features of bureaucratic hierarchies and market mechanisms— in short the normal processes of everyday life. (Young I. M., 1988, p. 6)

Philosophical and sociological approaches to oppression often highlight three distinctive features of oppression without a tyrant, and these will guide my arguments in this paper. The first distinctive feature is exclusion of the marginalized from networks of power and cooperation and the centrality of others in these networks. Such a relative network position and the resulting marginalization are often used to explain the emergence of an oppressive social order (see for example Anderson, 2010; Young, 1988). The second distinctive feature is the importance of both between and within group relations especially because in oppressive systems without a tyrant, group dynamics "do not always fit the paradigm of conscious and intentional oppression of one group by another" (Young I. M., 1988, p. 6). In fact, groups in such a system are very heterogenous and dynamic to the extent that assuming a coherent set of intentions or goals for everyone in a group or for the group as a whole becomes unjustified. The third feature is the consensus that disrupting such systems of oppression without a tyrant requires mobilization to reform the structure of social relationships.

What I call strategic injustice is a system of formal and informal conventions that results in oppression and derives from individuals' strategic choices about their social interactions. Strategic injustice is a form of oppression without prior agreement or centralized power. I call this form of injustice *strategic* to highlight the role of individuals in forming the exclusionary networks and discriminatory conventions that maintain them. Strategic injustice is best understood as a continuation of strategic justice that highlights the importance of *network effects* on individual incentives through encountering individuals' ability to control aspects of their social environment. Thus, to distinguish strategic injustice and its contribution from its ancestor,

I compare below their answers to the three questions that I mentioned in the previous section: what the necessary circumstances of justice are, what explains the emergence of a just social order, and how social injustices can be addressed.

4. The First Question: What Are the Necessary Circumstances of (In)justice?

The third chapter of *Strategic Justice: Conventions and Problems of Balancing Divergent Interests* starts with the claim that "unless parties are constrained by certain limiting conditions, one might doubt that justice can exist between these parties" (p. 85). The idea behind this claim is that some minimal conditions should be satisfied before we can meaningfully talk of justice or injustice. In this chapter, Vanderschraaf lays out the "Circumstances of Justice," which are a set of conditions that makes cooperation possible from a proper conventionalist standpoint. He also provides an alternative to the "Standard Account" by which moderate scarcity, moderate selfishness, and rough equality are the three necessary Circumstances of Justice.[3] Vanderschraaf suggests that the right background conditions for justice needs to simply summarize the essence of what it takes for individuals to feasibly expect cooperation from one another. He also argues that the structure of conflictual coordination games fits the bill.

> … parties have the right background conditions for justice when (i) they have available to them a variety of conventions over which their preferences differ to some extent, (ii) they can by working together to generate a cooperative surplus characterized by some of these conventions, but (iii) each is also vulnerable to being taken advantage of by others who aim for outcomes better for themselves that result in their fellow parties suffering relative losses. (Vanderschraaf, 2018, p. 116)

Conflictual coordination problems fit the bill because they can be resolved through a variety of conventions, allow players to create a cooperative surplus, and formalizes the vulnerabilities of individuals in their interactions with others. The problem is that, at least in the way discussed in the literature adjacent to concerns about justice, conflictual coordination problems do not need to account for individuals' choices about their associates. In other words, the only tools available to individuals in dividing their resources, for instance, is to choose among the strategies available

[3] Vanderschraaf (2018) explains these conditions in details and analysis their adequacy in chapter 3.

to them, not choose with whom they want to divide their resources. However, as Skyrms and Pemantle (2000) put it, even "A child who is being bullied learns either to fight better or to run away" (p. 9340). Similarly, in addition to changing our strategies, we often have the option of changing our associates to avoid unsatisfactory results in our interactions with others. Moreover, this change in association does not require prior communication and can emerge simply from repeated interaction (Skyrms B. , 2004).

For explaining the emergence of common and stable forms of unjust rather than just interactions, accounting for individuals' ability to control their social environment is necessary. In other words, individuals' discretion over some aspects of their interactions with others, which I call "Discretion" for the sake of brevity, is a necessary background condition for instances of strategic injustice. My argument to support this conjecture is as follows. Neither Vanderschraaf's Circumstances of Justice nor the conflictual coordination games assume Discretion. In fact, the no Discretion assumption is common among standard treatments of evolutionary game dynamics that forms the explanatory foundation of strategic injustice (see, for example, Hofbauer & Sigmund, 1988; Weibull, 1997). However, without Discretion, these evolutionary models often imply that injustice is an anomaly. I provide a representative evolutionary model that leads to this conclusion about the anomaly of injustice, which is at odds with individuals' testimonies and empirical data about the pervasiveness of exploitation, discrimination, and other forms of injustice. I also provide an alternative evolutionary model that considers Discretion and confirms the pervasiveness of unjust but cooperative equilibrium states. I conclude that no Discretion is a faulty assumption, and that Discretion is a necessary background condition for strategic injustice.

My argument above can be summarized as follows.

1. No Discretion (assumption)
2. No Discretion → Injustice is an anomaly. (standard evolutionary models)
3. It is not the case that Injustice is an anomaly. (empirical evidence)
4. Discretion → Injustice is not an anomaly. (alternative evolutionary models)

   Conclusion 1: From (1), (2), and (3) we can conclude that dismissing Discretion is problematic.

   Conclusion 2: From (2) and (4), we can conclude that Discretion is a necessary feature of a population in which injustice is not an anomaly.

Premise 1: No Discretion

A representative example of evolutionary models that implies injustice is an anomaly is the dividing of a chocolate cake. Brian Skyrms (1996) describes this example:

> Here we start with a very simple problem; we are to divide a chocolate cake between us. Neither of us has any special claim as against the other. Our positions are entirely symmetric. The cake is a windfall for us, and it is up to us to divide it. But if we cannot agree how to share it, the cake will spoil, and we will get nothing. (Skyrms, 1996, pp. 3-4)

Of course, there are many ways to divide the cake between two agents. But demanding only half of the cake is a unique solution because both agents can do the same and successfully divide the cake. Any other strategy, say asking for 1/3 of the cake, is only successful when the other player exactly asks for the other 2/3. Thus, without any information about the strategy of the other player, the more often people ask for half of the cake, or ask for a fair distribution, the more often they can divide the cake successfully. So, in a population of dyads, in which people are randomly coupled to divide a cake, individuals learn, over time, that demanding only half of the cake is the best strategy. Thus, if somehow agents who successfully divide a cake could duplicate, we would have a growing population in which fair players, who only ask for half of the cake, would be dominant.

What O'Connor (2019) calls an "evolutionary push for fairness" in fair division problems is a well-established phenomenon. For instance, Skyrms (1994) explains that in a population composed of random pairs the possibility for miscoordination pushes the evolving population to reach fair outcomes. This push for fairness is even more forceful when people meet the same neighbors repeatedly (Alexander & Skyrms, 1999). Such repeated interaction can correct disparities in bargaining power and background conditions (Alexander, The structural evolution of morality, 2007). And, of course, communication between neighbors is conductive to fair demands and, by extension, to fair norms of division (Skyrms, 2014). In fact, fairness is the evolutionary stable strategy even when we consider the stochasticity of individual actions (Sugden, 1986; Young, 1993). The problem is, however, none of these models take Discretion into account.

Premise 2: Without Discretion Injustice Is an Anomaly

As you can see in the graph below, simulating the dynamics of the games described above shows that almost every time a population engages in the dividing the chocolate cake game, playing fair is the evolutionary successful strategy. The model I used to create the information in the graph assumes that individuals are randomly moving around the "world"[4] and that whenever exactly two of them exist at the same place at the same time, they can share a cake. Individuals have one of the three strategies to do so: playing fair and claiming half the cake, playing modest and claiming one third of the cake, and playing greedy and claiming two-thirds of the cake. Individuals randomly bump into each other and if their strategies match, they get to invite a new agent with the same strategy to the game. If their strategies do not match, then they have to exit the game. Matching strategies requires the two individuals to claim the whole cake. Thus, successful encounters are either between two fair players who each claims half or between a greedy player who claims two-thirds and a modest player who claims one-third of the cake. In this model, the strategy that wins most often will have the largest population at the end of the game.

As Figure 3 indicates, at the end of the game, the population size for each group varies every time we run the model. However, every single time fair players end up with the largest population size. In other words, playing fair is clearly a winning strategy. Therefore, individuals have rational and self-interested reasons to be fair. More importantly, other strategies like being greedy or modest are not only irrational but also unlikely to be a dominant strategy evolutionary speaking. In other words, it will be an anomaly for these unfair strategies to become established evolutionary strategies or a convention.

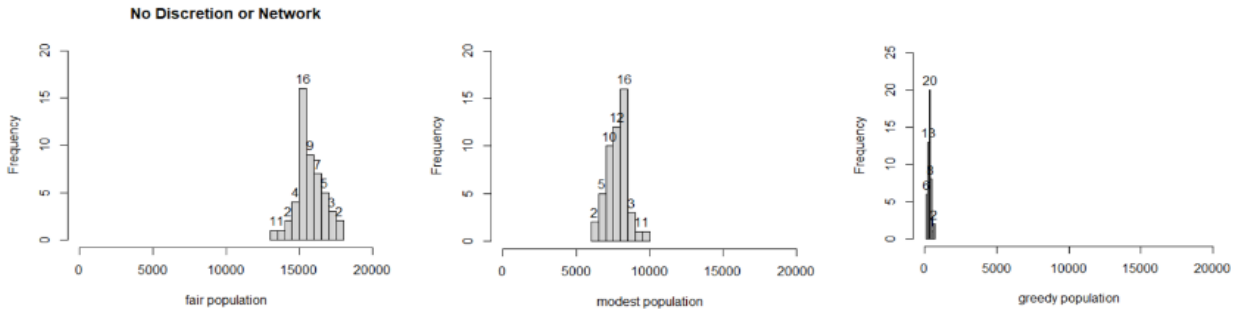


Figure 3 Simulation for dividing the cake with three strategies and no Discretion for agents to choose their future partners. The Y-axis, labeled as frequency, represents the number of times that the number of agents with a given strategy ended up in a particular range of population by the end of the game. Although there is some variation in the end-game population, the fair population is clearly the winner every time.

[4] World in many agent-based models simply refer to the space in which agents exist.

The dominance of the fair strategy in the dividing the chocolate cake simulation does not assume any prior communication or agreement among individuals. It also does not require any central authority or a sovereign to enforce cooperative behavior. Yet this simulation can successfully show that self-interested individuals can cooperate and in fact be consistently fair. However, this success implies that injustice is either an anomaly, as the simulation suggests, or requires prior communication or agreement among individuals or even demands the central authority of a tyrant that would impose an unjust order. Otherwise, the dividing the chocolate cake model implies that being unjust is an irrational strategy that leads to lower payoff over time and is not reinforced through interaction between individuals.

Premise 3: Injustice Is Not an Anomaly

As non-ideal theorists as well as feminist and anti-racist scholars argue, not only injustice is not an anomaly but also every aspect of our social world is influenced by these injustices one way or another.[5] Moreover, the consensus among these scholars is that a great number of these injustices are a direct result of individuals' positions in social hierarchies and their degree of integration in networks of power and cooperation (Anderson, 2010; Young 1988, 2012). Empirical evidence also confirms the prevalence of injustice especially for marginalized communities. Women's lower bargaining power in the workforce or in their households (Rosenfeld & Kalleberg, 1991) (Folbre, 2021), discriminatory practices in the labor market (Neumark, 2018), the labor exploitation of highly marginalized populations like victims of sex trafficking or undocumented immigrants (Schauer & Wheaton, 2006), the underfunded schools in the over policed "inner-city" neighborhoods, and the numerous cases of domestic violence (Boserup, McKenney, & Elkbuli, 2020) are just a few examples of well-studied but pervasive cases of injustice.

Racial segregation and its resulting harms are clear examples in which exclusion from networks of power and cooperation explain the dominance and resilience of social injustices individuals endure. In her book, *The Imperative of Integration*, Elizabeth Anderson (2010) goes

[5] For a similar argument see, for example, Mills (2008) idea of a racialized society.

to great lengths to explain the injustices that follow such exclusion and marginalization. Moreover, as Schelling (1971) argues, at least hypothetically, segregation can emerge without any communication among members of a group or a prior agreement to exclude others. Although this hypothetical scenario can be very far from reality, Schelling's model explains how coalition-like groupings emerge out of interaction among independent individuals. In combination with Anderson's analysis of segregation, this model shows that individuals' actions can play a role in the emergence of structural features that maintain stable forms of injustice and unfairness.

It is worth noting that the examples I offer above often require clear group membership and distinct features that allows individuals to *label* each other and themselves. This form of group membership based on shared features is a standard and well-studied gateway to explaining injustice and discrimination. It also conveys a similar idea with classic and recent work by game theorists and philosophers who are concerned with different realizations of injustice, such as exploitation or discrimination (for examples see Axelrod & Hamilton, 1981; Axelrod, 1984; Rubin & O'Connor, 2018). The common idea is that labeling players based on their strategies or group memberships provides some added information, allowing individuals to develop group-specific strategies. For instance, members of a racial group may play fair among themselves and greedy when they interact with others. Thus, one might argue that instead of adding Discretion to our models, all we need to do is to consider such forms of labeling and formalize the added information.

It is undeniable that membership in different social categories, especially when this membership is represented by a visible label like skin color or gender, plays a significant causal role in the emergence of discriminatory practices or unjust social orders. However, this undeniable significance does not undermine the importance of Discretion as a necessary background condition for injustice. First, this is because, as I will show shortly, even without the existence of labeling, Discretion can alter the likelihood of the dominance of fairness in a population. Second, as Axelrod (1984) famously argues, we often assume some social structure to explain the evolution of cooperation, and labeling is simply some added layer of structure that alters the dynamics of interaction. Hence, capturing the role of Discretion in social interaction provides a more general model that allows us to explain the formation of structures of any sort. I discuss this general model in the next section.

Premise 4: Discretion Explains the Pervasiveness of Injustice

Altering the dividing the chocolate cake simulation to capture Discretion shows that the mere ability of individuals to choose their associates explains why justice is not an anomaly. The original form of this simulation that I discussed earlier in this section, like many other well-studied evolutionary models, assumes "mean-field" interactions, by which individuals have the same likelihood of interacting with everyone else in a population. In other words, these models assume that individuals often do not have any control over their interactions and are randomly assigned to their next partner. In addition to this assumption being foreign to how we actually interact with one another, the need to explain the emergence of unjust but stable patterns of behavior gives us good reasons to doubt it.

The graph below shows the result of an adjusted simulation for dividing the chocolate cake in which individuals start with a set of random links with others and their links determine whether they can interact with one another or not. If the result of an interaction is successful, then they maintain their link with their partner and reproduce by, for example, inviting a new player with the same strategy to the game. If the interaction is not successful, players cut their link, which prevents them from engaging in future interactions with the unmatching partner. To remain close to the result of the previous model, this simulation adds a cost for unsuccessful interaction. Thus, having two unsuccessful interactions in a row makes the agents leave the game or "die." Applying this added cost to the original model does not change the outcome of the simulation but allows us to formalized the role of Discretion in social interactions.

This altered model considers that although many of our interactions might occur randomly, we do not have the same likelihood of interacting with everyone else in our social world. Also, the fact that we can change our associates to control our outcomes is built into this model by agents being allowed to sever their links when their interactions are unsuccessful. As the graph indicates, this alteration reduces the success of fair strategies by 36%. In this model, the greedy agents have a fair chance of becoming the group with the dominant strategy. In other words, if we understand the dominance of the greedy strategy as unfairness and indicative of injustice, then injustice is no longer an anomaly. It is worth noting that this dominance does not require any prior agreement or central authority to facilitate cooperation. In this simulation, individuals'

control over some aspects of their interactions allows them to survive and reproduce by simply learning with whom they should associate in order to maintain their payoff.[6]

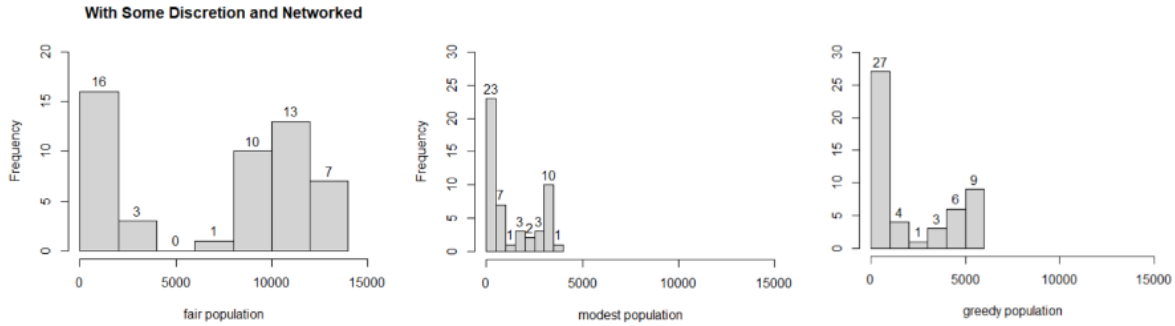


Figure 4 Dividing the chocolate cake when we allow individuals to have some discretion over who they want to interact with. You can see that the fair players are no longer the clear winners.

If we allow individuals to learn from past experiences and label others based on their strategies, the likelihood for unfair treatments to become dominant grows even larger (for further examples and discussions of this kind of model see Skyrms 2004; Alexander, 2007). A comparison of the three models shows the extent to which Discretion and labeling are important for the emergence of unjust patterns of behavior. In the first model of dividing the chocolate cake, without any network or Discretion, the chance for the fair strategy to become dominant is 100%. In the second model, when we add some Discretion that generates a network, the fair strategy becomes dominant in 64% of the cases while the greedy strategy gains a 36% chance to be dominant. Finally, in the third model, when we allow individuals to learn from past experiences and connect to the right kind of agents with a higher likelihood, the chance of the fair strategy becoming dominant is even smaller, 54%, and the dominance chance for the strategy of being greedy grows to 46%.[7]

[6] You can find an interactive version of this model here:

[7] I will provide all three models upon request.

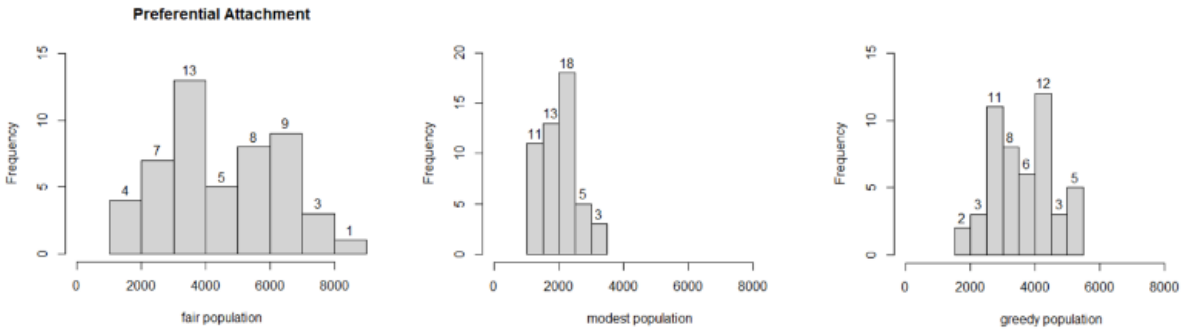


Figure 5 Dividing the chocolate cake simulation results when we allow the players to learn from their past experiences and make targeted links in the future.

Conclusion: Discretion Is a Necessary Condition for Injustice

In sum, the evolutionary models of fair division that support the main ideas of strategic justice do not assume Discretion for individuals. But as the simulations for dividing the chocolate cake show, without Discretion, there is a strong evolutionary pressure for fairness. In other words, with no Discretion injustice is an anomaly, which is simply not true in real world. Therefore, for a conventional approach to *in*justice, having no Discretion seems to be a faulty assumption. However, by comparing three simulation models of a fair division problem, I showed that in the absence of any information about the strategies of other players, Discretion can explain the dominance of unfair strategies. Therefore, Discretion, or individuals' ability to choose their associates to some degree, is a necessary background condition that allows the dominance of unjust strategies in a population.

The significance of Discretion is that it adds a dynamic network formation process to the evolution of cooperation. This dynamic process changes the outcome of social mechanisms that would otherwise ensure the emergence of fair cooperation in the population. However, this change is often not random. In the following section, I discuss the mechanisms through which a population arrives at a(n) (un)just convention and the ways through which the dynamic network formation can alter the results.

5. The Second Question: What Explains the Emergence of a(n) (Un)just Social Order?

Vanderschraaf anticipates the possibility of the emergence of unjust systems of conventions. He argues, just conventions "must lead all to an outcome better for all than a free-for-all

outcome, and they must not allow any individual parties to benefit too much at the expense of others, since otherwise they could not plausibly be considered the conventions of justice" (p. 116). Following Lewis, Vanderschraaf argues that chance, agreement, or some intrinsic feature of equilibrium solutions that makes them more salient to the players can explain why we end up with unfair and unjust conventions. He also provides a standard to distinguish just conventions from unjust, according to which, a convention is unjust when a hypothetical change in the background conditions provides incentives for individuals to renegotiate its terms.

It is a convenient and intuitively plausible move to explain the emergence of unjust conventions as a product of random background features like chance, intrinsic salience of an equilibrium turned convention, or even explicit agreements. However, if I am right about the role of Discretion in the emergence of "bad equilibrium" states, then individuals' decisions about their associates at least partially determines whether bad equilibrium states become a convention or not. For instance, in dividing the chocolate cake example, the possible equilibrium states are fair-fair and greedy/modest-modest/greedy. As I described in the previous section, there are clear reasons why the former equilibrium state becomes dominant. But my altered simulation models for dividing the chocolate cake shows that the latter state, which fits the description of a "bad equilibrium," becomes dominant not due to chance, agreement, or intrinsic salience (more than the first state), but when individuals gained Discretion or the ability to control some aspects of their interactions. In a sense, individuals learned to create a network of interaction that *excluded* players with whom cooperation is unsuccessful.

Explaining the emergence of an unjust convention due to some sort of agreement is not relevant to my project. And if Lewis is right, the alternatives to some sort of agreement are chance or intrinsic salience. However, both these alternatives are outside of individuals' control and incompatible with the idea of strategic injustice as defined earlier in this paper. In fact, by definition, strategic injustice requires the possibility of an unjust social order such that its emergence is a result of non-cooperative but nevertheless *strategic* interactions among the members of a population. In sum, if strategic injustice is possible, then the emergence of unjust conventions cannot be a coincidence and vice versa.

In this section, I prove by contradiction that the emergence of injustice is not a coincidence and that individuals' Discretion plays an important role in the existence and persistence of injustice. My argument is as follows. Even if coincidence and chance lead to unjust conventions,

there should be some kind of mechanism that explains how a population arrives at such stable but suboptimal equilibrium states. However, these explanatory mechanisms often assume structured interactions that are not a fixed characteristic of the real world. These structures are just a graph-theoretic representation of individuals' access or ability to interact with one another, and their formation process relies on individuals' Discretion or their ability to change their social environment. In fact, the structure/network formation models suggest a feedback loop between structures and individuals' Discretions that explains the emergence of unjust but stable treatments and conventions. Therefore, even coincidental cases, in which unjust conventions emerge because of some "random" background conditions, involve individuals' strategic choices about their interactions with others. It would be a stretch to call a mechanism that involves such choices and decisions a coincidence. Hence, I conclude that the emergence of unjust conventions is not accidental and that individuals' strategic choices about their interactions generates them.

Premise 1: If coincidence explains the emergence of unjust conventions, our explanatory models should account for the possibility of that emergence.

Note that salience of an equilibrium state as Lewis suggests on its own does not explain how a solution to a coordination problem turns into a convention. We need mechanisms like learning, natural selection, or cultural evolution over repeated strategic interactions to specify how a population arrives at a just or unjust convention. These mechanisms often require individuals, or generations of them, to learn from past experiences and adjust their strategies in the future. This ability to learn from past experiences leads to the ruling out of options that are unsuccessful or uniquely bad. These learning processes also allow individuals to develop some punishment strategies that ensure cooperation. The possibility of this phenomenon in the real world is supported by empirical evidence suggesting that many individuals have the tendency to voluntarily cooperate when they are treated fairly and to punish noncooperators even when such a punishment is costly. This tendency is what Vanderschraaf calls "strong reciprocity," and it ensures universal cooperation in a population.

Premise 2: The existence of some social structure/networks is necessary to explain the establishment of unjust conventions.

The function of strong reciprocity and other similar punishment strategies is to establish an equilibrium state by protecting it from invasion of potential mutations in strategies. But for these

strategies to be effective, individuals need to either interact repeatedly with the same players or rely on the engagement of a third player who would punish the wrongdoers. Axelrod (1984) in his book *The Evolution of Cooperation*, famously showed that interacting with the same players allows individuals to develop strategies like "TIT FOR TAT," which ensures cooperation. However, he also argues that some social structure is necessary to make the dominance of the strategy immune to invasion from potential mutations. In other words, completely isolated actors who adopt "TIT FOR TAT" can neither maintain cooperation in their population nor can convince a population of non-cooperators to change their strategies. Therefore, an unfair equilibrium state is immune to invasion by an alternative strategy when the potential invaders lack any social structure and interact with others randomly. However, "if the invaders had even small amount of social structure," (p. 147) for instance if "they came in cluster so that they had even a small percentage of their interaction with each other," (p. 147) then they can replace the dominance of the current strategies with alternatives.

The structure of interactions is important, even when individuals do not repeat interactions with the same players. For instance, strong reciprocators, individuals who voluntarily punish wrong doers on behalf of victims, at least need to know about the previous wrongdoings of an agent. In other words, structures become relevant when other than the two individuals who play the game in each round uninvolved agents, their information, or their strategies become relevant to the outcome of individuals' decisions. The influence of uninvolved agents is commonly formalized as a "network effect," which is when the architecture of the network in which individuals are embedded influences their incentives and behavior.

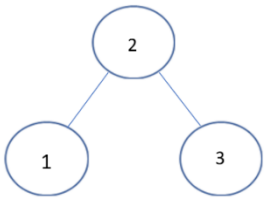
The main takeaway of these two punishment strategies, with or without repeated interaction with the same agents, and their efficiencies is that the structure of social interactions or the networks in which individuals are embedded determines whether and which equilibrium states become established enough to constitute conventions. For instance, Raub and Weesie (1990) as well as Ali and Miller (2009) show that although completely connected networks are indeed efficient in punishing deviations, in networks with lower levels of connectivity, players only observe their close neighbors' behavior. Thus, there will be some network architectures in which interpersonal punishment strategies are not a reliable and corrective force to maintain cooperative behavior. Even if these strategies can maintain cooperative behavior, the disparities

of exposure to these punishments can create disparities in enforcement and outcome of cooperation. In some other networks, the challenge to enforce individual cooperation is the time it takes for the bad behavior to reach one's neighbor through contagion. Moreover, as Burt (2001) suggests, *structural holes* can slow down the travel of information. Thus, depending on the context or position in the network, the extent to which norms of cooperation or fairness are enforced can vary. Thus, individuals can learn over time that violating norms of cooperation in interaction with some has fewer repercussions than it would if done with others.

Premise 3: The formation of social structures and networks is dependent on individuals' Discretion.

Networks can represent anything from formal or informal ties, friendships, and relations of influence to information channels or even the mere possibility for interaction, none of which are given or fixed features of the social world. Thus, although there is utility in studying networks as fixed characteristics of the environment, the most general and realistic way to capture their effect requires modeling their dynamic formations. And the most comprehensive models of network formation rely on individuals' strategic choices about their interactions with others. In fact, networks, or as Myerson (1976) calls them, "cooperative structures," are simply the graph-theoretic representation of individuals' willingness and ability to interact with others.

To see the importance of thinking about individuals' interactions in terms of networks Myerson (1976) invites us to consider a scenario in which "player 1 wants to cooperate with player 2 but not with player 3, 3 wants to cooperate with 2 but not with 1, and 2 wants to cooperate with both 1 and 3" (p.1). The combination of these preferences does not result in a stable "coalition"[8] because each possible combination suppresses at least the preferences of one player.[9] However, reimagining the connections and preferences of the players as a system of "bilateral cooperation links" creates a structure that accommodates all the players' preferences and facilitates cooperation. The following graph represents such a structure.



[8] Myerson's model was originally introduced as a reformulation of coalitional games. However, it does not require the common features of cooperative games.
[9] The possible partitions are {{1,2}, {3}} or {{1}, {2,3}} or {{1,2,3}}.

Figure 6 The picture represents the cooperation link between agents 2 and 3 as well as agents 1 and 2, with no cooperation between agents 1 and 3.

Myerson's representation of individuals' Discretions allows the "cooperative structure to depend endogenously on choices by players" (p.1). In this representation, two players form a link if and only if they both want it to exist, and they can simultaneously express their willingness to form a link without prior information about each other's preferences. Thus, each player can choose a set of players with whom they want to interact. And the graph that results from all cooperation links and players' choices is the "cooperative structure" of the game. More importantly, players choose their links based on the way the resulting cooperative structure affects their payoff.

Myerson provides a function that maps the cooperation structure into the allocation of the cooperative surplus. In this function, if two players are a connected component of the structure, then they must divide the total wealth they can earn. This function reduces complex problems with multiple players into a system of two players' games for which we can consider one link at a time without losing the influence of uninvolved agents and their payoffs in our models. Because both agents need to agree to form a link, Myerson argues that it is only fair if they both benefit from it equally. He further shows that there is a unique fair allocation rule for each game.

As the structure of the cooperation varies in topology and connectivity level, we can see that the allocations of the cooperative surplus and individuals' incentives vary as well. In fact, the total share of each agent from the cooperative surplus is dependent on the graph in which their link is embedded. For example, player 1 gains more in g1 compared to g2 because in g1 this player is more essential for the coordination of others or simply has more links. Moreover, two unconnected players cannot form a functioning coalition because they simply cannot interact.

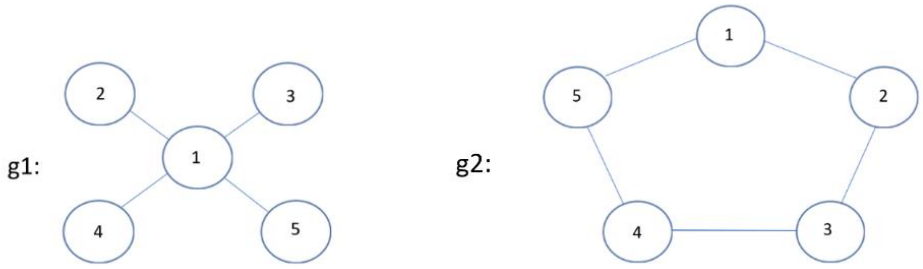


Figure 7 Comparing two network topologies.

Premise 4: The feedback loop between structures and Discretion explains the emergence of injustice and unfairness.

A similar logic about cooperative structure that considers the barriers of interaction explains the emergence of unfair cooperation without prior agreement or communication and its stability. The structure that results from individuals' strategic choices explain the mechanisms by which networks influence individuals' payoffs and their incentives. For example, these structures allow uninvolved individuals to impose social pressure on the interacting agents or tax the payoff of their cooperation. Individuals also can manage social pressure and the influence of uninvolved agents through altering their social interactions. In a sense, individuals' Discretions that shapes the cooperative structure and the cooperative structure that influences individuals' incentives create a feedback loop. In this feedback loop networks emerge from individuals' ability to control their interactions and alters individuals' incentives about interactions with others.

The feedback loop between individuals and their networks creates an environment in which an interaction between two agents can "be affected by and affect each of their relations with others" (Basu, 1986, p. 259). For example, Jackson, Rodriguez-Barraquer, and Tan (2012) formulate the influence of networks on the emergence of cooperation by examining cases in which individuals can ask their neighbors for a costly favor. They show that when two individuals have some common connections, the likelihood for them to ask for or accept the request for favor is higher. The idea is that the risk of losing common connections or of acquiring a bad reputation provides enough incentive for individuals to bear the cost of doing the favor. The more individuals in a network have common connections, the tighter will be the graph of

their social relations, something that is commonly referred to as a clustering in the network. Jackson and his team show that clustered networks impose a higher *social pressure* on individuals to cooperate, by which they mean to do favors when called upon. Individuals also have incentives to manage their social pressure by adjusting their links, which in turn alters the structure of cooperation.

The feedback between Discretion and networks also creates variation in individuals' relational features such as connectivity level. The fact that most large socially generated networks are neither random nor fully connected implies that individuals' access to other cooperators is limited and that this limitation is not equally distributed. Such an unequal distribution is very common and can result from a simple mix of random connections and local adjustments in individuals' strategic choices of their interactions. For instance, if in addition to making random connections, individuals form and sever links mostly with the friends of their friends, over time we end up with a great disparity among agents' connectivity level (Jackson & Rogers, 2007). In this network, some nodes have a very high and growing number of connections while others have very few, which they might lose over time without replacement (Jackson & Rogers, 2007). Consequently, the density and clustering of relations can vary in different parts of a social network.

The difference in connectivity level explains the pervasiveness of unjust treatment and even the emergence of an unjust social order. For example, due to no fault of their own, individuals may occupy parts of the network that make them "relationally vulnerable,"[10] in the sense that their network fails to provide them the support they need to survive. This lack of support is not a product of individuals' unwillingness to support one another. It is instead a product of the number people being available in their surrounding that they can feasibly ask for support and the extent to which those people can provide support based on their own level of connectivity or vulnerability. For example, even "strong reciprocity" would not be helpful if one is wronged by someone outside of their immediate network or by someone with whom others desperately need to maintain connection. Recall that strong reciprocity is helpful because "a party might be prevented from exploiting other parties currently unable to protect themselves by third parties who are capable of exercising some restraining force" (Vanderschraaf, 2018, p.114). However,

[10] Gordon-Bouvier (2020) introduce and discuss this idea in her book, *Relational Vulnerability: Theory, Law and the Private Family*.

as I discussed earlier "structural-holes" or mere fragmentation in the network can make this punishment strategy futile. In other words, individuals' position in social network exposes them to exploitation when the wrong doers realize that exploiting ones in such a position will not have any repercussions.[11]

In addition to unjust interpersonal treatments, the disparity in network effect can explain the emergence of unjust systems of conventions. These systems successfully facilitate cooperation but at the expense of some. Remember that for Vanderschraaf, just conventions "must not allow any individual parties to benefit too much at the expense of others, since otherwise they could not plausibly be considered the conventions of justice" (p. 116). However, various network formation processes indicate that individuals' rational decisions about their interactions can overtime create structures that benefits some too much and it does so at the expense of others. These structures also impose constraints on individuals' actions that promote further conventions such as norms against intermarriage or exclusionary rules of etiquette, that maintains the disparity in outcomes.

A famous example of a network formation process that leads to gross disparities in outcome is the "preferential attachment network." This dynamic model of network formation explains what is commonly known as the "rich get richer" phenomenon. In this process, the network grows gradually, and newcomers have a higher incentive to connect with better connected individuals. Over time, this process inevitably creates an environment in which the more connected a player is the faster the number of their connections grows. Thus, when more connections mean higher payoff, the payoff of highly connected individuals grows much faster than others. Longitudinal studies of wealth distribution in the United States and many other countries suggests that the resemblance of disparity in connectivity to disparity in wealth is not a coincidence. The fact that disparity in connectivity grows over time matches the widening wealth gap in the United States, as well as the gaps in social mobility and life expectancy (see, for example, Venkataramani, Daza, & Emanuel, 2020).

[11] The variation in network effects on individuals can also exacerbate individuals' differences "when persons must decide whether to adopt beneficial practices; when network externalities, social learning, or normative pressures influence adoption decisions; and when networks are homophilious with respect to individual characteristics that predict such decisions" (DiMaggio & Garip, 2012, p. 93), to name just a few situations.

The mere fact that individuals can choose their connections and have rational reasons to associate themselves with better-connected individuals creates a system that allows some parties "to benefit too much at the expense of others." Such a system most likely will not have adequate tools to ensure fair cooperation and distribution in interpersonal interactions. But even if, somehow, it does have such tools, disparities in connectivity level and other features of the network can deprive individuals from the minimum levels of connection and integration in the overall network that they need to survive.

Conclusion: Injustice is not a coincidence and emerges from individuals' strategic choices about their social environment

Social networks are not born; they are created by individuals who strategically alter their social environment to achieve their ends. These networks are dynamic and can vary in density, connectivity, and their effect on individuals' actions and decisions. They also interact with mechanisms, such as learning and punishment that explain how a population arrives at a convention. Thus, there should be no surprise that when things go wrong, a probable candidate to blame is the interaction between the networks and individuals' decisions about their links. However, explanatory models that dismiss the role of such an interaction might transfer the responsibility to background conditions or random features of the environment. On the other hand, accounting for this interaction shows that unjust conventions can emerge from individuals' strategic choices and without any prior agreement or communication. I call such conventions strategically unjust.

For example, the conventions and regulation that maintain racial segregation in the United States are strategically unjust. Segregation excludes people of color from the networks on which their survival is dependent. These networks regulate access to jobs and material resources, support and protection from violence, political and cultural influence, and bargaining power in interpersonal interactions (Anderson, 2010). Exclusion from these networks puts vulnerable populations in positions such that their exploitation does not lead to repercussions and seems even rational and profitable. Unsurprisingly, the product of this lack of access and isolation is stigmatization and discriminatory conventions that rationalizes their undesirable conditions.

Exclusion from networks of power, influence, and cooperation not only exposes individuals to the relational vulnerability I discussed in the previous section, but it also weakens the

connections among the members of marginalized communities. For example, exclusion decreases the resources and opportunities that circulate in an excluded network. Disadvantage multiplies in a tight community and provides incentives for its members to disconnect and weaken their ties. Stigma often follows exclusion and accumulated disadvantage in a community. Stigma also taxes the few beneficial connections that individuals might have. Anderson describes this process:

> Employed blacks are less likely than whites to recommend their unemployed male relatives and friends for a job because they do not trust them to do a good job and worry that this will damage the goodwill they have established with their employers. To save face, unemployed black men adopt a posture of "defensive individualism," avoiding the pain of rejection by not even asking their employed connections to provide a referral. Thus, even when Black men have connections, they are not utilized properly. (Anderson, 2010, p. 34)

5. The Third Question: How Can Social Injustices Be Addressed?

The closest that Vanderschraaf gets to addressing strategic forms of injustice is in his response to the exclusion objection for his conventionalist theory of justice. In this conventionalist account, justice is simply "rational prudence pursued in contexts where the cooperation (or at least forbearance) of other people is a condition for our being able to get what we want" (Barry, 1989, p. 7). However, when justice is the minimum constraint individuals accept to benefit from the fruit of cooperation with others, it excludes parties who are "unable to enforce others' compliance with requirements of justice themselves" or "vulnerable individuals" (p. 7) who cannot contribute to the cooperative surplus.

To address this objection, Vanderschraaf shows that we can expand the scope of justice by granting immunity rights to all and providing everyone "some means to sustain themselves" (2018, p.293). He argues that granting such immunity provides incentives for the powerful to not add to the vulnerable population since otherwise, for them, there "will be more mouths to feed." As he puts it, "there will be no temptation to make targeted members of society vulnerable in a system that makes one responsible for the care of those vulnerable" (2018, p.293). He concludes that an all-inclusive society is in fact compatible with a conventional account of justice in which mutual advantage incentivizes compliance with justice conventions.

The problem is, if I am right, that granting the means for individuals to sustain themselves is at least partly relational. Everyone seems to need some level of connectivity to satisfy their minimal needs. But maintaining that level of connectivity requires a decentralized coordination among many in the population. In other words, addressing injustice by granting some immunity to everyone is in itself a multi-agent coordination problem. Otherwise, individuals' uncoordinated actions to resolve disparities in connectivity will be ineffective and costly because, as I discussed in the previous sections, social structures involve various conventions that maintain their boundaries, as well as social mechanisms that tax crossing them.

Addressing the wrongs of a convention or a system of convention is not possible if everything else that leads to disparities in power and influence remains untouched. In what follows, I suggest that social movements involving "networks of interaction" are necessary to disrupt the mechanisms through which power is maintained. Such networks play an important role in the evolution of movements, and movements alter the network of social relations. The former effect allows movements to amplify and engage a wider portion of society. The latter effect allows movements to fill structural holes and disrupt the structures of cooperation through which the powerful maintain their dominance and exclude others from access to resources.

Some scholars view social movements as a particular form of social organization that emerges out of repeated and patterned interactions among multiple actors (Mische 2008, Diani 2015). Arguably, movements "destabilize a given situation and contribute to reorganization of positions and relations within a field" (Diani & Mische, 2015, p. 3). Moreover, they facilitate social change by creating a higher likelihood for contingent interactions and thereby changing the topography of the cooperative structures. For example, by participating in a movement, individuals provide the movement with access to their extended social network. Movements also change the network of social relations for their participants. For instance, joining movements allows individuals to encounter previously unknown people, ideas, tactics, and networks. In fact, social movements are particularly effective at increasing the number of weak ties that help to "bridge structural and cultural holes" (Diani & Mische, 2015, p. 13). Such restructuring of social relations provides opportunities for coalition building across social groups and networks that were not in contact before.

Social movements are also effective in creating strong ties that persist over a long time. In later stages of a movement, stronger ties are formed that cross the "structural holes" and break down the categorical cultural boundaries. These ties can generate new and alternative networks of cooperation and communication among individuals and groups. In this stage movements create cross-network negotiation and coordination. In other words, the change in the structure of social network in society disrupts the preexisting norms of collaboration and coordination. Formerly excluded individuals can gain access to various resources through their newly made connections or through the connections that people in their networks have formed with others. Such connections change the fabric of the society and provide personal and structural incentives for individuals to endorse alternative modes of interaction and cooperation.

6. Conclusion

My goal in this paper was to show that an oppressive social order can emerge from individuals' rational decisions without a need for any explicit agreement or promise in the form of coercion or collusion and without any centralized authority or sovereign that acts as a tyrant. I called this form of oppression strategic injustice because although it does not need agreement or central authority, it emerges from individuals' strategic decisions about aspects of their social interactions. I defined strategic injustice as a continuation of Peter Vanderschraaf's theory of strategic justice. Thus, I defended the possibility of a strategically unjust social order by comparing the answers to three main questions of a conventional theory of (in)justice that allows for this possibility and one that does not.

The first question was about the necessary background conditions for (in)justice. I argued that in addition to the evolutionary models that explain the emergence of correlative equilibrium solutions, the conditions that Vanderschraaf enumerates as the background conditions necessary for a just cooperation in addition to the explanatory models of justice as conventions make injustice an anomaly. However, not only is injustice not an anomaly, but there are also alternative evolutionary models that account for its pervasiveness. I provided agent-based simulations of a famous fair division problem to show what it takes for these models to account for the pervasiveness of unfair strategies and unjust treatments. Based on these models, I argued that individuals' discretion over at least some aspects of their interactions with others is a

necessary background condition to explain the pervasiveness of injustice. This kind of injustice is by definition strategic, and I showed that it can be a common phenomenon.

The second question was about the emergence of unjust conventions or systems of conventions. I relied on Vanderschraaf's distinction between just and unjust conventions to show that the emergence of unjust conventions is not always the coincidence he suggests. I argued that even if coincidence is always the cause, arriving at a particular convention requires mechanisms that are sensitive to structures and network effects. Moreover, I showed that structures are not fixed features of the environment, and the most general way to formalize their formation is by accounting for individuals' strategic choices about their interactions. In fact, these strategic choices and the structure of social interactions create a feedback loop that explain the emergence of unjust treatments and even an unjust social order. In sum, assuming coincidental emergence of unjust conventions seems to involve individuals' strategic choices about their social relations, which is a contradiction. I concluded that unjust conventions that are not the result of an agreement are not coincidences and that individuals' ability to control their interactions with others, even in the absence of information about others' strategies, explains their emergence. In other words, strategic injustice is possible, and modeling its emergence requires the involvement of the feedback loop between individuals' strategic choices about their social environment and their resulting structure of cooperation.

The fact that in addition to individuals, the topology of the network they form determines the outcome of their interactions implies that individuals on their own cannot address strategic injustices. In fact, for network formation processes like "preferential attachment" in which disparity in connectivity is inevitable, individuals are bound to make connections that feed into greater advantage for some at the expense of others. Thus, maintaining a rough equality, something that would guarantee enjoying a minimal immunity for everyone, requires solving a multi-agent coordination problem. In other words, an alternative network of social relations that can coordinate actions among individuals across groups, formal and informal organizations, and hierarchies is necessary to integrate marginalized individuals into the networks of cooperation, power, and support. Scholars of social movements have long argued that these movements provide such alternative networks through various mechanisms that incentivize interaction among distant parts of the social networks.

References

Alexander, J. M. (2007). *The structural evolution of morality.* Cambridge: Cambrdige University Press.

Alexander, J. M., & Skyrms, B. (1999). Bargaining with neighbors: Is justice contagious? *The Journal of Philosophy*, 588-598.

Anderson, E. (2010). *The Imperative of Integration.* Princton; Oxford: Princeton University Press.

Axelrod, R. (1984). *The Evolution of Cooperation.* New York: Basic Books, Inc., Publishers.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *science*, 1390-1396.

Banerjee, A., Breza, E., Chandreasekhar, A. G., Duflo, E., & Jackson , M. O. (2021). Changes in social network structure in response to exposure to formal credit markets. *National Bureau of Economic Research*.

Barry, B. (1989). *Theories of justice.* University of California Press.

Basu, K. (1986). One Kind of Power. *Oxford Economic Papers*, 259-282.

Boserup, B., McKenney, M., & Elkbuli, A. (2020). Alarming trends in US domestic violence during the COVID-19 pandemic. *The American journal of emergency medicine*, 2753-2755.

Costa, D. (2019, August 28). *Employers increase their profits and put downward pressure on wages and labor standards by exploiting migrant workers*. Retrieved from Economic Policy Institute: https://www.epi.org/publication/labor-day-2019-immigration-policy/

Cudd, A. (2005). How to Explain Oppression. *Philosophy of the Social Sciences*, 20-49.

Diani, M., & Mische, A. (2015). Network Approaches and Social Movements. In D. Della Porta, & M. Diani, *The Oxford Handbook of Social Movements.* Oxford University Press.

DiMaggio, P., & Garip, F. (2012). Network Effects and Social Inequality. *Annual Review of Sociology* , 93-118.

Folbre, N. (2021). *Gender inequality and bargaining in the U.S. labor market.* Economic Policy Institute.

Foucault, M. (1977). Historia de la medicalización. *Educación médica y salud,*, 3-25.

Gordon-Bouvier , E. (2020). *In: Relational Vulnerability.* Palgrave Macmillan.

Haslanger, S. (2015). What is a (social) structural explanation? *Philosophical Studies*, 113-130.

Hofbauer, J., & Sigmund, K. (1988). *The Theory of Evolution and Dynamical Systems.* Cambridge, UK: Cambridge University Press.

Hume, D. (1777/1975). Enquiries Concerning Human Understanding and Concerning the Principles of Morals. In L. A. Selby-Brigge. Oxford: Clarendon Press.

Jackson , M. O., & Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of economic theory*, 44-74.

Jackson, M. O. (2008). *Social and Economic Networks.* USA: Princeton University Press.

Jackson, M. O., & Rogers, B. W. (2007). Meeting strangers and friends of friends: How random are social networks? *American Economic Review*, 890-915.

Levando, D. (2017). Formation of coalition structures as a non-cooperative game. . *Higher School of Economics Research*.

Lewis, D. (1969). *Convention.* Cambridge: Harvard University Press.

Mills, C. (2008). Racial Liberalism. *PMLA*, 1380-1397.

Myerson, R. B. (1976). Graphs and cooperation in games. . *Mathematics of operations research*, 225-229.

Neumark, D. (2018). Experimental research on labor market discrimination. *Journal of Economic Literature*, 799-866.

Rosenfeld, R. A., & Kalleberg, A. L. (1991). Gender inequality in the labor market: a cross-national perspective. , 34(3), . *Acta Sociologica*, 207-225.

Rothstein, R. (2017). *The Color of Law: A Forgotten History of How Our Government Segregated America.* Liveright Publishing Corporation.

Rubin, H., & O'Connor, C. (2018). Discrimination and collaboration in science. *Philosophy of Science*, 380-402.

Schauer, E. J., & Wheaton, E. M. (2006). Sex trafficking into the United States: A literature review. *Criminal Justice Review*, 146-169.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of mathematical sociology*, 143-186.

Skyrms, B. (1994). Sex and justice. *The Journal of Philosophy*, 305-320.

Skyrms, B. (1996). *Evolution of the Social Contract.* . Cambridge: Cambridge University Press.

Skyrms, B. (2004). *The stag hunt and the evolution of social structure.* Cambridge: Cambridge University Press.

Skyrms, B., & Pemantle, R. (2000). A dynamic model of social network formation. *Proceedings of the National Academy of Sciences*, 9340-9346.

Sugden, R. (1986). *The Economics of Rights, Co-operation, and Welfare.* New York: Palgrave Macmillan.

Szilagyi, M. N. (2003). An investigation of N-person Prisoners' Dilemmas. *Complex Systems*, 155-174.

Vanderschraaf, P. (2018). *Strategic Justice : Conventions and Problems of Balancing Divergent Interests.* Oxford: Oxford University Press.

Venkataramani , A., Daza , S., & Emanuel , E. (2020). Association of Social Mobility With the Income-Related Longevity Gap in the United States: A Cross-Sectional, County-Level Study. *JAMA Intern Med*.

Weibull, J. (1997). *Evolutionary Game Theory.* Cambridge, MA: MIT Press.

Young, H. P. (1993). An Evolutionary Model of Bargaining. *Journal of Economic Theory*, 145-168.

Young, I. M. (1988). Five Faces of Oppression. *Philosophical Forum*.

Young, I. M. (2012). *Responsibility for justice.* New York: Oxford University Press.