1 Introduction

Over the past century, collaborative work has become increasingly common in academia [61]. In the sciences, it has continued to spread as a practice across disciplines and countries [50, 51]. In fact, in most scientific fields, the majority of published work is now co-authored [40]. The social sciences, likewise, have seen an increase in collaborative work, particularly in fields like economics and psychology [40, 32, 37]. Even in the humanities, collaboration, though not the norm, is increasingly practiced [37].

There are a number of reasons why this model of academic production is so popular. In some fields, increasing specialization, and the high cost of, for example, expensive instruments, mean that collaboration is necessary to produce meaningful research [40]. Many authors have argued, furthermore, that collaborating provides myriad benefits to individual researchers even when not strictly necessary. For example, collaboration can lead to greater academic productivity, which benefits individual researchers in hiring and promotion decisions [40, 36, 24, 38]. Co-authored papers in some areas are more likely to be accepted to top journals and are more likely to be cited [35, 28, 4]. Furthermore, it has been argued that collaboration can be useful to academics in defraying risk. When numbers of publications are important, it may make more sense to invest smaller amounts of time in many papers than to invest heavily in one paper that might not be accepted to a good journal [32]. Lastly, collaboration is important to academics for social reasons. Collaborators create stronger social networks in their fields, and may be viewed by peers as insiders [39].

These clear cut advantages to collaborative work raise ethical questions, however [46, 25, 33, 23, 64]. The structure of academia is such that rewards for successful work, such as promotion, prizes, and salary increases, tend to accrue to individuals. Furthermore, author order, which is the main proxy used to determine credit for academic collaborations, is inherently unequal. Co-authors must occupy different places in the author ordering. This means that academics must decide who will be best positioned to receive these individual rewards. Further benefits to collaboration, such as opportunities to present at prestigious conferences, must likewise be divvied up. How are the fruits
of collaboration divided in academic communities? Are these divisions equitable? Are there ways in which these divisions systematically disadvantage certain groups?

There are, at least, two types of potential disadvantage along these lines that academics might worry about. The first occurs between peers when one peer belongs to an underrepresented group in a field. For example, West et al. [61] and Sugimoto [56] have shown that in many fields women are proportionally less likely than men to hold coveted first and last author positions. A second potential disadvantage might arise as a result of the hierarchical structure of academia [27, 33]. It is common for professors to collaborate with graduate students or post-docs, tenured professors with non-tenured professors, and the like. In these situations, one member of a collaboration may have significantly more power than other members, and may take advantage of this position.

It has become increasingly clear that human behavior is deeply influenced not just by individual psychology, but by the networks of social interaction that humans engage in [18, 17]. It can be very difficult, however, to empirically observe causal patterns in the dynamic world of social engagement. In particular, for the phenomena we are interested in, dividing the fruits of collaborative labor, empirical evidence suggests that inequities arise, but does not necessarily explain the sources of these inequities. For this reason, we investigate possible causal pathways through which patterns of disadvantage can arise for certain groups in academic collaboration using formal models. We present two sets of evolutionary game theoretic models corresponding to the two types of possible disadvantage described above. The first set of models are from O’Connor and Bruner [44] who examine disadvantages that underrepresented groups may undergo as a result of the dynamics of social interaction when a minority group learns to interacts with a majority group. We also present a set of novel models looking at evolving bargaining scenarios where some actors are more powerful than other actors. We consider three possible types of power inequality and show that in each case the less powerful partners can end up at a bargaining disadvantage. As we will argue, in both sets of models inequities in bargaining over collaboration can arise even in populations of well-intentioned individuals.

The paper will proceed as follows. In section 2, we briefly discuss the formal framework we employ and its relevance for the phenomena under investigation. In section 3 we discuss the work of O’Connor and Bruner [44] and related work. In section 4, we present original results on effects of power in evolving collaborative scenarios. We conclude by discussing the relevance of these results to real world academic communities.

---

2It may be that women in these fields are contributing less than men in collaborative research projects, and so no inequity is occurring. Likewise, one possible explanation for the overrepresentation of men in last author positions could be due to the fact that older generations of academics simply have fewer women. It may be, however, that either male authors are more successfully negotiating for prestigious author positions, or journals are rejecting papers where women are first and last authors disproportionately.
2 Models and Methods

Evolutionary game theory is a formal framework for modeling how the behaviors of groups of actors engaged in strategic interactions change over time. These methods were first developed to model change by natural selection, but were quickly adopted by scientists and social scientists to model change due to cultural learning in humans. Evolutionary game theory is a good framework for investigating the phenomena of interest here because collaboration, and bargaining over collaboration, are strategic scenarios—each player’s outcome is influenced by the behavior of his or her interactive partners. Furthermore, behaviors in these scenarios are learned both from community members and through individual experience with past interactive partners. In other words, these behaviors change over time in systematic ways, making it an appropriate target for evolutionary analysis.

An evolutionary game theoretic model starts, unsurprisingly, with a game. A game is defined by a set of players, a set of strategies, or possible actions for each player, information each player has about the game, and payoffs to each player for every possible combination of strategies chosen. Evolutionary game theoretic models usually assume that a population of actors is playing a game, and then use dynamics to model how strategies in such a population will change over time. The particular dynamics we will employ here are the replicator dynamics. This model of change is extremely general. It assumes that the proportion of a population playing a certain strategy will change as a function of how successful that strategy is for those using it. In particular, a strategy that is more successful than average will proliferate in a population and one that is less successful will be outcompeted.\(^3\)

Notably, the replicator dynamics have been shown to bear deep formal similarities to dynamics for learning by imitation in groups (or cultural evolution) [60], and to dynamics for individual learning [31, 12]. In this paper, we will use the replicator dynamics to represent these types of change. We assume that academics, like other actors, learn to behave in their own best interest through individual experience and by mimicking successful peers.\(^4\)

In order to build an appropriate model of populations of collaborating academics, we first must consider the strategic scenarios that occur when academics collaborate. The first thing to note is that collaboration is a joint endeavor where two actors decide to cooperate in order to reap a reward that is either greater than that for individual work, or is unobtainable as an individual. The second thing to note is that in the process

\[^3\]The discrete time version of the replicator dynamics will be employed here as we present simulation results. The continuous time version of the replicator dynamics can be used to obtain analytic results, but these are difficult to calculate in many cases and unnecessary for our investigation. The discrete time dynamics give updates of a population distribution at discrete steps or generations. They are formulated as

\[
x_i' = x_i \left( \frac{f_i(x)}{\sum_{j=1}^{n} f_j(x) x_j} \right)
\]

where \(x_i\) is the proportion of a population playing strategy \(i\), \(f_i(x)\) is the fitness of type \(i\) in the population state \(x\) and \(\sum_{j=1}^{n} f_j(x) x_j\) is the average population fitness in this state.

\[^4\]This is not to imply that these are the only processes by which academics learn or choose to behave, but research on human learning in games indicates that in general humans, like all animals, can be depended on to repeat actions that benefit them [45].
of collaborative labor, the actors must decide who will be better positioned to reap the fruits of this labor. As discussed in the introduction, author position in academia is used as a proxy for credit by hiring bodies, peers, and academic institutions, often in ways that significantly impact academics. A higher number of first authorships, for example, can lead to perks like promotion, prizes, tenure, and higher salary [25]. Academics in collaboration must determine author order and thus choose which collaborative partners will be most likely to benefit most from the endeavor.

The game we employ as our basic model of collaboration in academia captures the strategic scenario just described. In the collaboration game, two actors first decide whether or not to cooperate, and then, if they choose to cooperate, decide how to divide the fruits of cooperation.\(^5\) The collaboration game can be thought of as an amalgamation of two commonly studied games. Actors first play what is referred to as a stag hunt. The stag hunt is a model of cooperation under risk. It is motivated by the following scenario. Two hunters can either choose to hunt stag or hare. The payoff for successfully hunting hare is lower than the payoff for successfully hunting stag (less meat). But two hunters are needed to take down a stag, while one alone can catch a hare. If one actor hunts stag and her partner does not, she gets nothing. Figure 1 shows a payoff table for this game, which includes payoffs to each player for any possible combination of strategies.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
\textbf{Player 1} & \textbf{Stag} & \textbf{Hare} \\
\hline
\textbf{Stag} & 3,3 & 0,2 \\
\hline
\textbf{Hare} & 2,0 & 2,2 \\
\hline
\end{tabular}
\caption{A payoff table for the stag hunt. Rows list strategies for player 1. Columns list strategies for player 2. Payoff for player 1 is listed first and player 2 second.}
\end{table}

Figure 1: A payoff table for the stag hunt. Rows list strategies for player 1. Columns list strategies for player 2. Payoff for player 1 is listed first and player 2 second.

This first part of the collaboration game models whether academics will collaborate. As pointed out, the payoff for doing solo research is often less than that for collaborative research, and so deciding to cooperate in academia can be thought of as hunting a stag, rather than hunting small game alone. There is some inherent risk to collaborative effort, however. As many academics can attest to, if a co-author fails to perform his or her duties on a joint project, fellow collaborators are left in the lurch. On a more insidious note, a researcher who falsifies data or inadequately supports research claims may put her co-authors at risk of academic dishonesty [54, 26]. These possibilities are captured in the game by the fact that a stag hunter with a hare hunting partner will receive a low payoff.

In the collaboration game, if both agents choose to hunt stag (cooperate), they then decide how to divide the resources gained through this cooperative endeavor (in this

\(^5\)This game is discussed at length by Wagner [59]. It is equivalent to a Nash demand game where players have an outside option, which will be discussed in section 4.
case, to decide author order) by playing a Nash demand game. The Nash demand game is a model of simultaneous bargaining.\(^6\) Two actors want to divide a resource, and each may demand some portion of it (50%, for example, or 10% or 92%). If the demands do not exceed the total resource, both actors receive their demand. If the demands are too high, though, the actors receive payoffs called the disagreement point (often nothing). This corresponds to a scenario in which bargaining fails. Figure 2 shows a payoff table for a Nash demand game with restricted payoffs. In this game actors can demand either 4, 5, or 6 of a total resource of 10 (we refer to these demands as Low, Med, and High). In the rest of the paper, we will model bargaining with a restricted game of this sort.\(^7\) It is tempting to interpret these demands literally when thinking about bargaining in collaborative scenarios. High is first author, Low is second author. Given that levels of effort tend to differ in collaborative work, it is more fruitful here to think of these demands as relative to effort. The High demand can be thought of as a demand for a relatively prestigious author position compared to effort exerted on the paper. The Low demand can be thought of as agreeing to a relatively poor author position compared to effort. The Med demand can be thought of as an attempt to distribute credit equitably.

<table>
<thead>
<tr>
<th>Player 2</th>
<th>Low</th>
<th>Med</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>4,4</td>
<td>4,5</td>
<td>4,6</td>
</tr>
<tr>
<td>Med</td>
<td>5,4</td>
<td>5,5</td>
<td>0,0</td>
</tr>
<tr>
<td>High</td>
<td>6,4</td>
<td>0,0</td>
<td>0,0</td>
</tr>
</tbody>
</table>

Figure 2: A payoff table for the Nash demand game. Rows list strategies for player 1. Columns list strategies for player 2. Payoff for player 1 is listed first and player 2 second.

The full payoff table for one version of the collaboration game is shown in figure 3. As is clear from the figure, the actors may hunt hare, or may hunt stag and make either a low, medium, or high bargaining demand of their opponent. It should be noted that in this game, as in the other two just described, the payoffs themselves are less important than the relationships between payoffs. What unites a set of different payoff tables as one ‘game’ is the strategic scenario these relationships induce. As such, in our analyses of these games, we will sometimes vary payoffs to investigate how these changes influence evolutionary outcomes.

We introduce the collaboration game as a good model of bargaining over collaboration. In some cases, though, the basic Nash demand game may be more appropriate. The collaboration game assumes that players can opt out of collaborative arrangements (by hunting hare), but, as noted, in some areas of academia this will not be possible. In particular, in the lab sciences collaboration between professors and graduate students

\(^6\)This game is so-called because it was introduced by Nash [43, 42]. It is also sometimes called the ‘Nash bargaining game’ or the ‘divide the dollar’ game.

\(^7\)Similar restrictions have been employed in previous evolutionary analyses of the Nash demand game [48, 47, 49, 63, 8, 2, 1].
Figure 3: A payoff table for the collaboration game. Rows model strategies for player 1. Columns model strategies for player 2. Payoff for player 1 is listed first and player 2 second.

<table>
<thead>
<tr>
<th>Player 1</th>
<th>Stag - Low</th>
<th>Stag - Med</th>
<th>Stag - High</th>
<th>Hare</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stag - Low</td>
<td>3,3</td>
<td>3,5</td>
<td>3,7</td>
<td>0,4</td>
</tr>
<tr>
<td>Stag - Med</td>
<td>5,3</td>
<td>5,5</td>
<td>0,0</td>
<td>0,4</td>
</tr>
<tr>
<td>Stag - High</td>
<td>7,3</td>
<td>0,0</td>
<td>0,0</td>
<td>0,4</td>
</tr>
<tr>
<td>Hare</td>
<td>4,0</td>
<td>4,0</td>
<td>4,0</td>
<td>4,4</td>
</tr>
</tbody>
</table>

is essential for all involved. Furthermore, increasing specialization means that collaboration between lab groups is sometimes necessary to complete a project. Throughout the paper, we discuss evolutionary analyses of both sorts of games and the relevance of these analyses to bargaining in academia.

Before moving on, it will be useful to say something about the Nash equilibria of the games just described. A Nash equilibrium for a game is a set of player strategies where neither player can deviate and improve her payoff. These strategy sets are thought of as special because once players arrive at a Nash equilibrium, they will not be incentivized to change. In evolutionary analyses, Nash equilibria are important because populations tend to evolve to them. The Nash demand game has three pure strategy Nash equilibria—Low v High, Med v Med, and High v Low. At any of these outcomes, the entire resource is divided. If either player deviates she either demands less (and so receives less) or demands more and reaches the disagreement point. The version of the collaboration game shown here has two pure strategy Nash equilibria—the strategy sets where both actors play Stag-Med and where both actors play Hare. Altering the payoffs for this game, though, will alter the set of equilibria. If, for example, the Hare payoff is less than the low bargaining demand, the equilibria will be Stag-Low v Stag-High, Stag-Med v Stag-Med, Stag-High v Stag-Low and Hare v Hare. If the payoff for Hare is higher than that for the medium bargaining demand, only Hare v Hare will be an equilibrium.

We have now described the basic games employed throughout the paper. In section 4 we will look at some variations of these games, which will be described, and justified, there.

---

8Graduate students can, of course, switch laboratories. The point is that collaboration itself is unavoidable.

9To be more specific, stable outcomes for the replicator dynamics are always Nash equilibria, but not all Nash equilibria are evolutionarily stable.

10Pure strategies are those where players always choose the same action. Alternatively, game theorists consider mixed strategies, where players probabilistically decide how to behave. We only discuss pure strategy equilibria here as these are more relevant to our evolutionary analyses.
3 Minority Populations and Bargaining

O’Connor and Bruner [44] investigate how underrepresented groups may be disadvantaged in academia as a result of the dynamics of social situations. These authors employ models that represent two interacting populations where members from each population can condition their behavior on the population their interactive partner hails from (henceforth referred to as the ‘type’ of the agent). In academia, such a set-up might represent a field in which both male and female researchers behave differently when interacting with men than with women. There is empirical evidence that real world researchers are prone to such conditional strategic behavior. For example, there is often backlash or anger against women (but not men) who behave assertively in bargaining scenarios [58, 15]. It has been shown that researchers, when presented with otherwise identical male and female academic job candidates, are more likely to hire the male and more likely to offer him a higher salary [41, 53]. Similar bias has been found to be directed at both black and gay job candidates [57, 7]. And the list goes on.

In two-population models, like those analyzed by O’Connor and Bruner, evolutionary outcomes can be significantly different than in one-population models. In other words, allowing agents to condition their behavior on type (male/female, white/black, gay/straight) can significantly influence how groups of people learn to behave in strategic scenarios. In part this is because in a game with this sort of conditioning, a player’s strategy set changes. In the Nash demand game with types, strategies are no longer ‘Med’ or ‘Low’, but are now of the form ‘play Med with in-group members and play Low with out-group members’. For simplicity sake, we will refer to these strategies using an ordered pair like <Med, Low> where the first place represents an in-group strategy and the second an out-group strategy. Significantly, with these new strategies, new equilibria where actors reach a different equilibrium with in- than with out-group members are possible. Consider, for example, a situation where women play <Med, Low> and men play <Med, High>. Here both groups divide resources equitably with in-group members, but when men and women bargain, men take home a greater portion of the resource. Under our interpretation, this would correspond to a scenario where same gender co-authors
tend to negotiate equitable authorship agreements, but between genders men negotiate for more favorable authorship positions. This second type of equilibrium will play an important role in the results we discuss throughout the paper.

Axtell et al. [3] look at simulations of such two-population models of bargaining and find that this type of equilibrium, which they interpret as reflecting ‘discriminatory norms’, arises. In particular, situations like the one just mentioned, where both groups are fair with their own type, but discriminate against the other, commonly occur in their evolving populations. It is worth noting that these outcomes cannot occur in single population models in which conditional strategies cannot be employed. In single populations, fair outcomes, i.e., those where the entire population evolves to demand 5, are most common in evolutionary models [47, 48, 22, 63].

Bruner [16] extends Axtell’s results to show that when one of two groups playing a Nash demand game is in the minority, it is significantly more likely that this group will end up disadvantaged when bargaining with majority types. (They demand Low of the majority, who demand High in response.) In general, the smaller the minority, the greater the chances that the group will be disadvantaged. The reason for this relies on the dynamics of social interaction when a small group interacts with a larger one. Neither minority nor majority types interact with minorities very often as a result of their small numbers. For this reason, minority types are strategically relatively unimportant. For minorities, on the other hand, the majority group is very important. Minorities quickly learn to respond to what majorities are doing. In bargaining scenarios, this often means learning to make small demands, rather than risk reaching the disagreement point again and again. Once minority members learn to make small demands, majority types slowly learn that they can take advantage of this situation by demanding more. O’Connor and Bruner observe that in these types of outcomes, minority members are doubly disadvantaged because they almost always meet majority types for interaction. This means that they are receiving lower payoffs both in cooperative and bargaining scenarios during the majority of their interactions. Note that the smaller the population, the more likely the disadvantaged outcome and the more detrimental this outcome.

O’Connor and Bruner [44] investigate whether similar effects can occur in the collaboration game. These authors present simulations of majority-minority population

---

15These authors do not use the replicator dynamics, but instead consider actors who choose best responses to expected opponent behavior based on some limited set of memories about past opponent behavior. Similar results are found for the replicator dynamics [16]. A more detailed comparison of these two dynamics is outside the scope of this paper.

16Sometimes a portion of the population will evolve to demand Low while another portion demands High, but in this case there is no particular discrimination [47]. Those demanding High do so to everyone, even though they get nothing when paired with others who demand the bulk of the resource.

17See Bruner [16] for a more detailed description of this effect. In some situations, Bruner found that minority populations actually were advantaged by the same effect. This could occur if, for example, it benefited minorities to quickly learn to demand more of majority types because of the particular details of the game they played.

18This may be particularly bad for intersectional groups that are very small (such as, for example, black women in a field where people of color and women are underrepresented). Complicated dynamics may be occurring for such groups, though, making this an area for further research.
models and find that minority populations are more likely than majority populations to end up at a bargaining disadvantage when they collaborate with out-group types. In particular, they find that when the payoff to playing hare (non-collaborative work) is low, collaboration is ubiquitous but minority types are much more likely to arrive at a bargaining disadvantage than majority types. These results, and those from Bruner, indicate that minority groups may be more likely to end up systematically disadvantaged when bargaining in academia, and in particular when bargaining over collaboration.

O’Connor and Bruner also investigate whether these types of inequitable outcomes disincentivize collaboration for minority members. As noted, collaboration can greatly benefit academics, and so if underrepresented groups are disincentivized from collaboration, this could lead to negative career outcomes. Assuming that collaborative efforts will result in a bargaining disadvantage for minority types, O’Connor and Bruner indeed find that minority types are less likely to engage in between-group collaboration in the first place (by instead choosing to hunt hare). This is increasingly likely the worse their bargaining disadvantage.19 Furthermore, this effect is more pronounced when the minority population is smaller. Again, the disadvantage to minority populations at these equilibria are especially problematic because minority types tend to frequently meet majority types. This means that in these simulations minorities are very likely to either collaborate with majority types and end up disadvantaged, or else avoid such collaborations and take the lower hare hunting payoff.

One thing to note about this latter set of models is that they can also be thought of as representing situations where factors outside both collaborators’ control ensure that the advantages that accrue to one type are greater than to those of another. Consider, for example, a case where a white academic and an academic of color collaborate. Now suppose that those evaluating this work assume that the academic of color contributed less than his collaborator (irregardless of author ordering) and, as a result, provide fewer benefits to him (promotions, tenure, salary, prizes, etc.) The results just cited indicate that in cases like this the disadvantaged types will be less likely to engage in collaboration since they are less likely to obtain the usual benefits from such an arrangement.

One might point out that if minority types in an academic community are disadvantaged when collaborating with majority types, they can just collaborate with other minorities. In fact, there are empirical results suggesting both that women are less likely to enter collaborations in many fields [61] and that when they do so, they tend to choose female collaborators [39, 24].20 Del Carmen and Bing [20] show that black criminologists are significantly less likely to collaborate than white ones. Furthermore, it has been observed that underrepresented groups tend to cluster in academic subdisciplines [14]. This may be due to an effect like those noted by O’Connor and Bruner. For small minority groups, though, it will be unrealistic, and potentially harmful, to restrict collaborative endeavors to like types. McDowell and Smith [39], for example, argue that observed

---

19This was modeled by assuming that the demands available to players in the Nash demand game varied. Actors could always choose an equal split (5), but the unequal choices ranged from 5.1 and 4.9 to 7.8 and 2.2. The more unequal the split, the more likely it was that minority groups would avoid collaboration.

20This latter observation is from economics, and is slightly outdated.
productivity gaps between male and female economists might, in part, be due to gender sorting in collaboration. They argue that if females tend to co-author with females only, they lose chances to collaborate, especially in departments with few women.

We will now move on to the second set of models on disadvantage in bargaining over collaboration. In the conclusion of the paper we will further discuss the implications of the models just described, their fit with real world phenomena, and their explanatory power.

4 Power and Bargaining

We present a novel investigation of another type of situation in which certain academic groups might be disadvantaged in collaborative work. Academic communities are arranged in an explicit hierarchy. Full professors outrank associate professors who outrank assistant professors. Professors outrank post-docs who outrank graduate students. Even within these categories, there are unofficial hierarchies where famous or highly successful academics outrank the less famous, professors at research universities outrank those at teaching schools, community colleges, etc., and tenure track professors outrank those in non-tenure track positions. Beyond hierarchical considerations, it is also worth noting that academics from wealthier schools and nations, or with more substantial grants may have unequal relationships with academics from poorer schools and nations and without impressive funding.

It is no surprise that power can lead to bargaining advantages (as we will discuss in more detail below), and authors have discussed how such power imbalances can lead to unethical behavior in academia particularly [54]. In many academic disciplines, explicit norms of authorship have been proposed and adopted to help mitigate this problem [21, 25, 52, 23]. For example, in many STEM fields, it is assumed that PIs (principle investigators) will be the last authors listed on a paper. First author will be reserved for the graduate student or post-doc who did most of the work, and the remaining collaborators will be arrayed in between. This convention is intended to ensure that a graduate student is given proper credit for joint work with a professor. Such conventions do not, however, prevent all types of potential injustice in the division of credit. Even in the lab sciences, there are scenarios where norms of convention (about co-authorship) are silent, such as when two PIs equally collaborate on a paper, or two graduate students share the majority of work. Outside STEM fields, conventions of authorship are often less established, and it is unclear how and whether rank plays a part in determining who will take what author role.

There is evidence that even when clear conventions governing author order exist, there is flexibility in how said order is determined. For example, Nobel laureates have been found to be less likely to be first author on papers after receiving the prize, pre-

\footnote{We do not mean to imply that these hierarchies totally order all academics. For example, the hierarchical relationship between a promising assistant professor at a prestigious school and a less successful associate professor at a less prestigious school may be ambiguous. What matters here is that in many cases these hierarchical relationships do exist and are completely clear to both parties.}
sumably because they feel comfortable giving this privileged position to a more needy colleague [64]. Similarly, senior faculty are less likely to be first author when co-authoring with graduate students than junior faculty [19]. The upshot of all this is that there is reason to explore how certain types of hierarchical privilege can lead to bargaining power in academic collaborations, even though norms governing co-authorship have been established in some cases.

To assess how the dynamics of hierarchical interactions, such as those in academia, can result in such systematic inequalities, we return to our model of bargaining. We explore three ways of capturing how strategic bargaining evolves in scenarios with power imbalances between the interacting parties. Note that ‘power’, as we use it here, does not necessarily refer to a unified concept. Instead, we think there are a cluster of ways that certain academic positions are more ‘powerful’ than others. Here we explore how some of these can translate into an improved bargaining position.

4.1 Different disagreement points

Recall the notion of a disagreement point. We have been assuming until now that both parties in our bargaining interactions have the same disagreement point. This assumption can of course be relaxed, the idea being that if a compromise cannot be reached, one party is left in a better position than her counterpart. This could be the case for a number of reasons in an academic scenario. For example, publishing yet another paper is of little consequence to an established professor (tenure is achieved, reputation is established). For an unestablished researcher, on the other hand, a single paper might make the difference on the job market, or the tenure chopping block. Intuitively, one would expect that when two agents do not have the same disagreement point the individual with the higher one has a bargaining advantage. In other words, the fact that she fares better than her counterpart when bargaining breaks down can be leveraged to ensure she receives a majority of the contested resource. This intuition is born out in both the theoretical and empirical literature. The famous Nash bargaining solution, for example, stipulates that, all else being equal, an increase in one agent’s disagreement point simultaneously increases the amount she should expect when a compromise is made. Likewise, recent work in experimental economics has shown that the agent with the higher disagreement point is at a slight advantage.

Consider the game presented in Figure 4. This is nearly identical to the bargaining game of Figure 2 with the important difference being that Player 1 and Player 2 no longer have the same disagreement point. We assume that \( D > d \). In other words, Player 1 (the established agent) does better than Player 2 (the unestablished agent) when their demands are incompatible.

Assume that \( d < 4 \), i.e., that the unestablished researcher always benefits from collaboration. If \( D > 4 \) then the situation in which Player 1 demands Low and Player 2 demands High is no longer an equilibrium. Similarly, if \( D > 5 \) then the sole pure-strategy

---

22To clarify, if bargaining breaks down the joint project is not completed and both agents are stuck with the unequal status quo.
23See Binmore et al. [11] and Binmore et al. [9].
Figure 4: Payoff tables for a Nash demand game with different disagreement points. Rows model strategies for player 1. Columns model strategies for player 2. Payoff for player 1 is listed first and player 2 second.

<table>
<thead>
<tr>
<th>Player 1</th>
<th>Low</th>
<th>Med</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>4,4</td>
<td>4,5</td>
<td>4,6</td>
</tr>
<tr>
<td>Med</td>
<td>5,4</td>
<td>5,5</td>
<td>D,d</td>
</tr>
<tr>
<td>High</td>
<td>6,4</td>
<td>D,d</td>
<td>D,d</td>
</tr>
</tbody>
</table>

The equilibrium of this game is for Player 1 to demand the High and for Player 2 to concede. So one tentative conclusion we can draw from this is that for certain disagreement points the established agent is only willing to collaborate if she receives the bulk of the benefits produced from said collaboration (a good author position relative to effort). We now go on to detail how such systematic inequalities can naturally emerge for less extreme values of $D$.

Figure 5: Simulation results for the game presented in figure 4. The y-axis shows the proportion of simulations that result in senior academics demanding Low, Medium, and High as $D$ increases (x-axis) and $d = 0$.

We consider a two-population model where the two groups are of equal size. We

All of the simulation results presented in this paper were conducted in MatLab 2012. For each set
once again assume that individuals interact with both types of agents and the strategies employed by established and unestablished types change as determined by the replicator dynamics. Figure 5 details our results. In particular we track the proportion of simulations that result in one of three strategy pairings between populations (High v Low, Med v Med, and Low v High) as the disagreement point of the established population increases. When \( d = D = 0 \) it is just as likely that a norm favoring the established emerges as it is for a norm favoring the unestablished. Yet as the disagreement point of the established population increases the proportion of simulations that head to the equilibrium in which the established demands High increases. Not surprisingly, as \( D \) nears 5 it becomes overwhelmingly likely that the established population will demand the high amount when interacting with the unestablished population. The number of equal splits plummet at this point because achieving the equal split no longer significantly benefits the established player. These results indicate that if established academics suffer less than unestablished research partners when collaborative efforts break down, they may learn to demand and receive better author positions.

It is worth noting that a different interpretation is available for the levels of the disagreement points in these models. Nash [42] argued that the value of disagreement points could be established not only through exogenous forces, but through threats made during pre-play discussion in bargaining scenarios. In other words, one actor could threaten to mete out punishment to the other and thus lower the other’s payoff should bargaining fail. These ‘threatpoints’ fell out of favor because other theorists challenged the ability of most bargainers to make plausible threats of this sort. In cases where one actor is clearly more powerful than the other, however, it is plausible that credible threats could be made. Academia is such a case. An established academic need not even explicitly threaten graduate students, post-docs, or early career colleagues with some sort of punishment (poor reviews, failure to mentor, refusal to write letters of recommendation, etc.) in the case of bargaining breakdown in order for such an implicit threat to exist. A graduate student collaborating with a powerful academic may be extremely conscious of the fact that should she anger this powerful colleague things will likely go poorly for her. Such threats (or perceived threats) can induce a payoff table like the one in figure 4, even if there are no exogenous factors creating it.

4.2 Different background payoffs

Another way of modeling the hierarchical structure of academia is to assume that established academics have multiple projects they are working on simultaneously, while those new to the profession are banking on successfully completing a small handful of papers (some of which may be co-authored). What this means is that when two individuals come to negotiate the terms of a collaborative endeavor, the more senior academic will have a higher ‘background payoff’. Furthermore, for the less experienced academic, collaboration with his more senior colleague could constitute a large chunk of his early career research. Figure 6 presents a game which captures this scenario. \( B \) and \( b \) refer
to the background payoff of established and unestablished researchers, respectively. In
other words, these are the payoffs both agents get regardless of whether or not collabora-
tion is successful. We assume that any benefits derived from collaboration will be in
addition to these background payoffs. Finally, we make the reasonable assumption that
$B > b$.

<table>
<thead>
<tr>
<th>Player 1</th>
<th>Low</th>
<th>Med</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>4+B, 4+b</td>
<td>4+B, 5+b</td>
<td>4+B, 6+b</td>
</tr>
<tr>
<td>Med</td>
<td>5+B, 4+b</td>
<td>5+B, 5+b</td>
<td>B, b</td>
</tr>
<tr>
<td>High</td>
<td>6+B, 4+b</td>
<td>B, b</td>
<td>B, b</td>
</tr>
</tbody>
</table>

Figure 6: Payoff tables for a Nash demand game with different background payoffs.
Rows model strategies for player 1. Columns model strategies for player 2. Payoff for
player 1 is listed first and player 2 second.

One may wonder whether this is just a different way of representing the game from 4.1
(fig 4). This is not the case. In the game just presented both parties stand to gain equally
(in absolute terms) from successful collaboration. In the previous game, the junior
researcher benefits more from collaboration than the senior researcher. Furthermore, for
this new game, the senior researcher stands to gain from collaboration for all parameter
values. Recall that this was not the case in section 4.1. If $D > 6$, the senior researcher
would actually receive lower payoffs for any collaborative outcome than for a failure to
collaborate. For this reason, one might expect that this new scenario would be less likely
to yield unequal outcomes, as both researchers always care about collaborating and all
collaboration outcomes are Nash equilibria for all parameter values.

Turning to the replicator dynamics we can again track how likely a norm favoring
senior academics is to emerge. Figure 7 presents simulation results. It is clear that
the magnitude of the background payoff significantly effects outcomes. The larger the
difference in background payoffs, the more likely the established type is to end up de-
manding High.\textsuperscript{25} The explanation for this is similar to that provided in Bruner [16]
and O’Connor and Bruner [44]. As described, these authors found that in majority-
minority bargaining scenarios the larger group often ends up demanding High of the
minority. This is due, as discussed, to an asymmetry in the learning environments of
the two groups.\textsuperscript{26} In other words, one population (the minority population) is much
more sensitive to strategic behavior of the other population (the majority group). This

\textsuperscript{25}The proportion of simulations that result in the senior academics demanding High continues to
increase as their background payoff, $B$, increases. Eventually, however, this proportion of simulations
levels off at around 60. For $B = 200$ 57.6 percent of simulations resulted in established agents demanding
high, and this percentage increases to 58.5 percent when $B = 500$.

\textsuperscript{26}Relatedly, see Bergstrom and Lachmann [6] who argue that similar effects occur in biology when two
species are interacting and one evolves at a slower rate than the other (this is the so-called Red King
effect).
means the minority quickly adapts to the behavior of the majority while the majority is by and large insensitive to the minority population’s behavior (see the previous section for more on this). In such cases, the majority group can be thought of as having more inertia than the minority. An asymmetry of inertia is also present when the two-populations have different background payoffs. Recall that strategies which do better than the population average grow in size under the replicator dynamics. Furthermore, strategies that do significantly better than the average grow rapidly, while those that only do slightly better grow at a slower rate. Thus background payoffs play a large role in determining the speed at which change occurs. If the background payoff is small then slight payoff differences between strategies are magnified and the population undergoes change rather quickly. Alternatively, when the background payoff is large, small payoff differences between strategies result in only minimal change from generation to generation. Established academics are slower to adapt, which in turn bestows upon them a bargaining advantage. These results indicate that if certain collaborative efforts are relatively unimportant for established academics, as compared to unestablished academics, systematic bargaining disadvantages may develop for early career researchers.

![Simulations of the Nash Demand Game with Different Background Payoffs](image)

Figure 7: Simulation results for the game presented in figure 6. The y-axis represents the proportion of simulations that result in senior academics demanding Low, Medium, and High as $B$ increases (x-axis).
4.3 Different outside options

We consider one final way in which an established academic may have bargaining power over an unestablished collaborative partner. Suppose that either party may choose to opt out of collaborating before bargaining begins and choose to engage in individual work instead. Formally, such a scenario is identical to a collaboration game (where players can either choose to bargain or to hunt hare). It may be, in such a case, that one player’s outside option (hare payoff) is significantly better than the other’s, and that this inequity can lead to unequal bargaining positions. If one player has a good enough outside option, cooperative scenarios where she does not receive a significant portion of the resource are no longer Nash equilibria, meaning that she will be disincentivized from accepting a low bargaining demand (or, in our interpretation, a poor author position). It is likely, for example, that an established academic can expect more significant payoffs from individual work—acceptance to respected journals, citations, invitations to present, etc.—than an unestablished partner. Alternatively, it could be the case that when a more experienced researcher collaborates with an unestablished academic the senior researcher must exert effort to mentor her counterpart. This is an extra cost that the established agent takes on, and thus the net benefit of collaboration is smaller for her. It is also an added benefit for the unestablished researcher making her net benefit to collaboration higher. We model these scenarios with a collaboration game where the payoff to hunting hare is greater for the established academic. The payoff table for this game is shown in figure 7.

<table>
<thead>
<tr>
<th>Player 1</th>
<th>Stag - Low</th>
<th>Stag - Med</th>
<th>Stag - High</th>
<th>Hare</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stag - Low</td>
<td>4,4</td>
<td>4,5</td>
<td>4,6</td>
<td>0,h</td>
</tr>
<tr>
<td>Stag - Med</td>
<td>5,4</td>
<td>5,5</td>
<td>0,0</td>
<td>0,h</td>
</tr>
<tr>
<td>Stag - High</td>
<td>6,4</td>
<td>0,0</td>
<td>0,0</td>
<td>0,h</td>
</tr>
<tr>
<td>Hare</td>
<td>H,0</td>
<td>H,0</td>
<td>H,0</td>
<td>H,h</td>
</tr>
</tbody>
</table>

Figure 8: Payoff table for the collaboration game with unequal payoffs for Hare. Rows model strategies for player 1. Columns model strategies for player 2. Payoff for player 1 is listed first and player 2 second.

There are some similarities between this game and the one presented in section 4.1. In the bargaining game with unequal disagreement points, if the disagreement point for the established researcher is high, she will be disincentivized to collaborate unless she receives a significant portion of the resources generated. In this scenario, likewise, if \( H > 4 \) the outcome where the established researcher receives the low bargaining demand will no longer be an equilibrium. If \( H > 5 \), the only collaborative outcome that remains an equilibrium is the one where the established researcher demands 6. (Of

\[27\] For an analysis of equilibrium selection in these types of games see Harsanyi and Selten [29].
course, if \( H > 6 \), there will be no collaborative equilibria.) Unsurprisingly, results of our evolutionary simulations are consistent with this (see figure 9). As \( H \) increases, outcomes where the established academics demand Low disappear, followed by Med, and then High. At the same time, the probability that the established academic chooses to go it alone instead of collaborating increases and is inevitable once \( H > 6 \).

Perhaps surprisingly, when both outside options are low, even if one is significantly higher than the other, the population with the lower outside option is not disadvantaged in the subsequent bargaining. As is seen in figure 9, the established academic is just as likely to demand the majority of the resource as her unestablished peer while \( H \) is low. This can be explained in the following way. In this game, if players decide to collaborate, the structure is the same as a Nash demand game with identical disagreement points. In other words, if the hare payoff does not deter collaboration, it becomes irrelevant to the bargaining interaction. To give a clearer understanding of the evolutionary process occurring, the evolution of hare hunters can be thought of as separate, in a way, from the evolution of the other players. The hare hunters do not bargain with other players and so influence the payoffs of other types equally. (I.e., a hare hunter will equally disadvantage the Low, Med, and High players when she meets them.) Hare players will either disappear if the other players outperform them, or take over the population if they outperform the rest. Whichever of these they do, they influence the other three types equally and so do not significantly alter the evolutionary process happening between these other types.

These last results indicate that if senior academics reap greater benefits than junior colleagues from individual labor, they may learn to only collaborate in cases where they are well compensated for their efforts, i.e., only cases where they receive prestigious author positions.

5 Conclusion

We will start with separate discussions of the two sets of models presented, and what they tell us about the target phenomena. We will then discuss the explanatory force of the type of formal methods we use here to explore bargaining disadvantage in collaboration.

The results from O’Connor and Bruner [44] indicate that social dynamical forces may lead to disadvantage for minority groups in collaboration and in bargaining over collaboration. It is worth noting that the results cited need not be the effect of any sort of pernicious behavior or intent on the parts of the involved parties. They arise on three conditions: 1) that individuals condition behavior based on type of interactive partner, 2) that one type of interactive partner is in the minority, and 3) that individuals evolve their behavior in ways that benefit themselves. Arguably the only morally questionable acts among these involve conditioning behavior on the type of partner, though even this is not universally thought to be morally wrong. Of course, one could also argue that individuals always have a moral obligation to protect colleagues from their own self-promotion, but the point is that actors in these models need not be trying to disadvantage underrepresented groups or even be aware that they are.
The set of models on power in bargaining over collaboration, likewise, indicate that social dynamical effects can lead to systematic patterns of bargaining that disadvantage unestablished academics such as graduate students, post-docs, and early career researchers. These effects should not be terribly surprising given that, as mentioned, previous research has indicated that in non-evolutionary scenarios, i.e., one time interactions where rational actors make decisions, powerful positions can lead to bargaining advantages. Once again, though, it is significant to note that in our results, actors need not actively decide to pursue their own interest at the expense of an interacting partner. Instead, processes of learning through experience and from peers can lead to these inequitable outcomes. For these sorts of processes to occur there are now only two necessary preconditions 1) actors condition strategy on interactive partner and 2) actors learn to behave in their own best interest. In this case, arguably, neither of these behaviors is ethically problematic. It is impossible to avoid conditioning behavior on hierarchical positions in academia. Indeed, this sort of conditioning is expected in almost every sort of interaction across the explicit hierarchies of academia. And, obviously, learning to do self-benefical things is as natural to humans, and indeed almost every type of animal, as breathing. Once again, the effects we describe can be seen when actors are not behaving
in ways that would usually be considered objectionable, and may not even be aware of
the systematic disadvantages occurring in their fields. Of course, this is not to imply
that the resulting behaviors in both sets of models are unproblematic. Rather our mod-
els show that from a starting point without bias, populations can arrive at behaviors
consistent with bias and discrimination simply by slowly learning to benefit themselves.

The results of both sets of models fit with certain empirical observations. In regards
to underrepresented groups, Wilkie and Allen [62] find that women sociologists tend to be
credited as equal collaborators when co-authoring with women (and men with men) but
that in male-female collaborations men are more likely to be credited as the principal
contributors. Furthermore, there are indications that women may be disincentivized
from collaborating by disadvantages of these sorts [61, 39, 24, 13]. There is also an
extensive literature on differences in bargaining behaviors and outcomes for men and
women in a wider context. The results in this literature are sometimes contradictory, but
it is agreed that women tend to secure less favorable bargaining outcomes in situations
such as salary negotiations in part due to gender conditional behaviors by men and
women in these scenarios [55, 15].

Empirical work also shows that ethical concerns over
co-authorship with senior colleagues do arise in academic communities [27, 19, 26, 5, 21].
In some areas, there has been increasing concern over senior faculty who pressure junior
faculty for undeserved authorship positions [5, 34]. Such practices have been argued to
damage junior faculty who then receive less credit from hiring and promotion bodies
[26, 54].

How well do the models we present capture the phenomena of interest? Obviously
evolutionary game theoretic models are extremely simple, and so can only be thought
of as idealized representations of complex processes such as those occurring in the social
sphere. In these particular models, for example, actors choose strategies concurrently. In
reality complete breakdowns of bargaining are presumably rare because actors can engage
in an ongoing discussion of the situation at hand. Another simplification that bears
mentioning is the assumption that academics learn from peers and through experience
to do whatever benefits them. Human behaviors are, of course, influenced by many
complex factors, including overt ethical obligations, and our models do not account for
these. Despite simplification, however, these sorts of models provide insight into causal
processes in these very complex situations that is difficult to obtain otherwise. As just
discussed, they allow us, for example, to list preconditions for these effects to occur that
would not have been obvious without these models.

There are several important things to be said about the results just described which
can mitigate worries about the explanatory power of these models. First, these results
certainly provide evidence that such social dynamical effects are possible. At very least,
they ought to give reason for academics to be on the look out for the occurrence of these
sorts of bargaining disadvantages even in groups of well-intentioned peers. Second, both

\[28\] One might point out that bargaining inequities occurring between groups in academia are likely
related to larger societal phenomena such as these. We agree. We have neither the time nor the
expertise to try to analyze how patterns of bargaining in larger society relate to patterns of bargaining
in academia, but it should be noted that our results can be applied to bargaining in a larger community
as well as to bargaining in academia.
sets of results are very robust. In other words, they are not relics of a particular model, but correspond to deep mathematical regularities in the sorts of set-ups we examine. Two-population outcomes arise for all sorts of games, in all sorts of parameter setting. We have already explained why. Two-population models have entire sets of new Nash equilibria that become evolutionarily relevant. As long as groups of interacting partners are conditioning behavior on types, these effects can and will arise. The minority effects observed in O’Connor and Bruner [44] are furthermore extremely robust. Robust results, like these, are usually thought to be more likely to provide actual (instead of just possible) explanations of real world phenomena as the chances that the phenomena will fall under the set of conditions under which these outcomes occur are higher.

Third, the aspects of these mathematical models that lead to disadvantaged outcomes are easy to understand, and easy to verify empirically in the real world. It is clear that actors condition behavior on interactive partners. It is clear that real world actors learn strategic behavior from peers and from past experience. And it is the case that in many academic fields women and minorities are indeed underrepresented, and that established academics have certain types of power when compared to unestablished academics. These observations increase confidence that the effects seen in these models could actually be occurring in academic communities.

Now it seems appropriate to ask: Given these results, what should be done? Are there policy recommendations that these results support? Perhaps the most important take-away from our work is one we have already emphasized. As we have seen, norms which disadvantage the vulnerable are likely to naturally emerge in the absence of formal or explicit rules regarding how credit should be allocated among co-authors. For this reason, we follow authors such as Erlen et al. [23], Fine and Kurdek [25], and Strange [54] in recommending explicit communal policies be adopted governing co-authorship. Such policies will hopefully decrease chances of inequity in determining authorship. This, as argued, will not be a universal solution, as there will always be ambiguous or unusual cases where it is not clear how the collaboration fits with discipline standards. But we believe such measures are better than nothing. Importantly, work by West et al. [61] indicates that some norms are better than others. These authors find that women are (or have been) underrepresented in first and last author positions in a number of sciences. In mathematics, however, where authorship is alphabetic, this effect is not observed.29 (Of course, it is not clear that this is an entirely satisfactory practice in cases, for example, where one author did significantly more work.) This is particularly true because different practices across disciplines may lead to confusion in assigning credit.

More generally we suggest that our results should help provide fodder for what ought to be an ongoing discussion in academia—how to ensure fairness in collaborative work. If academics are aware of the ways that inequities can arise as a result of the dynamics of social interaction, either between minority and majority groups or between groups in hierarchies, they will hopefully shape discussions of policy matters related to collaboration appropriately.

29 Other conventions such as, for example, randomly determining authorship would similarly prevent such inequities.
References


[44] Cailin O’Connor and Justin Bruner. Modeling the underrepresentation of women in academia.


