

The Emergence of Intersectional Disadvantage

Cailin O'Connor, Liam K. Bright, Justin P. Bruner

Intersectionality theory explores the peculiar disadvantages that arise as the result of occupying multiple disadvantaged demographic categories. Addressing intersectionality theory through quantitative methods has proven difficult. Concerns have been raised about the sample size one would need in order to responsibly tease out evidence for the claims of intersectionality theorists. What is more, theorists have expressed concern about our ability to formulate novel intersectional hypotheses in a non-ad-hoc manner. We argue that simulation methods can help address these, and other, methodological problems, because they can generate novel hypotheses about causal dependencies in a relatively cheap way, and can thus guide future empirical work. We illustrate this point using models which show that intersectional oppression can arise under conditions where social groups are disadvantaged in the emergence of bargaining norms. As we show, intersectional disadvantage can arise even when actors from all social categories are completely identical in terms of preferences and abilities. And when actors behave in ways that reflect stronger intersectional identities, the potential for disadvantage increases. As we note, this exploration illustrates the usefulness of idealized models to real world inquiry.

1 Introduction

Intersectionality theory is a branch of social theorizing focused on the effects of occupying multiple demographic categories.¹ It arose from the work of black feminists who highlight ways in which theories of oppression (and policies based upon them) can fail to take into account the peculiar circumstances of black women in America.² This focus on aspects of social life pertaining to oppression, disadvantage, and power imbalances remains central to intersectionality theory. As Collins and Chepp put it in their textbook entry: “[t]he first core idea of intersectional knowledge projects stresses that systems of power [...]

¹This is a longstanding intellectual tradition, with basic ideas going back to at least the 19th century [Cooper, 1892/1988]. See [Gines, 2011] for a summary history.

²For instance, take the case highlighted by Crenshaw, in a foundational essay [Crenshaw, 1989]. Black women employees of General Motors found that they were not legally protected from ‘last hired, first fired’ forms of discrimination. For, when plaintiffs took a grievance to court, the court ruled that GM’s hiring and retaining white women meant that they could not be charged with violating laws prohibiting gender based discrimination, and likewise would not treat black men and black women as separate classes.

cannot be understood in isolation from one another; instead, systems of power intersect and coproduce one another to result in unequal material realities and the distinctive social experiences that characterise them” [Collins and Chepp, 2013, 60].

There are several branches of intersectionality theory. One project focuses on the, “phenomenological experience of a specific group with intersecting identities” [Parent et al., 1995, 642]. Another branch of study has focused on metaphysical claims or “the ontological claim that intersectionality can theorize the convergence, co-constitution, imbrication, or interwovenness of systems of oppression” [Carastathis, 2014, 307]. Both lines of research tend to place great emphasis on displaying the full complexity of social life [McCall, 2005]³

The project we will discuss here, on the other hand, uses highly simplified, idealized models of human behavior to investigate the emergence of specially intersectional disadvantage. This might seem to be in tension with the aforementioned approaches to intersectionality. There is, though, a further tradition of experimental or large scale sociological research explicitly attuned to intersectionality theory. It is this tradition that our work here is intended to most closely interact with and contribute to. In particular, we will argue that such simplified models can help address a number of methodological issues that have arisen in this field.

We will focus on four such methodological issues. First, sociologists have worried that, “to address intersectional questions, it is necessary to develop complex designs involving prohibitively large samples” [Cole, 2009, 170].⁴ The study populations for women, black people, and the disabled, for instance, are much larger than the study population of black disabled women, which poses pragmatic research challenges. Second, theorists have complained that this field of study lacks a well defined method, meaning that it is often hard to know what explorations count as intersectional [Nash, 2008]. For instance, when Patricia Hill Collins and her students were called to pick out intersectional work for a conference they found that “we thought we “knew” intersectionality when we saw it but couldn’t quite define what it was” [Collins, 2015, 2].

The third critique goes farther, citing methodological difficulties that can lead to theoretical vacuity. The idea is that in defining intersectional populations, theorists sometimes appeal to behaviors or features that are themselves to be explained. For instance, Tommy Curry argues that when discussing the social life of black men, intersectional theorists’ explanation of alleged “violence is rooted in a theoretically justified circularity” [Curry, 2017, 200]. The charge is that, for lack of good means of investigating their claims, these theorists generate pseudo-explanations by baking common essentialist assumptions into their definitions of identify categories. Thus the apparently empirical explanations or accounts of events that intersectionality theorists offer end up being mere analytic entailments.⁵

³Notice that this reflects Longino’s account of what feminist sciences tend to value [Longino, 1995].

⁴In the quoted passage Cole is reporting on this concern without endorsing it herself.

⁵For a helpful elaboration and contextualisation of Curry’s charges against intersectionality theory in this vein, see [Táiwò, 2018, §3]. The influential critique contained in [Kwan, 1996] also involves the charge that intersectionality theory too often works by assuming empirical details and complexities that really ought be empirically evinced and explained.

Last, relatedly, one may worry that the characteristic claim of intersectionality theory—that some groups suffer what is called “non-additive” disadvantage—lacks empirical content (at least when applied outside the initial stock of examples drawing from experiences of African American women). The specter of this worry is raised in [Mutua, 2013]. Mutua argues that the most obvious extension of this claim was simply false when applied to black men [Mutua, 2013, 345], and, as she describes them [e.g. on 348], subsequent attempts to develop new hypotheses extending the theory seem ad hoc.⁶

Hence, while intersectionality theory has been (and continues to be) influential in both the study of social life and activist responses to oppression in America, it has had a number of serious charges laid against it in the literature. It has been alleged that it is too difficult to feasibly study, that it is unclear what its claims amount to, that it only gains apparent empirical content by rendering its claims analytic or trivial, and that in investigating non-additive disadvantage it only avoids false claims by being somewhat ad-hoc in what hypotheses it generates.⁷

Our work is a contribution to this methodological literature. We present models which show that intersectional oppression can arise under conditions where social groups are disadvantaged in the cultural emergence of bargaining norms, either by dint of small size or by disempowerment. As we demonstrate, this intersectional disadvantage can arise even when actors from all social categories are completely identical in terms of preferences and abilities. And when actors behave in ways that reflect stronger intersectional identities, the potential for disadvantage increases.

In deriving these results we illustrate a relatively under-used method for exploring intersectionality which is especially well suited given the critiques listed above. First, simulation methods avoid problems associated with sample size. Second, as we will illustrate, this approach generates theoretically well-grounded hypotheses about a source of characteristically intersectional oppression. In doing so it addresses a worry common to Collins, Curry, and Mutua, by providing a source of truly intersectional hypotheses that may then be precisified and subject to test in empirical study.

Third, by generating these hypotheses in a formal framework that explicitly makes assumptions about the homogeneity of the groups modeled, the modeling framework avoids Curry’s critique that intersectionality theory only generates apparent empirical claims by baking substantive empirical assumptions into its description of demographic categories. By explicitly assuming away culturally or historically laden descriptions of

⁶These considerations prompted Mutua (and other theorists she discusses in her essay) to develop an alternative theory of social identities and their interaction. Whether or not one takes up Mutua’s preferred theory, it illustrates the difficulties theorists have in bringing to bear intersectionality theory on new areas of social life, and thereby fruitfully guiding our responses to oppression. For an illustrative discussion of the epistemic problems that can arise when social researchers have too many degrees of freedom in deciding what their hypotheses shall be, see [Simmons et al., 2011].

⁷Methods have been advocated for experimentalists (e.g. Cole [2009]) or sociologists (e.g. Bright et al. [2016]) that would address some of the peculiar challenges faced by work on intersectionality theory. (In addition, Gines [2011] argues that intersectionality theory does not need a method peculiar to it.) However, these responses have focused on ways one can precisify and test an outstanding intersectional hypothesis. As we understand what is common to the criticisms by Collins, Curry, and Mutua is that it is often not clear what sort of hypotheses intersectionality theorists should be offering in the first place.

the demographic groups in question we have tied ourselves to the mast, rendered ourselves unable to subtly presuppose they will behave as intersectionality theory predicts. And yet, we shall argue, we have done this in a way that is none the less useful for identifying ways unequal social power can lead to inequity.

Last, as we will elaborate, the models provide a framework that helps makes precise several notions of non-additive disadvantage, addressing some of Collins' and Mutua's concerns. In this vein, the methodology can supplement the richer, descriptive empirical studies typical of intersectionality theory. We shall return to this, but one of our primary hopes for this work is that it shall facilitate cross- or inter-disciplinary work.⁸

One final contribution these models can make to intersectionality theory, as we will elaborate, is in showing how little is needed to generate specifically intersectional disadvantage, which may have implications for activist or political interventions on inequitable societies. This seems especially important in the context of intersectionality theory, where it is expected that theorists retain an ongoing connection with the work of activists [Collins and Bilge, 2016, 198].

We take this work to be relevant to philosophy of science as a methodological discussion of an important area of social science. However, we also observe that this investigation makes a meta-point for philosophy of modeling. Philosophers of science have worried that simplified models may be unable to tell us much about the world.⁹ We think this project neatly demonstrates how highly simplified models can play multiple, useful epistemic roles even in one project, especially in concert with empirical investigation.¹⁰

Our paper will proceed as follows. In section 2, we introduce and justify the modeling framework which we use in the rest of the paper. In section 3 we show how actors in multiple minority groups can be specially disadvantaged simply by dint of their double-minority status. In section 4, we show how power differences between groups can lead to intersectional disadvantage, and how both power and minority status can interact in the emergence of bargaining. In 5 we elaborate how this framework allows us to say more about what one might mean by 'non-additive disadvantage'. And in 6, we conclude, returning to our main points about methodology.

⁸For a discussion of the kind of modeling/empirical cross disciplinary work we have in mind, that also vitally includes negotiation and cooperation with those whose lives are directly affected by the research topic at hand, see [Poteete et al., 2010].

⁹For example, Fumagalli [2015, 2016] argues that 'minimal models', which he takes to be those that "lack any similarity, isomorphism or resemblance relation to the world" [Grüne-Yanoff, 2009, 83], cannot prompt learning about the world.

¹⁰Jhun et al. [forthcoming] argue that we should think of models (even a single model applied to a single target) as tools that can play multiple roles in arguments as does [O'Connor, 2017c]. Previous authors have argued for broad pluralism with respect to how models explain [Downes, 1992, O'Connor and Weatherall, 2016], but few philosophers have focused on the pluralism of modeling explanations within single projects.

2 Power, Minority Status, and The Evolution of Bargaining

When it comes to the cultural emergence of bargaining norms, recent work suggests that various features of social groups can impact the chances that such groups are advantaged or disadvantaged. In particular, it has been shown that minority status can lead to disadvantage under a number of conditions [Bruner, 2017, O’Connor and Bruner, 2017, O’Connor, 2017a], and also that various forms of economic, or political power can likewise improve bargaining outcomes for groups [Gallo, 2014, Bruner and O’Connor, 2017, LaCroix and O’Connor, 2018]. Here, we will use these results as a framework for exploring intersectional effects. In particular, we look at how intersectional identities can impact the emergence of such discriminatory bargaining norms. In order to do this, we must first introduce this framework. We begin by introducing models of the cultural evolution of bargaining between those in different social groups. We then describe the cultural Red King Effect, by which minority groups can be disadvantaged in the emergence of bargaining conventions, and move on to show how one form of economic power can likewise impact such conventions.

2.1 The Evolution of Bargaining Between Groups

Evolutionary game theoretic models, in general, represent the evolution or emergence of strategic behavior in populations of actors, either biological or social. In particular, they have been of great use to philosophers in understanding the emergence of norms and conventions in human groups.¹¹

These models start with a *game*—a representation of strategic interaction that specifies *players* (who interacts), *strategies* (what strategic actions they can take), and *payoffs* (their outcomes from various sets of strategies). In the *Nash demand game*, two players each demand some portion of a resource [Nash, 1950]. If their demands are compatible, they each receive what they asked for. If their demands are incompatible in that they exceed the resource, actors are unable to agree on a division and receive a poor payoff often called a *disagreement point*.

Figure 1 shows what is called a payoff table for a mini-version of the Nash demand game.¹² Rows represent possible strategies for player 1 and columns for player 2. Entries to the table list payoffs for each player for each combination of strategies with player 1 listed first. The available strategies are to make a Low, Med, or High demand. For simplicity sake, we assume that the total resource is 10, that $L + H = 10$, and that $L \leq 5 \leq H$. This means that the demands always include 5, plus two compatible demands that perfectly divide the resource like 1 and 9, or 4.5 and 6.5. The disagreement point, which actors get for combinations of incompatible demands like Med v. High and High v. High is chosen to be 0.

¹¹For examples see Skyrms [1994, 2004], Alexander and Skyrms [1999].

¹²The mini-game approach to bargaining, which looks at a small but representative subset of possible strategies for a more complex game, is commonly employed in models of the evolution of bargaining for tractability reasons. (See, for example, Young [1993], Skyrms [1994].)

		Player 2		
		Low	Med	High
Player 1	Low	L,L	L,5	L,H
	Med	5,L	5,5	0,0
	High	H,L	0,0	0,0

Figure 1: Payoff table for a general Nash demand mini-game.

One of the most important solution concepts in game theory, which will also be relevant to our social evolutionary approach, is that of a *Nash equilibrium*. Nash equilibria are sets of strategies in a game where no actor can unilaterally switch and improve their payoff. Because of this they are often stable in the sense that actors playing them will have no incentive to change. There are three Nash equilibria of this mini-game—Low v. High, Med v. Med, or High v. Low.¹³ At these pairings, if an actor demands more, they get the disagreement point, if they demand less, they get less.

To be clear, this model of interaction, though highly simplified, captures the features we will need to represent discriminatory bargaining norms. Individuals divide resources, and are incentivized to come to a bargaining agreement. And this division can either be equal, or can advantage one of the two participants.

Besides the game, the other feature of an evolutionary game theoretic model is the *dynamics*. These are rules that specify how the behaviors of some group of actors engaging in a strategic scenario will change over time. The most commonly used dynamics in evolutionary game theory are the *replicator dynamics* which suppose that successful strategies become more prevalent in a population while those that do poorly die out. These dynamics have been shown to bear deep formal relationships to explicit models of individual learning [Börger and Sarin, 1997, Hopkins, 2002] and learning by cultural imitation [Björnerstedt et al., 1994, Weibull, 1997, Schlag, 1998] so we will employ them throughout the paper as a model of cultural evolution in populations developing conventions of bargaining. We can imagine it representing a situation, for example, where individuals repeatedly bargaining and periodically copy bargaining strategies that seem to be working well for others. Over time, given this type of imitation, we see group-level patterns of behavior emerge.¹⁴

There is one last element to the framework we employ, and this is that we focus

¹³To be completely clear, there are three pure strategy Nash equilibria. We will not address mixed strategy Nash equilibria because they are not relevant to our evolutionary results.

¹⁴The version of the replicator dynamics we employ is similar to the discrete time two population replicator dynamics, but where individuals interact with members of both populations and where one population may be smaller than the other. Let x and y represent the two populations so that $x + y = 1$ and $x \leq y$. Strategies for population x update according to $x'_i = x_i \left(\frac{f_i(x,y)}{\sum_{j=1}^n f_j(x,y)x_j} \right)$ where x_i is the proportion of the x population playing strategy i , $f_i(x,y)$ is the fitness of type i in x given the population states of x and y , and $\sum_{j=1}^n f_j(x,y)x_j$ is the average population fitness for x given the states of x and y . Strategies for population y update according to the analogous dynamics.

on populations divided into social groups—women and men, or white people and black people, for example. We assume that everyone in the population recognizes these social categories as the relevant ones for choosing strategies and for picking imitative role models. Either type may be in the minority, or else the two types could be equally prevalent.

If we evolve such models we find that separate conventions emerge for in-groups and out-groups. In particular, when members of either type interact with their in-group they come to do one of two things. Most commonly, they all eventually learn to demand Med, or to treat each other fairly. Sometimes they develop a fractious pattern of behavior where some actors demand High and some Low. Of more interest for our purposes are the conventions that emerge between the two types. These reflect the three Nash equilibria described above—either one type demands High and the other Low, or else they make Med demands of each other. Axtell et al. [2000], looking at similar models, describe the two sorts of outcomes where one type demands High of the other as discriminatory in the sense that actors treat out-group members differently than in-group members to the detriment of the out-group. This is also how we will capture discrimination in this project.

2.2 The Cultural Red King Effect

The Red King effect occurs in biology when two mutualistic species co-evolve, but at different rates. Bergstrom and Lachmann [2003], using an evolutionary game theoretic framework, show that under some such conditions the slow evolving species is more likely to end up getting more of the products of the mutualism.

Bruner [2017] replicates this effect in a cultural context when a minority group interacts with a majority one. As noted, there are three possible outcomes for the evolutionary models we employ: one group gets more, the other does, or they make fair demands. The *cultural* Red King effect occurs when a type that is in the minority becomes more likely to end up being discriminated against. Under the replicator dynamics, this occurs when that discriminatory outcome has a larger set of starting population states that evolve to it, and thus becomes more likely to emerge.

Why might we see this effect? While minority types meet majority types commonly, as a result of their prevalence in the population, the reverse is not true. This means that the majority is very significant to the minority from a payoff standpoint, leading to a situation where minority members are more reactive to the majority group and learn more quickly how to interact with them. In cases of bargaining and division of resources, the minority can end up at a disadvantage as a result [Bruner, 2017, O’Connor and Bruner, 2017, O’Connor, 2017a].¹⁵

Figure 2 shows this effect. Results are proportions of 10k runs of simulations that end up at each of the three possible outcomes between types.¹⁶ Let $L = 4$ (so that the

¹⁵Rubin and O’Connor [forthcoming] also find a similar effect in actors of two types bargaining on a network.

¹⁶We use the discrete time replicator dynamics to generate these simulation results. For each trial the proportions of strategies in each type are randomly chosen and we evolve populations until they are

three demands are 4, 5, and 6). Let p_1 be the proportion of the majority type in the population, and p_2 the minority, so that $p_1 + p_2 = 1$ and $p_2 \leq p_1$. As is evident from the figure, when $p_1 = .5$, meaning that the two types are equally prevalent, the most common outcome is the one where the two types make fair demands of each other. The two discriminatory outcomes are equally likely to emerge. As p_1 increases, three things happen. First, fair outcomes become increasingly unlikely. Second, the majority group becomes increasingly likely to end up demanding High. And third, the minority group becomes increasingly unlikely to demand High.

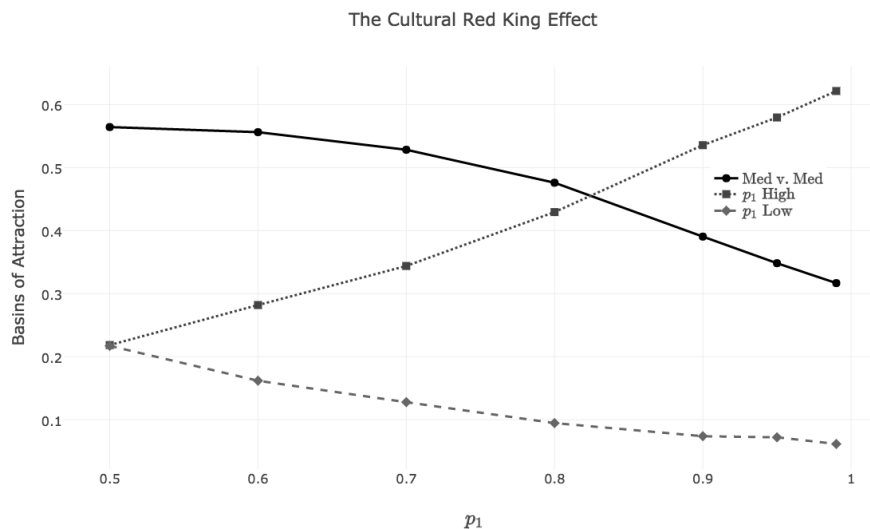


Figure 2: Basins of attraction for groups playing a Nash demand game as p_1 varies.

In other words, simply by dint of small numbers alone, members of a minority group can be disadvantaged in the emergence of bargaining. The cultural Red King effect is documented in much greater detail by Bruner [2017], O’Connor [2017a]. As these authors point out, in some models a cultural Red Queen effect can also occur, where the majority instead gains an advantage. This happens if, because of the details of their strategic situation, minority members quickly learn to make High demands and, as a result, gain an advantage from their reactivity. O’Connor [2017a], though, argues that if actors are risk averse or show in-group preference, common assumptions that hold of many real-world actors, the Red King is a much more powerful effect in the cultural realm.

2.3 Power

Nash [1953] used an axiomatic approach to argue that in rational choice scenarios of two player bargaining, a more powerful player can gain resources for themselves via a converging to one of the three equilibria.

credible threat. Bruner and O’Connor [2017] explore not how actors can wield power to their benefit in a rational choice scenario, but how power differences between social groups can translate into an advantage with respect to emerging bargaining norms.¹⁷

We will use a notion of power drawn from Nash’s original work and employed by Bruner and O’Connor [2017] as well—that of asymmetric disagreement points. Suppose that when members of social groups bargain, and when bargaining breaks down, the fall back position for one group tends to be better than for the other.¹⁸ Figure 3 shows a payoff table for a game integrating this assumption. This is the same game shown in figure 1, but where the disagreement point for player 1 is D , assuming that $D \geq 0$.

		Player 2		
		Low	Med	High
Player 1	Low	L,L	L,5	L,H
	Med	5,L	5,5	D,0
	High	H,L	D,0	D,0

Figure 3: A payoff table for a mini Nash demand game with different disagreement points.

Suppose two groups evolve bargaining norms and one receives a higher disagreement point, D . Although we will not show a figure outlining these results, as D increases the likelihood that the powerful group ends up at the equilibrium where they demand High also increases. The likelihood that they end up demanding Low or Med drops [Bruner and O’Connor, 2017].¹⁹ In other words, their powerful situation translates into advantage vis a vis the emergence of bargaining norms. This is because groups with higher disagreement points are less incentivized to learn lower demands, since these demands do not yield outcomes that are significantly better than the disagreement point.

We will now employ these two sets of results to look at intersectional effects in populations where norms and conventions of bargaining emerge. In particular we will ask: How do intersectional identities impact results in these models? And, are small intersectional groups especially disadvantaged by the cultural Red King or power disparities?

¹⁷These authors employ their results to explore how power differences in the academic hierarchy can translate to credit advantage in co-authored papers. Their results, though, are in fact quite general.

¹⁸In economic work on the family, this assumption has been employed in models of household bargaining. Suppose that men tend to have a generally higher level of wealth, and better access to gainful employment, than women. When men and women bargain over household labor, then, the idea is that should the bargain break down (leading to household dissolution), men will tend to end up better off [Manser and Brown, 1980, McElroy and Horney, 1981].

¹⁹The robustness of this result has been verified by O’Connor [2017b], LaCroix and O’Connor [2018], who show how the same effect occurs in agent-based models.

3 Minority Status and Intersectional Oppression

We start with the cultural Red King and move through three sets of models where the importance of intersectional identity as a determinant of both behavior and cultural adaptation increases at each step. First, we compare two models where the second incorporates intersectional identity in a more full-blooded way than the first. As we will show, this leads to more serious disadvantage for small intersectional types. Then we present a model with even stronger intersectional assumptions. Although the results cannot be directly compared to the first two models (as we will elaborate), we see strong disadvantage for the smallest intersectional population. Across all three models, we find that the smallest intersectional population is specially disadvantaged solely by dint of their size.

3.1 Minimal Intersectionality

We start with a model that incorporates intersectional identities, but in a minimal way. Suppose that there are two dimensions of identity along which a population is split. For simplicity sake, call the prevalence of the majority and minority types along one dimension p_1 and p_2 and the prevalence of the types along the other dimension q_1 and q_2 . The two dimensions of identity yield four resulting types. We assume that the two dimensions perfectly crosscut each other so that the proportions of the four types are $p_1 * q_1$, $p_1 * q_2$, $p_2 * q_1$, and $p_2 * q_2$. Figure 4 shows what these proportions might look like.

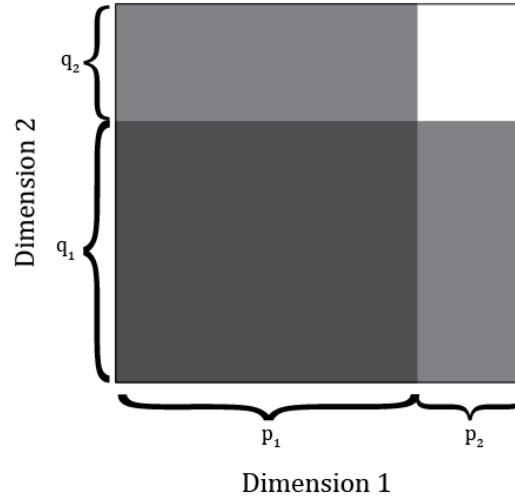


Figure 4: Diagram of proportional representation of intersectional types.

Because the models we will look at can get a bit complicated, we will use an example to help keep track of what is going on. The two relevant dimensions of social identity

for bargaining behavior in this example will be gender—men (p_1) or women (p_2)—and race—white (q_1) or black (q_2). We assume that men and white people will be the larger types. The four intersectional types are then white men ($p_1 * q_1$), black men, ($p_1 * q_2$), white women ($p_2 * q_1$), and black women, ($p_2 * q_2$). Of course, our models are highly simplified, and there is nothing in particular that captures the subtle (and not-so-subtle) differences between different social categories in them. The example is simply to help us keep the groups straight.

Suppose that this population regularly engages in two sorts of bargaining scenarios, for each of which only one of their identities becomes salient. For example, one arena of bargaining could occur in the workplace over salary, benefits, or workload and, in this arena, race could be particularly salient to the actors. Another arena of bargaining could occur in the marketplace over the cost of goods, and for this gender could be more salient.

In this model, we expect that two separate processes will occur where bargaining norms conditioned on race emerge in the workforce and norms conditioned on gender emerge in the marketplace. We do not expect these processes to impact each other, but in each arena we do expect the minority group to be disadvantaged in bargaining as a result of their size, i.e., for women to be more likely to get the short end of the stick in the market and black people in the office.

To keep things tractable, we will focus on an even smaller version of the Nash demand game, where actors may only demand High or Low, with payoffs otherwise identical to the game pictured in figure 1. There are four possible joint outcomes of the evolutionary processes in this model. First, both men and white people can demand High when interacting with women and black people respectively, who demand Low in response. Second, men and black people can demand High. Third, women and white people can demand High. And last, women and black people can demand High. Figure 5 shows the proportions with which these outcomes emerge in the model as the size of the majority types increases. We assume that the two majorities are the same size, or that $p_1 = q_1$.²⁰

As is evident from the figure, as the sizes of the larger populations increase it becomes increasingly likely that men and white people demand High. It also becomes very unlikely that both women and black people demand high. And last, it becomes slightly less likely that the outcomes where women and white people or men and black people both demand High occur. These outcomes are a result of two cultural Red King effects, one for each of the evolving bargaining scenarios.

In this case it is most likely that members of the largest intersectional group end up making High demands in both arenas, and least likely that members of the smallest intersectional group do. This is arguably not truly an intersectional effect, though, as the outcomes are straightforwardly derived from two processes by which minority groups are disadvantaged. (In such a scenario, more robustly intersectional effects might occur if there are ways these processes combine to impact the lives of those involved. This

²⁰This assumption allows us to easily display our results. We investigated models where p_1 and q_1 varied independently and generally the greater either one was, the greater the disadvantage for the small types.

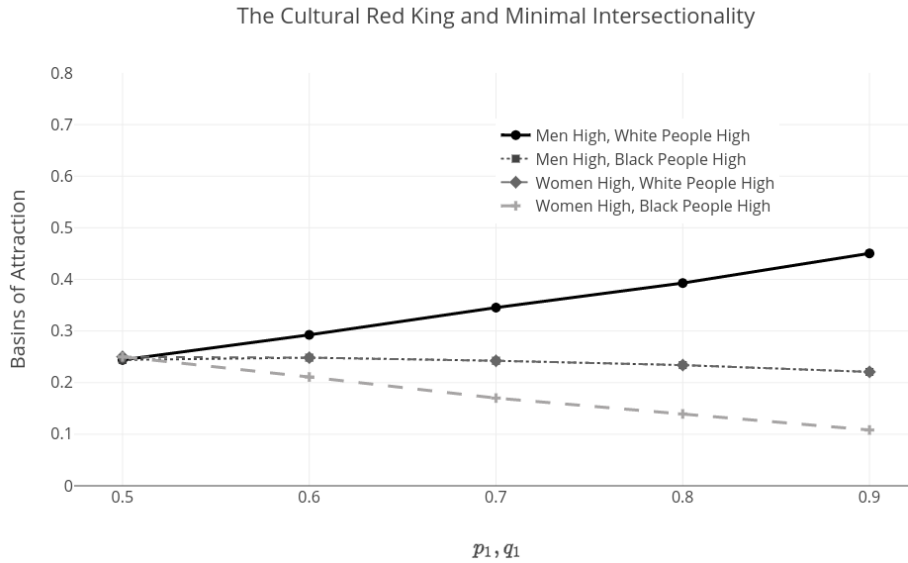


Figure 5: Proportions of four possible joint outcomes for minimally intersectional populations.

could be the case if, for example, workplace bargaining determines what actors have to bargain over in the marketplace.)

3.2 Moderate Intersectionality

Now we consider a model with a stronger sort of intersectionality. Imagine a scenario as above where there are still two arenas in which bargaining occurs, and that each sort of identity is salient in one arena. Now assume, though, that actors only learn socially from those in their intersectional type. At the market, for example, white men do not assume that any man is a good role model for them (even though gender is salient for interaction), but only adopt role models who also share their race.

This small assumption greatly alters the cultural evolutionary process in these models. In particular, it means that the difference in learning speed between the largest and smallest types is much more significant in both arenas of interaction. This results in a much stronger Red King effect as is evident in figure 6. In the previous model, at the most extreme values of p_1 and q_1 , the largest population ended up always demanding High about 45% and the smallest population 11% of the time. In this model these values are now 76% and 1.5%.

In this case, the stronger intersectional assumptions in the model translate directly into a more significant disadvantage for the smallest intersectional group, and a stronger advantage for the largest group. In the next section, we look at a model with even stronger intersectional assumptions.

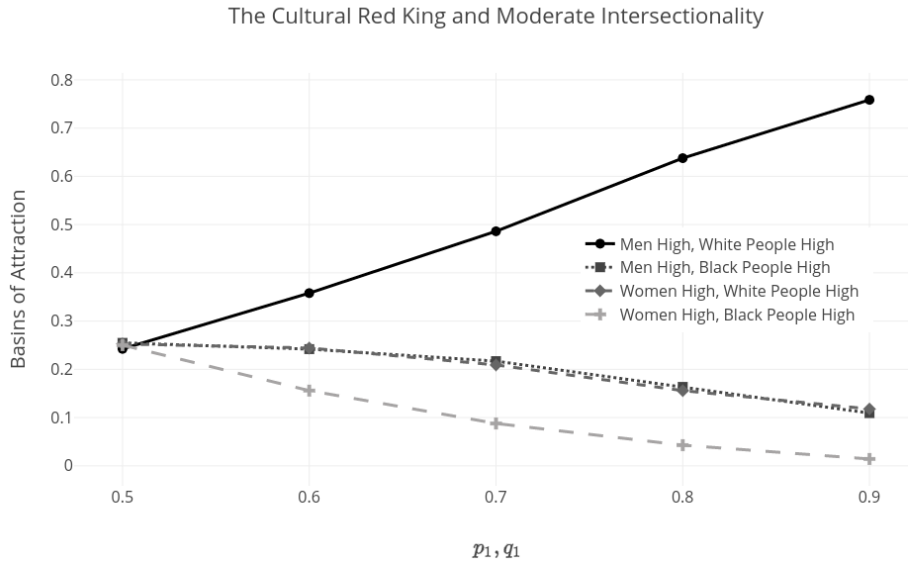


Figure 6: Proportions of four possible joint outcomes for moderately intersectional populations.

3.3 Strong Intersectionality

We now assume that there is just one arena of interaction for our intersectional groups, and that within this arena only intersectional identities are salient for interaction. In other words, the individuals pay attention to intersectional identities in determining both 1) how to interact with bargaining partners and 2) which role models to choose.

The outcomes of this model are quite different from those presented in the last section. In- and out-group conventions now develop between *each* pair of the four possible intersectional types. This means that for each outcome of the model there are six types of out-group conventions, each of which may go one of two ways. Rather than report the proportions of runs that end up at each of these outcomes, we instead calculate the proportion of time each type spends demanding High of out-group members averaged over simulations.

Figure 7 shows results for this model. As in the last two models, we see that as the proportions of types grows more disparate, the disadvantage for the smallest type becomes more significant, while the advantage for the largest type increases. In this model, again, the effect is fairly dramatic. When $p_1 = q_1 = .9$ white men demand High 90% of the time, compared to 9% of the time for black women. This occurs because in a single arena, again, the size difference between the largest and smallest types is significant, leading to a strong cultural Red King.²¹

²¹Note that there is something unrealistic about this model, which is that the conventions developed need not be particularly consistent. For example, it could be the case that black women discriminate

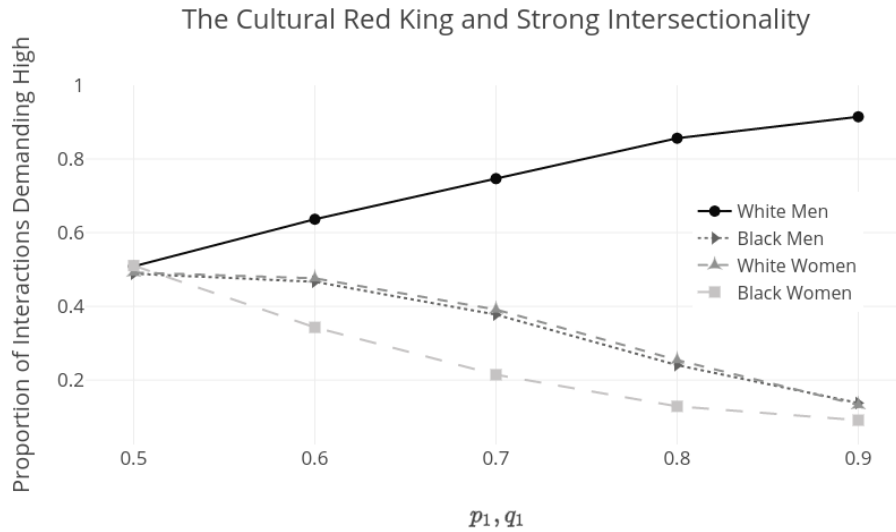


Figure 7: Average proportion of high demands for strongly intersectional populations.

4 Power and Intersectionality

The models just presented tell us that under conditions where actors are members of two sorts of minority groups, and where actors take their intersectional identities seriously, members of small intersectional groups are at special risk of disadvantage as a result of dynamical effects.

In this section of the paper, we will now look at situations where there are power differences between groups. The goal here is twofold. First, we provide a robustness check on the sorts of intersectional effects we have been describing—similar patterns emerge even under different modeling choices. And second, we give examples of models that are perhaps a bit more realistic.

4.1 Minority at Work, Powerless at Home

Let's first consider another two arena model like those discussed in sections 3.1 and 3.2. Suppose now that the two arenas of interaction are the workplace, where race is more salient, and the home, where gender is more salient.

As before, let's assume that in the first arena of interaction—the workplace—black people tend to be in the minority so that the prevalence of white people, p_1 , ranges from .5 to .99. And in the home, let's assume that men tend to have greater economic privilege, meaning their disagreement point for household bargaining, D , ranges from

against white men, who discriminate against white women, who discriminate against black women. This sort of circularity does not occur in the two arena models.

0 to 3.9 while women’s disagreement point is always 0. (We’ll assume that $L = 4$ and $H = 6$ for this set of models, as in the previous ones.)

As in the previous set of models, we look at a case where actors display minimal intersectional identities when it comes to choosing role models (i.e., copy those of their race at work, and those of their gender in the home), and also a case where actors use role models only from their intersectional groups.

We find that the disadvantage generated by a cultural Red King effect at work, and by a power disadvantage at home can compound to create special intersectional disadvantage. Figure 8 shows results from the moderately intersectional models. We show a subsection of parameter values intended to capture the joint effects of power and minority status.²² Again we show basins of attraction for the four possible joint outcomes of the model as we increase the disadvantage for both black people with respect to proportional representation at work and women with respect to power at home. We use a bar graph here rather than a line graph because the x-axis represents two changing parameter values with no straightforward relationship to each other.

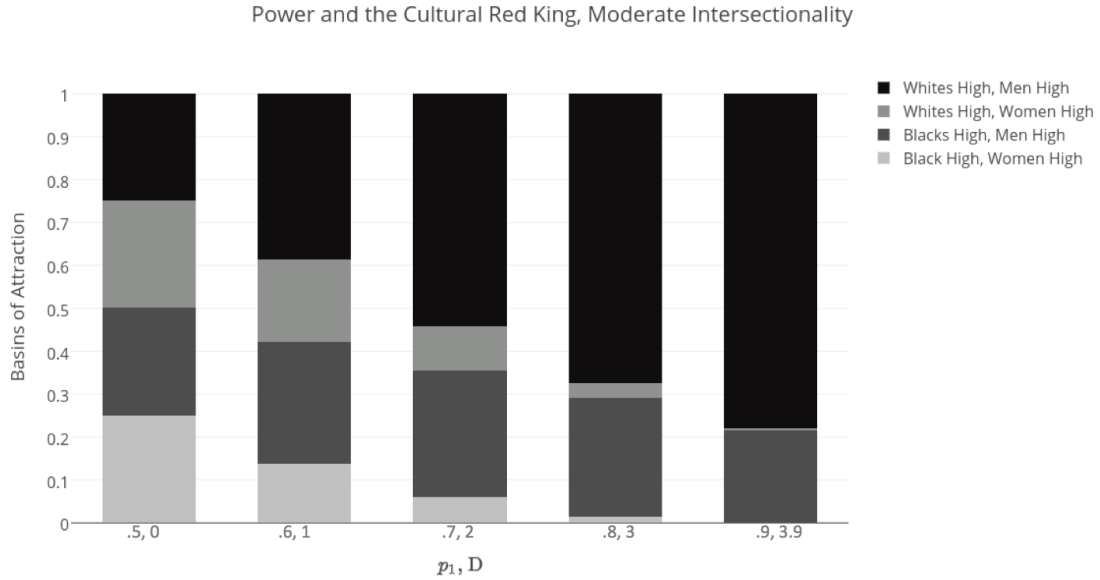


Figure 8: Basins of attraction for moderately intersectional populations with power and minority status.

As the figure makes clear, these variations lead to increasing disadvantage for women and black people, but especially for black women. At the most extreme values for disadvantage, they always demand Low in almost every simulation. White women are also strongly disadvantaged in these simulations because the effects of power are stronger than those of the cultural Red King.

²²We ran many other parameter values. In general, minority status and lower disagreement point led to increased disadvantage.

As in the models solely relying on the cultural Red King, we also see a stronger intersectional disadvantage here when actors choose role models only from their intersectional types. For example, in figure 8, at the most extreme values white men always demand High in almost 80% of simulations. For the same parameter settings in the model with non-intersectional role models, white men always demand High in 63% of simulations. (We do not display these results for space purposes.)

4.2 Strong Intersectionality and Power

We also repeat the strong intersectionality model presented in section 3.3 but using power asymmetries, rather than size asymmetries, between groups. Remember that now we simply have four intersectional groups that interact and develop bargaining norms. Actors pay attention to intersectional types both in choosing their strategies, and in choosing role models to imitate.

We now assume that the four types are equally prevalent in the population. But, because of the background conditions in their culture, there are power differences between them. We assume that white men have the most power so that in general when bargaining breaks down they will have a good fall back position and that black women are especially disempowered. We also assume that white women and black men occupy a middle ground, and for simplicity sake always give them the same amount of power. This is operationalized by making the disagreement point for white men always 3, for black women always 0, and for white women and black men varying it between 0 and 2.5.²³

This might initially sound like a strange model. We start by assuming that black women are disadvantaged and then observe a disadvantage for them. But notice that we start with differences in power, and then look for subsequent disadvantage with respect to resource division.

Figure 9 shows results from this model. The x-axis tracks the disagreement point for black men and white women. On the y-axis we see the proportion of time that each group demands High when meeting out-group members, averaged over all runs of simulation at that parameter value. When D for black men and white women is 0, all the intersectional groups (besides white men) are equally disadvantaged. As the disagreement points for the mid-power groups increases white men become slightly less likely to demand High in interaction, because now they are slightly more likely to end up disadvantaged with respect to these groups. Black women, though, become very unlikely to ever demand High. When $D = 2.5$, they only do so about 1% of the time.

Notice here that the trends related to disadvantage for black people and women generally do not capture how black women are disadvantaged in particular. When D is higher, in this scenario, while black men and white women do better on average, black women do worse. Part of the observation here is that in cases where intersectional types are truly all that matter, disjoint aspects of identity may have little power to predict

²³We looked at many other values for these parameters. In general, the higher the disagreement point for a group, the better for them, and vice versa.

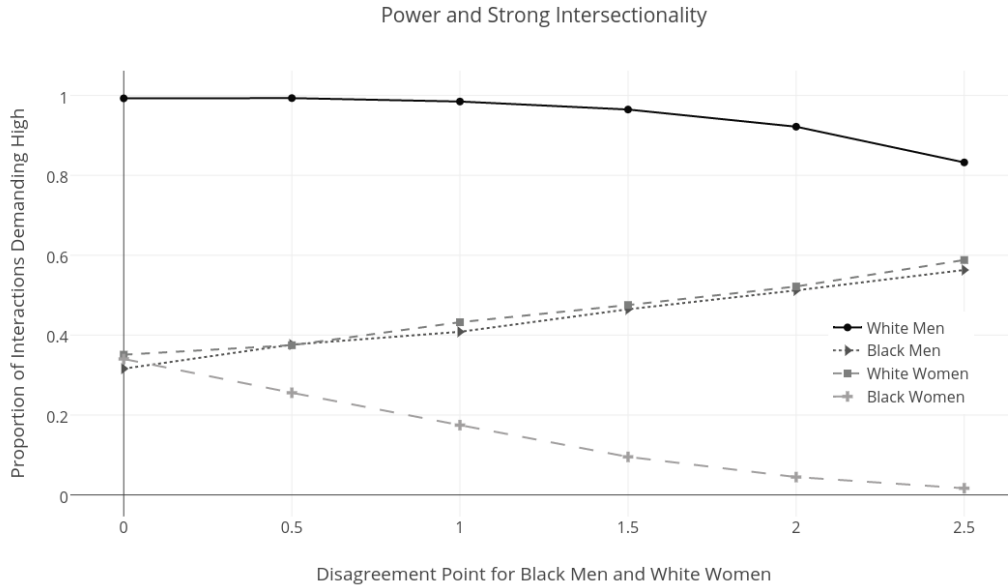


Figure 9: Proportion of interactions demanding High for four strongly intersectional groups.

disadvantage on their own.²⁴ Of course, in this model we freely vary the disagreement points of the four groups. If we assume that the disagreement point for black women is somehow a combination of those of black people and women, things will be different.

Before moving on to the discussion, we will pull out the sense in which these models display minimal conditions for the emergence of intersectional disadvantage, as described in the introduction. The generation of disadvantage here does not depend on representations of psychological features usually associated with discrimination. For example, the agents do not exhibit in-group preference and out-group bias, or implicit bias, or stereotype threat. We think these sorts of effects are important to full explanations of real world oppression, but our models tell us that we can get it even without them. The preconditions necessary for oppression in our model are 1) intersecting minority demographics (or intersecting groups with power differentials) 2) actors who learn to behave in their best interest and 3) a strategic interaction that involves resource division. These minimal conditions tell us that intersectional disadvantage is easy to generate, underpinning claims by intersectional theorists that systems of power and oppression cannot be understood without understanding intersectional effects.

²⁴Note that in the strong intersectionality model in the last section, this was not the case because assumptions in the model meant that when the small intersectional group got smaller, this was a result of decreasing size of the two single dimensional groups making it up.

5 Quantifying Non-Additive Disadvantage

To summarize the last two sections, we found that the cultural Red King effect can produce special disadvantage for intersectional types, and that this disadvantage is worse when stronger assumptions of intersectional identity are made in the models. We also found that power and the cultural Red King can combine to produce intersectional disadvantage, and that power differences alone can lead to such disadvantage.

In order to fully understand how these results bear on intersectionality theory, though, we need to look more carefully at the theory itself. As mentioned in the introduction, one of its core ideas is that in some sense or another oppression is not additive [Bright et al., 2016, 63]. Broadly speaking, this means that the combined effect of being oppressed along multiple axes is worse than one might expect if one were to naively sum the consequences of each form of oppression considered individually. As noted, Mutua’s concerns about intersectionality theory first arose specifically when attempting to apply claims about non-additive disadvantage to the situation of African American men.

We now investigate the sense in which the disadvantage we have observed in our models meets this criterion. As it turns out, though, this is not a straightforward task. To accomplish it, one first needs sufficient theoretical background to generate a null hypothesis—by which we mean a theory of what it would look like for the oppressions in question to have a ‘merely’ additive effect. However, theories of oppression are not often formulated in sufficient detail to allow for the clear prediction, or even statement of, such a null hypothesis. This is broadly in line with Collins’ concerns in [Collins, 2015]. It also relates to Curry and Mutua’s respective concerns about whether intersectionality theory has anything substantive to say about novel empirical scenarios that it was not explicitly formulated to deal with. To address her concerns Collins advocated greater clarity about the central notions of intersectionality theory. To gain this clarity, we think it is first necessary to precisify what a *non*-intersectional type of oppression, i.e. an additive one, would look like, and we can use the modeling framework presented here to do so.

We will present here two plausible precisifications of additive oppression to contrast non-additive effects with. The first criterion can be illustrated with a story. Suppose an employer encounters a job applicant who is black and a woman. Suppose further that they discriminate against black folk with probability P_b and women with P_w . If they are not being intersectional it may be as if they flip a coin (or are moved by what they had for breakfast that morning, or what the radio was discussing on the drive in to work, or whether or not the sun is presently in their eye, etc) to decide which is more salient to them, race or gender. They then focus on that aspect of the applicant in deciding whether to discriminate. In this case one would expect a level of discrimination somewhere in between P_b and P_w , because the combined chance of being discriminated against is a convex combination of the two chances of discrimination. With this understanding of non-intersectional oppression, we can say that there is non-additive intersectional oppression when the probability of being discriminated against given that you are a black woman is greater than both P_b and P_w .

A second, and quite different, null criterion is given by the following. In deciding

whether to discriminate, it is as if the employer draws from first P_b and then P_w . If either proscribes discrimination, the employer discriminates. So if $P_b = .2$ and $P_w = .2$, the employer would discriminate against black women with probability .36. This is a second, stronger but plausible ‘non-intersectional’ level of discrimination.²⁵ Notice that we could apply these two standards of non-additivity, for example, to measure special intersectional disadvantage for black women when it comes to firing practices. Is their rate of firings higher than that for both women and black people? Is it higher than the combined probability of firing for women and black people?

In the rest of this section we will use these two possibilities to discuss the results we have presented. For our models, we can calculate, given a set of parameter values, the probability that members of each group are discriminated against over all simulation outcomes. (This probability is slightly different from the ones just described in motivating our criteria because it is averaged over a set of mutually exclusive outcomes. It does, however, capture something about the general chances of disadvantage for each group.)

For all the models we have presented, except when group sizes and disagreement points are equal, the most disadvantaged intersectional group suffers from non-additive oppression in the first sense. I.e., the probability of being discriminated against if you are a black woman is always greater than the probabilities of discrimination for one’s two component identities. In general, as minority status becomes more dramatic, or as disagreement points for disadvantaged groups become lower, this difference in probability of discrimination for the intersectional group and the component identity groups also becomes more dramatic.

Some parameter settings of our models also show non-additive intersectional disadvantage in the second sense. Figure 10 shows various probabilities of discrimination for the moderate intersectionality Red King model. As we can see, for all values of majority size, $p_1 = q_1$, the probability of being discriminated against as a black woman, $P(D|Bl\&W)$, is greater than or equal to that for black people or women, $P(D|Bl)$ and $P(D|W)$, meaning it meets the first criterion as just described. When $p_1 = .9$, $P(D|Bl\&W) > P(D|Bl)\&P(D|W)$, or the probability of discrimination as a black woman is more than that expected if one met discrimination at the rate of both black people and women combined.

Likewise, for the moderate intersectionality power models, when $D = 3.9$ and $p_1 = .9$, the probability of discrimination for women is .3 and for black people is .55. The probability of discrimination for both of these identities combined in the second sense would be .69, but the actual experienced probability of discrimination by black women is .75, meaning there is again special intersectional disadvantage in both of the sorts of ways we identify.

The take-away from this discussion is two-fold. First, we find that the sorts of effects we have been outlining can generate non-additive intersectional disadvantage in either of the two ways we describe. On the less demanding criterion, all the models we consider

²⁵Note that these two criteria will not capture null predictions for all cases of intersectional oppression. Consider empirical results showing that black women earn a lower wage than either white women or black men. Because these are not probabilistic results, neither of our senses of additive discrimination apply.

Non-Additive Intersectional Disadvantage in Moderate Intersectionality, Minority Model

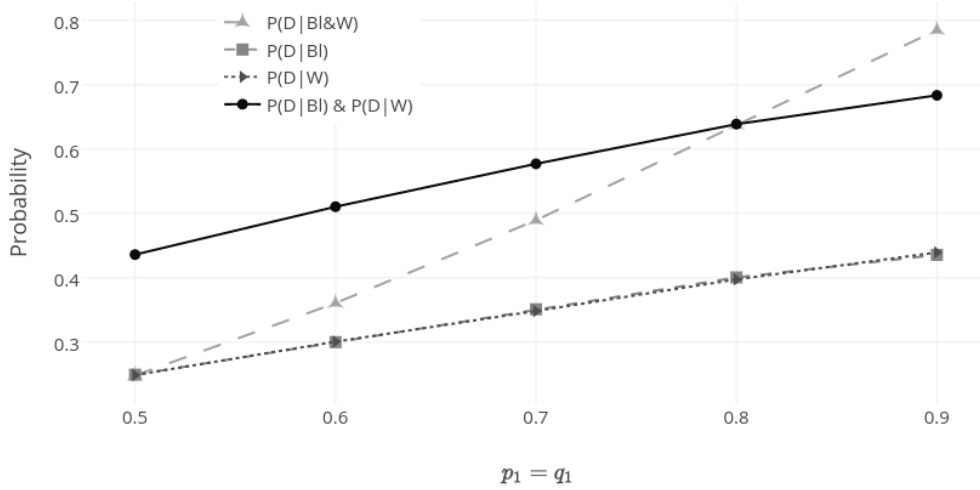


Figure 10: Probabilities of discrimination for different groups across parameter values.

generate such a disadvantage. Second, the modeling framework here allows us to generate very precise notions of what additive (and thus non-additive) disadvantage entail for some sorts of situations, which is responsive to extant critiques in intersectionality theory.

6 Conclusion

We would now like to return to our claims in the introduction that the simulations here can help address methodological critiques in intersectionality theory. First, it is easy to see how the models here avoid the pragmatic trouble of testing large enough sample populations. Because ‘populations’ in simulations are completely unconstrained in size there is no problem testing intersectional effects, even if one wanted to look at the intersection of many identities. (Of course, this is a general virtue of using modeling methods in intersectionality theory even if one is critical of our particular models for whatever reason.)

Second, as indicated, the models here provide well-ground empirical hypotheses for causal processes that can lead to genuinely intersectional disadvantage. In cases where members of small intersectional populations (or especially disempowered ones) divide resources with larger groups, and are at a bargaining disadvantage, theorists now have a new explanatory possibility. In general wherever we see resource division between intersectional social groups with some size/power asymmetries, our models predict that at least one sort of non-additive disadvantage should arise. If it should turn out that our models’ predictions are not borne out in such cases this may be problematic for

intersectionality theory more generally. The fact that our models point to a potential set of empirical results that would be troubling for intersectionality theory is actually a boon to the theory, seen against backdrop of the concerns of Collins, Curry, and Mutua.

When it comes to experimental tests, these models suggest studies, using methods from experimental economics, that have the potential to show these effects in groups of real humans. In particular, experimentalists can have groups of individuals in the lab engage in bargaining games to see whether especially small intersectional groups, or disempowered ones, tend to take home less. Along these lines, Mohseni and O'Connor [2018] have used previous work on the cultural Red King effect to direct a laboratory experiment which has produced evidence that the effect occurs among human subjects. Notice that the hypotheses generated here, as noted, help address worries from Curry about 'vicious circularity'. In these models there is no doubt that it is the fact of existing group identity, combined with minority status and power disadvantage, that causes economic inequity. In other words, we see clearly that real intersectional disadvantage can emerge from facts about identity, rather than acting as a suppressed definition of identity.

While some might worry that our models are simplified, and highly idealized, this fact helps them address Curry's critique. As we have emphasized our models do not display many of the characteristics that contribute to real-world inequity, such as psychological features, or society-wide features like segregated neighborhoods, police violence, and gendered pay inequalities. Removing these contextual details makes clear that we are not building in the intersectional effects, but rather showing how they can fall out of basic assumptions about group membership and strategic interaction.

This minimality also means the models can play another explanatory role, in showing how surprisingly little is needed to generate intersectional effects.²⁶ Our first set of models show how groups can end up disadvantaged merely by dint of being particularly small. The second set shows how one aspect of empowerment can compound to further economic disadvantage. In other words, very small asymmetries between groups can start processes of intersectional disadvantage. This sort of observation is particularly important in thinking about interventions on inequitable systems. For example, our models show that even if we managed to completely eradicate out-group bias in social groups, such as might be the idealised outcome of corporate seminars on implicit bias, we should still expect simple processes of learning to divide resources to nonetheless generate special disadvantage for intersectional groups.²⁷

The last thing we will highlight here is the conceptual contribution made in section 5 that we hope goes some way towards addressing Collins' concerns. Simply trying to assess modeling outcomes helped generate measures that may be useful to empirical studies of intersectionality as well. Which of the two senses of additive non-intersectional disadvantage one thinks is the more plausible contrast case (if either) will presumably depend

²⁶See O'Connor [2017c] for more about this particular explanatory role in simple models.

²⁷This seems to support a greater concentration on institutional arrangement and power relations therein when one is making interventions, suggesting a similar moral to one of those drawn in [Mohanty, 2013].

on contextually salient facts about the particular domain of potentially intersectional disadvantage one wishes to study. To this extent, as we noted in the introduction, our use of minimal models to study intersectional disadvantage would be best deployed in tandem with more historical or richly descriptive empirical studies of particular domains of application.

From a philosophy of modeling standpoint, this entire exploration demonstrates how highly simplified models can play many important epistemic roles in arguments about real world systems (even within one modeling project), contra some philosophical viewpoints. In general, the models here play a role somewhat reminiscent of that outlined by Morgan and Morrison [1999]—that models can mediate between theory and data. For example, simulations of particles in physics have been used to generate physical equations of motion. Here theoretical work on intersectionality, including critiques of empirical methods, inspired modeling work that can, in turn, help drive further empirical investigations, and well as theoretical ones.

Bibliography

- Jason Alexander and Brian Skyrms. Bargaining with neighbors: Is justice contagious? *The Journal of philosophy*, pages 588–598, 1999.
- Robert Axtell, Joshua M Epstein, and H Peyton Young. The emergence of classes in a multi-agent bargaining model, 2000.
- Carl T Bergstrom and Michael Lachmann. The red king effect: when the slowest runner wins the coevolutionary race. *Proceedings of the National Academy of Sciences*, 100(2):593–598, 2003.
- Jonas Björnerstedt, Jörgen W Weibull, et al. Nash equilibrium and evolution by imitation. Technical report, 1994.
- Tilman Börgers and Rajiv Sarin. Learning through reinforcement and replicator dynamics. *Journal of Economic Theory*, 77(1):1–14, 1997.
- Liam Kofi Bright, Daniel Malinsky, and Morgan Thompson. Causally interpreting intersectionality theory. *Philosophy of Science*, 83(1):60–81, 2016.
- Justin P Bruner. Minority (dis)advantage in population games. *Synthese*, 2017. DOI: 10.1007/s11229-017-1487-8.
- Justin P Bruner and Cailin O’Connor. Power, bargaining, and collaboration. In Conor Mayo-Wilson Thomas Boyer and Michael Weisberg, editors, *Scientific Collaboration and Collective Knowledge*. Oxford University Press, 2017.
- Anna Carastathis. The concept of intersectionality in feminist theory. *Philosophy Compass*, 9(5):304–314, 2014.

- Elizabeth R. Cole. Intersectionality and research in psychology. *American Psychologist*, 64(3):170–180, 2009.
- Patricia H. Collins and Valerie Chepp. Intersectionality. In Johanna Kantola Georgina Waylen, Karen Celis and S. Laurel Weldon, editors, *The Oxford Handbook of Gender and Politics*, chapter 2, pages 57–87. Oxford University Press, 2013.
- Patricia Hill Collins. Intersectionality’s definitional dilemmas. *Annual Review of Sociology*, 41(1):1–20, 2015.
- Patricia Hill Collins and Sirma Bilge. *Intersectionality*. John Wiley and Sons, 2016.
- Anna Julia Cooper. *A Voice From The South*. Oxford University Press, 1892/1988.
- Kimberlé Crenshaw. Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *The University of Chicago Legal Forum*, 139(1):139–167, 1989.
- Tommy J Curry. *The Man-Not: Race, class, genre, and the dilemmas of Black manhood*. Temple University Press, 2017.
- Stephen M Downes. The importance of models in theorizing: a deflationary semantic view. In *PSA: Proceedings of the biennial meeting of the philosophy of science association*, volume 1992, pages 142–153. Philosophy of Science Association, 1992.
- Roberto Fumagalli. No learning from minimal models. *Philosophy of Science*, 82(5):798–809, 2015.
- Roberto Fumagalli. Why we cannot learn from minimal models. *Erkenntnis*, 81(3):433–455, 2016.
- Edoardo Gallo. Communication networks in markets. Technical report, Faculty of Economics, University of Cambridge, 2014.
- Kathryn T. Gines. Black feminism and intersectional analyses: a defence of intersectionality. *Philosophy Today*, 55(Supplement):275–284, 2011.
- Till Grüne-Yanoff. Learning from minimal economic models. *Erkenntnis*, 70(1):81–99, 2009.
- Ed Hopkins. Two competing models of how people learn in games. *Econometrica*, 70(6):2141–2166, 2002.
- Jennifer Jhun, Patricia Palacios, and James Weatherall. Market crashes as critical phenomena? explanation, idealization, and universality in econophysics. *Synthese*, forthcoming.
- Peter Kwan. Jeffrey Dahmer and the cosynthesis of categories. *Hastings Law Journal*, 6(48):1257–1292, 1996.

- Travis LaCroix and Cailin O'Connor. Power by association. 2018.
- Helen Longino. Gender, politics, and the theoretical virtues. *Synthese*, 104(3):383–397, 1995.
- Marilyn Manser and Murray Brown. Marriage and household decision-making: A bargaining analysis. *International economic review*, pages 31–44, 1980.
- Leslie. McCall. The complexity of intersectionality. *Signs: Journal of Women in Culture and Society*, 3(30):1771–1800, 2005.
- Marjorie B McElroy and Mary Jean Horney. Nash-bargained household decisions: Toward a generalization of the theory of demand. *International economic review*, pages 333–349, 1981.
- Chandra Talpade Mohanty. Transnational feminist crossings: On neoliberalism and radical critique. *Signs: Journal of Women in Culture and Society*, 38(4):967–991, 2013.
- Aydin Mohseni and Cailin O'Connor. when size really does matter. 2018.
- Mary S Morgan and Margaret Morrison. *Models as mediators: Perspectives on natural and social science*, volume 52. Cambridge University Press, 1999.
- Athena D. Mutua. Multidimensionality is to masculinities what intersectionality is to feminism. *Nevada Law Journal*, 2(13):341–367, 2013.
- Jennifer C. Nash. Re-thinking intersectionality. *Feminist review*, 1(89):1–15, 2008.
- John Nash. Two-person cooperative games. *Econometrica: Journal of the Econometric Society*, pages 128–140, 1953.
- John F Nash. The bargaining problem. *Econometrica: Journal of the Econometric Society*, pages 155–162, 1950.
- Cailin O'Connor. The cultural red king effect. *Journal of Mathematical Sociology*, 2017a. DOI: 10.1080/0022250X.2017.1335723.
- Cailin O'Connor. *Dynamics of Inequity*. 2017b.
- Cailin O'Connor. Modeling minimal conditions for inequity. 2017c.
- Cailin O'Connor and Justin P Bruner. Dynamics and diversity in epistemic communities. *Erkenntnis*, 2017. DOI: 10.1007/s10670-017-9950-y.
- Cailin O'Connor and James Owen Weatherall. Black holes, black-scholes, and prairie voles: An essay review of simulation and similarity, by michael weisberg. *Philosophy of Science*, 83(4):613–626, 2016.

- Mike C. Parent, Cirleen DeBlaere, and Bonnie Moradi. Approaches to research on intersectionality: Perspectives on gender, lgbt, and racial/ethnic identities. *Sex Roles*, 68(11-12):639–645, 1995.
- Amy R Poteete, Marco A. Janssen, and Elinor Ostrom. *Working together: collective action, the commons, and multiple methods in practice*. Princeton University Press, 2010.
- Hannah Rubin and Cailin O'Connor. Discrimination and collaboration in science. *Philosophy of Science*, forthcoming.
- Karl H Schlag. Why imitate, and if so, how?: A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*, 78(1):130–156, 1998.
- Joseph P. Simmons, Leif D. Nelson, and Uri Simonsohn. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11):1359–1366, 2011.
- Brian Skyrms. Sex and justice. *The Journal of philosophy*, pages 305–320, 1994.
- Brian Skyrms. *The stag hunt and the evolution of social structure*. Cambridge University Press, 2004.
- Olúfẹ́mi O. Táíwò. The Man-Not and the dilemmas of intersectionality. *APA Newsletter on Philosophy and the Black Experience*, 17(2):6–10, 2018.
- Jörgen W Weibull. *Evolutionary game theory*. MIT press, 1997.
- H Peyton Young. An evolutionary model of bargaining. *Journal of Economic Theory*, 59(1):145–168, 1993.