Ergo AN OPEN ACCESS JOURNAL OF PHILOSOPHY

On the Stability of Racial Capitalism

LIAM KOFI BRIGHT
London School of Economics and Political Science

NATHAN GABRIEL
University of California Merced

CAILIN O'CONNOR
University of California Irvine

OLÚFÉMI O. TÁÍWÒ
Georgetown University

What is the connection between capitalism and racial hierarchy? In line with the tradition known as 'the theory of racial capitalism' we show that the latter can functionally support the former. As a social construction, race has just those features which allow it to facilitate the sort of stable, inequitable distributions of resources that tend to emerge in capitalist systems. We support this claim using techniques from evolutionary game theory and cultural evolutionary theory, and end by discussing the normative political consequences of this relationship.

1. Introduction

The theory of racial capitalism proposes an origin story for how it is that the global economy came to be racially stratified and (in the main) organised along capitalist lines. The proposal is that the very same events led to both—Europe was already organising its workforces along proto-racial lines at about the time it was spreading its economic form through colonialism. As such, European expansion ended up simultaneously bringing capitalism and racial organisation in its wake. The evidence for this has been accumulated by historians, and we will discuss it in Section 2.

Contact: Liam Kofi Bright <liamkbright@gmail.com>

https://doi.org/10.3998/ergo.7965

However, many scholars make a somewhat stronger claim than noting the mere historical contingency that racism and capitalism co-occurred. Many argue that this coincidence is functional: the development of racial forms of social organization helped the capitalist mode of production survive and perpetuate itself. This is because capitalism will inevitably generate an unequal distribution of control over factors of production and (perhaps therefore) division of the resultant social surplus. Some means of explaining, justifying, and continuing this rampant and easily observed inequality is required, and, in particular, one which allows elites to retain their place. Race and racialism, by being easily observable, hard to change, and passed down across generations, worked nicely.

But why do these features of race work to stabilise capitalist systems? Using modelling techniques from evolutionary game theory, and drawing on some previous results, we show how oppressive schemes employing race are especially well-suited for underpinning stable and highly unequal systems of dividing labour and reward. We show how these models provide an explanation for the co-occurrence of race and capitalism that vindicates arguments from racial capitalist theory.

The rest of the paper will proceed as follows. Section 2 discusses theories of racial capitalism, especially those arguing that race plays a functional role in capitalist systems. Section 3 discusses the sort of explanation we take this paper to provide, that is, how we use these models to address the topic at hand. In §4 we describe the modeling paradigm we draw on here— evolutionary game theoretic models of the cultural evolution of inequitable systems of behavior. We discuss why this is a useful framework for analyzing theories of racial capitalism. Then, in §5 we describe in detail several different models intended to illuminate the functional role that various aspects of race play in capitalist systems. We start with the fact that race is hard to change or imitate. That is, it is fairly inflexible. We then discuss the fact that race is often fairly easy to identify compared to alternative tags or markers. And last we discuss the heritability of race. In each case we show how these features underpin systems of inequality. We argue that if powerful groups were to select some categorical system to ground inequality, they benefit themselves by focusing on race. And we show that in models powerful groups do indeed tend to culturally evolve to focus on race for this reason. Section 6 concludes.

2. Racial Capitalism

The theory of racial capitalism has multiple intellectual origins. The term "racial capitalism", and analyses stressing the inter-dependence of systems of race and

class, were developed by radical intellectuals in eastern and southern Africa, including African Marxists like Harold Wolpe and Neville Alexander (Al-Bulushi 2020). The term and associated analysis was later popularised in the US by Cedric Robinson (Kelley 2017; Robinson 2020). Robinson united the term, and some of the insights of the theory, with a related tradition of thought known as "world-systems theory" pioneered by African thinkers like Samir Amin as well as Robinson's Trinidadian-American predecessor Oliver C. Cox (Al-Bulushi 2020; Wallerstein 2000).

World-systems theory inspired theorists of racial capitalism to focus on the global political system as a whole rather than the social structure of particular nations (Hudson 2018; Táíwò 2022). It is this holistic version of racial capitalism that we focus on here. It involves two contentions about global social structure. First, racial divisions play a functional role in the social order built around capitalist production. Second, these divisions do so on a world scale (Cox 1948). Our paper attempts to provide evidence for the first claim, using models which are in principle consistent with the second.

Recent literature adopts these claims, but takes them to be well established. Perhaps as a result, few have directly argued for the claimed causal relationship between race and capitalism, namely that the former in some sense supports or props up the latter. Accounts tend to oscillate between descriptions where race is posited as logically or conceptually necessary for capitalism and descriptions where race is just contingently linked to capitalism's development and stability (Go 2021). One contribution of our paper, then, is to provide explicit arguments for a version of the existing view that race and capitalism are contingently but functionally linked (Dawson 2018).

Some accounts of racial capitalism do clearly assign racism a functional role in the maintenance of capitalism, but take that role to be primarily justificatory and ideological. On these accounts capitalism requires inequality or exploitation in some form or other, and race and racism provide a justification for inequality (Camp et al. 2019; Go 2021; Taylor 2016). A second contribution of our paper, then, is to show how racial stratification functionally contributes to the maintenance of capitalism in a different way than these authors contend. Categorical systems are crucial to grounding inequities, as we will discuss in §4. They allow unequal systems to persist by structuring the decision making environments for people who live within them. In particular, they prevent those who are oppressed from simply making behavioral choices that undermine structures of inequality, thereby solidifying essential features of any class hierarchy (Cicerchia 2021; Táíwò 2018). And they allow oppressors to easily identify an oppressed class and treat its members in such a way as to reap material rewards (Táíwò et al. 2021). Thus something like race is necessary to facilitate rules about who is oppressed by, and who benefits from, capitalist systems.

Once stabilised with racial stratifications, capitalist nations proved very good at generating surplus wealth for the powerful, and thus facilitating imperial conquest. This, in turn, spread the racist-capitalist system around the globe. Racism was thus good for capitalism, and capitalism in turn spread racism globally. This symbiotic relationship of functional assistance is core to the overall thesis of racial capitalism.

Our investigation also lends insight into why specifically racial forms of social organization proliferated out of the various possible forms of inequality that might prop up capitalism. Cedric Robinson argues that prototypical forms of racial organisation predated the development of modern capitalism, and thus were already available to the empires who built capitalism (Robinson 2020). In addition, the particular features of race are especially well suited to underpin capitalist inequality. Racial markers are relatively easy to observe, hard to change, and heritable, meaning that those caught in oppressive racialised systems cannot easily escape. Although other systems of inequity might have played a similar role, the availability of proto-racial concepts, and functionality of race in underpinning capitalism, help explain the swift spread of capitalist/racist systems.

In §§4 and 5 we discuss the models supporting this claim. But first we turn to a discussion of explanation in the social sciences to defend our use of models for this purpose.

3. Social Functional Explanation

According to theories of racial capitalism, a key part of what allowed capitalism to develop, persist, and spread via imperialism is the strategy of racial organisation of production and reward. This is a central claim we seek to defend in this paper, and to do so we use cultural evolutionary models in order to produce a type of functional explanation (Cohen 1978: ch. 9). However, in addition to defending the specific claims we do in the models below, we are aware that functional explanation in general has proven somewhat controversial in the social sciences (Kincaid 2007). We hence devote this short section to being clear about what we are and are not committed to in this paper.

Controversies about functional explanation have often centred on whether it is viciously circular, or somehow a-causal in a way that makes it inappropriate for naturalistic social science. Functional explanations can look like they posit something mysteriously teleological: because such and such would be good for a given system or institution, it therefore comes about. By itself that does not seem like enough to explain the original occurrence of the phenomenon in question; a mechanism is missing (Van Riel 2020). In fact, functionalist explanation has even been explicitly contrasted with game theoretic explanation by critics—the point

being to highlight that the latter, unlike the former, comes attached with plausible mechanisms which explain how its predicted equilibria may be expected to arise (Elster 1982).

Evidently we will not be contrasting game theory with functional explanation. Rather, our use of evolutionary game theory to model cultural evolution provides plausible causal mechanisms for the functional explanation we are supporting (Birch 2017; Lewens 2015: ch. 8). While there have been some doubts as to whether a selection mechanism is compatible with properly functional explanation, we set aside such demarcation problems as unimportant for our purposes (Jackson 2002: 173).

What matters for us is that we can clearly answer the question—why would the fact that racism stabilises capitalism explain the fact that capitalist social orders produce racially stratified labour forces? The answer is that in societies with racial divisions, stable capitalist inequalities can persist and be reinforced via cultural learning and evolution. The wealthy and powerful (and their progeny), tend to remain wealthy and powerful. As such, they are inclined to learn to pay attention to racial divisions, and disinclined to move their social form away from capitalism/racism. What is more, though this is not explicitly modeled here, they can use their riches to displace the ruling classes of societies that did not adopt as efficient a means of generating social surplus, and adapt the conquered territory in the conquerors' image. This overall story, then, provides a functional explanation for why capitalist societies have been racially organised societies: the stablisation provided by racial stratification encourages growth and spread, and so successful capitalist societies that survived tended to be the ones which made use of racial stratification.

Hence, by bolstering theories of racial capitalism with the evolutionary game theoretic functional explanation we offer here, we allow these theories to better explain the spread and the resilience of racist capitalist social organisations (Pettit 1998). As a social marker, race has just the features most useful for stabilizing hierachical societies. Capitalism is one such hierarchical society, and perhaps especially in need of such a stabilisation mechanism, as there are natural inclinations for the oppressed to disrupt the system. To see how racial organisation can work in this way, we now turn to explaining and analysing our models.

4. Modeling Inequity

In this paper, we draw on models of inequitable, discriminatory, or oppressive systems to think about why race has played such a key role in the history of the expansion of capitalism. The question we address now is: what sort of modeling framework can capture the phenomenon we are interested in?

We draw on a framework where social groups learn, or culturally evolve, to bargain with each other. Bargaining, as a broad phenomenon, is at the heart of how humans divide resources, labor, and the fruits of production. Sometimes bargains are explicit, sometimes they are implicit. Whenever humans work together to create goods and services, though, they must somehow come to an agreement about who will do what labor, how much of it, and who will reap what rewards from their joint production. We take this as a minimal model of the sort of productive enterprise where racial systems can be used to determine roles and rewards.[1]

A number of authors have looked, in particular, at bargaining models where social categories or social identity groups are present. In these models, it is assumed that individuals may use the identities of others in deciding how to bargain. For instance, a member of one race may learn to make fair bargains with those in their in-group, while bargaining aggressively with those in their out-group. A key finding in this literature is that the presence of identity groups, plus the ability to condition bargaining behavior on others' social identities, deeply impacts what sorts of social systems can emerge. In uniform groups, fair bargaining tends to be the norm (Alexander & Skyrms 1999; Skyrms 1994). In groups with social categories, inequitable systems commonly emerge where one group systematically gets more, and the other less (Axtell et al. 2001; O'Connor 2019). Thus this sort of bargaining model can serve as a simple representation of systems like those that have arisen during the spread of capitalism, where race plays a key role in organizing who does what labor and how much, and who reaps the rewards.

Let us go into more detail describing how models in this paradigm usually work, and why social categories contribute to the emergence of inequity. These models typically represent bargaining using what is called the *Nash demand game*.[2] In this game, two actors must divide some resource. Each can make a "demand" for some portion of it. The game assumes that if their demands are compatible in that they do not over-demand the resource, each actor gets what they request. If the demands are mutually too aggressive it is assumed that the bargain fails and each actor receives a poor payoff labelled the "disagreement point". Thus there are a mix of considerations for the actors. They would each like more resource, but must take care not to be so aggressive that they are unable to come to an agreement.

Nash equilibria in game theory refer to strategy pairings where no actors can improve payoffs by unilaterally changing strategies. For this reason, these equilibria

1. Notice that bargaining models can be used to represent even highly oppressive systems. We might not usually think of a chattel slave, or a woman in a repressively patriarchal society, as able to bargain. But even in these most extreme situations people can withhold labor and otherwise work towards their own interests. For a dramatic historical illustration of this see Du Bois (1935: ch. 4).

2. This game has its roots in the bargaining problem introduced by Nash (1950).

tend to be stable, and thus predict behavior in games. The equilibria of the Nash demand game correspond to strategy pairings where actors perfectly divide the resource (like 55% and 45%, for instance). At these pairings, no one wants to increase their demand, because they would then receive the disagreement point, while a lower demand will simply yield a lower payoff.

This model has been used across a wide range of applications—bargaining over salary, division of labor, conventions for dividing crops between sharecroppers and landowners, etc. Importantly, it captures scenarios where individuals divide some resource, but there are many ways that this division could go. One could get more, the other could get more, the division could be approximately fair, or it could be extremely uneven.

Of interest to us here are cultural evolutionary models that incorporate Nash demand games. These models track wide social patterns where groups interact, bargain, and learn bargaining strategies over time. What happens in models of this sort?

As briefly described, when actors in a single group learn to play the Nash demand game with each other, they tend to learn "fair" demands, or to request half the resource (Alexander & Skyrms 1999; Alexander 2007; Skyrms 1994; Sugden 1986; Young 1993). There are many ways to model cultural evolution—actors can imitate successful group members, repeat their past successful behaviors, try to respond well to behaviors they observe from interactive partners, etc. This finding is stable across all these sorts of models because 50/50 splits are special. This is the only symmetric equilibrium of the Nash demand game. This means it is the only equilibrium an entire group can adopt and guarantee that each pairing of individuals will efficiently divide the resource.[3] Whenever some members in a group make demands other than the fair one, they are guaranteed to sometimes fail to split resources efficiently with others. Those with low demands will sometimes leave resource behind, those with high demands will sometimes get the disagreement point.

Things change in models with social identity groups. These models typically incorporate what are called "tags"—arbitrary labels that do not initially signal anything meaningful about individuals, but that do create an arbitrary distinction between them. Tags can act as bare bones representations of social categories or identity groups like genders, races, ethnics groups, religious groups, etc.[4]

3. This is not the only stable evolutionary outcome. Some models of this sort see the emergence of "fractious" outcomes where some actors make high demands and others make compatible low demands, but these are always less likely than "fair" outcomes. They are also relatively inefficient (Skyrms 2014).

4. In principle they could represent divisions as arbitrary as folks who like Nicolas Cage versus the strange ducks who do not. But in general tags are useful to modelers in representing divisions that matter to social behavior, like those listed here.

Importantly, actors can condition their behaviors based on these tags. If we label two tags "yellow" and "blue", actors can choose strategies like "make aggressive demands of yellows and accommodating demands of blues".

As noted, this is the small shift that makes possible a whole new set of cultural evolutionary outcomes. In these outcomes, social groups can systematically treat each other unfairly. Equilibria can emerge where, say, all yellows make high bargaining demands when they meet blues, while blues make low demands in response. This new possibility emerges because tags provide an arbitrary asymmetry on which to condition strategies (Axtell et al. 2001; O'Connor 2019). With no tags, actors must treat all others equally in order to efficiently divide resources. With tags, they can learn conditional rules, and social identity can become relevant to how bargains happen (Bowles & Naidu 2006; Cochran & O'Connor 2019; Hoffmann 2006; O'Connor et al. 2019; Poza et al. 2011; Rubin & O'Connor 2018; Stewart 2010). For instance, if all yellows make aggressive demands of blues, and all blues concede, the presence of tags allows for an efficient, inequitable system that is not possible in a tag-free society.

Once these sorts of discriminatory outcomes emerge, they are stable in a wide range of models. This may sound unintuitive—why would actors adhere to an outcome that disadvantages them? Why not just refuse to comply? Remember that over-aggressive pairs of demands in bargaining scenarios yield poor outcomes. This means that disadvantaged players cannot change strategies on their own without lowering their own payoffs. If blues always meet aggressive demands from yellows, their best response is to concede. In this way, these modeling outcomes mimic inequitable conventions in the real world, such as those that stipulate lower pay for Black workers.[5] Black people in such arrangements cannot univocally demand higher wages and expect to receive them. Instead, they will experience bargaining failures with employers who expect to pay low wages to Black people. This helps explain how race stabilises capitalist systems, where advantaged classes profit from the effort of the less advantaged. Race creates an asymmetry where those associated with a less advantaged social position receive their best payoffs by complying with the system. Those in advantaged positions can profit from this fact.

The most advantaged position one can hold in these models is membership in a group that discriminates against another. Such agents reap systematic rewards because they get high payoffs from bargains with their out-group, while the out-group is systematically disadvantaged. For this reason, such groups have strong incentives to perpetuate discriminatory systems. But what tends to predict

5. Wilson and Rogers III (2016), for example, find that Black workers systematically receive lower pay in the United States even when controlling for variables like education, experience, and region.

which groups end up advantaged? This kind of advantage can arise randomly as an accident of history, but a series of models also show that power can act as a symmetry breaker which tends to advantage groups in evolved bargaining conventions (Bruner & O'Connor 2017; LaCroix & O'Connor 2021). While power can refer to many different sorts of things, most of these models operationalise it using the disagreement point of the Nash demand game, following Nash (1951).[6] The disagreement point, remember, is what actors expect to get should bargaining fail. The idea is that one is in a more powerful bargaining position if the bargain is less important to them. If bargaining breaks down, they are still in a relatively good position compared to those with lower disagreement points.

One can then incorporate power into cultural evolutionary bargaining models by supposing that one social group tends to have higher disagreement points than another. This sort of power translates to a bargaining advantage. Groups with high disagreement points tend to end up at conventions that advantage them. This is the last important part of the modeling framework we use here to address racial capitalism. It allows us to consider questions like: what sorts of tags allow powerful groups to gain the most from discriminatory bargaining outcomes? And: what sorts of tags can powerful groups use to ground oppressive systems?

At this point, it is hopefully becoming clear how the modeling paradigm presented here supports some key claims from racial capitalist theories. These models show that if there are fault lines of social difference, these can become loci of social inequity. Societies without such fault lines, on the other hand, do not tend to evolve towards such inequalities. In other words, those who seek to gain from capitalism do well to use social identity or social categories to organise their systems of production. In the next section, we will introduce some other modeling work from the literature, and several new models, to ask: why race in particular as an organizational system for capitalism?

5. Modeling Racial Capitalism

In the development of capitalistic oppression, why has race been so systematically employed as a key locus of oppression as opposed to other loci? In the last section, as we saw, categorical differences can play a crucial role in grounding inequality. But not all categorical differences are alike. In particular, we will now consider some of the features of race as a social category that make it particularly useful for powerful individuals seeking advantage through production. The three features

6. There is also a long history of using the disagreement point to represent power in models of household division of labor. See, for example, Manser and Brown (1980) and McElroy & Horney (1981).

we focus on here are, 1) the relative stability and inflexibility of race as a tag or social marker, 2) the relative reliability of race as a tag, and 3) the heritability of race.

5.1. Flexible Tags

Although there are different systems for race, those developed along with capitalism in the western world tend to sort people into distinct categories that are considered innate, that are treated as heritable, and that are inflexible (Appiah 1998; Robinson 2020; Schachter et al. 2021). These systems tend to associate racial classification with biological markers (e.g., skin color, hair texture) which are hard to change. Moreover, historical evidence shows deliberate efforts to increase or safeguard the reliability of racial tags. European controlled colonies, for example, often adopted "sumptuary laws" that prevented racially dominated groups from using forms of self-presentation that might lower the reliability of markers of racial classification, such as wearing jewelry and adopting certain styles of dress associated with more racially advantaged groups (Earle 2003; Pastore 2002).

This first set of models considers the emergence of inequity when categories are inflexible, as these sorts of racial categories tend to be, versus in cases where actors have some choice of category membership. The suggestion here will be that employing inflexible categories in the creation of inequitable systems yields particular benefits for powerful actors. For this reason, these systems tend to emerge over evolutionary time.[7]

Critically, in the sorts of models here, inequitable bargaining conventions should be unstable if actors can easily change their tag. If this were possible, then upon the emergence of an unfair outcome, we would expect members of the disadvantaged group to simply switch tags. Why be part of a disadvantaged group if you can join an advantaged one?[8] Upon doing so, they could garner bargaining advantages for themselves. But if all disadvantaged members make the switch, the model reverts back to a single-group model, where fairness tends to be the rule.[9]

7. Mills (2014) makes a similar, informal argument about the shift in focus on Muslims as the "infidel" outsider of the European world, to "savages". As he says, "[r]ace gradually became the formal marker of this differentiated status, replacing the religious divide (whose disadvantage, after all, was that it could always be overcome through conversion)" (23).

8. Of course in reality there may be many reasons individuals are unwilling to abandon their social identities. The point is that when one can switch groups, there is a material incentive to do so. And absent reasons for staying, we should expect individuals to adopt advantaged identities.

9. In general, tag flexibility can greatly alter the evolutionary outcomes of game theoretic models. For instance Bruner (2015) shows how flexibility to alter tags can lead to greater tolerance in cooperative scenarios.

Popa et al. (2021) consider a model along these lines. Their key intervention is to compare two conditions. In the first, individuals in each group are assigned tags that match their group—call these α and β to match groups A and B. These tags are immutable, and immediately recognizable by those in both groups (as in the models from the last section). In the second condition, individuals are able to update their tags, and imitate both strategy and tag from successful members of their group.[10] They find that in the latter condition, fairness is basically guaranteed to emerge, even when they incorporate power imbalances between the groups. This happens for the reason just described—if disadvantage begins to emerge for one group, members of that group adopt tags that make it impossible to identify them. This ensures fair treatment.

Popa et al. (2021) consider a further pair of models. In the first model, only one powerful group has the ability to flexibly change tags. In the second, only one less powerful group has this ability. When the disempowered group is unable to change, they find that discrimination regularly emerges. The powerful group often adopts a distinctive tag for themselves, and uses the inflexible tag of the other group to organise bargaining inequity. Furthermore, this is more likely when the powerful group is particularly powerful (i.e., has a very high disagreement point). When the disempowered group can change tags, on the other hand, they "camouflage" by mimicking the tag of the powerful group, and preventing discrimination.

We extend this exploration by considering a few other possibilities for flexible markers.[11] We assume, as in previous models, a group of individuals divided into two groups—A and B. We also assume these individuals regularly engage in bargaining scenarios represented by a simplified Nash demand game with three demands for a high, medium, or low amount of some resource of value 1. The medium demand always represents the fair split, or 1/2, while the high and low demands represent compatible but unequal splits. In considering this simplified model, we follow previous authors. Figure 1 shows the *payoff table* for this game with low and high demands of 1/3 and 2/3.[12]

The three Nash equilibria are bolded. As in the more complex game, the equilibria are the strategy pairings where the resource is perfectly divided. These will also, generally, track the stable outcomes that will emerge between groups in our models. In other words, when we culturally evolve groups playing this game

10. They do so using pairwise proportional imitation dynamics (PPI) which will be introduced shortly.

11. All reported results in this paper were produced twice, independently to ensure replicability. Codes are available at https://github.com/cailinmeister/racialcapitalism and https://github.com/NattyGabe/racial_capitalism.

12. Previous work suggests that other divisions should yield qualitatively simliar results (O'Connor 2019).

		Player 2		
		Low	**Med**	**High**
Player 1	**Low**	1/3,1/3	1/3,1/2	**1/3,2/3**
	Med	1/2,1/3	**1/2,1/2**	0,0
	High	**2/3,1/3**	0,0	0,0

Figure 1: Payoff table for a simplified Nash demand game.

with each other, they will tend to end up at one of three states. Either 1) group A demands high, and group B low, or 2) both groups demand medium, or 3) group A demands low and group B high. Within each group, fair behavior will tend to emerge since in-groups act like unitary populations.[13] Thus despite its simplicity this game has the representative power to capture a situation where 1) discrimination is not inevitable, that is, there is a fair, non-discriminatory equilibrium, but 2) it is stable, that is, it is the endpoint of many cultural evolutionary models.

There are many rules for learning one could employ to represent cultural evolution in this model. We use the pairwise proportional imitation dynamic (PPI) introduced by Schlag (1998).[14] The important feature of this dynamic is that successful strategies tend to spread in proportion to their success. Agents in the two subgroups interact over time and tend to imitate successful members of their own subgroup. In particular we consider an agent-based model where in each round of simulation each agent meets every other for interaction. At the end of the round, each agent is paired with a member of their own subgroup for possible imitation. The likelihood that imitation occurs tracks the payoff difference between the two individuals, where high payoff strategies tend to be copied. In particular the likelihood that agent i copies their imitation partner j is 0 if i outperformed j. But if j did better the probability is:

$$p_{i,j} = (u_j - u_i)/u_m \tag{1}$$

where u_i is the total payoff of agent i in the last round of interaction. The term u_m refers to the largest possible payoff difference between the players.

In our first extension of this model, we add the simple possibility that agents are sometimes able to experiment by changing tags and strategies. This is included by adding a mutation rate, μ. In each round of simulation with probability μ each agents mutates. In doing so, they randomly select a tag (either the group A tag, α,

13. As we will note, in some cases our models do not approximate these outcomes because of effects of drift and mutation.

14. It shares characteristics with the replicator dynamics—the most commonly used dynamics in evolutionary game theory (Izquierdo et al. 2019). We follow Popa et al. (2021) in this choice.

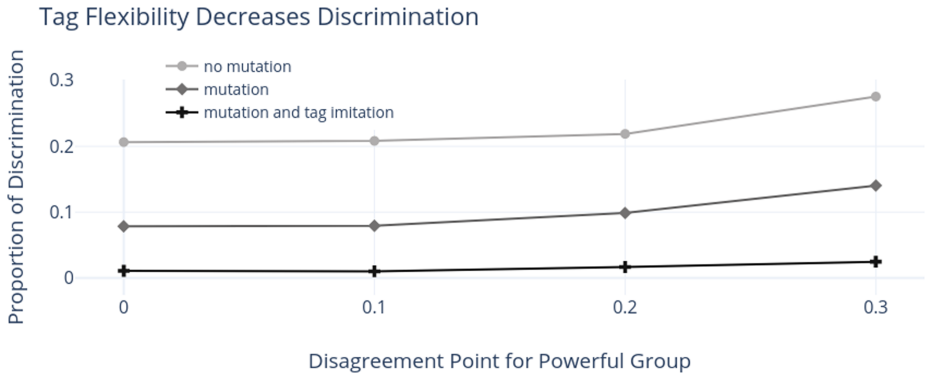

Figure 2: Adding the ability to randomly adopt new tags dramatically decreases incidences of discrimination. If actors can imitate group members who do so, this further decreases discrimination. Figure shows $N = 20$, equally sized groups, $\mu = .001.$, $\rho = .01$.

or group B tag, β) as well as strategies to employ.[15] We always start simulations with tags that "match" each agents' group membership. We run the simulation multiple times and observe what happens. As Figure 2 shows, mutation dramatically decreases the emergence of discriminatory behavior in the model.[16] Here discrimination tracks the proportion of interactions between individuals with different tags where one makes a high demand.[17] Note that even when the disagreement point for both groups is zero, the existence of inflexible social categories

15. In this model with no mutation the stable states that emerge tend to mimic the equilibria of the models described in §4. Within in-groups behavior is mostly fair (and sometimes fractious) because in-groups function like single populations. Between group behavior is either fair or discriminatory. Without mutation, our agent based models with small populations can sometimes drift into non-equilibrium behavior and get stuck there (i.e., all demanding low within one in-group, despite the fact that medium would do better). This is because strategies change by imitation, so if a strategy randomly dies out it cannot re-enter the population. Once mutation is added, this does not happen—successful strategies that randomly die out can reappear and spread. With mutation there are also no fully stable outcomes because actors periodically adopt new, random strategies. Groups tend to play strategies that are close to some equilibrium, with occasional mutations away from it. In general for both models with and without mutation, smaller populations are less likely to end up at outcomes approximating equilibrium as a result of random events.

16. Results are averages across 2k runs of simulation. Simulations were run for 1k time steps without mutation, and 3k with mutation. With no mutation the shorter simulation was enough to guarantee convergence to a stable outcome. With mutation outcomes are less stable, and longer runs were necessary.

17. We use this measure since, given the details of the model, we cannot report on the prevalence of various equilibria at the end of simulation. Notice that the amount of discrimination on this measure is influenced by the prevalence of different tags. If all actors adopt the same tag, no discrimination is possible. If just one actor adopts a different tag, their behavior strongly determines the level of discrimination. This means the measure will sometimes fail to disambiguate between cases where a few individuals with a different tag discriminate and where there is a wider pattern of between-group discrimination. In general, though, we find that the measure intuitively tracks cases we would deem as discriminatory. In addition, we calculated a different measure: the proportion of

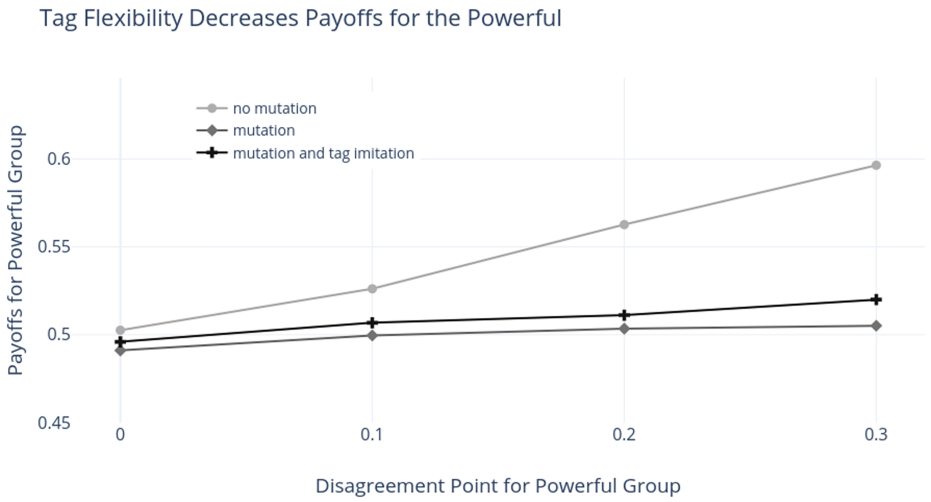

Figure 3: Adding the ability to adopt new tags dramatically decreases the average payoff of a powerful group. Figure shows $N = 40$, equally sized groups, $\mu = .001$.

(no mutation) allows discrimination to emerge. Increasing the disagreement point for the powerful group exacerbates the propensity for discrimination to emerge.

The reduction in discriminatory behavior here stems from the way that flexibility in tags reduces their correlation with behavior. For instance, the group A tag stops being able to dependably carry the meaning "this individual will demand low" or "I can safely discriminate against this individual". When this correlation is broken, systemic discrimination does not work, and cannot benefit one dominant group.

We also consider the possibility that agents can imitate tags from successful group members, as in Popa et al. (2021). That is, if some other group member experiments and adopts an out-group tag, others in their group will copy this strategy if it yields high payoffs. This extra addition lowers the chances that discrimination emerges beyond that from mutation alone (Figure 2). Once agents can experiment and try an out-group tag, imitation allows successful experiments of this sort to spread through the group.[18]

We can use these results to ask: if a powerful group were to choose a tag system, what would they prefer? Would something relatively inflexible like race be a preferred system? Or something more flexible? Figure 3 shows the average payoffs to a powerful group in these models with and without the ability to switch

high demands against out-group out of all interactions between agents. All qualitative results were the same.

18. The degree to which mutation and imitation reduce discrimination varies across different parameter values for these models, and, in particular, group size. But for these results, and others in the paper, reported trends are always stable across parameters.

tags (and to imitate tags from successful group members). In particular, we focus on the payoffs agents from the powerful group are getting when interacting with agents from the less powerful group. The highest possible average payoff in this scenario would be 2/3, if the powerful agents always successfully demanded high, whereas if all outcomes were fair payoffs would be 1/2. As we see, with inflexible tags, powerful groups can expect to end up with a higher average payoff. This suggests that the strategic use of inflexible tags is relevant to those building systems of oppression.[19]

Another way to incorporate tags flexibility is to allow imitation between groups. In this next model, we assume that most of the time agents imitate within their own groups. With a small probability, ϕ, they imitate an agent in the other group instead. This allows agents to adopt both tags and strategies from those outside their identity group. (Though we assume that power is a property of the individual, and not changeable.) In this model with even a very small probability of out-group imitation, we find that either 1) groups evolve fair treatment, and a diversity of tags are preserved, or else 2) if unfair treatment emerges the entire group evolves to use the same tag, thus making their payoffs equivalent.[20] As we see in Figure 4 adding even a small probability of out-group rather than in-group imitation entirely eliminates out-group discrimination.[21] This, again, suggests that a powerful group should prefer tags that cannot be imitated by an oppressed group in order to preserve their bargaining advantages.

It is the ability to change tags that prevents discrimination in this model. In particular, similar to findings from Popa et al. (2021), we find that it is the ability of the less powerful group to imitate the tags of the dominant group that protects them from discrimination. As one group becomes more powerful, it becomes more and more likely that members of the entire population mimic their tag in these models, as is evident in Figure 5. This figure shows the percentage of agents on average who, at the end of simulation, use the tag initially associated with the powerful group. In order to preserve their bargaining advantage it is thus incumbent on the powerful group to choose tags that are hard

19. Notice that in this figure the powerful group does slightly better with mutation and tag imitation than mutation alone. This is true even though tag imitation helps reduce discrimination. The reason is that with mutation only, the models tend to end up in suboptimal states more often, and all players receiver lower payoffs as a result.

20. Interestingly, this model sometimes evolves to an outcome where all agents identify as members of one group, but where all the powerful agents demand High and all the less powerful agents demand Low. That is, demands are still correlated to original group identity. In these cases, the entire population mimics a fractious, single group, and all agents expect the same payoffs, but original group identity is relevant in determining how the fractious strategies are distributed.

21. We assume no mutation of strategies in this model. We assume that when actors imitate their in-group, they copy tags. Without this assumption, out-group imitation still significantly decreases the prevalence of discrimination.

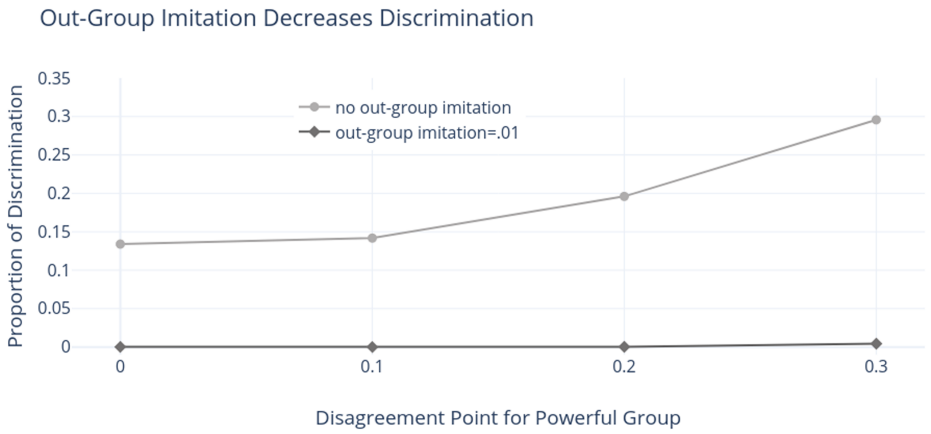

Figure 4: If individuals can imitate their out-group with respect to tags and strategies, discrimination is eliminated. Figure shows $N = 40$, equally sized groups, $\phi = .01$, $\mu = 0$.

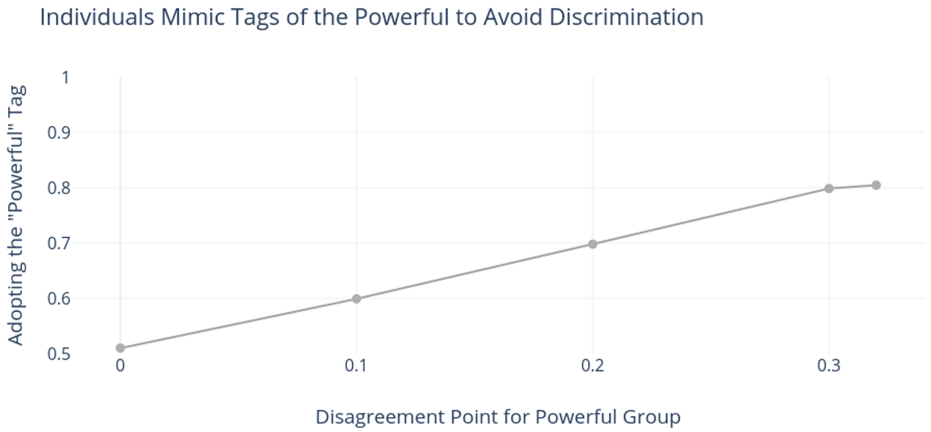

Figure 5: Individuals tend to imitate the tags of powerful groups to avoid discrimination. Figure shows $N = 40$, equally sized groups, $\phi = .01$.

to imitate, or else to actively shape tag systems that are inflexible. Race is such a marker. Cultural rules like sumptuary laws remove further flexibility from racial tags.

We take the observations in this section to provide an explanation for why capitalist systems are racialised. Inequitable systems, like capitalism, require loci of inequity in order to emerge and stabilise. In this section we have shown that race, in particular, is especially useful to the beneficiaries of capitalist systems. This suggests that these beneficiaries should prefer to focus on tags like race for this purpose, rather than flexible alternatives. We have not provided a specific mechanism by which powerful groups actually do come to attend to tags like race, though. Given the payoff benefits of doing so, this could happen via many

processes of cultural learning and evolution. We return to this point at the end of the next section.

5.2. Unreliable Tags

Some categorical systems depend on markers that are easily recognizable, as with, for instance, gender systems built on biological sex differences. While these biological differences need not entirely determine sex categorization (i.e., how the sex of an individual is perceived), they strongly constrain this process. Other markers may be less dependably identifiable. For instance, categories are sometimes built upon ethnic distinctions associated with small, or hard to detect, phenotypic differences. In this section we consider models where tags are only partly recognizable. The goal will be to compare systems where tags are easily perceived vs. those where they are not, and, again, to ask which sort of system a powerful group might prefer, and how this impacts cultural evolution.

Again, Popa et al. (2021) consider a model along these lines. They suppose that the members of each group are seen as belonging to their own group with some probabilities p_A and p_B, and as belonging to the other group with probabilities $1-p_A$ and $1-p_B$. The question is then: if tags are only partially informative of group membership, does inequality emerge? They find that inequality is much more likely for highly informative tags, than for partially informative ones. With only partially informative tags, an advantaged group can only sometimes effectively discriminate, and in other situations will miscoordinate if they try.[22]

If we consider this from the perspective of an advantaged or powerful group, we can see that they reap the greatest benefits when they settle on a tag system that allows them to dependably identify those in the less powerful group. Otherwise, they end up making more fair agreements, and sharing payoffs.

For this reason, we extend this model by considering individuals who culturally evolve to attend to some tags and not others. We assume that there is a powerful group and a less powerful group and that they are engaged in bargaining interactions. We assume there are two types of tags associated with the less powerful group—these might be skin color and height for example—and that the powerful group can choose to pay attention to either of these. One tag identifies members of the less powerful group 100% of the time. The other is unreliable with probability p_t, and so powerful agents who attend to it sometimes mistakenly identify their interactive partner as an in-group member.

22. Bruner (2019) looks at a similar model with interacting groups evolving to play bargaining games. He is interested in whether a particular effect that disadvantages agents in a minority group will hold up when tags are only partially informative. He finds that it only does when tags still are largely dependable, that is, identify group members more than 95% of the time.

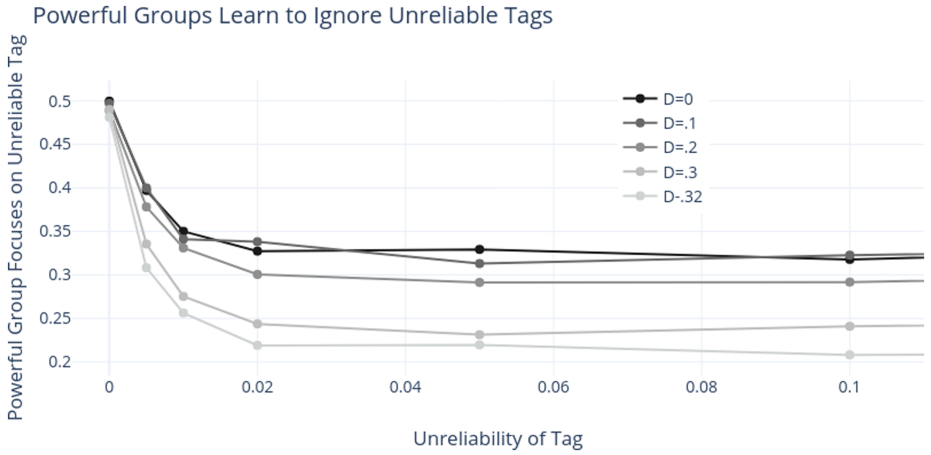

Figure 6: If given a choice of tags to attend to, powerful groups tend to choose reliable markers of out-group identity. Figure shows $N = 40$, equally sized groups, mutation rate = .001.

The strategy of the powerful group then involves choices of 1) what strategy to play against perceived in-group members, 2) what strategy to play against out-group members, and 4) which sort of out-group tag to attend to.

We run simulations of this model, varying both the disagreement point of the powerful group, and also p_t, or how unreliable the unreliable tag is.[23] We ask, across these variations, how often do members of the powerful group end up attending to these different tags? Do they learn to choose the reliable one, allowing them to successfully identify out-group members? Does doing so facilitate discrimination?

Increasing the disagreement point of the powerful group increases the emergence of discrimination, as in other models. As one tag becomes less reliable, there is no change in this level of discrimination. This is because powerful agents instead learn to attend to the reliable tag, and use this marker to ground discrimination. Figure 6 shows the percentage of powerful agents on average who pay attention to the unreliable tag at the end of simulation. As is clear from the chart, the more unreliable a tag, the less likely agents are to pay attention to it. This effect is stronger the more powerful one group is. In particular, notice that when $p_t = 0$, or when the second tag is always reliable, powerful agents are equally likely to pay attention to either tag. Both tags work to identify out-group members. As

23. All simulations were run 2k times, and for 3k timesteps to ensure they reached approximately stable end points. We included a small probability of mutation of strategies (not tags) for all simulations. This prevented groups from getting stuck at inefficient outcomes as a result of random drift. We did not allow for any tag imitation or mutation. We assumed that members of the less powerful group always attend to the reliable tag, that is, they always correctly identify group members. Only members of the powerful group evolve their recognition strategy.

one tag becomes more unreliable, the rate at which it is attended to drops and the levels off.[24]

We also considered similar models, but where powerful actors learn to either attend to a flexible tag as in the last section (one that members of the less powerful group can mutate and imitate) or an inflexible one. Again we find that powerful groups learn to prefer inflexible tags because they can better ground inequitable systems. We do not present these results for space reasons.

The take-away here is that powerful groups can learn to attend to tags like race rather than others that are harder to identify. Doing so allows these groups to better benefit from systems of bargaining. Once again, we see why a tag like race, constructed so as to render it relatively easy to identify, plays a functional role in hierarchical capitalist societies.

5.3. Inheritance

One property of many systems of discrimination involves inheritance. That is, children inherit their social identity based on their parents' social identities. In some cases, these inheritance systems are specifically created in ways that allow powerful groups to take advantage of disempowered groups. A notable example is the "one drop rule", from the United States, where any Black ancestry was taken to confer status as a Black person. This rule was used to prevent mixed raced children of slave owners and other Whites from inheriting property or status from their white parents.

A question arises: as with inflexible markers and informative tags, are there reasons for the powerful to choose tags so that they will be inherited? Does this confer a payoff advantage? The answer is yes.

We do not build new models in discussing this point, but start by reinterpreting the models from §5.1. In those models, agents imitate those in their groups with successful strategies. In some cases, agents were not able to change tags. In other cases they were, either through experimentation or through imitation. We might instead think of this model as tracking not imitation, but generational turnover in the population. On this reading, strategies are changing because some agents die, and new ones are born. If the new agents always adopt the tags of their in-group, we have a model where tags are strongly heritable. If new agents are sometimes able to to adopt tags from a different group, or to mutate tags, tags are less heritable and discriminatory outcomes are less viable.

24. This leveling off is because there are always some simulations that end up with fair treatment between the groups. If so, it does not matter which tag is attended to because payoffs are the same.

We can also imagine new individuals born with tags that are potentially unreliable, as in the models from §5.2. This is of interest in the generational interpretation, since children can have parents from different racial groups. Rules for racial categorization—the "one drop rule" but also other more intricate systems of racial classification—create situations where discrimination is stabilised. Before instigating such rules, one might not know how to categorise a mixed race person, and thus how to treat them. The rule transforms this situation to one where most tags are reliable indicators of group membership.[25] Again, it is a transformation that make inequitable outcomes more likely.

6. Discussion

Our models show how race can undergird the emergence and stability of capitalist systems. This supports central claims of racial capitalist theory. Social categories can act as asymmetries that ground oppressive systems by enabling coordination around inequitable distributions. These systems can emerge via cultural learning, often taking advantage of and reinforcing prior existing modes of domination (Cicerchia 2022). Once they do, they are stable because even disadvantaged individuals do best to choose behaviors that conform to the patterns of behaviour which constitute and uphold that social system. As we show at length in §5, powerful groups have incentives to use tags like race, which are relatively inflexible, easy to identify, and heritable, to ground these systems. In addition, they can learn to do so in simple models. For these reasons, we suggest, capitalist systems have tended to use race as a locus for oppression. As discussed in §3, this explanation is a functional one, but also one that describes causal pathways by which this sort of functional system employing race to underpin capitalism can emerge in human societies.

Broadly, our analysis suggests there could be benefits if philosophers and social scientists engage further with the considerable literature on racial capitalism. These authors have something to offer by way of further developing racial capitalist theory, as our paper demonstrates. Further, philosophers and social scientists working on inequitable economic systems, and their emergence and stability, should take into account the role of racial markers in these systems. Racial capitalist theory provides deep insights into this role.

25. Our treatment here uses models that split the population into two groups (the racially advantaged and racially disadvantaged). This most closely matches a "one-drop rule" scenario. But in models with multiple groups, inequity is also stable (O'Connor et al. 2019). Future work could investigate differences between systems of racial organization, including kinds that involve three or more racial groups, which may more closely match cases where mixed-race groups form their own racial strata.

An important avenue for further analysis concerns the central point about system stability. If racial social systems are stable because they contribute to a payoff structure that no one has immediate incentive to deviate from, this invites the question of how to change such a social system.

These models suggest that attempts to change the social structure are likely to be ineffective if they focus only on moral persuasion. Individuals in this sort of system are making the best moves available to them, even those who reproduce the conditions of their own oppression as they do so. Those benefiting from the system have every incentive to resist moral education, perhaps by developing justifications for or ideologies supporting the social order (Kinney & Bright 2023). Those oppressed by the system may not be able to afford the sacrifices necessary to change it.[26]

On the other hand, interventions that facilitate collective bargaining and social movements by lowering their costs may be effective. In practice this could mean, for instance, creating organizations or spaces wherein members of an underclass can privately pool information and strategise (Goodin & Spiekermann 2015). Intervention that aims at changing features of the overall incentive structure, such that a critical mass of individuals face different incentives than they otherwise would, may likewise prove successful. For instance, the models indicate that power imbalances tend to lead to bargaining inequity. Then interventions that attempt to redistribute power, in the form of wealth redistribution, or the strengthening of democratic functioning, may help facilitate cultural movement towards fair norms of bargaining.

In addition, while race is a useful marker for capitalist exploitation, perhaps the flexibility inherent in a social construction may allow us to modify race itself such that it can no longer serve this purpose. The models in §5 show how interventions that erode the reliability and heritability of racial markers can disrupt unfair systems. As is emphasised by theorists such as Mills (2000), the construct of race has sufficiently many and varied moving parts that it is conceivable that by strategically insisting upon some elements rather than others we could over time break the associations that render race heritable or a reliable marker. This sort of ameliorative social ontological project has recently gained much currency in analytic philosophy (e.g., Dutilh-Novaes 2020; Haslanger 2005), and we hope that models such as ours can help direct attention to the sort of changes that would actually disrupt race's ability to play its functional role in oppressive systems.

26. Of course, moral norms do often successfully convince people to go against their own material interests (Bicchieri 2005). We do not mean to imply that ethical systems are unimportant here, but that moral education alone might not be enough to disrupt inequitable systems.

One further thing our models show is that there is a real risk that attempts at positive change might reproduce a similar order. Take, as a hypothetical example, an extreme form of credentialism. In such a system we sort people at a young age based on standardised test performance—and there is no widely available opportunity to go back and redo these tests once they have been carried out. The results of these tests are made easily accessible to potential employers, to societies one might wish to join, etc.

Suppose test performance is strongly correlated with what school one's parents could get one into, which is itself downstream of how well placed they are in the credentialing system. This hypothetical system could appear meritocratic. However, it functionally could play the same role race now does. One's scores are hard to change. They are easily observed by relevant decision makers. And via an inegalitarian education system they are passed down from generation to generation. Race, or at least something playing a similar functional role, would have found its way back into our mode of social organizaton.

This hypothetical highlights an important insight from our work. In changing the social structure one must attend to the particular features of race as a social construction which have allowed it to stabilise inequality. Otherwise one might simply reproduce its functional role in another guise. The hypothetical example was chosen for its resonance with features of contemporary social life in the U.K. and U.S.A. But there are many ways of producing systems of social markers that are hard to change, easily identifiable, and heritable. Any of these could serve as the means of upholding an inegalitarian social order.

7. Conclusion

Racial hierarchy can function to maintain capitalist hierarchies. That is not to say their connection is necessary or that no other way of upholding such a system could be devised. This is only to say that given features of the social construct of race, and given the distribution of power and resources capitalism in fact generates, the former complements and supports the latter.

We end with a few ideas for further work. An obvious next step from the modeling reported in this paper would be to consider more explicitly intersectional models—for instance ones where individuals have multiple identity markers. Another extension might consider gradient tagging systems and how they might relate to the fine grained colorism that many parts of the colonial Americas developed. There is much to learn here that could help us both better understand how our world came to be and the levers by which it may be changed.

Acknowledgements

Many thanks to Travis LaCroix, Kathleen Creel, Kevin Zollman, and Mihaela Popa for early comments on this manuscript. Special thanks to Nick Makins for invaluable critical feedback on an earlier iteration of the paper's argument structure. Thanks to our reviewers for their time and helpful comments. Dedicated to the memory of Charles Mills, friend and mentor to so many of us.

References

Al-Bulushi, Yousuf (2020). Thinking Racial Capitalism and Black Radicalism from Africa: An Intellectual Geography of Cedric Robinson's World-System. *Geoforum*, *132*(1), 252–262.

Alexander, J. McKenzie (2007). *The Structural Evolution of Morality.* Cambridge University Press.

Alexander, Jason and Brian Skyrms (1999). Bargaining with Neighbors: Is Justice Contagious? *The Journal of philosophy*, *96*(11), 588–598.

Appiah, Kwame Anthony (1998). *Race, Culture, Identity: Misunderstood Connections.* Princeton University Press.

Axtell, Robert L., Joshua M. Epstein, and H. Peyton Young (2001). The Emergence of Classes in a Multiagent Bargaining Model. In *Social dynamics* (191–211). MIT Press.

Bicchieri, Cristina (2005). *The Grammar of Society: The Nature and Dynamics of Social Norms.* Cambridge University Press.

Birch, Jonathan (2017). *The Philosophy of Social Evolution.* Oxford University Press.

Bowles, Samuel and Suresh Naidu (2006). Persistent Institutions. *Working Paper*, *08-04*(1), 1–26.

Bruner, Justin and Cailin O'Connor (2017). Power, Bargaining, and Collaboration. In T. Boyer-Kassem, C. Mayo-Wilson and M. Weisberg (Eds.), *Scientific Collaboration and Collective Knowledge* (135–159). Oxford University Press.

Bruner, Justin P. (2015). Diversity, Tolerance, and the Social Contract. *Politics, Philosophy & Economics*, *14*(4), 429–448.

Bruner, Justin P. (2019). Minority (Dis) advantage in Population Games. *Synthese*, *196*(1), 413–427.

Camp, Jordan T., Christina Heatherton, and Manu Karuka (2019). A Response to Nancy Fraser. *Politics/Letters*, *15*(1), 1–6.

Cicerchia, Lillian (2021). Why Does Class Matter? *Social Theory and Practice*, *47*(4), 603–627.

Cicerchia, Lillian (2022). Structural Domination in the Labor Market. *European Journal of Political Theory*, *21*(1), 4–24.

Cochran, Calvin and Cailin O'Connor (2019). Inequality and Inequity in the Emergence of Conventions. *Politics, Philosophy & Economics*, *18*(3), 264–281.

Cohen, G. A. (1978). *Karl Marx's Theory of History: A Defence.* Princeton University Press.

Cox, Oliver Cromwell (1948). *Caste, Class and Race: A Study in Social Dynamics.* Monthly Review Press.

Dawson, Michael (2018). Racial Capitalism and Democratic Crisis. *Insights from the Social Sciences*, *4*(1), 1.

Du Bois, William Edward Burghardt (1935). *Black Reconstruction in America: Toward a History of the Part Which Black Folk Played in the Attempt to Reconstruct Democracy in America, 1860-1880.* Routledge.

Dutilh-Novaes, Catarina (2020). Carnapian Explication and Ameliorative Analysis: A Systematic Comparison. *Synthese*, *197*(3), 1011–1034.

Earle, Rebecca (2003). Luxury, Clothing and Race in Colonial Spanish America. In Maxine Berg and Elizabeth Eger (Eds.), *Luxury in the Eighteenth Century: Debates, Desires and Delectable Goods* (219–227). Palgrave Macmillan UK.

Elster, Jon (1982). Marxism, Functionalism, and Game Theory: The Case for Methodological Individualism. *Theory and Society*, *11*(4), 453–482.

Go, Julian (2021). Three Tensions in the Theory of Racial Capitalism. *Sociological Theory*, *39*(1), 38–47.

Goodin, Robert E. and Kai Spiekermann (2015). Epistemic Solidarity as a Political Strategy. *Episteme*, *12*(4), 439–457.

Haslanger, Sally (2005). What Are We Talking About? The Semantics and Politics of Social Kinds. *Hypatia*, *20*(4), 10–26.

Hoffmann, Robert (2006). The Cognitive Origins of Social Stratification. *Computational Economics*, *28*(3), 233–249.

Hudson, Peter James (2018). Racial Capitalism and The Dark Proletariat. *Boston Review*, *20*(1), 1.

Izquierdo, Luis R., Segismundo S. Izquierdo, and William H. Sandholm (2019). An Introduction to ABED: Agent-based Simulation of Evolutionary Game Dynamics. *Games and Economic Behavior*, *118*(C), 434–462.

Jackson, William A. (2002). Functional Explanation in Economics: A Qualified Defence. *Journal of Economic Methodology*, *9*(2), 169–189.

Kelley, Robin D. G. (2017). What did Cedric Robinson Mean by Racial Capitalism? *Boston Review*, *12*(1), 1.

Kincaid, Harold (2007). Functional Explanation and Evolutionary Social Science. In Stephen P. Turner and Mark W. Risjord (Eds.), *Philosophy of Anthropology and Sociology: A Volume in the Handbook of the Philosophy of Science*, Chapter 5, 213–247. North-Holland.

Kinney, David and Liam Kofi Bright (2023). Risk Aversion and Elite-Group Ignorance. *Philosophy and Phenomenological Research*, *106*(1), 35–57.

LaCroix, Travis and Cailin O'Connor (2021). Power by Association. *Ergo*, *8*(1), 163–189.

Lewens, Tim (2015). *Cultural Evolution: Conceptual Challenges.* Oxford University Press.

Manser, Marilyn and Murray Brown (1980). Marriage and Household Decision-Making: A Bargaining Analysis. *International Economic Review*, *21*(1), 31–44.

McElroy, Marjorie B. and Mary Jean Horney (1981). Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand. *International Economic Review*, *22*(2), 333–349.

Mills, Charles W. (2000). 'But What Are You Really?' The Metaphysics of Race. In Andrew Light and Mechthild Nagel (Eds.), *Race, Class, and Community Identity: Radical Philosophy Today* (23–51). Humanity Books.

Mills, Charles W. (2014). *The Racial Contract.* Cornell University Press.

Nash, John (1951). Non-cooperative Games. *Annals of Mathematics*, *54*(2), 286–295.

Nash, John F. (1950). The Bargaining Problem. *Econometrica: Journal of the Econometric Society*, 155–162.

O'Connor, Cailin (2019). *The Origins of Unfairness: Social Categories and Cultural Evolution.* Oxford University Press.

O'Connor, Cailin, Liam Kofi Bright, and Justin P. Bruner (2019). The Emergence of Intersectional Disadvantage. *Social Epistemology*, *33*(1), 23–41.
Pastore, Chaela (2002). Consumer Choices and Colonial Identity in Saint-Domingue. *French Colonial History*, *2*(1), 77–92.
Pettit, Philip (1998). Functional Explanation and Virtual Selection. *The British Journal for the Philosophy of Science*, *47*(2), 291–302.
Popa, Mihaela, Roland Muhlenbernd, and Jeremy L. Wyatt (2021). Fairness and Signaling in Bargaining Games. unpublished.
Poza, David J., Félix A. Villafáñez, Javier Pajares, Adolfo López-Paredes, and Cesáreo Hernández (2011). New Insights on the Emergence of Classes Model. *Discrete Dynamics in Nature and Society*, *2011*(1), 915279.
Robinson, Cedric J. (2020). *Black Marxism, Revised and Updated Third Edition: The Making of the Black Radical Tradition.* UNC Press Books.
Rubin, Hannah and Cailin O'Connor (2018). Discrimination and Collaboration in Science. *Philosophy of Science*, *85*(3), 380–402.
Schachter, Ariela, René D. Flores, and Neda Maghbouleh (2021). Ancestry, Color, or Culture? How Whites Racially Classify Others in the US. *American Journal of Sociology*, *126*(6), 1220–1263.
Schlag, Karl H. (1998). Why Imitate, and If So, How?: A Boundedly Rational Approach to Multi-armed Bandits. *Journal of Economic Theory*, *78*(1), 130–156.
Skyrms, Brian (1994). Sex and Justice. *The Journal of Philosophy*, *91*(6), 305–320.
Skyrms, Brian (2014). *Evolution of the Social Contract.* Cambridge University Press.
Stewart, Quincy Thomas (2010). Big Bad Racists, Subtle Prejudice and Minority Victims: An Agent-based Analysis of the Dynamics of Racial Inequality.
Sugden, Robert (1986). *The Economics of Cooperation, Rights and Welfare.* Blakwell.
Táíwò, Olúfémi O. (2018). The Empire Has No Clothes. *Disputatio*, *10*(51), 305–330.
Táíwò, Olúfémi O. (2022). *Reconsidering Reparations.* Oxford University Press.
Táíwò, Olúfémi O., Anne E. Fehrenbacher, and Alexis Cooke (2021). Material Insecurity, Racial Capitalism, and Public Health. *Hastings Center Report*, *51*(6), 17–22.
Taylor, Keeanga-Yamahtta (2016). *From# BlackLivesMatter to Black Liberation.* Haymarket Books.
Van Riel, Raphael (2020). In Search of the Missing Mechanism. Functional Explanation in Social Science. In Rebekka Hufendiek, Daniel James and Raphael van Riel (Eds.), *Social Functions in Philosophy*, Chapter 4, 70–92. Routledge.
Wallerstein, Immanuel (2000). Oliver C. Cox as World-Systems Analyst. *Research in Race and Ethnic Relations*, *11*(1), 173–183.
Wilson, Valerie and William M. Rogers III (2016). Black-White Wage Gaps Expand with Rising Wage Inequality. *Economic Policy Institute*, *12*(1), 1.
Young, H. Peyton (1993). An Evolutionary Model of Bargaining. *Journal of Economic Theory*, *59*(1), 145–168.