Priority and privilege in scientific discovery

Mike D. Schneider and Hannah Rubin

Abstract

The priority rule in science has been interpreted as a behavior regulator for the scientific community, which benefits society by adequately structuring the distribution of intellectual labor across pre-existing research programs. Further, it has been lauded as part of society's "grand reward scheme" because it fairly rewards people for the benefits they produce. But considerations about how news of scientific developments spreads throughout a scientific community at large suggest that the priority rule is something else entirely, which can disadvantage historically underrepresented or otherwise marginalized social groups.

1 Introduction

In scientific practice, discoveries of sufficient impact appear to generate certain quantities of prestige, which are to be bestowed upon particular scientists by the scientific community at large. The priority rule is a broad descriptive norm concerning the proper allocation of that prestige. The inference toward such a norm comes via an observed phenomenon in the history of science whereby, in situations of *multiple discovery*, disputes about who deserves the prestige that comes associated with a discovery are often fought by way of assertions about who was first to make it.

That the history of science is laden with instances of multiple discovery has been known since at least the early 1900s. Ogburn and Thomas [1922] document 148 such cases in a variety of high-profile domains of scientific inquiry, and mention in a footnote there there are disputes over priority of discovery in many cases. (In one example, the authors report that the discovery of the cellular basis of animal and vegetable tissue was claimed by seven independent researchers in or around the year 1839.) The particular investigation of these so-called "priority disputes" was eventually taken up (famously) by Merton [1957]. Merton sought to explain, by appeal to the institutions and norms of science, how otherwise dignified and reserved scientists would, when involved in instances of multiple discovery, often fight tooth and nail against the possibility that anyone but them bore responsibility for the discovery.

What Merton [1957] observes is that in cases of multiple discovery, the prestige that results from the discovery is generally awarded to just one of the independent parties involved. This state of affairs naturally engenders disputes over which party ought to enjoy access to that new prestige. In the event of such a dispute, Merton argues convincingly that the rhetoric of the dispute often centers on the lack of sufficient time-stamped evidence about which of a set of claimants was first to make the key insights that drove the content of the multiple discovery, as well as accusations about whether those key insights were indeed made independently of knowledge of the work done by any of the other claimants.

Merton makes sense of this state of affairs in terms of the perception amongst scientists of intellectual property rights, for which it is the case that the first person responsible for the invention of a new bit of intellectual property ought to enjoy the profit that comes from it. This feature of the account— that the prestige that is generated by a particularly impactful discovery is taken to belong to whoever is truly first to produce it— is what has come to be known as the "priority rule" in science.

Strevens [2003], following a lead from Kitcher [1990], famously formalizes the priority rule in science as a particular manifestation of a "grand reward scheme" (p. 76) that exists in our society, which allocates prestige associated with new discoveries in such a way that incentivizes scientists to optimally distribute themselves across a variety of programs of research. Moreover, he argues that this epistemic role of the priority rule provides extrinsic justification for the continued existence of the rule. In addition to this epistemic justification, Strevens argues that the priority rule rewards people based on the benefit they confer – since only the first to make the discovery produces a benefit to society, it is regarded as fair that only the first discoverer is rewarded.

However, in Strevens's account, it is assumed that whoever makes the discovery first will be assigned priority by the scientific community (or, at least, scientists believe that this is how priority assignments work). While it may often be clear who is ultimately responsible for a particular discovery, the history of science is punctuated by the (perceived) failure of the community to adequately implement a priority rule like that which Strevens formulates. Merton's original evidential impetus for talking about priority was that it is priority *disputes* which are commonplace in the history of science and which are fought so animatedly, but these disputes are idealized away in Streven's priority rule. In light of this substantial gap between Strevens's idealization of scientific practice and Merton's original evidence for the priority rule, the first goal of the present paper is to offer a different formulation of the priority rule than that which Strevens provides, based on taking seriously considerations about the social mechanisms by which credit is assigned by individuals in the scientific community who may not always agree on who should be given priority.

Having argued for the appropriateness our new formulation of the priority rule as descriptive of scientific practice, the second goal of the present paper is to evaluate the consequences of scientific communities dispersing prestige. We introduce a model in which inequities in the underlying social network of the scientists in a community can allow prestige to accumulate in the hands of those historically well-positioned within the scientific community. We show how historically underrepresented or otherwise marginalized minority groups can suffer in the context of receiving prestige for particular discoveries. Notably, this disadvantage arises due to facts about the social structure of scientific communities, rather than due to any differences in skill or achievement, or any bias against the minority population.[1]

2 From priority disputes to the priority rule and back

The operational notion of credit in priority disputes is not as immediate payment for services rendered. What is at stake is the accumulation of prestige, or wide-scale credit, in the eyes of the community (considered as a single body). Merton [1957] lists several examples of how the community confers prestige to individuals, including eponymy (p. 643), honorifics such as the Nobel Prize, introduction into "honorary academies" like the Royal Society and the French Academy of Sciences (p. 644), and posthumous recognition by historians (p. 645). To this list of examples, Strevens [2003] adds "reputation, a sizable office, the rapt attention of graduate students and the like" (p. 57). Based on these examples then, prestige appears to be a retrospective quantity, conferred on individuals when the community as a whole has come to associate them with the corresponding discoveries.

A key observation is that in most cases of disputes about who ought to receive the prestige that corresponds to a particular discovery, at least one of the parties involved perceives there to be a great injustice afoot: they are being denied access to newfound prestige because another party involved is, for whatever reason, unrelated to the content of the discovery, in a better position to win the prestige instead. As Merton argues, in cases of multiple discovery, a particular kind of dispute can emerge that is focused on the precise times of the scientists' respective discoveries. As Merton [1957], quoting Arago, points out: " 'about the same time' proves nothing; questions as to priority may depend on weeks, on days, on hours, on minutes" (p. 658). Merton proceeds to explain how such priority disputes emerge as unsurprising artifacts of the institutions and norms that govern scientific practice. And so is born the notion of a 'priority rule' in science.

Strevens takes Merton's work as motivation to formalize the priority rule in such a way that he can assess its impact on scientific inquiry. In particular, he models the priority rule as a reward system characterized by two "parts" (p. 56):

> First, rewards to scientists are allocated solely on the basis of actual achievement, rather than, for example, on the basis of effort or talent invested. Second, no discovery of a fact or a procedure but the first counts as an actual achievement.

Strevens then argues that, from the perspective of a central planner, although the priority rule may seem harmful initially (as compared to, e.g., a rule that

---

[1] For other cases where minority disadvantage can arise in absense of biases, see e.g. Bruner [2019], and Rubin and O'Connor [2018].

rewards scientists for hard work and talent), it is actually beneficial for scientific inquiry. In his formal setup, it is by virtue of the priority rule that intellectual labor is efficiently distributed across various research programs with differing odds of success. The basic idea is that scientists balance the odds that a research program will be successful and the odds that, if their research program is successful, they will be the one to make the discovery. Therefore, scientists do not all abandon other lines of research to join the most promising-looking research program. Instead, they distribute themselves among research programs (with more scientists working within more promising-looking programs). Others have since used this same basic modeling framework to discuss, for example, incentives to publish early and frequently [Heesen, 2018] or share intermediate results [Heesen, 2017], disincentives to replicate previous findings [Romero, 2017], and how scientists are motivated by a combination of truth and credit [Zollman, 2018].

It seems odd that in this formulation of the priority rule, Strevens excludes the possibility of genuine conflict in cases of multiple discovery. After all, it was a litany of such conflicts in Merton's article that gives rise to an inference toward the priority rule in the first place. This gap between Merton's treatment of priority disputes and Strevens's account of the priority rule is highlighted by the shift in Strevens's article away from Merton's historical cases of priority disputes and toward scenarios in which different research groups approaching scientific problems are in a winner-takes-all race toward the resolution of those problems. Unlike in Merton's work, "discoveries" are now the resolutions of those problems and prestige is doled out to whoever wins the race.

While such scenarios certainly occur in science, they are not the scenarios in which Merton's work about priority is most relevant. Where is the disconnect? One answer can be found in an ambiguity in Strevens's initial framing of the matter (p. 58): "The question with which I am concerned, then, may be put thus: why does the scientific community disburse prestige in accordance with the priority rule rather than... some alternative scheme?" Nestled implicitly in the statement of this question is the descriptive plural action claim:

> The scientific community disburses prestige.

Following the terms provided by Ludwig [2016], this claim can either be read as a distributive action sentence or as a collective action sentence. Moreover (following the form of Ludwig's argument, p. 131), we may understand the ambiguity between these two readings to wholly consist in an ambiguity of scope:

> On the distributive reading, we mean that each of [the scientists within the community] were separately sole agents of [the disbursal of prestige] in a certain way. On the collective reading we mean that for [the disbursal of prestige] each of [the scientists within the community] (and no one else) was an agent of it in a certain way.

As an example, we may read the sentence "Two people built boats" as claiming either that two people independently built boats (distributive reading) or that two people came together to build boats, e.g., with one sawing and the other hammering pieces together (collective reading). Similarly, scientists could separately take part in the community's disbursal of prestige by each independently engaging in a prestige-relevant task— e.g. individually attributing credit (distributive reading), or they could come together as a whole to disburse the prestige as a group (collective reading). Strevens seems to take for granted the collective reading of this plural action claim when he idealizes away the matter of how that quantity of prestige comes to be disbursed by the community, considered as a singular body.

Meanwhile, the distributive reading suggests a rich story to be hypothesized about the internal social dynamics that go into the disbursal of that quantity. It is therefore only on the distributive reading that we are subsequently free to study how the implementation of a priority rule in the case of multiple discovery can give rise to priority disputes. A formulation of the priority rule that better coheres with Merton's work, i.e. which gives rise to morally-charged priority disputes in cases of multiple discovery, should involve considerations about precisely those social mechanisms that are implied by the distributive reading that Strevens avoids.

One fears that whatever Strevens's formulation of the priority rule does to distribute intellectual labor theoretically across a variety of competing research programs, it is simply not the priority rule for which we have thorough evidence is present throughout the history of science. The history of science is replete with individual mis-attributions of credit that result in the "wrong" person winding up with prestige. In the case of eponymy, Stigler's Law, autologically coined by Stigler [1980] himself, is the descriptive rule that (humorously) captures this phenomenon: no scientific discovery is named after its original discoverer (p. 147). In section 3, we argue that the social structure of science plays a role in determining who individuals assign credit to, and thus who receives the prestige. But first, we provide evidence for a distributive reading of the priority rule.

2.1 Evidence for the distributive reading

The scientific community, though in some respects quite hierarchical, is not centrally governed. The awarding of prestige, though associated with individual accomplishments, often happens at no particular moment in time. That individuals come to be awarded prestige at no particular time by no particular decision made on behalf of the community strongly suggests that the disbursal of that prestige is not a single act of which each scientist was an agent. Instead, it seems more plausible to view the disbursal of that prestige as following from the members of the community separately taking actions that need not individually resemble the awarding of prestige. That is, individual scientists assign credit, and along the way prestige is conferred by the scientific community. (No one gains the respect of a community because one person gives them credit for a discovery, but if the members of a community all generally associate a particular scientist with a discovery, that scientist will receive the associated reputational benefits, etc.)

To support this alternative view, it is helpful to recall cases from the history of science in which the prestige disbursed by the community as a whole follows as a consequence of a sufficient number of individuals each engaging in practices other than prestige-disbursing. The three cases we will offer in this regard are drawn from the (computational) social sciences, evolutionary biology, and quantum mechanics. In discussing each of them, our goal will be to emphasize just one feature common to all three: *individuals attributing credit as they go about their ordinary affairs are what, in aggregate, determines to whom prestige comes to be awarded.*

James M. Sakoda beat Thomas C. Schelling to the public invention of computational models of segregation: the latter's model amounts to a special case of one of the former's, which was printed in the previous issue of the same journal in which the latter's model appeared. Moreover, Sakoda's model had its origins in his dissertation work twenty years earlier, furthering his claim to priority on the subject [Hegselmann, 2017]. Nonetheless, Schelling undoubtedly enjoys the prestige that surrounds the subject: the basic checkerboard model of segregation is equivalently called the "Schelling model" (an instance of eponymy), the so-named Schelling model is taught in any introductory course on the subject of agent-based computational models in the social sciences, and Schelling won a Nobel prize in economics in part for his work on such models (a prime example of an honorific).

But as Hegselmann [2017, p. 5-6] argues, "No crime happened, no conspiracy was involved. No discrimination whatsoever was at work.... as to the main actors, nobody did anything wrong." Instead, one deciding factor it seems was Schelling's subsequent decision to write a book, developing many of the ideas in his paper, accessible for much broader audiences than just those computer scientists who happen to additionally be interested in modeling social dynamics. Those broader audiences were encouraged to try out small, table-top examples of the checkerboard models under scrutiny, whereupon: "They all had experienced how surprisingly fast, right before their eyes, certain unexpected, dramatic macro structures evolved, generated by fairly innocent looking micro-motives–an eye-opening phenomenon par excellence" [Hegselmann, 2017, p. 87]. This got those broader audiences talking about the demonstrative power of these simple computational models, and they were talking about it in the context of Schelling's work. Whatever reasons we give for why Schelling enjoys the prestige, an indispensable part of the story is that he enjoys it because individuals separately began to associate him with the discovery, rather than because the scientific community wholesale decided that it was he who was responsible for the discovery.

One finds a situation of a similar form in evolutionary biology, wherein the distinction between proximate and ultimate causes in biology is attributed to Ernst Mayr. As Laland et al. [2011] point out, this distinction was made well before Mayr wrote about it in the 1960s (they cite an article by J. Baker written in the 1930s). Nonetheless, Mayr's article is what ostensibly led to the distinction's widespread acceptance in evolutionary theory. Moreover, even as Laland et al. [2011] flag the trivia that Mayr was not the first to make the distinction, they consistently refer to it as “Mayr’s distinction”. This is illuminating because, as an evidently conscious decision of the authors in that article, it demands explanation. One prudent explanation is that of historical usage: since people started citing Mayr when talking about the distinction, it subsequently became known as his distinction. Hence, even when the pre-history of Mayr’s work is acknowledged, Mayr’s legacy continues to enjoy the prestige.

A final case worth mentioning is one in the history of physics, particularly in the development of mature quantum mechanics. In 1932, von Neumann published an alleged “no-go” proof of the viability of hidden-variables underpinning quantum mechanical behavior in an ultimately deterministic theory. Grete Hermann evidently discovered a flaw in the scope of the proof in 1935, yet this discovery was “not widely known at the time, and her criticism had no impact whatsoever” [Seevinck, 2016, p. 107]. In 1964, John Bell happened on the same such discovery, in the aftermath of the development of Bohmian mechanics (whose success as a deterministic, hidden-variables alternative to quantum mechanics clearly stood as proof of the alleged impossible).

There are a few details to note in this case: Hermann, though formally trained as mathematician under Emmy Noether and in contact with the physicists of note in the day, nonetheless identified primarily as a philosopher in the neo-Kantian tradition [Hansen-Schaberg, 2016]. She was also a woman (as well as political activist), working in a time and discipline that was hostile and exclusionary. Finally, Bell enjoyed the benefit of offering Bohmian mechanics as a demonstration of his point (whereas in 1935, Hermann could only conclude that the possibility of such a hidden-variables theory was left open). For some combination of these reasons, and perhaps others, Hermann’s contributions at the time were overlooked. Our point here is that following this neglect, when most scientists were for the first time ready to credit *someone* for identifying the ostensible flaw in von Neumann’s proof (that is, following the development of Bohmian mechanics), Bell’s independent study of the subject led him to be the recipient of that credit, and, eventually following, prestige.

In each of these historical anecdotes, there are myriad reasons one could give for why a particular person enjoys the prestige. The common feature across each of these histories is that the recipient of the prestige associated with a discovery of some or other form is not the same individual as who was first responsible for the relevant discovery, but rather was the same individual as who first became largely known to be associated with the discovery. We take this as evidence from the history of science for a distributive reading of the claim “the scientific community disburses prestige”: in each of these cases, each of the scientists within the relevant community happened to attach credit for the relevant discovery to some or other scientist. Meanwhile, by virtue of each such scientist acting in this way separately, i.e. as a sole agent, prestige came eventually to be disbursed by the community to the individual that most of the scientists credited.

2.2 Credit attribution contests

Taking a lead from the historical cases just discussed and the distributive reading of the disbursal of prestige that those cases support, we offer the following perspective. Scientific developments are not immediately known by everyone in a scientific community. News of them spreads throughout the community, rather, through an informal social network that spans the community. Insofar as that development proves valuable to the community, the party responsible for that development enjoys a corresponding quantity of prestige. Of course, it need not be the case that *every* individual in the community has assigned credit; the community may just as well bestow prestige when the vast majority has learned about the development (whereupon any stragglers learn about the development as do those outside of the community, as discussed below).

In cases of multiple discovery, just as in any other case, when individuals within the community go looking for news of pertinent developments, they generally come to associate a development with whomever they first learn is responsible for it. The difference is that in the case of multiple discovery, a large majority of the community may all come to associate the development with one of the parties involved, such that prestige comes to be bestowed upon that party (as it is normally) while any other party that independently produced the same development is neglected.

At risk of being denied the prestige that they believe they are due on the basis of their work, any party that does not enjoy the support of the majority will protest that they ought to receive the prestige that in other circumstances would have been awarded to them, for instance had the social network just happened to have been structured differently. As Merton [1957] notes, those in the minority who believe the wrong person received the prestige will often protest, too. In other words, the function of such a protest is obvious: to change individuals' associations as to who gets credit for a discovery. This is because the conferring of prestige by the community occurs when individuals within the community associate a particular person with a discovery.

Finally, we have arrived at a priority rule that falls directly out of the nature of priority disputes, such that the priority rule is what gives rise, in instances of multiple discovery, to priority disputes. We state the rule as follows:

> **The Priority Rule**: Prestige associated with a particular scientific development is bestowed wholesale upon a particular party as a consequence of a large majority of the surrounding community having all individually come to associate that development with them.

Given this rule, in instances of multiple discovery, we may imagine 'credit attribution contests' between individuals within the community who independently produce similar developments at the same time.

One might object that a large majority is not enough – surely all or nearly all of the community must agree on a discoverer for that person to get the prestige. There are reasons we think a large majority would suffice. First, in the instance of a priority dispute between the two competing parties, one might image that those who enjoy the support of the large majority are more likely to win the priority dispute: along the lines of the observation by Merton just mentioned, those who enjoy the support of the large majority also enjoy a larger collection of possible defenders, ready to fight against an instance of perceived injustice.

Second, even without a priority dispute, a large majority of scientists coming to associate one person with a discovery might be sufficient for prestige to be bestowed on that individual. For instance, when using reference to a name as short-hand for an idea that one's interlocutor will understand (e.g. using "Schelling's segregation model" to refer to a mathematical model whereby minimal conditions for segregation are demonstrated), the most effective name to choose is the name most well-known in connection to that idea (one's interlocutor is less likely to know of Sakoda's work). One might imagine that there is some tipping point, or threshold, at which it becomes prudent for individuals wishing to communicate in this way to defer to the use of the more well-known name, irrespective of who they individually have associated with it.[2] Similarly, new members of the community, hoping to signal their understanding of the field, will think to provide the name most people within their community associate with the discovery, lest they be thought ignorant. Finally, it is reasonable to expect that the more well-known name would be used in review articles written by members of the community, to communicate the results of their sub-field to a wider audience (and to the few members of the community who may not have already heard of the discovery), and so both the broader scientific community and, eventually, the public at large and history books would come to recognize the more well-known name.[3]

In the next section, we will formalize these credit attribution contests so as to study the influence of network structure on the awarding of prestige in instances of genuine multiple discovery. This model is meant to be descriptive of scientific practice, not prescriptive of how a central planner ought to distribute prestige. As such, we do not compare the efficacy of different *mechanisms for dispersing* prestige; rather, we take prestige to simply be that which is disbursed in the way we have so far described. We make two arguments based on this model. First, we argue that this more descriptively accurate picture of scientific priority does not necessarily align with our ideas of fairly rewarding people for benefits they confer (as in society's grand reward scheme). Second, as we will argue in section 4, what is incentivized by our version of the priority rule is not the optimal division of cognitive labor that Strevens envisages.

---

[2] Thanks to REMOVED FOR REVIEW for discussions on this point. In the next section we provide a particular tipping point of 2/3. While this number is arbitrary, the existence of a tipping point is not arbitrary, and the particular number chosen will not significantly affect results. For instance, similar results have been obtained with 3/5.

[3] Thanks to REMOVED FOR REVIEW for encouraging us to think about the role of review articles.

3 The basic model

We formalize scientific communities as networks of agents, where nodes of the network represent individual scientists and edges, or links, represent regular information channels between them. These links are bidirectional and can be thought of as representing people who talk to each other when working on a new project, or who ask each other if looking to reference a paper on some or other topic.

We model credit attribution contests where information about discoveries spreads on these networks as individuals go looking for particular new results that they happen to care about in the moment.[4] In the model, we start with an instance of multiple discovery. Two scientists each independently make some discovery, and not knowing about the other, they each believe themselves to be the discoverer. In the second time-step, we pick a third node at random to go looking for news of the discovery. Whoever they are closest to in the network is who they are going to get the news from first,[5] in which case they then associate that person with the discovery, cite them with respect to that idea, etc. Then in the next timestep, another random node goes looking for a discovery. This fourth scientist gets news of that discovery from whoever is closest to them in the network that already has some belief about who made the discovery. (The idea here is that when one goes looking for a paper on some topic x, for instance so as to cite a discussion of that topic, one might ask their friends if they know about a paper on x, or alternatively, one might go looking to trusted collaborator's papers for references, or so on.) This fourth scientist then has a belief about who deserves credit for the discovery. This process continues until all scientists have attributed credit. If there is a super-majority of 2/3 in favor of one discoverer over another, we say the former wins the credit attribution contest. If not, there is a tie.

As mentioned, this information spread occurs over networks of scientists and the structure of the network will impact who receives credit for a discovery. At first, we will consider basic Barabási-Albert networks [Barabási and Albert, 1999], though in later sections, we will consider alterations to that model, which include social identity types and network change over time. As explained below, these networks capture several important features of scientific communities.

Barabási-Albert networks are formed in the following way. First, we start with a small number of fully connected nodes (nodes with all possible links between them), m_0. Then, new nodes are added one by one until the network

---

[4] This modeling choice is motivated by the historical examples above, wherein it seemed that the scientists in the community first had to care about learning of the discovery before credit actually came to be attributed. One might instead think information spreads more like a diffusion process, with discoverers advertising their results to those they are connected to on the network. The way news about a discovery spreads is likely a mix of processes like these— particularly once one introduces self-promotion, as an incentivized strategy, to win potential priority disputes when one perceives there to be the threat of a loss. We will note some results from a diffusion model where relevant, and discuss the incentivization of self-promotion in section 4.

[5] If there is a tie, a node is chosen at random.

reaches a designated size. Each time a new node enters, it forms a set number of links, m. For the results presented here, $m_0 = m = 4$. New links are formed via preferential attachment. That is, the more links a node already has (i.e., the higher its *degree*), the more likely it is that an entering node will form a link with it. The probability p_i that the new node is connected to node i is proportional to:

$$p_i = \frac{d_i}{\Sigma_j d_j}$$

where d_i is the degree of node i, and $\Sigma_j d_j$ is the sum of the degrees of all nodes in the network (not including the new entering node). The higher the degree a node already has, the more likely it is the new node will connect to it.

These networks have a couple of important features. First, in these networks, the 'rich get richer': nodes that already have many links are more likely to get new links. This captures a scenario seen in many real-world networks where the oldest members of the community tend to be the most central and well-connected individuals in the network. Moreover, this model takes on a natural interpretation in the context of scientific research communities: as new researchers enter the community (e.g. as graduate students), they often seek out social relationships with the more well established members of the community, the oldest of which are often the most esteemed.

Second, they are scale free, meaning their degree distribution follows a power law. In other words, there are many nodes with a few links and a few nodes with a large number of links; there are a few 'hubs' in the network. Many real world networks are (approximately) scale-free, including many types of social networks and collaboration patterns. Among the scale-free networks, Barabási-Albert networks are particularly useful for our purposes. There is evidence that collaboration networks are formed via preferential attachment, similar to the method of preferential attachment used in the formation of Barabási-Albert networks [Newman, 2001, Barabási et al., 2002]. Additionally, since in later sections we will be discussing how networks evolve over time, Barabási-Albert networks are useful because they already stipulate what should happen when new nodes enter the network.

For this basic model, we formed a network of 100 people, then ran 1000 contests on each network to estimate the likelihood of each person getting the credit for their discovery, then performed 100 replications (i.e., we formed 100 different networks of 100 people, and ran 1000 contests on each). Figure 1 shows the likelihood of winning for each node in the network. That is, of the contests a node was a part of, it shows the percent of contests that node won. In this figure, the lower number a node is, the older it is (nodes 1-4 are all the same age, as the network started with 4 fully connected nodes). We found that, as one might expect, older nodes were more likely to win credit attribution contests because they tended to have a higher degree. Intuitively, when there is an instance of multiple discovery, those scientists who are more well connected are more likely to be given credit for their discovery because the news of their discovery travels faster.

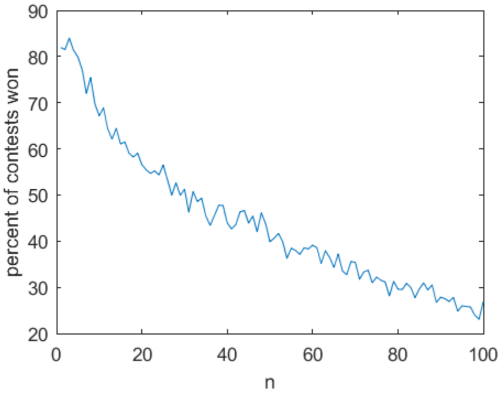

Figure 1: Likelihood of winning for each node in the network

3.1 The model with types

This points to a possible disadvantage for historically underrepresented groups (HUGs). Since older nodes tend to be members of the historically entrenched group (HEG), the HEG members will tend to be better connected, even when members of a HUG begin to enter the community at an equal rate later on. This means that HEG members tend to receive credit for their discoveries at a higher rate, even when they make the discovery at the same time as a member of the HUG.

In order to model this intuitive argument, we introduce types into the basic model: HEG members will be type 1 and HUG members type 2. There is nothing intrinsically important about these types; they are not related to scientific competence or likelihood of producing a scientific discovery. They are, however, socially relevant, in that type 1 enters at a higher rate earlier in time (i.e. they are historically entrenched, and type 2 is historically underrepresented).

In particular, we used the following logistic growth equation

$$P(\mathit{type}\ 2) = .5 \times \frac{1}{1 + 10 \times e^{-.05t}}$$

to determine how likely it was a new node was type 2. This represents a case where the HUG finds it hard to enter the scientific community at first, but once there is a sufficient number of them it becomes much easier. By the end, the HUG enters at roughly the same rate as the HEG.[6]

We add one more feature of this model to reflect aspects of real world scientific communities: *homophily*, or the preference for linking to members of one's own social identity group. People are homophilic in a variety of contexts (e.g.

---

[6] Nothing hangs on using this particular equation. We also used an equation where the probability of type 2 increases quickly then asymptotes at .5, $P(\mathit{type}\ 2) = 1 - \frac{1+2e^{-t/20}}{2+e^{-t/20}}$, and results were similar.

when forming friendships [Currarini et al., 2009]), and scientific communities are homophilic as well, especially when it comes to co-authorship patterns [Ferber and Teiman, 1980, McDowell and Smith, 1992, Boschini and Sjögren, 2007, del Carmen and Bing, 2000, West et al., 2013]. We implemented homophily in the model by having agents place some weight, H, on their similarity to a node in addition to their degree. We used the following to determine how much the incoming node values linking with each of the existing nodes:

$$v_{ij} = H \times \frac{s_{ij}}{\sum_k s_{ik}} + (1 - H) \times \frac{d_{ij}}{\sum_k d_{ik}}$$

where $s_{ij} = 1$ if nodes i and j are of the same type and 0 otherwise.[7] The probability that the new node links with a particular node is proportional to:

$$p_i = \frac{v_i}{\Sigma_j v_j}$$

The likelihood a node is chosen is thus determined by its value to the new individual, including both homophilic preferences and degree, rather than just its degree.[8]

For 100 of these networks, we look at 1000 credit contests where a type 1 and type 2 individual are competing for credit. Out of all these contests, type 1 wins about 42.3% of the time, type 2 wins about 16.7% of the time, with the remaining contests consisting of ties. That is, being part of the HEG confers a distinct advantage. More specifically, we can define *HEG advantage* as the probability a HEG member wins minus the probability a HUG member wins the credit attribution contest in an instance of multiple discovery, where the two discoverers belong to different social identity groups. In this model, with no homophily, the HEG advantage is 25.6%.[9] In this model, homophily does not significantly affect the results – the likelihood of each type winning stays relatively constant as we vary homophily.

3.2 Evolving networks

Of course, scientific communities change over time. Older members of the community retire and new scientists enter the community. After a time, if the

---

[7] We incorporate homophily in this way because its influence on an incoming node's linking choices remains constant. Some authors incorporate homophily as a weighting of the degree of a node, and find, similar to our results here, that homophily "makes the rich even richer" [Kim and Altmann, 2017] and that homophily can lead to minority group members occupying less important places in the network [Karimi et al., 2018].

[8] One might also think that winning a credit attribution contest would make an existing node more "prestigious" and thus would increase its value in the eyes of incoming nodes. This is not included in the model, but would likely intensify the effects reported here.

[9] This HEG advantage exists in a model where information spreads by diffusion as well, in a simple contagion model where each discoverer tells their neighbors, who tell their neighbors that have not already heard of the discovery, and so on (with people receiving conflicting information choosing to assign credit randomly). However, there are less often super-majority winners in this case. If we look at cases where one discoverer has over 50% of the community assigning them credit, we see that type 1 wins about 54% and type 2 wins about 46% of the time.

HUG enters the community at a rate equal to its size in the population, it will eventually achieve proportional representation in the scientific community. Our questions then are: Will the HEG advantage over the HUG ever go away? If so, how long will it take?

We incorporate network change over time into the model in the following way. There is a maximum network size of 100 scientists, so as a new scientist enters beyond the first 100, the oldest node is removed from the network (it "retires") along with all its links. When a new node enters, it then forms a set number of links, as in the Barabási-Albert model. Additionally, in order to capture the consequences on the social network structure of a scientific community that trains its young, when the network grows beyond 50 nodes, incoming nodes also choose an 'advisor'. For each node the advisor is linked to, the new node has a chance of linking with that node as well (a 50% chance for the results below) in addition to its Barabási-Albert links. This captures a scientific field that grows to a certain size, becomes established, then begins to adopt practices to train new generations of scientists.[10]

How long will the HEG advantage persist when the network evolves? There are many factors to consider. We will talk about HEG advantage over *generations* of the scientific community. A generation is defined as the time it takes to have a complete turnover of scientists. Since in each time-step the oldest scientist retires and a new scientists enters, with 100 scientists a generation is 100 time-steps.

Let us first consider a case where the HUG approaches 50% of the population, where by timestep 100 (after 1 generation) they are entering in equal proportions and after 2 generations they have achieved equal representation. For each level of homophily, we formed 250 networks. For each of these networks, we performed 250 credit attribution contests (where a type 1 and type 2 individual were competing for credit) every 25 rounds to get an idea of how likely it was that each social identity group would get credit for their discovery, and how these chances changed over time. Figure 2 shows what happens in this case.

We find that the HEG advantage disappears quickly over time. Interestingly, the HUG has an advantage for a short period of time (that is, the HEG advantage goes negative). As homophily increases, this temporary HUG advantage increases. This is likely because, when there are very few members of this group for the initial time period, the one or two that exist serve as focal points for the incoming members. As the HUG starts entering at higher rates, these focal points become highly connected to all the new people such that if a member of their social identity group makes a discovery, they will know about it. Since

---

[10] Nothing depends on this particular way of doing things. For instance, we obtained similar results when new nodes used the following copying mechanism to create new links, modified from Kumar et al. [2000]. When a new node enters the network it chooses a 'prototype' (its advisor) via the Barabási-Albert procedure and forms a link with their advisor. Then, it copies $0-3$ randomly chosen links from its advisor (how many is chosen randomly, subject to the condition that the advisor had enough links). Then, if the new node still had less than m links, they formed more links according to the Barabási-Albert procedure (up to $m=4$ total). The results of this model are very similar to the ones presented below, although homophily has slightly less of an effect on the HEG advantage.

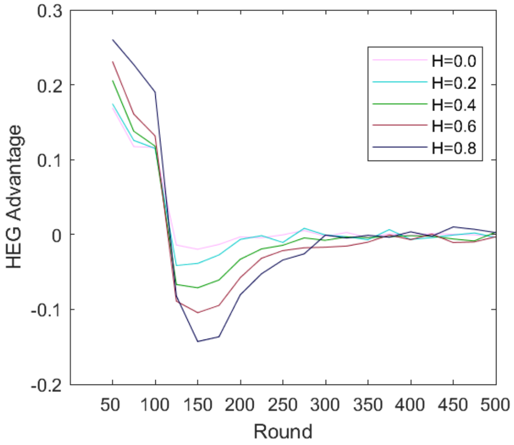

Figure 2: HEG advantage over time, for different levels of homophily, where the HUG is 50% of the population.

these focal individuals are so highly connected, they are likely to have at least some connections with new HEG members despite homophily, because people still do care to some extent about forming links with highly connected people. So news of the HUG members' discoveries can spread.

That the HEG advantage disappears quickly in this case looks somewhat promising — if we can get equal representation, eventually no group is disadvantaged.[11] A situation like this is achievable if we are thinking about men and women, but not if we are thinking about minority groups like racial minorities, people with disabilities, etc. In these cases, we can talk about what happens if proportional, rather than equal, representation is achieved.

The situation is different when the HUG is a minority group. In this case, even when the HUG achieves proportional representation within two generations, the HEG advantage can be more severe and last longer, especially in the presence of strong homophily. As seen in figure 3, when homophily is low, the HEG advantage disappears around when the HUG achieves proportional representation (with no real period of advantage for the HUG). But as homophily increases, the HEG advantage can persist over time, meaning the HUG is less likely to get credit for their discoveries.

4 Concluding remarks

Strevens [2003] interprets the priority rule in science as a behavior regulator for the scientific community, part of a grand reward scheme which benefits

---

[11] It has been argued that in other situations, such as when there is a discriminatory bargaining norm in a community of scientists, merely increasing representation of the minority group will not eliminate inequalities [Schneider et al., 2019].

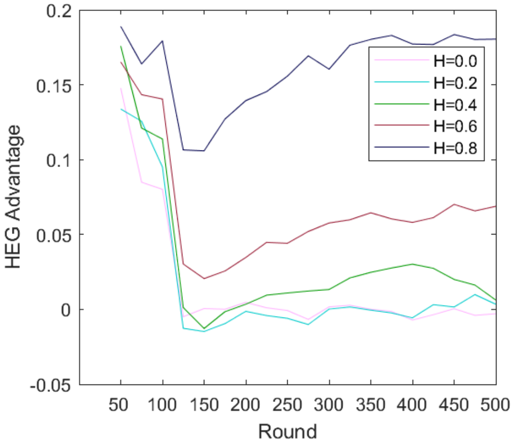

Figure 3: HEG advantage over time, for different levels of homophily, where the HUG is 10% of the population.

society by adequately structuring the distribution of intellectual labor across pre-existing research programs. We have demonstrated that prestige does not necessarily go to who deserves it. People are not always rewarded based on the benefit they confer, and so it is difficult to regard the priority rule as an instance of a grand reward scheme that resonates with our notions of fairness.

There is also the further question of whether the priority rule can serve the function of dividing cognitive labor in the way Strevens suggests.[12] In order for the priority rule to provide this benefit, it must be the case that scientists believe they will receive credit for their discovery, so that they decide which research program to join based on the likelihood of making a discovery. Merton [1957] already provides some evidence that scientists believe recognition is not automatic after discovery, and (to the contrary) is something that is more likely for some than others. He quotes Norbert Weiner explaining:

> I was competitive beyond the run of younger mathematicians, and I knew equally that this was not a very pretty attitude. However, it was not an attitude which I was free to assume or to reject. I was quite aware that I was an out among ins and I would get no shred of recognition that I did not force (p. 649).

If instead scientists believe the correct picture of credit assignment is closer to our version of the priority rule (and consequently that news of their discovery must spread through the scientific network, with more well-connected scientists more likely to receive credit when there is a close race), Strevens's optimal division of labor may not be what is incentivized. For instance, scientists motivated by credit might be incentivized to link with well-connected people – this might

---

[12] Thanks to REMOVED FOR REVIEW for this point.

be better achieved by joining a research program that already has a lot of followers, even if the chance of making a discovery in the other research program is better. Alternatively, scientists might be incentivized to pick a research program with fewer well-connected people, so as to decrease the likelihood that the prestige associated with their potential discovery will be "scooped" by another who is in a better position to capture it.

In the case of socially diverse communities, per the considerations in section 3, members of minority groups should, at first pass and in ignorance about their precise location in the social network, be disincentived to join research programs that feature large numbers of scientists from the majority group. This disinsentive may often be in tension with those same scientists' considerations of the "intrinsic potentials" of those programs, as considered by Strevens. These competing incentives may eventually lead to clustering into sub-disciplines according to social identity. Following the arguments given by Schneider et al. [2019], there is some historical precedent to suggest that such clustering is ultimately detrimental to the general state of our scientific knowledge across disciplines.

Absent further study, it is difficult to discern which of these various factors (or others) clearly dominates in the decision-making of scientists, and even more difficult to suss out the epistemic consequences of such decision-making; we leave this investigation for future work. Our point is that these considerations are absent in Strevens's model, and that the extent to which any such strategies are motivated at all can only be evaluated in a model of the priority rule which takes into account individual scientists' credit attributions. Altogether, the epistemic benefits of the priority rule that Strevens sees, in terms of maximizing the likelihood of a discovery, are not guaranteed when we take seriously the implications of a distributive reading of the statement: The scientific community disburses prestige.

References

A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

A.-L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, and T. Vicsek. Evolution of the social network of scientific collaborations. *Physica A: Statistical mechanics and its applications*, 311(3-4):590–614, 2002.

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

J. P. Bruner. Minority (dis) advantage in population games. *Synthese*, 196(1): 413–427, 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. del Carmen and R. L. Bing. Academic productivity of African Americans in criminology and criminal justice. *Journal of Criminal Justice Education*, 11 (2):237–249, 2000.

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.

I. Hansen-Schaberg. A biographical sketch of Prof. Dr. Grete Henry-Hermann (1901–1984). In *Grete Hermann-Between Physics and Philosophy*, pages 3–16. Springer, 2016.

R. Heesen. Communism and the incentive to share in science. *Philosophy of Science*, 84(4):698–716, 2017.

R. Heesen. Why the reward structure of science makes reproducibility problems inevitable. *The Journal of Philosophy*, 115(12):661–674, 2018.

R. Hegselmann. Thomas C. Schelling and James M. Sakoda: The intellectual, technical, and social history of a model. *Journal of Artificial Societies and Social Simulation*, 20(3), 2017.

F. Karimi, M. Génois, C. Wagner, P. Singer, and M. Strohmaier. Homophily influences ranking of minorities in social networks. *Scientific reports*, 8, 2018.

K. Kim and J. Altmann. Effect of homophily on network formation. *Communications in Nonlinear Science and Numerical Simulation*, 44:482–494, 2017.

P. Kitcher. The division of cognitive labor. *The journal of philosophy*, 87(1): 5–22, 1990.

R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal. Stochastic models for the web graph. In *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on*, pages 57–65. IEEE, 2000.

K. N. Laland, K. Sterelny, J. Odling-Smee, W. Hoppitt, and T. Uller. Cause and effect in biology revisited: is Mayr's proximate-ultimate dichotomy still useful? *Science*, 334(6062):1512–1516, 2011.

K. Ludwig. *From individual to plural agency: collective action I.* Oxford University Press, 2016.

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.

R. K. Merton. Priorities in scientific discovery: a chapter in the sociology of science. *American sociological review*, 22(6):635–659, 1957.

M. E. Newman. Clustering and preferential attachment in growing networks. *Physical review E*, 64(2):025102, 2001.

W. F. Ogburn and D. Thomas. Are inventions inevitable? a note on social evolution. *Political Science Quarterly*, 37(1):83–98, 1922.

F. Romero. Novelty versus replicability: Virtues and vices in the reward system of science. *Philosophy of Science*, 84(5):1031–1043, 2017.

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

M. D. Schneider, H. Rubin, and C. O'Connor. Promoting diverse collaborations. In G. Ramsey and A. De Block, editors, *The Dynamics of Science: Computational Frontiers in History and Philosophy of Science*. 2019.

M. Seevinck. Challenging the gospel: Grete Hermann on von Neumann's no-hidden-variables proof. In *Grete Hermann-Between Physics and Philosophy*, pages 107–117. Springer, 2016.

S. M. Stigler. Stigler's law of eponymy. *Transactions of the New York academy of sciences*, 39(1 Series II):147–157, 1980.

M. Strevens. The role of the priority rule in science. *The Journal of philosophy*, 100(2):55–79, 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.

K. J. Zollman. The credit economy and the economic rationality of science. *The Journal of Philosophy*, 115(1):5–33, 2018.