Promoting Diverse Collaborations

Mike D Schneider[1], Hannah Rubin[2,3], and Cailin O'Connor[1]

[1]University of California, Irvine
[2]University of Groningen
[3]University of Notre Dame

Abstract

Philosophers of science and social scientists have argued that diverse perspectives, methods, and background assumptions are critical to the progress of science. One way to achieve such diversity is to ensure that a scientific community is made up of individuals from diverse personal backgrounds. In many scientific disciplines, though, minority groups are underrepresented. In some cases minority members further segregate into sub-fields, thus decreasing the effective diversity of research collaborations. In this paper, we employ agent-based, game theoretic models to investigate various types of initiatives aimed at improving the diversity of collaborative groups. This formal framework provides a platform to discuss the potential efficacy of these various proposals. As we point out, though, such proposals may have unintended negative consequences.

1 Introduction

Philosophers of science and social scientists have argued that diversity in scientific communities is critical to the progress of science, and have explored initiatives that might help diversify science. However, there has been much less work done on promoting diversity in collaborative teams, where scientists are actually interacting and working with those unlike themselves. Rubin and O'Connor [2018] use evolutionary game theoretic models to show that when members of one social group tend to get more credit for collaborative endeavors this can disincentivize collaboration between groups, leading individuals to mostly collaborate with those like them. This sort of process can negatively impact the progress of science whenever collaboration benefits from diversity.

In this paper, we use a similar evolutionary framework to explore the conditions which promote diverse collaborations in scientific disciplines. In particular, we employ agent-based, game theoretic models of actors in collaboration networks to test how discriminatory norms interact with individual decisions to collaborate across identity groups. We consider various types of policy proposals aimed at improving diversity, including measures to promote the representation of minority groups and active incentives for diverse collaboration, focusing on the latter. As we will outline, a tension arises—some policies that could successfully increase the diversity of scientific collaborations will also increase the level of inequity experienced within the community. In other words, we identify cases where policies to promote epistemic goods and policies to promote social goods in scientific communities can come apart. However, as we further argue, segregation into scientific subfields based on social identity can have negative effects if and when subfields associated with minority groups lose standing.

We will begin by briefly highlighting the ways that personal diversity might matter to the success of scientific collaboration, using a few historical examples. In section 3 we describe the modeling framework we will employ here, which uses bargaining models from game theory to explore the emergence of patterns of collaboration. We will also discuss a few relevant previous results. Section 4 describes different proposals for increasing diversity in collaboration networks. Section 5 looks at one of these in more detail, exploring the possibility of directly rewarding diverse collaborations. In section 6, we again appeal to the history of science to discuss how a prolonged lack of diverse collaborations in an otherwise diverse scientific community might itself lead to inequities.

2 Why Diverse Collaborations Might Matter

Diversity has been championed as an important feature of successful academic communities both by those in feminist epistemology/philosophy of science, and by those doing formal work in social epistemology. As mentioned, though, this paper addresses diverse collaborations, not simply diverse communities. That is, we are interested in what might cause collaboration networks to be *homophilic*, segregated along social identity lines, and what interventions might break these patterns. One question of obvious relevance to this exploration is: does the type of homophily we discuss actually impede epistemic progress? Presumably, we should be most worried about homophily in epistemic groups if it hurts scientific inquiry.

So long as diverse ideas are present somewhere in a community, we might ask, why should it matter whether collaborations themselves are diverse? One might imagine a circumstance in which a researcher from one social identity group is likely to figure out A and a researcher from another group B. If they collaborate, then they might also together conclude C, which follows from A&B. However, the community has another route to concluding C: A and B are published separately, whereupon the community as a whole has access to these ideas and any one of them can conclude C. This is somewhat similar to the picture Okruhlik [1994] has in mind—diverse researchers will generate and test diverse hypotheses which will then be assessed by the usual scientific methods. In such a case, diversity within a community matters, but diversity within collaborative groups does not.

There are few reasons to think that diversity within collaborations themselves might be important as well, though. First, it is possible that independent discoveries or ideas (accessible to members of different groups) which together would generate outcomes worthy of publication, are on their own relatively insignificant. In the toy example from above, A and B might not individually warrant publication and only be valuable realizations insofar as they jointly imply C. If this is the case, it is reasonable to think that A and B would never be published on their own. This might be especially likely if members of one group struggle to publish in top journals due to reputational effects. (This possibility will be discussed more in section 6.)[1]

In addition, there is some evidence that the actual process of group deliberation and decision making is altered, and sometimes improved, by personal diversity. Diversity seems to prompt individuals to state their background assumptions and beliefs, rather than assuming they are shared, and to challenge others' assertions more readily. This is sometimes called information elaboration. For example, Sommers [2006] finds that racially diverse juries deliberate differently by sharing more information.[2] Phillips et al. [2006] find that small, racially diverse groups tend to solve problems better than homogenous ones. And Van Dijk et al. [2012] in a meta-analysis argue that objective measures of performance show general benefits of diversity in teams. More specific to academia, it has been shown that culturally diverse collaborative teams tend to be more productive, arguably due to the presence of diverse skills, experiences, and cognitive frameworks [Barjak and Robinson, 2008].[3] Additionally, Campbell et al. [2013] find that in ecology, gender diverse collaborations generate work that is cited more by peers, and argue that this is an indication of its higher quality. Though Bear and Woolley [2011], Eagly [2016], who give literature reviews of work investigating gender diversity and group performance, find contextually sensitive and mixed results.

Lastly, information spreads more slowly through homophilic networks [Golub and Jackson, 2012], so a homophilic epistemic network will be less efficient in that it will take longer for the community to reach various conclusions, assuming (as we think is reasonable to do) that people are more likely to engage with their co-authors, in terms of discussing relevant research and spreading ideas. Under certain assumptions, such homophilic networks can also prevent the spread of new and better scientific practices throughout the community as a whole [Schneider, working paper].

In order to elucidate the sorts of cases where diversity might improve collaboration, we will now pull out details from a few cases in the history of science. Without regular collaboration between women and men, sexology at the turn of the twentieth century simply would not have prospered as it did. Leng [2013] illustrates how women who participated in the predominantly masculine discourse concerning human female sexuality helped improve the state of the field at the time:

> As the British physician Havelock Ellis pointed out, the female sex drive was an 'elusive' phenomenon, a 'mocking mystery' even, because social prohibitions against female sexual expression made it extremely difficult to acquire accurate and comprehensive information – for male physicians, at any rate (Ellis, 1902, p. 47). (p. 132-133)

Although the male sex drive was well-studied, to have a general theory of sexology, the men who dominated the community needed to collaborate with women

[1] Thanks to Liam K. Bright for input in this discussion.

[2] This echoes claims from Longino [1990] about the ability of diverse scientific groups to challenge the biased claims of their peers.

[3] Freeman and Huang [2015] similarly find that ethnically diverse collaborative papers tend to be published in more prestigious journals and cited more often, possibly due to diversity of knowledge though, in this case, also possibly due to network effects. That is, the authors might belong to different social networks, so, together, can inform more people about their work, resulting in increased citations.

who were in the position to identify which of their conclusions were not, in fact, general ones, but were specific to the male sex drive. Women researchers, on the other hand, were for obvious reasons less able to generate successful insights into aspects of male sexuality. It was by virtue of there being active collaborations across gender lines (even if they did not actually co-author – the academic environment generally promoted sole-authored publications) that the field of sexology improved (c.f. Leng [2013, p. 147]).

Another historical example can be found in the early history of public museums, in which collaborations across gender lines were essential to the success of a new field of museum pedagogy. In the United States at the turn of the 20th century, museum educators, predominately women, were experts on innovative pedagogical techniques appropriate for educating the public and, in particular, for younger audiences. These women brought their expertise in pedagogy and library science to newly public-facing natural history museums, run by men with advanced degrees in the natural sciences [Kohlstedt, 2013]. Productive collaborations between men and women whose expertise differed according to their respectively gendered occupations allowed museums to emerge as institutions of public (and often hands-on) learning, and allowed the field of museum pedagogy to thrive.

3 Bargaining and Discrimination

The last section made the case for the potential importance of diverse collaborations to science. We will now use evolutionary game theory to explore possible avenues for promoting such diverse collaborations. To do so, we draw on previous results regarding the dynamics of discrimination and collaboration. Rubin and O'Connor [2018] introduce a framework with two elements—a network, where each researcher is a node and each collaborative engagement a link between nodes, and a bargaining game representing collaborative interactions between the individuals on the network. Agents update their collaboration strategies, as well as their network links, creating a dynamic system where the way academics treat each other can influence who they decide to work with.

It may sound unintuitive to use a bargaining game as a representation of academic collaboration, but, in fact, collaboration is a strategic interaction where researchers have to bargain, whether implicitly or explicitly, to decide 1) how much work each will do and 2) how to divide credit in the form of author order. The particular bargaining game they use, which we will also employ, is a mini version of the Nash demand game.[4] In this model, two actors each demand some portion of a resource – a low, medium, or high amount. When applied to scientific collaboration, these demands are best understood as requests for author position relative to amount of work done. An academic who does the majority of work on a project, and is first author, is demanding a medium amount, whereas a first author who did little work makes a high demand.[5] If their demands are compatible in that they do not exceed the resource (credit per

[4]This game is also sometimes referred to as the Nash bargaining game, bargaining game, divide the pie, or divide the cake.

[5]Notice that this model assumes a 'credit economy' for academics, where they are motivated to seek credit, just as others might be motivated by money. This is a standard assumption in formal social epistemology.

	Player 2			
		Low	Med	High
Player 1	Low	4,4	4,5	4,6
	Med	5,4	5,5	0,0
	High	6,4	0,0	0,0

Figure 1: A payoff table for a mini Nash demand game. Rows represent strategies for player 1, and columns for player 2. Entries list payoffs for combinations of strategies with player 1 first.

time, in this case), they each get what they demand. If they are incompatible, the assumption is that the actors cannot peaceably reach an agreement and so they instead get a poor payoff called the 'disagreement point'. For simplicity sake, we assume the resource has a value of 10, the medium demand is always 5, and that the other two demands are compatible (3 and 7, for example, or 1 and 9). Figure 1 shows a payoff table of this game where Low and High are 4 and 6 and the disagreement point is set to 0.[6]

This game has three 'Nash equilibria', or strategy pairings where actors have no incentive to change what they are doing. Because no one can get a higher payoff by deviating from such an equilibrium, they tend to be stable, and, in particular, to show up as the end point of evolutionary processes. In particular, this game has one fair equilibrium where both actors demand 'Med', and two unfair or inequitable ones where one actor demands 'High' and the other 'Low'. We call these latter two equilibria inequitable because one collaborator does more work per credit received, and the other less.

In the models we will consider, and those considered by Rubin and O'Connor [2018], agents play this Nash demand game on a network and may be of two types which might represent two races, or genders, or cultural groups. Individuals can condition their demands based on the type of their partner.[7] Note that this model has the capacity to represent something like a discriminatory norm or convention. When members of each group tend to make fair demands within their own type, but between groups one side tends to demand High and the other Low, we call this discrimination.[8] Rubin and O'Connor [2018] show that under many conditions discrimination in this sense can emerge in such models.[9]

[6]This model collapses two aspects of bargaining—over work done and over credit received—into one payoff, creating a mismatch between the model and the target system. Cochran and O'Connor [2019] look at related models, but where bargaining over work and credit are explicitly separated. They find that inequity emerges commonly in these joint models. Although they do not directly investigate network structure, the general finding from Rubin and O'Connor [2018]—that agents will learn to avoid those discriminating against them—should hold for the more complicated model. Later, in footnote 18, we discuss possible ramifications of this simplification for our findings.

[7]In focusing on two types, their models ignore the possibility of intersectional identities. For instance, a researcher might be both a woman and black, or both a man and disabled. O'Connor et al. [2019] look at related, non-network models where there are multiple, intersectional identities. They find that inequitable conventions can emerge across any division into types. Whenever that happens, the work of Rubin and O'Connor [2018] predicts that homophily should emerge. Furthermore, our analysis of interventions to increase out-group collaboration should apply to these more complicated scenarios directly.

[8]This is similar to usage of the term in Axtell et al. [2001].

[9]See also Poza et al. [2011].

In particular, they find that a group is more likely to end up being discriminated against when they are in the minority.[10] Furthermore, when members of one group face discrimination, they tend to break out-group links and collaborate with those like themselves. This results in a homophilic network, rather than one where collaborations tend to include diverse members.[11]

In real epistemic groups, notably, similar patterns have been observed. Studies have found that women are less likely to hold prestigious first and last author positions [West et al., 2013, Sugimoto et al., 2013]. [Feldon et al., 2017] find that women students in biology labs tend to put in more work, but are less likely to be granted authorship on papers. Additionally, in some disciplines, researchers have found that women are less likely to collaborate in general, and more likely to collaborate with other women [Ferber and Teiman, 1980, McDowell and Smith, 1992, Boschini and Sjögren, 2007, West et al., 2013]. Botts et al. [2014] also find that black philosophers tend to cluster into subfields.

It should be noted that a similar effect is expected when external forces distribute credit inequitably, even if academics are treating each other fairly. For example, in economics author order is alphabetical, but women who co-author are much less likely to receive tenure than men who co-author (holding all else fixed) [Sarsons et al., 2017].[12] The models from Rubin and O'Connor [2018] would predict that in such a case women would tend to learn to stop collaborating with men. (And indeed, economics is one of the disciplines where this very pattern has been observed.)

4 Diversity Initiatives

When inequity or discrimination disincentivize collaboration between members of different social identity groups, what steps can be taken to re-incentivize collaboration among socially diverse partners? Here, we will discuss some possible policy proposals aimed at improving the diversity of individual collaborations using the framework just described.

4.1 Improving Minority Representation

As mentioned, when one group is in the minority, the chance that they end up discriminated against in network models is higher. Other results on the emergence of bargaining norms suggest that, in general, minority status may lead to similar disadvantage.[13] Perhaps, then, a solution is to try to increase the prevalence of members of minority groups in order to promote fair bargaining norms (and diverse collaborations).

[10] This is a result in asymmetries in their number of between group links. In particular, between any two groups there is some number, n, of collaborative engagements. This means that when one group is in the minority, they will have more between group collaborative links on average than the majority group. As a result, their strategy updates will tend towards more conservative, lower demands, which means that the chance they end up receiving less credit at equilibrium is increased. See Rubin and O'Connor [2018] for further details.

[11] O'Connor and Bruner [2017] also show in a population level model how discrimination can disincentivize individuals from collaborating at all.

[12] This effect is ameliorated when women co-author with other women.

[13] This has been called the cultural Red King effect. For more, see Bruner [2017], O'Connor and Bruner [2017], O'Connor [2016, 2019].

In the modeling framework described, though, the addition of individuals to a minority group in an existing community will not change existing patterns of bargaining. This is because new members of a group will be in an environment where those around them already tend to adhere to some norm, meaning that their best response will involve adhering to the same norm. Imagine, for example, a community where women are in the minority, everyone makes fair collaborative demands of their in-group, and when men and women collaborate, men demand High and women Low. If more women are added to this network, they will meet men who demand High of them, meaning that Low is their most successful response. These new members of the community will learn to demand Low of men, and eventually to avoid collaborating with them, leading to a perpetuation of homophily and non-diverse collaborations. In communities where there is not a stable norm, but some variety of behaviors, we should still expect all those incoming academics who meet discriminating out-group members to learn to avoid them.

This is not to suggest that there are not good reasons to promote the presence of underrepresented minorities in epistemic groups. The point here is merely that we should not expect the simple addition of minority individuals to change inequitable patterns of behavior to equitable ones, or to decrease collaborative homophily.

4.2 Special Grants for Diverse Collaborations

Another suggestion might be that grant-giving agencies create special initiatives to promote diverse collaborations. There are a few ways to do this. In Rubin and O'Connor [2018], academics only have a certain number of collaborative links available to them. This makes sense, since no one has an infinite amount of time and resources for academic work. Grant agencies, then, might introduce initiatives to clear up the schedules of academics who are interested in engaging in a new collaborative project with an out-group member. This could involve, for example, a paid course release. Another possibility is special money to hire research assistants who can lighten workload, and so create more time for those interested in an out-group collaboration. These interventions might be thought of as creating new links for academics, but ones that can only be used for diverse collaborations.

As a result, we should expect between group collaborations to increase under this sort of initiative. There is a possible downside, however. Under this type of initiative, academics will choose to collaborate with out-group members even in cases where they are being discriminated against. After all, some amount of academic credit is better than none. If there are norms and patterns of discrimination in an academic community, then out-group links tend to involve someone being taken advantage of. The gain in diversity within collaborations is a loss in equity.

Another similar initiative might increase the credit granted to collaborations between scientists in different groups. I.e., instead of increasing links, increase the size of the credit-pie that collaborators share. This could be achieved by making it more likely that diverse groups win grants or by giving larger grant amounts to projects with diverse investigators. (Scientists with such grants can publish more, generating more credit.) It is less immediately clear what the effects of this sort of initiative might be, as researchers still must choose whether to collaborate with in- or out-group members, only now with an added incentive to out-group collaboration. Therefore, we provide a model in the next section in order to evaluate the possible consequences of this sort of initiative.

5 Modeling Increased Credit

As in Rubin and O'Connor [2018], our agents can learn to update both their bargaining strategies (collaborative behavior) and their network structure (who they collaborate with). We start with an empty network, with each agent's bargaining strategy randomly determined. In each round, there is some small probability that each agent will take an action. If an agent takes an action, there is a chance they will update their set of collaborators and a chance they will update their bargaining strategy (agents do not update both at once).

Each agent's bargaining strategy consists of two parts: a demand when interacting with an in-group member and a demand when interacting with an out-group member. Agents receive payoffs from each successful collaboration.[14] Agents update their strategies by using what is called myopic best response: the strategy an agent picks is the one that would have gotten them the best payoff in the last round, given the demands of their collaborators.[15] This captures the idea that agents are trying to choose a strategy that is likely to result in them getting the most out of a successful collaboration, while avoiding the poor payoff from a failed collaboration.

The evolution of the collaborative network is slightly more complex. We employ a model similar to Watts [2003] in which agents can choose to form or break links with other agents based on their payoffs from bargaining with those other agents. A player can unilaterally sever a link, but both players must consent to a new link being formed. This represents the fact that all the researchers involved in a collaboration must consent to be part of the collaboration. Additionally, agents have a maximum number of links, capturing the fact that there are a limited number of projects academics can work on.

As mentioned, we begin with an empty network (there are no links between any nodes). At each time-step, two nodes are chosen at random. One of these is an agent who will update their links and the other is either a potential or current collaborator of the agent. If it is a potential collaborator, we determine whether both parties will consent to form a new link between them – each will consent if they either do not already have the maximum number of links or they can increase their payoff by breaking a link with another collaborator in order to form this new link.[16] If both agents consent they will form the link, otherwise no links are formed or broken. By contrast, if we have chosen a current collaborator, the agent has an option to break the link and form a link with a new randomly chosen collaborator. Again, both the agent and the new potential collaborator must consent to form the link, otherwise no links are formed or broken.[17]

[14]Their total payoff then is just the sum of all these payoffs. Agents who either have not formed collaborative links yet or who are only part of collaborations in which the parties' demands exceed the whole will have a total payoff of 0.

[15]In the event of a tie for best, one of the best responses is chosen at random.

[16]If an agent chooses to break a link, they break the one that gives them the lowest payoff (chosen randomly in the case of a tie).

[17]See Rubin and O'Connor [2018] for more details, as we use the same network dynamics

To give a big picture of what this model represents—imagine a community of researchers, some men and some women. They have regular collaborative partners, with whom they divide labor and credit. Sometimes they realize that their collaborative strategies are not working (because, for example, they are too aggressive about credit, or maybe are doing too much work), and try something that will be more successful given what their partners are doing. Sometimes they encounter a new collaborative partner who is less demanding, and decide to drop an older working partner. Over time, the whole group settles down into stable partnerships, and stable demands for work and credit.

As mentioned, we are particularly interested in the effects of changing the payoffs for between group collaboration, under the idea that this might promote diverse collaborations. In the Nash demand game we presented, actors divide a resource of size 10. Now, we vary the size of the 'pie' for between group collaborations by multiplying the total available payoff (credit) by some amount, π. We look at results for π ranging from .5 to 2, in intervals of .125. That is, we look at cases where between group collaborations are half as valuable to twice as valuable as within group collaborations.[18]

5.1 Results

A preliminary result is that varying the payoffs for between group interactions does not change the relationship between minority group size and majority discrimination discussed in Rubin and O'Connor [2018]. Majority group members are still more likely to discriminate when the minority group is smaller, no matter the amount of credit allocated to between group collaboration.[19]

We were also interested in how changing the value of between group collaborations affected the between group collaboration levels. In order to quantify this sort of effective diversity of the network, we use the following measure of homophily, called *inbreeding homophily*:

$$I_i = \frac{H_i - w_i}{1 - w_i}$$

where H_i is the proportion of a group i's links that are within group links and w_i is the fraction of the population that group comprises [Currarini et al., 2009]. This measure takes into account what level of between group linking would be expected given the relative sizes of the groups, and then yields a number that is positive when between group linking is less than expected (i.e., when there is homophily), and negative when it is greater than expected.

Figure 2 shows that varying π affects homophily in a predictable way. When between group collaborations are less valuable than within group collaborations ($\pi < 1$), homophily is high. As these between group collaborations become

they do.

[18]Data were collected for networks of 100 agents with a high demand of 6, probability of taking an action set to 10%, a probability of updating a link rather than a strategy of 20%, minority size ranging form 10% to 50% of the population in intervals of 10%, and maximum number of links set to 3 or 9. Each combination of parameters was run 100 times, and for 10,000 rounds.

[19]In fact, the proportion of majority expected to discriminate (i.e. the number of majority group members whose out-group strategy is to demand High) is very similar to what Rubin and O'Connor [2018] find, ranging from about .4 when the minority is 10% of the population to about .1 when the minority is 50% of the population.

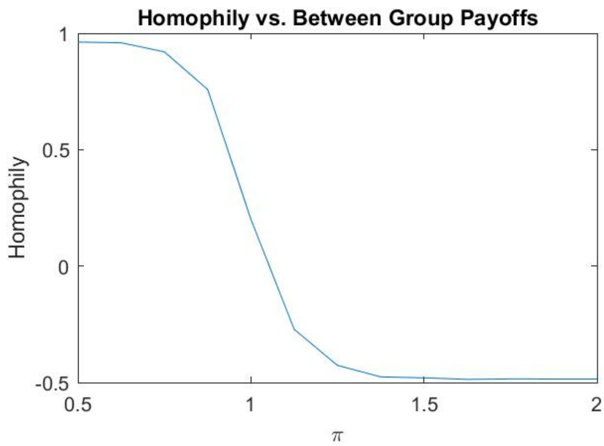

Figure 2: Level of homophily, measured using inbreeding homophily, over changes in the payoffs for between group collaborations.

more valuable, homophily decreases and becomes negative, i.e. there are more between than within group collaborations.

So the results suggest that special incentives increasing the credit available to members of diverse collaborations could, indeed, improve the diversity of collaborative scientific groups. But this diversity comes at a cost. Rubin and O'Connor [2018] found that even when many members of the majority group have underlying discriminatory strategies, there was not much actual discrimination occurring in their models since minority group members were able to break links with discriminators. One concern is that increasing π will lead to more discrimination against the minority group: minority group members will be incentivized to accept discriminatory collaborations because they are worth more than fair collaborations with in-group members (i.e. $\pi \cdot L > M$).

Figure 3 shows how varying π affects discrimination. We look at the *proportion of discrimination*, which is the proportion of between group collaborations that involve a majority member who demands high. In addition, we look at the *instances of discrimination*, found by simply adding up the total number of collaborations in which a majority member discriminates against a minority group member.[20]

We can see from figure 3 that increasing π only slightly increases the proportion of discrimination.[21] This might seem encouraging but we should not be too optimistic. The counted up instances of discrimination increase precipitously as π increases, as shown by the dashed line in figure 3. This is because,

[20]Note that this is averaged over every run of the simulation, so it averages over different minority sizes. The same general trend is observed for any size of the minority, but the numbers will differ.

[21]The one exception to this is at $\pi = 1.125$, where there is a dip in actual discrimination. This is because both majority and minority members prefer fair between group collaboration to fair within group collaboration ($5 \cdot 1.125 > 5$), while both (but most notably, the minority) still prefer fair within group collaboration to receiving the low payoff from between group collaborations ($5 > 4 \cdot 1.125$), meaning that they break off between group links with discriminators. Note that even though the proportion of discrimination decreases at this point, the instances of discrimination increase.

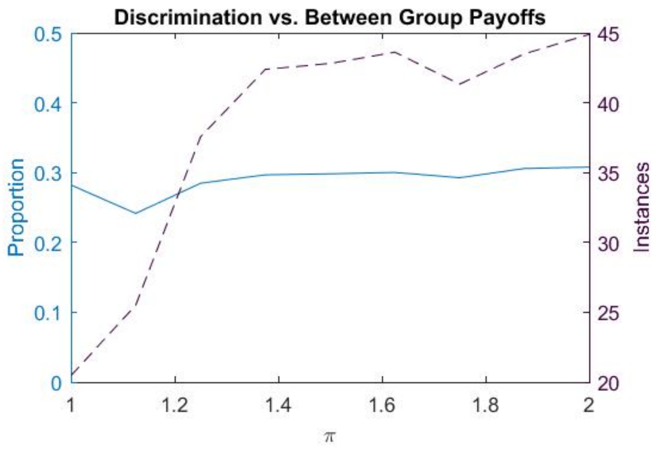

Figure 3: Majority discrimination over changes in the payoffs for between group collaborations. The line in blue is given in terms of the scale on the left-hand y-axis and the line in purple is given in terms of the scale on the right-hand y-axis.

as shown in figure 2, increasing the size of the pie increases the total number of between group links. Since the between group links tend to be discriminatory at a higher rate, this means that individuals are now incentivized to accept these discriminatory interactions. To be completely clear, the majority group is not acting any worse than before (i.e. demanding High more often), but the policy brings minorities in contact with existing discriminatory behavior.

If all we were concerned with were the proportion of between group collaborations that were discriminatory, then this policy might seem beneficial: it increases diversity, and majority group members are about as likely to discriminate as before. Furthermore, the policy as modeled improves outcomes for everybody by increasing between group credit. However, the minority group is now receiving a greater number of inequitable outcomes as a result of the increase in out-group collaboration. And, on average, there is far more discriminatory behavior across the entire community. Furthermore, as we will explain in the next section, the situation is worse when we more accurately represent the conditions under which a diversity initiative like this would typically be implemented.[22]

5.2 Non-Random Starting Points

The models discussed in the last section start with a random distribution of strategies and no network structure. But diversity initiatives tend to be implemented in communities where there is already a lack of diversity, and where discriminatory behavior is present. To more realistically represent these starting

[22]In footnote 6, we point out that this model collapses bargaining over workload and over credit. The incentive scheme examined here only actually increases credit—workload is not effected. If we consider a more complex model where both interactions are represented, however, we should expect similar findings. Increasing credit should increase between group linking, whether or not discrimination occurs in workload or credit sharing or both.

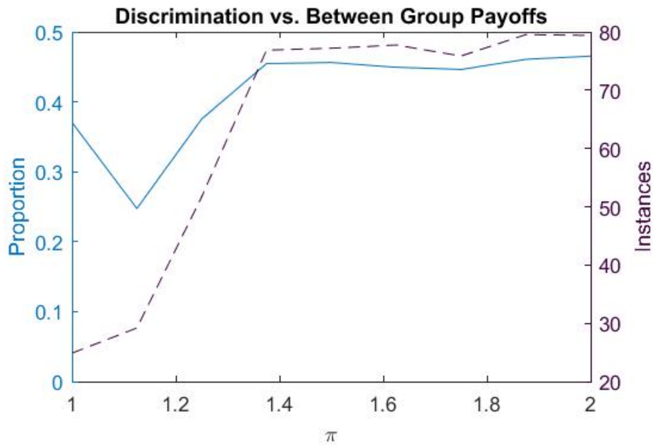

Figure 4: Level of majority discrimination over changes in the payoffs for between group collaborations. The line in blue is given in terms of the scale on the left-hand y-axis and the line in purple is given in terms of the scale on the right-hand y-axis.

conditions, we alter the above model in two ways.

First, instead of beginning with an empty network, we start with a homophilic network. In particular, we make use of multi-type random graphs—networks which are used to model populations with multiple social identity groups [Golub and Jackson, 2012]. Each agent has some probability of forming a link with each in-group member, p_{in}, and some probability of forming a link with each member of the other group, p_{out}. When $p_{in} > p_{out}$, the network is homophilic. Here, we set $p_{in} = .1$ and $p_{out} = .05$. After we form the network, we then ensure that no agent has more than the maximum number of links. For agents that have more links than maximum, we randomly choose links to break until they are at the maximum number.[23] This procedure results in networks that, on average, have an inbreeding homophily of .18, which roughly matches the average level of homophily when $\pi = 1$.[24]

Second, in order to represent the fact that minority groups often receive less from collaborations in real academic communities, we do not start with a random distribution of strategies. Instead, we alter the probabilities with which between group strategies are assigned initially, so that there is a 45% chance that a majority group member will demand High against a minority group member and, in turn, a 45% chance of a minority group member demanding Low against a majority group member (each of the other strategies are employed with equal probability). Additionally, to capture the fact that the fair outcome is most common within groups, there is initially a 90% chance of demanding Med of in-group members. These choices are arbitrary, but capture the sort of case where discrimination and homophily are already occurring in the community.

[23]This means not all the agents will have the maximum number of links at the start of the simulation, but those agents can form links up to the maximum as they update their collaborations.

[24]The average homophily was estimated by forming 10,000 networks for each possible minority group size, then averaging over all data points.

In this altered model, we see interesting changes in how varying π affects discrimination.[25] As figure 4 shows, increasing π leads to more discrimination, now both in terms of overall instances of discrimination *and* the proportion of between group collaborations that are discriminatory.[26] This occurs because as minority members seek to form between group links, it is more likely these links will be discriminatory for two reasons. First, it is less likely they will find majority group members demanding Med. Second, it is more likely they will be able to successfully collaborate with a majority group member demanding High (because the minority group members are more likely to demand Low between groups). That is, minority members are both more likely to encounter and to accept inequitable demands than when we looked at random starting conditions in section 5.1.

So, as it turns out, more accurately representing the conditions under which we believe these initiatives might be implemented only makes the negative consequences of the policy more prevalent. Note that these results are for only mild increases in the likelihood of majority members demanding High (an increase from 33% under the assumption of random starting points to 45%). More initial potential discrimination would only lead to greater increases in actual discrimination.

5.3 Implications

We have argued that diversity initiatives which promote between group collaborations may achieve their goal, while unintentionally fostering inequity in the credit awarded to minority group members. So, in the short term at least, epistemic and social good come apart. This is pertinent in light of a common argumentative technique in discussions about increasing diversity in academic communities that goes like this: here are X, Y, and Z reasons why increasing diversity will promote various social goods. If these fail to convince; here are the purely epistemic benefits as well. Call these 'private sins as public goods' arguments[27] because they convince those only interested in epistemic gains to incidentally promote socially beneficial policies.[28] Unfortunately, this argumentative strategy is not available if an epistemically beneficial policy turns out to have unintended negative social consequences. The instinct to give 'private sins as public goods' arguments backfires in these situations, insofar as the default expectation is that all initiatives ought to carry explicit epistemic benefits, while social benefits are an afterthought.

There are further implications when thinking about the long term, after the initiative ends. If the initial lack of diverse collaborations is due to minority group members breaking links with majority group members in order to protect

[25] The first two results from section 5.1 are affected in predictable ways. The minority group is still more likely to be discriminated against the smaller it is, but a greater proportion of the majority discriminate at every group size. This should be unsurprising since we started with the majority more likely to discriminate. For the effect of π on the amount of homophily, the shape of the line is the same as in figure 2, but homophily varies from about .65 for $\pi = .5$ to about -.15 for $\pi > 1.25$. This is because we started the network with homophily around .18 and simulations were only run for 10,000 rounds.

[26] This, again, is modulo a downswing when $\pi = 1.125$.

[27] Thanks Jan-Willem Romeijn for this phrase.

[28] The arguments for the epistemic benefits of diverse collaborations in section 2 perhaps count as such arguments.

themselves from discrimination, then when the initiative ends, homophily will likely reappear. In the meantime, the initiative has not only perpetuated, but further entrenched inequity. This is because even though both majority and minority group members' absolute payoffs are increased during the initiative, the difference between the minority payoffs and majority payoffs increases.[29] So, the initiative only temporarily fosters diversity, while the majority group accumulates more power, prestige, etc., by collaborating with the minority and receiving an undue amount of credit for those collaborations.

It is important to note that our aim is not to argue against the implementation of these diversity initiatives, and we do not take our results to yield any specific policy recommendations. Rather, we think these results caution against naive instincts that any policy will be helpful. Some initiatives may need to be complemented by other policies in order to achieve their goals without further exacerbating the situation they are intended to amend. For instance, the diversity initiatives discussed in this section might helpfully be complemented with improved standards for awarding credit to ensure equitable collaborations.

6 A Contagion of Disrespect

In section 2, we argued that diverse collaborations can carry epistemic benefits. However, as the previous section showed, there may be circumstances in which policies to promote epistemic goods and policies to promote social goods in scientific practice come apart. One may worry that the epistemic benefits of diverse collaborations are insufficient to merit policies that entrench social inequity. Unfortunately, there may be further reason to think that homophily, by itself, will lead to a different sort of inequity, which, perhaps, diverse collaborative links might solve. That is, although we have argued that some initiatives to promote effective diversity may carry negative social consequences, there are reasons to think that we need to do *something*, because persistent lack of effective diversity can create new social and epistemic harms.

Homophilic scientific communities can come to champion different scientific sub-disciplines, or 'niches'.[30] The cases we are worried about are those where niches championed by minority groups become low prestige by virtue of their associations with marginalized researchers while, correspondingly, niches championed by the majority group enjoy elevations of prestige. That is, lack of effective diversity can create the conditions for particular domains of study to become ghettoized by virtue of their associations with particular marginalized social groups. They suffer a *contagion of disrespect*[31], where new results in these

[29] Consider a simplified example to demonstrate this point. A minority group and a majority group member each have a fair within group link, garnering a payoff of 5. Then, an initiative corresponding to $\pi = 1.5$ is put in place, and they form a between group discriminatory link. Now the minority group member receives a payoff of 6 and the majority group member a payoff of 9. Both agents receive higher payoffs than before, but inequity has increased—the majority group member benefits much more from the initiative.

[30] There is some evidence suggesting that minority groups may organize in this way. For instance, Botts et al. [2014] find that black philosophers tend to cluster in subfields. Schneider [working paper] also provides some reason to think that these niches will persist when there is little collaboration across social identity lines.

[31] Thanks to Liam K. Bright for this phrase. Such a 'contagion' is reminiscent of the well-studied 'devaluation view' in sociology of employment, which holds that "a change in the gender composition of an occupation will lead to a change in the valuation of the work being

niches are increasingly dismissed as unimportant to the production of scientific knowledge.[32]

In the modeling framework discussed, this sort of process will lead to a state where the total size of the pie for one in-group gets smaller. As evidence for this contagion of disrespect, we provide two examples from the history of science, in which niches that came to be associated with women declined in prestige as a result. We end by considering the the effects of such a contagion on individuals' choices to suffer inequitable collaborations across groups.

6.1 Child Study

Beginning in the 1870s and 1880s various 'men of science' interested in the emerging field of naturalized psychology began to focus on infant cognitive development [Lorch and Hellal, 2010]. These men of science began publishing detailed accounts of their children's psychological developments, and encouraging others to do the same [von Oertzen, 2013, p. 176]. As von Oertzen [2013] notes, however:

> Scientific empiricism of this kind...presented unforeseen obstacles. The intimate space of the nursery, widely regarded by contemporaries as a quintessentially female domain, restricted fathers' and other men's access to human offspring. (p. 176)

This created an opportunity for scientifically-minded women, such as Milicent Shinn, to boldly go where men were culturally barred. In addition to compiling extensive notes on the early development of her niece that became a widely circulated book in the new field of child study in the 1890s, she also trained and established a network of college-educated mothers and aunts to behave as citizen-scientists, who provided valuable findings (c.f von Oertzen [2013, p. 183]).

Experimental psychologists and social scientists, nearly all men, were quick to categorically dismiss the field of child study for its reliance on women's observations. von Oertzen [2013], quoting from one paper by Shinn, writes:

> One prominent critic, American psychologist James Mark Baldwin, asserted that 'only the psychologist can "observe" the child, and he must be so saturated with his information and his theories that the conduct of the child becomes instinct with meaning for mind and body. This is just the difference between the mother and the psychologist – she has not theories: he has. She may bring up a family of a dozen and not be able to make a single trustworthy observation; he may be able from one sound of one yearling to confirm theories of the neurologist and educator which are momentous for the future training and welfare of the child'. (p. 186)

performed, leading to a change in occupations' relative pay rates" [Levanon et al., 2009, p. 868].

[32] There is some evidence that such a contagion exists even before the niches become meaningfully segregated. For instance, Sugimoto et al. [2013] find that papers across many scientific disciplines for which women were the sole author, first author, or last author were cited less often.

Baldwin's criticism will strike contemporary readers as backwards. Nonetheless, such views were widely held in the scientific community of the period. Despite the women in the field having amassed a wealth of data and analysis on child development, the credibility of the field of child study dwindled over the next twenty years [von Oertzen, 2013, p. 190]. Eventually, the field became overshadowed by experimental psychology, a discipline whose methodology favored men who controlled access to labs suitable for experiments. Here we see a clear case where the fact that the practitioners of the field were women led directly to a devaluation of the work done.

6.2 Home Economics

At the time of her death in 1911, Ellen Swallow Richards was the head of the Department of Social Economics at MIT and president of the Home Economic Association. Richards, the, "'engineer' of the modern home economics movement...saw domestic science as a way to move women trained in science into employment in academics and industry" [Stage, 1997, p. 5]. Home economics achieved this goal as it "tied the kitchen to the chemical laboratory, emphasizing nutrition and sanitation" [Stage, 1997, p. 5].

However, over the next several decades, home economics transformed from an academic discipline to an "international conspiracy to keep women in the kitchen" [Silva, 1998].[33] What is responsible for the field's fall from grace? There are good reasons to think that home economics suffered a serious decline in prestige precisely because its practitioners were predominately women. As Rossiter [1997, p. 96] explains, in the 1950s and 1960s,

> [College] Administrators, all men then, held skeptical and hostile attitudes toward home economics, even as they expressed unabashed ignorance about the field. ... To them such female domination constituted proof that the field was out of date.

While home economics was "one of the primary areas in which educated women found professional employment in academia and business from the 1900s to the 1960s..." [Stage, 1997, p. 4], and enrollment of students in the field modestly grew throughout the first half of the 1900s, the number of faculty working in the field and funding for research declined at increasing rates [Rossiter, 1997, p. 98-99].

In the context of this decline, centers of research in the discipline during the 1960s and 1970s began to re-brand the subject as 'human ecology' [Stage, 1997, p. 6]. That is, a new field was created whose methods and domains of inquiry happened to coincide exactly with the older field because its reputation had so deteriorated.

6.3 Choosing Inequity

Our models would predict that this sort of contagion of disrespect would incentivize members of minority groups to collaborate with the majority group,

[33] When the activist Robin Morgan spoke to the American Home Economics Association convention in 1972, she declared that "As a radical feminist, I am here addressing the enemy" [Stage, 1997, p. 1].

despite receiving inequitable amounts of credit.[34] If niches become low-prestige, the total payoff awarded to pairs of researchers working in these niches (the size of the pie) shrinks. At a certain point, the payoff for equitable collaborations in low-prestige niches will be less than the payoff for minority individuals accepting inequitable bargaining outcomes when collaborating with majority members in high-prestige niches.

At this point, minority researchers should come to prefer the unfair collaboration with a majority group member over the fair collaboration with a minority group member. This is similar to what was described in section 5 when the size of the pie was increased for between group (diverse) collaborations. However, in that case diverse collaborations had higher payoffs than either the minority or majority group collaborations. In this case, it is only that the minority collaborations have lower payoffs.

The history of science provides countless cases of women who succeeded in making contributions to a variety of high-prestige fields of science, but who were not awarded a fair share of the credit. In other words, they elected to channel their efforts into highly inequitable collaborations, rather than, for instance, into equitable collaborations in low-prestige niches associated with their gender. In light of what has been said here, one explanation of this choice is that they elected to suffer inequitable bargaining norms so as to enjoy access to the sort of high-prestige research happening in the niche dominated by members of the majority group.

The suggestion here is that while initiatives aimed at promoting diverse collaborations may lead to inequities, other social processes associated with homophily can do likewise. This should complicate our thinking about the potential benefits and pitfalls of initiatives to promote diversity.

7 Conclusion

We have seen how a variety of good-faith policy proposals to improve effective diversity in the name of scientific progress might succeed, but might also risk further entrenching social inequity in scientific communities. Our aim is not to conclude with explicit policy recommendations on the basis of these simplified models, but rather to have identified risks for various sorts of diversity initiatives to carry unintended negative consequences for the scientific communities involved. In particular, our investigations have suggested how policies to promote epistemic goods and policies to promote social goods in scientific communities can easily come apart. On the other hand, we have also suggested how a lack of any sort of policy at all may itself carry negative consequences, both epistemic and social, in the long term.

References

R. L. Axtell, J. M. Epstein, and H. P. Young. The emergence of classes in a multiagent bargaining model. *Social dynamics*, pages 191–211, 2001.

[34] Thanks to Liam K. Bright again.

F. Barjak and S. Robinson. International collaboration, mobility and team diversity in the life sciences: impact on research performance. *Social Geography*, 3(1):23, 2008.

J. B. Bear and A. W. Woolley. The role of gender in team collaboration and performance. *Interdisciplinary science reviews*, 36(2):146–153, 2011.

A. Boschini and A. Sjögren. Is team formation gender neutral? evidence from coauthorship patterns. *Journal of Labor Economics*, 25(2):325–365, 2007.

T. F. Botts, L. K. Bright, M. Cherry, G. Mallarangeng, and Q. Spencer. What is the state of blacks in philosophy? *Critical Philosophy of Race*, 2(2):224–242, 2014.

J. P. Bruner. Minority (dis)advantage in population games. *Synthese*, 2017. doi: 10.1007/s11229-017-1487-8.

L. G. Campbell, S. Mehtani, M. E. Dozier, and J. Rinehart. Gender-heterogeneous working groups produce higher quality science. *PloS one*, 8 (10):e79147, 2013.

C. Cochran and C. O'Connor. Inequity and inequality in the emergence of conventions. *Politics, Philosophy, and Economics*, (forthcoming), 2019.

S. Currarini, M. O. Jackson, and P. Pin. An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4):1003–1045, 2009.

A. H. Eagly. When passionate advocates meet research on diversity, does the honest broker stand a chance? *Journal of Social Issues*, 72(1):199–222, 2016.

D. F. Feldon, J. Peugh, M. A. Maher, J. Roksa, and C. Tofel-Grehl. Time-to-credit gender inequities of first-year phd students in the biological sciences. *CBE-Life Sciences Education*, 16(1):ar4, 2017.

M. A. Ferber and M. Teiman. Are women economists at a disadvantage in publishing journal articles? *Eastern Economic Journal*, 6(3/4):189–193, 1980.

R. B. Freeman and W. Huang. Collaborating with people like me: Ethnic coauthorship within the united states. *Journal of Labor Economics*, 33(S1): S289–S318, 2015.

B. Golub and M. O. Jackson. Network structure and the speed of learning measuring homophily based on its consequences. *Annals of Economics and Statistics/ANNALES D'ÉCONOMIE ET DE STATISTIQUE*, pages 33–48, 2012.

S. G. Kohlstedt. Innovative niche scientists: Women's role in reframing north american museums, 1880–1930. *Centaurus*, 55(2):153–174, 2013.

K. Leng. An 'elusive' phenomenon: Feminism, sexology and the female sex drive in germany at the turn of the 20th century. *Centaurus*, 55(2):131–152, 2013.

A. Levanon, P. England, and P. Allison. Occupational feminization and pay: Assessing causal dynamics using 1950–2000 us census data. *Social Forces*, 88 (2):865–891, 2009.

H. E. Longino. *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton University Press, 1990.

M. Lorch and P. Hellal. Darwin's "natural science of babies". *Journal of the History of the Neurosciences*, 19(2):140–157, 2010.

J. M. McDowell and J. K. Smith. The effect of gender-sorting on propensity to coauthor: Implications for academic promotion. *Economic Inquiry*, 30(1): 68–82, 1992.

C. O'Connor. The cultural red king effect. *The Journal of Mathematical Sociology*, 2016. doi: 10.1080/0022250X.2017.1335723.

C. O'Connor. *The Origins of Unfairness*. Oxford University Press, 2019.

C. O'Connor and J. Bruner. Dynamics and diversity in epistemic communities. *Erkenntnis*, 2017. doi: 10.1007/s10670-017-9950-y.

C. O'Connor, L. K. Bright, and J. Bruner. The emergence of intersectional disadvantage. *Social Epistemology*, 2019.

K. Okruhlik. Gender and the biological sciences. *Canadian Journal of Philosophy*, 24(sup1):21–42, 1994.

K. W. Phillips, G. B. Northcraft, and M. A. Neale. Surface-level diversity and decision-making in groups: When does deep-level similarity help? *Group processes & intergroup relations*, 9(4):467–482, 2006.

D. J. Poza, F. A. Villafáñez, J. Pajares, A. López-Paredes, and C. Hernández. New insights on the emergence of classes model. *Discrete Dynamics in Nature and Society*, 2011, 2011.

M. W. Rossiter. The men move in: Home economics in higher education, 1950-1970. *Rethinking home economics: Women and the history of a profession*, pages 96–117, 1997.

H. Rubin and C. O'Connor. Discrimination and collaboration in science. *Philosophy of Science*, 85(3):380–402, 2018. doi: 10.1086/697744.

H. Sarsons et al. Recognition for group work: Gender differences in academia. *American Economic Review*, 107(5):141–145, 2017.

M. D. Schneider. The non-epistemic origins of research strongholds. working paper.

E. B. Silva. Review: Rethinking home economics: Women and the history of a profession by Sarah Stage; Virginia B. Vincenti. *Work, Employment and Society*, 12(3):570–572, 1998.

S. R. Sommers. On racial diversity and group decision making: identifying multiple effects of racial composition on jury deliberations. *Journal of personality and social psychology*, 90(4):597, 2006.

S. Stage. Home economics: What's in a name? *Rethinking Home Economics: Women and the history of a profession*, pages 1–14, 1997.

C. R. Sugimoto, V. Lariviere, C. Ni, Y. Gingras, B. Cronin, et al. Global gender disparities in science. *Nature*, 504(7479):211–213, 2013.

H. Van Dijk, M. L. Van Engen, and D. Van Knippenberg. Defying conventional wisdom: A meta-analytical examination of the differences between demographic and job-related diversity relationships with performance. *Organizational Behavior and Human Decision Processes*, 119(1):38–53, 2012.

C. von Oertzen. Science in the cradle: Milicent shinn and her home-based network of baby observers, 1890–1910. *Centaurus*, 55(2):175–195, 2013.

A. Watts. A dynamic model of network formation. In B. Dutta and M. O. Jackson, editors, *Networks and Groups*, pages 337–345. Springer, 2003.

J. D. West, J. Jacquet, M. M. King, S. J. Correll, and C. T. Bergstrom. The role of gender in scholarly authorship. *PloS one*, 8(7):e66212, 2013.