Scientific Sharing: Communism and the Social Contract

Michael Strevens

To appear in *Scientific Collaboration and Collective Knowledge*, edited by Thomas Boyer-Kassem, Conor Mayo-Wilson, and Michael Weisberg, Oxford University Press, New York

Abstract

Research programs regularly compete to achieve the same goal, such as the discovery of the structure of DNA or the construction of a TEA laser. The more the competing programs share information, the faster the goal is likely to be reached, to society's benefit. But the "priority rule"—the scientific norm according to which the first program to reach the goal in question must receive all the credit for the achievement—provides a powerful disincentive for programs to share information. How, then, is the clash between social and individual interest resolved in scientific practice? This chapter investigates what Robert Merton called science's "communist" norm, which mandates universal sharing of knowledge, and uses mathematical models of discovery to argue that a communist regime may be on the whole advantageous and fair to all parties, and so might be implemented by a social contract that all scientists would be willing to sign.

Acknowledgments

Thanks to Thomas Boyer-Kassem and to audiences at the PSA 2008, the conference on Collective Knowledge and Epistemic Trust at the Alfried Krupp Wissenschaftskolleg, Greifswald, the Colombian Conference on Logic, Epistemology, and the Philosophy of Science at the Universidad de los Andes, Bogotá, Carnegie Mellon, and the University of Barcelona.

1. Sharing, For and Against

Society has a clear and strong interest in the wide distribution of scientific knowledge. Such dissemination has direct and indirect social benefits. The direct benefits come through the public consumption of knowledge, in the form of miracle cures, magical gadgets, and eternal truths to contemplate. The indirect benefits come through other scientists' consumption of knowledge, since once digested such knowledge will tend to increase the rate, breadth, and depth of scientific discovery and so the magnitude of the direct benefits. (Qualifications should no doubt be made about weaponized anthrax, cobalt bombs, and so on.)

Scientists' own interest in sharing the knowledge they have themselves created is more circumscribed. Their desire to make the world a better place and, even if that desire is lacking, to receive credit for their discoveries militates in favor of sharing. But the threat of one scientist's finagling credit for another scientist's discovery militates, at least temporarily, in favor of secrecy.

The reward system in science—the "priority rule", according to which only the first scientist or team of scientists to make a discovery receives credit for the discovery (Merton 1957)—increases the strength of the arguments both pro and con. On the one hand, the priority system gives you reason to publish your results as early as possible, in order to maximize your chances of being recognized as the first discoverer (Dasgupta and David 1994). On the other hand, the same considerations give you a powerful incentive not to share your results before you have extracted every last publication from them, to minimize the chances that someone else takes advantage of your research to gain credit that might otherwise have gone to you.

These pro and con arguments apply to largely non-overlapping parts of the research and publication cycle. Suppose, somewhat artificially, that the two parts are entirely distinct, that is, that every research program can be divided into phases before and after discovery. The priority rule motivates a scientist to keep all data, all technology of experimentation, all incipient hypothesizing secret before discovery, and then to publish, that is, to share widely, anything and everything of social value as soon as possible after discovery (should a discovery actually be made). The interests of society and the scientist are therefore in complete alignment after discovery, but before discovery, they appear to be diametrically opposed.[1]

The resulting moral and political conflicts are dramatically portrayed in case studies such as James Watson's (1968) account of the discovery of the structure of DNA, in which Watson wheedles a crucial X-ray diffraction image of the molecule from the Wilkins lab without the crystallographer Rosalind Franklin's knowledge, or Wade's (1981) account of Roger Guillemin and Andrew Schally's race to sequence the brain-signaling hormone TRH (Schally's policy on sharing: "Don't talk to the enemy!"), or Collins's (1974) account of the development of the TEA laser, in which rivalry among competing labs discourages communication that might have sped everyone to their joint goal—with one scientist memorably confessing to the following policy on sharing research: "Let's say I've always told the truth, nothing but the truth, but not the whole truth" (p. 180).

This is all quite entertaining, but it prompts a serious question. Are there circumstances, or can circumstances be created, under which information will be shared in a way that precipitates, rather than merely promulgating, discoveries? Let me begin by surveying some of science's current pro-sharing arrangements.

Certain grant agencies require a degree of data sharing as a condition for accepting their funds. In the United States, both the National Institutes of Health (the NIH) and the National Science Foundation (the NSF), two of the largest government grantors, have explored this approach. Typically, the data to be shared has been used as the basis for publications (there is no obligation

1. Many writers have discussed the conflict between sharing and scientists' narrow self-interest; for a sample of approaches see Dasgupta and David (1987), Arzberger et al. (2004), Jasanoff (2006), and Resnik (2006).

to share before publication), but further significant publications might be extracted from the same body of information; this is the sort of case where the "before" and "after" phases distinguished above are intermingled. Some publicly minded publishers, such as PLOS, have similar policies.[2]

In special cases, notably in the life sciences, rules for the sharing of certain classes of data have been agreed upon by funding agencies, industry representatives, and the heads of research projects working together. An early and influential example was the negotiation in 1996 of the "Bermuda Principles" mandating the immediate publication of DNA sequences obtained by the Human Genome Project (Contreras 2011).

The power of such requirements is limited, however, so long as the organizations in question are fighting against individual researchers' own aims and desires: the NIH planned a demanding data-sharing protocol in the early 2000s, but this was watered down in 2003 after objections from many scientists.[3] Further, the publication trigger may come far later than is socially optimal, and in any case, data is only one aspect of the epistemic and material goods generated by a research program's progress that may be fruitfully shared before discovery—think of software and computational methods, genetically engineered lab animals, the sort of instrumental or surgical know-how that has to be learned in person, the loan of equipment or technicians, and so on.

A more radical strategy is to abandon priority as a principle for allocating credit throughout science, thus removing much of the disincentive to share. A step in this direction has been taken by Nancy Wexler's Hereditary Disease Foundation:

---

2. Though to sometimes limited effect (Savage and Vickers 2009).

3. The "Final NIH Statement on Sharing Research Data" reads: "Several groups and individuals objected to sharing of research data prior to publication… NIH recognizes that the investigators who collect the data have a legitimate interest in benefiting from their investment of time and effort. We have therefore revised our definition of 'the timely release and sharing' to be no later than the acceptance for publication of the main findings from the final data set. NIH continues to expect that the initial investigators may benefit from first and continuing use but not from prolonged exclusive use" (National Institutes of Health 2003).

> In larger projects which require many scientific lines of inquiry at once, HDF funded scientists work together toward the same end goal. Credit for breakthroughs is shared by all collaborators, regardless of which vein of research and scientist(s) "arrives first". The HDF community understands that complex problems are solved more rapidly when cooperation between those with complementary skills takes place. (Hereditary Disease Foundation 2014)

The priority rule, however, performs a valuable function in determining an efficient allocation of cognitive labor among scientific projects (Strevens 2003). It solves, in effect, an enormously complex and pressing coordination problem, and so—in spite of its often rough justice—it should be if at all possible retained.

Rather than forcing scientists to share their knowledge as a condition of their funding, then, it would be better to find ways to make them want to share, building if possible on preexisting inclinations. Are there, already ensconced in the scientific mind, any impulses toward the free dissemination of information?

A partial answer to this question is provided by Boyer's (2014) and Heesen's (2016) important arguments that it is rational for scientists solely concerned with the pursuit of personal credit to break up their work into "least publishable units" and to place these units in the journals as soon as possible, despite the collateral advantage it gives to their competitors. What looks altruistic—sharing early and widely—is thereby motivated by narrow self-interest, seldom in short supply. The benefits are circumscribed, however: as with data-sharing requirements, nothing will be shared until something relevant is ready for publication, and worse, it is only what characteristically goes into the journals that gets broadcast, so the other shareables mentioned above will remain hidden or otherwise unavailable.

Agreements to trade information for mutual benefit are another means of sharing. Like Boyer's and Heesen's atomized publication strategy, these contractual exchanges are pursued for private advantage; unlike that strategy, they are not confined to publishable information and goods.

All of this arm-twisting and self-regard is, however, nothing but an ugly and awkward prelude to a far sweeter song: an ideal of sharing already firmly established in science's guiding system of values.

2. The Communist Way

What Robert Merton provocatively but accurately called science's communist norm enjoins common ownership of scientific information: "secrecy is the antithesis of this norm; full and open communication its enactment" (Merton 1942, 274). To the extent that allegiance to the communist norm opposes the dictates of narrow self-interest, it drives sharing in science. When self-interest anyway points toward sharing, communism may help smooth the way with good feelings and the satisfying sense of duty done. (Merton focuses on information rather than on more tangible research-related goods such as bacterial strains or genetically modified mice. The communist norm appears to apply to some extent to these, but the boundaries are unclear. I follow Merton's lead in focusing on information, but you might also understand the term "information" in the remainder of this chapter as a catch-all for any kind of shareable.)

Merton offered little evidence for the existence of informational communism; he presumed, I suppose, that every scientist would recognize it as a part of their professional code. Later more empirically scrupulous sociological studies have vindicated his confidence. A survey conducted by Macfarlane and Cheng (2008) found 95% of respondents (of whom there were several hundred, mostly UK university professors) in favor of sharing research in progress. A similar result was found in a survey of nearly 2000 US geneticists and other life scientists, with 91% agreeing that they should "freely share information, data, or materials with all academic scientists" (Louis et al. 2002).[4] Even scientists working in private industry endorse unrestricted sharing, though they are according to the study somewhat less likely to realize the ideal in practice.

More evidence for the communist norm comes from the rhetoric surrounding the formulation of sharing agreements such as the Human Genome Project's Bermuda Principles: Contreras (2011) documents the importance of "strong open science norms" both in motivating the original agreement and in its subsequent implementation by other institutions (pp. 82, 88).

There is little reason, then, to question the existence of a communist norm. But the norm does not make scientists' decisions to share straightforward. The negative consequences of sharing are palpable: 35% of the surveyed life scientists reported that, as a result of sharing information, they had been "scooped" by other researchers, losing the opportunity to gain credit for work based on their own results. Perhaps unsurprisingly, then, scientists frequently fail to live up to the communist ideal. About 30% said that they had "within the past three years … withheld research results from other academic scientists prior to publication on at least one occasion"—not unlike Collins's laser scientist quoted in the previous section. Subscribing to and complying with a norm are, as we all well know, not the same thing.[5]

It may seem peculiar that the norm of communism exists at all, given scientists' powerful motivation to ignore its prescriptions. And you might wonder how long it can endure in the mercilessly competitive world of science today.[6]

4. These studies make clear something about which Merton was vague: the moral imperative to share applies to unpublished as well as published research.

5. Tenopir et al. (2011) provide some further statistics on the spotty sharing of data.

6. Humans are altruistic: they recognize a ceteris paribus obligation to take steps, when possible, to improve the world at large. Might the communist norm be explained, then, as an application of the general altruistic urge to the special case of unpublished scientific information? Might it be, not a norm particular to science, but a manifestation of the broadest moral norm of all?

The principal aim of this chapter is to provide what I will call a Hobbesian vindication of the communist norm, a vindication that might variously be used to justify, to shore up, or to explain the origin of the norm.

A Hobbesian vindication of a norm has two parts. The first part is a transformation of the norm into a social contract that behaviorally mirrors the norm. Rather than a moral obligation to share, then, we consider a contract in which the signer promises to behave just as the norm would oblige them: to share all pre-publication information unconditionally. (It is assumed in the Hobbesian spirit that the contract creates an irresistible, but in this chapter unspecified, means of enforcement.)

The second part is a rationale for signing the contract. Typically it is assumed that either everyone signs the contract—the "cooperation" scenario—or that no one does—the "state of nature". The cooperation scenario is shown to be superior, for everyone involved, to the state of nature. In the case of sharing, then, the aim is to demonstrate that it is better for all scientists if they share universally, as required by the communist norm, than if they do not share at all. An even stronger vindication would demonstrate that the communist contract is more desirable than any other sharing contract; that too I will attempt to show.

What is the point of a Hobbesian vindication of a norm? It concerns a fictional social contract rather than the actual moral precept in question; it tells (by implication) an origin story, a collective agreement to exit the state of nature, that is unlikely to have any precise historical parallel; and it has nothing practical to say about the means by which such an agreement might be negotiated or enforced.

Altruism certainly plays a role in motivating sharing; for example, it is explicitly cited as a rationale for institutional requirements and agreements such as the NIH data-sharing policy and the Bermuda Principles. But I do not think it can on its own explain the communist norm: we do not in general feel an obligation to do something just because it is socially beneficial, especially when there is substantial risk of personal harm—the optionality of charitable giving, lamented by Singer and Unger, being a notable example. Thus I will suppose in what follows that the communist ideal is its own thing, in need of its own special explanation.

Nevertheless, the vindication can be used, given the right sort of philosophical or empirical scaffolding, to do a number of things: to justify the norm, to give an evolutionary account of the origin of the norm, to give an account of the social or psychological stability of the norm, and as a tool for social engineering, by repairing or enhancing the stability of or (in some cases) dismantling the norm. With respect to the communist norm, I will not attempt any of this. But I take it that, by providing the Hobbesian vindication, I will have cast some light on the best way to go about achieving any or all of these aims.

Let me turn, then, to the question of contractual sharing. Consider scientists working in a near state of nature: no institutionalized agreements to share, no communist norm, no governments or corporations or other funders with their own agendas—nothing except the priority rule, providing the background against which the advantages or otherwise of sharing are to be assessed. In such a world, what exchanges of information might a self-interested scientist willingly sign on to?

3. Contractual Sharing

3.1 *One-Off Exchange*

In a world where information is not freely shared, two scientists or two scientific research programs might nevertheless agree to exchange certain specified pieces of information for narrow, self-regarding reasons—that is, because each thinks they can improve their prospects by making the trade.

It is easy to see how agreements to share might be mutually beneficial when the parties are pursuing somewhat different goals. Such contracts may involve a "no trespassing" clause: the parties agree not to compete in the future, or even more explicitly, agree who has the right to publish and receive credit for which discoveries.

How, though, can a sharing agreement seem desirable to two research programs that are competing to make the same discovery, most particularly in the extreme case where they have precisely the same goals? You might reason as follows. Any trading of information must lead to one of two outcomes: either the trade will favor one side, increasing the chance that the favored program attains its goal before the other—that is, increasing the chances that one program will "scoop" the other—or it will be neutral, making a scoop no more likely than before. In the former case, the disfavored program will of course have good reason to eschew the trade. In the latter case, it seems that neither program will have reason to take the time and trouble to enter into an agreement. (Blumenthal et al. (2006) discuss the considerable work involved in preparing information for sharing.)

There are four reasons that a contract might nevertheless be feasible. First, one or both of the programs may have false beliefs about the effects of the trade, so that each thinks it has the advantage over the other. Second, in the case where the trade advantages neither party but has some net social benefit, altruistic desires may motivate both parties to expend the effort to make the exchange. Third, even if you are not in the least altruistic, some swaps might seem desirable although they directly benefit only your rival, just because your rival's winning the race to discovery, though it entails your failure, nevertheless produces some good for you as it does for all of science and society, in the form of new knowledge to consume. Fourth, both parties may gain an advantage as a result of the trade, not relative to one another, but relative to their competitors and, as it were, to nature.

I will put aside the explanation from ignorance, assuming for now that researchers can reliably estimate the benefits that will accrue from an information exchange. I also put aside the second and third explanations, from altruism and from the desire to know, for tactical reasons: if a communist contract is attractive even to scientists who entirely lack these motivations—as I eventually hope to show—then it must surely be all the more desirable to those who feel their pull.

To the fourth explanation, then. I will focus on a simple case in which two research programs are attempting to make a certain specific discovery or to reach some other specific goal, such as determining the structure of DNA, explaining superconductivity, or finding the genetic basis of sickle cell disease. (These are dramatic examples, but for the purpose of understanding sharing in a world governed by the priority rule, "goals" and "discoveries" include anything that would earn scientific credit—in other words, anything that a scientist would consider worthy of publication for its own sake.) Each program has its own particular methodology, or guiding theory, or expensive detector, and so there is no sense in which they could join forces to constitute a single coherent research program. They are by their very nature rivals, engaged in a race that only one of them can win.

There are three possible outcomes to such a race: one program makes the discovery first, the other program makes the discovery first, or neither makes the discovery. Suppose that a trade of information would decrease the third probability, and so increase the probability that some program makes the discovery. Suppose further that this increase to some extent benefits both of the rival programs. Then as a result of the trade, each will see an increase in its probability of making and receiving credit for the discovery. Thus, the trade is to the advantage of both programs—it is what I will call *mutually advantageous*.

Will the programs make the trade? If their only consideration is to maximize their expected credit earned, they will.[7] But in reality, they might reject

7. What if other, longer-term consequences are taken into account? You might refuse to make a trade that is mutually advantageous in the technical sense just defined if it is better for you that both you and your rival fail to make a discovery than that your rival succeeds and you fail—as when, for example, the winner of a discovery race not only receives credit for the discovery but also the lion's share of the funding for the next race. In these and similar circumstances, then, narrow self-interest militates against sufficiently unbalanced "mutually advantageous" trades. But I will assume that there is no need to take such considerations into account explicitly, as the work they do to motivate balanced trades is fully replicated by the concern for fairness described in the following paragraphs.

such a trade if they perceive it to be unfair—if they think that their trading partner will benefit more (or considerably more) from the trade than they will. Or if they are able, they will negotiate a fairer deal. In short, they will arguably make only trades that are both mutually advantageous and fair. This concern for fairness is, I take it, a fundamental feature of human social psychology; it is to be accepted without further explanation as a fundamental constraint on contractual exchanges.

Many possible trades are mutually advantageous, increasing both trading partners' probabilities of discovery. But which of these will be regarded as fair?

Consider two proposals. First, a fair trade may be one that preserves the ratio of the programs' race-winning probabilities. Suppose, for example, that before the trade the probability that one program (call it the higher-power program) wins the race is 0.4, the probability that the other (lower-power) program wins the race is 0.2, and the probability that neither attains its goal is therefore 0.4. The higher-power program is, then, twice as likely to win the race as its lower-power competitor; the probability that the higher-power program wins, conditional on a discovery being made at all, is 2/3. Now suppose that the trade changes the probabilities to, respectively, 0.6, 0.3, and 0.1. Then the higher-power program is still twice as likely to win as before; thus, the ratio-preservation criterion counts the trade as fair.

A second notion of fairness defines a trade as fair if it increases both programs' expected utility—their probability of winning the race multiplied by the benefits they gain from doing so—by the same amount. The benefits comprise the credit for making the discovery and all its indirect effects, such as increased funding for the next round of research and the thrill of victory (as well as perhaps the darker thrill of seeing your rivals fail). I assume that they are more or less the same for both programs—that the discoverers of the structure of DNA, for example, receive the same amount of credit, and feel the same level of satisfaction, whoever they happen to be.[8] Equal increases in expected utility, then, correspond to equal increases in each program's probability of winning. In the previous paragraph's scenario, for example, the trade is unfair in the present sense because it increases the higher-power program's probability of winning by 0.2 and the lower-power program's probability of winning by only 0.1—so increasing the higher-power program's expected utility by twice as much as the lower-power program's.

The second view of fairness is, I suggest, the one more likely to govern information trading and other arrangements in science. Or at least, it is the notion of fairness that we humans take to apply to contractual arrangements freely entered into: each party should put up goods of equal value. Thus a fair trade of information is one that increases each program's probability of winning by the same amount.

Three remarks. First, a program may under some circumstances be willing to make an unfair trade. Such decisions will depend on many factors, perhaps most of all the question whether it is in the programs' power to make a fairer trade. In some cases—if the exchangeable information comes in large, indivisible chunks—no fair trade may be possible, in which case the programs may enter into a mutually advantageous but unfair contract without the burden of resentment.

Second, the treatment of one-off contracts is easily extended to the case where there are other research programs engaged in the same race. A neutral trade may be mutually advantageous to the two traders if it improves their prospects, not relative to one another, but relative to a third program.

Third, an information exchange might be desirable not only if it increases the probability of the traders' making a discovery but also if it decreases the expected time taken to make the discovery: the sooner the discovery, the sooner the arrival of the endowed chair, the devoted graduate students, the

8. I assume also, needless to say, that the utility of the scenario in which neither program makes the discovery is the same for both.

corner office, and the big prize. Two programs will want to engage in a fair trade, then, if it decreases their expected waiting times to discovery, even if their race-winning probabilities remain the same.

To summarize, there are a range of circumstances under which, even before discovery and publication, contractual exchanges of information among programs will be beneficial to all parties to the exchange, and will be considered fair.

3.2 *Open-Ended Exchange*

Suppose that two research programs consider abandoning a regime of one-off exchanges for something more committing: an open-ended agreement to share information. Under what circumstances might they enter into such a contract?

Assume as in the previous section that the two programs are locked in a race to make the same discovery. In a one-off contract, what information is traded for what can be to a great degree specified in advance; the trade can typically be calibrated, then, so that both programs find it fair and mutually advantageous. The open-ended contract will of necessity be far less specific about what will be shared. What forms might it take?

I will consider two possible arrangements:

1. *Balanced sharing:* The programs undertake to share about the same amount of information with each other, and

2. *Total sharing:* The programs undertake to share any and all relevant information.

The first of these requires a little more structure; how to arrange things so that the equality of exchange is maintained? But I will not fuss about the details, as it will turn out that the second kind of agreement—the contractual equivalent of Merton's communism—is fairer and hence more agreeable to all parties.

How to compare the two proposals? Either, if implemented, will result in a certain number of discrete acts of sharing, concluding when the race to discovery is won (or peters out in failure). Consolidate these, so thinking of the consequence of the open-ended agreement as a single exchange in which each program gives to the other a package containing the sum total of the information disseminated over the duration of the race. An open-ended contract is desirable if this consolidated exchange is mutually advantageous and fair, using those terms just as I did above (but ignoring for simplicity's sake improvements in expected time to discovery):

> *Mutual advantage:* Both programs' probabilities of winning the race go up at least slightly (at the expense of the probability that no program wins).
>
> *Fairness:* Each program's probability of winning the race goes up by approximately the same amount.

So conceived, the open-ended contract is structurally identical to a one-off contract. The difference lies in the uncertainty surrounding the consequences of the agreement. Programs signing an open-ended contract will have to use some rather general principles to predict whether they are getting a fair and advantageous deal. We social scientists of science will have to do the same.

Let me begin by assuming that any exchange of information between research programs racing to make a discovery will diminish the probability that neither program makes the discovery, and that the complementary increase in some program's making a discovery benefits both programs at least a little.[9] It follows that any open-ended exchange will be mutually advantageous to both programs. The remaining, and more difficult, question is which open-ended agreements will result in fair exchanges.

9. For circumstances under which, despite what I have said in the main text, one of the programs does not benefit from such an exchange, see note 11. See also note 7, although the concern discussed there is rendered moot by the fairness requirement.

Break down into two components the probability that a research program wins the two-way race to discovery. The first component is the probability that our program would make the discovery in question, were it allowed to continue until it had exhausted its intellectual and physical resources. The second component is the probability that, if both our program and its rival do succeed in making the discovery, our program does so first. Call the probabilities respectively our program's *discovery probability* and its *race-clinching probability*. To continue the race metaphor, the discovery probability is the probability that the program crosses the finishing line; the race-clinching probability is the conditional probability that, if both it and its competitor both cross the line, it crosses before its competitor.

Two remarks. First, do not confuse the race-clinching probability with the unconditional probability of winning a race. This latter, *race-winning probability* is a function of the discovery and race-clinching probabilities of both programs (as explained in section 6.1).

Second, in a scientific race, a program that is on track to make the discovery will normally abandon its research if it is scooped—if the discovery is made first by another program—since the priority rule awards no credit to runners-up. The event to which the discovery probability is attached, then, has a modal aspect, reflected in my language above: it is not the probability that the program actually makes the discovery (crosses the finish line), it is the probability that it would make the discovery (it would cross the line) if it were to stay in the race while remaining ignorant of its competitor's success.

The discovery and race-clinching probabilities may in principle vary somewhat independently of each other. There might be two strategies for realizing a goal, one slow and steady and the other risky but fast. A program adopting the first strategy will have a high probability of crossing the finishing line but will take a long time to do so; a program adopting the second strategy will have a low probability of crossing the finishing line, but if they get there at all they will get there quickly. Nevertheless, there are many factors that similarly impact both probabilities. The more numerous a program's scientific workers, or the more reliable its equipment, the more likely it is to reach its goal and to do so quickly. Similarly, programs that start out with some false assumptions are less likely to reach their destination, and even if they reach it, more likely to take their time, than programs that start out with mostly true assumptions (depending of course on the role of the assumptions). For these reasons, I will assume that a program's chances of success and its speed are correlated.

Suppose, then, using the same language as in the previous section, that each program has a certain degree of "power", and that a more powerful program is both more likely to make a discovery and more likely, if it makes a discovery, to make it relatively quickly.

Suppose also that if two programs exchange equal amounts of information, their respective discovery probabilities increase by the same amount.[10]

In the light of these assumptions, let me consider in turn the two possible open-ended information-sharing agreements laid out above: the balanced sharing agreement, according to which programs share equal amounts of information, and the total sharing agreement, according to which programs share all relevant information.

If programs share equal amounts of information then their discovery probabilities increase by equal amounts. You might think that this means that their probabilities of winning the discovery race increase by equal amounts too, in which case the balanced sharing rule results in trades that are fair for everyone. But this is not quite correct.

To see why, consider an especially simple model of a discovery race. The model assigns to each program a power between zero and one, and stipulates

10. There is reason to think that a higher-power program will benefit more from the same amount of information: it is in a better position to exploit that information. But there is also reason to think that a lower-power program will benefit more: it is less likely to have already in its possession an alternative solution to the problem solved by the information (a specific case of the phenomenon of diminishing marginal returns). Perhaps these roughly balance out, yielding something like the same information/same increase principle proposed in the main text.

the following consequences of power:

1. A program's discovery probability is equal to its power.

2. The higher-power program has a race-clinching probability that is greater than one-half; since race-clinching probabilities sum to one, the lower-power program has a race-clinching probability of less than one-half.

These assignments can be thought of as notionally effecting the following protocol for determining the winner of a discovery race:

1. Each program has a probability equal to its power of being placed into the pool of would-be winners.

2. From the pool of would-be winners, the actual winner is determined by the race-clinching probabilities.

Being placed into the pool represents, then, the outcome to which a program's discovery probability is attached: its making the discovery eventually under its own steam. If both programs make it into the pool, then being chosen from the pool as the winner represents the outcome to which a program's race-clinching probability is attached: its being the first to make the discovery, given that both programs make it (or rather, would make it) at all.

In what follows, I examine the consequences of an especially basic version of the simple model in which the higher-power program's race-clinching probability is equal to one, implying a zero race-clinching probability for the lower-power program. When both programs make it into the would-be winners' pool, then, the higher-power program invariably wins the race.

Suppose that two programs follow the balanced sharing agreement, exchanging equal amounts of information. By assumption, they gain equally as far as their discovery probabilities—their chances of admission to the would-be winners pool—are concerned. But what is the impact on what ultimately matters, their probabilities of winning the race? If the value of admission to the pool changed commensurately for both programs, or remained the same, then this equal gain in the chances of admission would bring about equal increments in the race-winning probabilities, and hence the expected utilities, of both programs. But the value of admission to the pool changes unequally: it stays the same for the higher-power program, but it decreases for the lower-power program.

Why? The value of admission to the pool does not change for the higher-power program because the higher-power program's race-clinching probability is equal to one: when it makes it into the pool, it always wins. The lower-power program wins, by contrast, only when it makes it into the pool and the higher-power program does not. Its expected value of admission to the pool decreases, then, as the higher-power program's probability of admission increases.

In short: as the discovery probabilities of the two programs increase, their probabilities of admission to the would-be winners' pool increase commensurately, but whereas the value of admission to the pool remains the same for the higher-power program, it decreases for the lower-power program. Consequently, equal increments of discovery probability do not benefit the higher- and lower-power programs equally; the higher-power program does better. A trade of equal amounts of information between the two programs is for this reason unfair; a fair trade would require the higher-power program to put up more information than the lower-power program.[11]

The same is true for more realistic parameterizations of the simple model in which the race-clinching probabilities lie between zero and one (section 6.2). What matters is that, as programs share information, there are more likely

11. When the discovery probabilities are already high, an equal exchange may even result in an absolute decrease in the lower-power program's race-winning probability. Exchanges of information that increase discovery probabilities are not always advantageous for both programs, then. In the main text I quietly pass over this possibility, though it provides a further consideration against balanced sharing.

to be multiple would-be winners. A high-power program, with its higher race-clinching probability, does relatively better in such situations than a low-power program, and so will tend to benefit more from an equal exchange of information. To share equally is not to share equitably; on the contrary, fair sharing requires more powerful programs to give more—an observation that I will call the *Marxian precept*. "From each according to its ability…"

Fair-minded researchers will, as a consequence of the Marxian precept, refuse to enter an open-ended contract that requires balanced sharing. How will they regard the other contract suggested above, the total sharing agreement that requires programs to exchange all information in their possession? Suppose that the amount of information possessed by a program is proportional to its power. Then the effect of a total sharing contract is that lower-power programs will share less information and higher-power programs more. But that is just to implement the Marxian precept, and so to do what is needed—qualitatively speaking—to effect an open-ended exchange that is not only mutually advantageous but fair.

3.3 Toward Universal Exchange

If an open-ended agreement to share any and all information can make sense between two research programs of unequal power in direct competition, why not among many research programs, some going head to head, some partially competing, and many related only in passing? Might not such an arrangement seem to most or all researchers likely to result in exchanges that on balance improve everyone's prospects while implementing the Marxian precept? If so, they should be willing to sign a social contract mandating total sharing—the Hobbesian equivalent of science's communist norm.

To get from the fairness of a two-way open-ended exchange to the fairness of a universal social contract, various bridges must be crossed. Does it make a difference, for example, that many programs participating in the great information exchange are no longer competing directly? Not obviously, since the less the competition, the less the worry that by giving away information, you are giving your rivals an unfair advantage. A more serious question concerns the enforcement of the contract—though the answer to this question depends rather a lot on what you aim to do with the Hobbesian vindication.

Concerning these issues, I have almost nothing further to say in this chapter. They must eventually be tackled, but it is far more important to examine the weak points in the argument for total sharing in the two-program case.

Let me review the assumptions driving that argument:

1. Higher-power research programs have higher discovery and race-clinching probabilities.

2. An information exchange increases a program's discovery probability by an amount proportional to the amount of information exchanged.

3. The simple model yields a realistic estimate of the changes in research programs' expected utilities effected by various trades.

4. The amount of information that a research program has to share is roughly proportional to its power.

The first three assumptions establish the Marxian precept, the qualitative claim that fairness requires higher-power programs to share more than lower-power programs, and the fourth takes you from the Marxian precept to the approximate fairness of total sharing.

The latter three assumptions are rather simplistic, but it is not ridiculous to suppose that (2) and (4) hold on average over a large number of information transfers, given a suitable notion of "amount of information", or that they are as good as anything available for the purposes of deliberating about whether to sign an open-ended contract. The third, however, looks to be badly wrong.

In the simple model, the two programs' race-clinching probabilities—their probabilities of winning the race, given that both make it into the pool of would-be winners—remain the same after the information exchange. (I set the probabilities equal to zero and one in the exposition in the main text; in the generalization in section 6.2 they may take on any values provided that the higher-power probability is greater than the lower-power probability, but still they are held constant through the exchange.) However, you might reasonably surmise that the higher-power program's race-clinching probability would decrease and the lower-power program's probability would increase as a result of the exchange, on the grounds that, if the two programs' discovery probabilities increase by the same amount, the ratio of the probabilities decreases, and thus the potential (in some sense) of the lower-power program increases relative to that of the higher-power program. Suppose, for example, that the discovery probabilities of the two programs go from 0.4 and 0.2 respectively to 0.6 and 0.4. Then the lower-power program goes from having half the potential (in some sense) of the higher-power program to having two-thirds its potential. Why not think that it enjoys also a commensurate increase in relative speed?

This would work to nullify at least partly the effect so important in the argument for the Marxian precept, in which entry to the pool of would-be winners becomes less valuable for the lower-power program as the discovery probabilities of the two programs increase, so would in part undo the rationale for total sharing.

To address this worry, and to better understand the structure of scientific races more generally, let me construct a more sophisticated model of discovery and information exchange. It will turn out that, though the simple model is itself too simple, its key property—the preservation of race-clinching probabilities necessary to secure the argument for the Marxian precept—is surprisingly robust.

4. A Waiting-Time Model of Discovery

4.1 The Discovery Density

The subject matter of the new waiting-time model is the same sort of idealized race to make a single discovery that was treated in previous sections. To review the essential assumptions about such a race: Each research program in the race has as its only goal to make the designated discovery. If and only if it does so before any other program, it "wins" the race. There are two ways the program might fail to win the race: it might because of faulty assumptions or some other mishap not be in a position to make the discovery at all, that is, it might fail to cross the finishing line (a possibility quantified by the program's discovery probability), or it might for systematic reasons or from sheer bad luck fail to make the discovery soon enough, crossing the finishing line behind some other program with the same goal (a possibility quantified by the program's race-clinching probability).

The waiting-time model associates with each research program a probability distribution representing the chance of its making the discovery in any given time period, which I call the program's *discovery distribution*. The discovery distribution can be represented by a probability density function $f(t)$, defined so that the probability that the program makes its discovery between any two given points of time is equal to the area under the density between the corresponding values of t (figure 1).[12]

The probability that the program makes its discovery before time T, then, is the area under the density function to the left of T. The probability that the program, if allowed to go on indefinitely, at some point makes its discovery—what I have been calling the program's discovery probability—is the total area under the density. This probability is typically, I will suppose, less than one.

12. I assume that the density satisfies the constraints obviously implied by this interpretation: it is zero for values of t less than zero and the area under the density is not greater than one. Because the area may be, and usually is, less than one, the discovery density is not strictly speaking a probability density. But in every other respect it behaves like one.

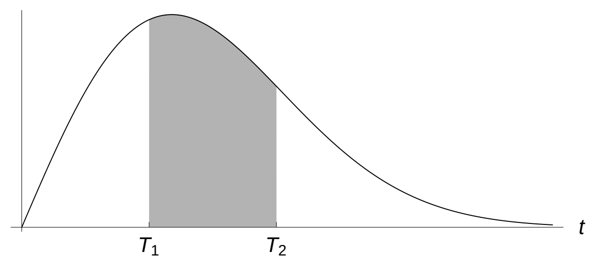

Figure 1: The discovery distribution: the probability that the program makes its sought-after discovery between T_1 and T_2 is the area under the density between those two points (shaded)

The model itself does not assume any particular interpretation of the discovery distribution, but since the probabilities are supposed to explain what exchanges of information scientists regard as mutually advantageous or fair, I will suppose that the distribution represents scientists' subjective probabilities at a point in time where a trade is being considered. For simplicity's sake, I will assume that all scientists engaged in a race see the situation in the same way, and so have the same subjective probabilities at any point in time. Nothing nearly so strong is necessary, however; scientists' conclusions about sharing will turn out to be much the same whatever their probabilities. There is no need to assume that the scientists' probabilities remain fixed over time, since the question examined by this chapter is of what trades seem mutually advantageous or fair at any given moment.

As remarked above, the discovery probabilities themselves are inherently subjunctive. The area under the density is the probability that a program *would* make its discovery, if allowed to go on until it reaches its natural end. When in a priority race one program makes the discovery, the other program, having nothing to gain from further research, typically gives up. In that case, we will never know whether the event of the losing program's making the discovery, if allowed to go on indefinitely, "occurred". But in the model, the probability of this event is perfectly well defined, and the supposition that scientists make use of such probabilities in their deliberations is eminently reasonable, I think, even though they only occasionally have the chance to calibrate them by observing the frequencies of the events in question.

Suppose that several research programs are competing to make the same discovery. Then, on the assumption that their discovery probabilities are stochastically independent (see note 14), their discovery densities determine the values of the race-clinching and race-winning probabilities. A program's race-winning probability, for example, is the sum, for every possible time, of the probability that the program makes the discovery at that time multiplied by the probabilities that each of the other programs fails to make the discovery by that time. (Such quantities are better expressed mathematically, for which see the mathematical development in section 6.3.2 and section 6 more generally.)

The discovery densities of the programs engaged in a particular race, then, comprise a simple but complete stochastic model of the race. It is this model that I will use to investigate the connection between boosts in discovery probabilities and race-clinching probabilities.

4.2 *The Effect of Inflation*

The argument for the fairness of total sharing turned on the argument for the Marxian precept, that is, the qualitative claim that higher-power research programs ought, if sharing is to be fair, to give away more information than lower-power programs. That argument proceeded by examining the fairness of an equal exchange of information, showing that higher-power programs benefit more from such exchanges than low-power programs.

Two assumptions did the work in this demonstration of the unfairness of equal sharing between unequally powerful programs:

1. The race-clinching probability is higher for higher-power programs.

2. When programs trade equal amounts of information, so boosting their discovery probabilities, their race-clinching probabilities do not change.

It is the second of these assumptions that I questioned above. It is true by fiat in the simple model, but that appeared to be a tendentious idealization. I now show that appearances are misleading; the assumption is quite broadly valid.

Suppose, then, that two programs are engaged in a discovery race, one of higher power than the other. They exchange equal amounts of information. The result of the exchange, I assumed above, is an equal boost in discovery probability, in absolute rather than relative terms; for example, both programs' discovery probabilities increase by 0.1. The question: do such boosts alter race-clinching probabilities?

The answer is that it depends on the details. Let me begin by examining what is perhaps the simplest form of boost, what I will call *uniform inflation* (or inflation for short). A uniform inflation increases a program's discovery probability by uniformly "blowing up" its discovery density, or in other words, by multiplying the density by some number greater than one, as shown in figure 2.

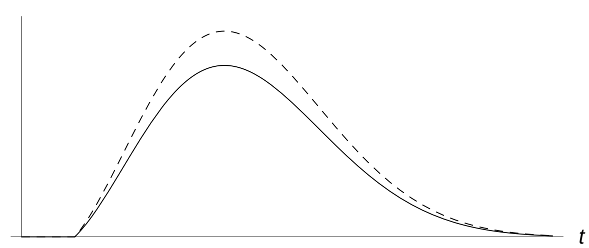

Figure 2: Uniform inflation. The original density is plotted with a solid line, the inflated density with a dashed line.

What effect does inflation resulting in equal boosts have on race-clinching probabilities? None whatsoever. Indeed, the size of the boosts is immaterial. Inflate or deflate two competing programs' discovery densities by any amounts—inflate one and deflate the other, if you like—and there will be no change to the race-clinching probabilities. A proof is given in section 6.3.3; the reason in words is that the clinching probabilities are conditional on both programs making their discoveries, which means that the magnitudes of the discovery probabilities are rendered irrelevant. All that matters is the way that one program's discovery probability is distributed in time relative to the other's, something that inflation leaves unchanged.[13]

Provided that information exchange results only in uniform inflation, the Marxian precept holds true. But why suppose that the effect of an information exchange is confined to inflation? Could what is learned from other research programs not move the discovery density's mass along the axis of time? The next section will be devoted to examining such operations.

4.3 Advancement and Compression

How might new information, or any other improvement in a research program's resources, affect its discovery density if not by inflation? Two simple possibilities are:

> *Advancement:* As a result of the new information, the density is shifted to the left.
>
> *Compression:* As a result of the new information, the density is "squeezed", into a smaller range, while preserving the total area under the density.

---

13. If a program's race-clinching probability does not depend on the area under its discovery density—on its overall probability of making a discovery—but only on the way that whatever probability it has is distributed relative to the probability of the other program, then why suppose, as I did in the argument for the Marxian precept, that the higher-power program has a higher race-clinching probability? I gave some answers in section 3.2, but the discussion of inflation provides good reason to reconsider these carefully.

One answer in particular withstands, I think, reevaluation in the light of the proof. Of two competing research programs, the higher-power program—the program with the higher probability of making its discovery—will tend to attract more scientists (Kitcher 1990; Strevens 2003). With more scientists at work, the higher-power program is likely to move through the steps on the road to discovery relatively more quickly than the lower-power program. Its expected time to discovery conditional on its making a discovery at all will tend to be less than that of the lower-power program, then, and so its race-clinching probability will tend to be higher.

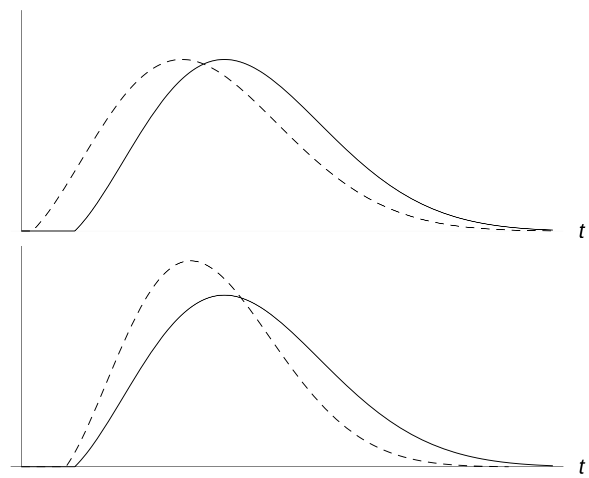

Figure 3: Advanced (top) and compressed (bottom) discovery densities. The original densities are plotted with solid lines; the transformed (i.e., advanced or compressed) densities are plotted with dashed lines.

Whereas inflation affects a program's probability of discovery, advancement and compression affect the expected speed of discovery. To say the same thing more formally, advancement and compression, unlike inflation, decrease a program's expected waiting time to discovery (conditional on a discovery's being made), while leaving the discovery probability itself untouched.

Suppose that, as a result of their exchanging equal quantities of information, two programs' discovery densities are advanced by the same amount, that is, shifted equal distances to the left. What is the effect on the race-clinching probabilities? Obviously, none: the two programs are in the same relative position as they were before. A formal argument is given in section 6.3.4.

The same is true for compression: equal compressions of two programs will leave their race-clinching probabilities unchanged (section 6.3.5). (By contrast with the case of inflation, note, it is crucial for the preservation of the race-clinching probabilities that the degree of advancement or compression is the same for both densities.)

It can be concluded not only that individual inflations, equal advancements, and equal compressions preserve race-clinching probabilities, but that any combination of these operations does the same. If an equal exchange of information results in both an equal compression and an inflation of the programs' discovery densities, for example, the race-clinching probabilities are unaffected and so the exchange benefits the higher-power program more—vindicating the Marxian precept. The precept holds true, then, for a wide range of exchange-induced transformations.

How might that range be expanded?

4.4 *A Multiple-Stage Model of Discovery*

Uniform inflation, advancement, and compression are all what you might call global operations on a discovery density, affecting all regions of the density equally. Might the effect of some information exchanges be more regional? What if a research program's path to discovery has multiple steps, and a certain body of information bears only on one of the steps?

To investigate this possibility, I introduce an enhancement of the waiting-time model that I call the multiple-stage model of discovery; I will use it to show that race-clinching probabilities are remarkably resistant to change even when it is supposed that new information has a purely regional impact.

Consider a research program with two stages. The first stage must be completed successfully in order for the second stage to begin. If the second stage is then completed successfully, the discovery is made.

The two stages are assigned completion densities; these are the analog of discovery densities, giving the probability that a stage is completed within any given period of time (relative to the time that work on the stage begins).

Assume that the two stages are stochastically independent (although of course the second stage will be attempted only if the first stage is successfully completed). Then the program's discovery density is determined by the stages' completion densities as explained in section 6.4.1.

Now suppose that an information exchange boosts the completion probability of the first stage only, by uniformly inflating its completion density. What is the effect on the discovery density for the program as a whole? It is uniform inflation by the same amount (section 6.4.2). Causally, the effect on the program is regional, but probabilistically, it is global, of a sort that we already know has no effect on race-clinching probabilities.

The same is true of advancement: advancing the completion density for the first stage by a certain amount advances the discovery density of the whole by the same amount (section 6.4.3). And also for compression: compressing the first stage's completion density by a certain amount compresses the discovery density by roughly the square root of that amount (section 6.4.4).

The result may be fully generalized: it holds for research programs with more than two stages, and for transformations applied not only to the first but to any such stage (section 6.4.5).

Consequently, inflating, advancing, or compressing the completion densities for single stages of two competing multiple-stage research programs yields the same result, relatively speaking, as applying the operations to the discovery density as a whole. Any inflation of single stages, and any equal advancement or compression of single stages, will therefore leave race-clinching probabilities unchanged.

5. Hobbesian Sharing

Self-interested scientists, I have argued, ought to see the value of participating in a (properly policed) regime of total information sharing. That is the Hobbesian vindication of the communist norm. The vindication relies on a number of assumptions, some prima facie plausible and one—the preservation of race-clinching probabilities through information exchanges—prima facie implausible but nevertheless correct.

Two pressing worries remain. First, I assumed back in section 3.1 that scientists were accurate in their estimates of research programs' prospects—that is, accurate in their estimates of the programs' discovery and race-clinching probabilities, and by extension, of their discovery densities. You might also see the need to assume agreement among scientists on how to measure information, and so on what constitutes an exchange of equal amounts of information, and agreement on the effects of information exchange on the densities. Are scientists really so well informed? Or at least, so well coordinated in their judgments?

The Hobbesian vindication of communism in fact requires very little agreement on these things. Scientists might have quite divergent views about the prospects of this or that program, or about what goes into a quantum of information. Or they might have no particular opinion at all. What my argument for a universal commitment to the Marxian precept requires is that each scientist is able to proceed from whatever assumptions they have to the conclusion that total sharing is considerably more fair than balanced sharing, and considerably more advantageous than no sharing. For that, they need only agree to the premises used in my argument above. I assumed no particular form for the discovery densities, and no particular scheme for measuring information. Rather, I called upon certain considerably more abstract propositions: that discovery probability is correlated with race-clinching probability; that exchanges will result in some combination of inflation, advancement, and compression; that an information exchange that increases the participating programs' discovery probabilities by about the same amount will compress their discovery densities (if at all) by about the same amount; and so on. That is all the scientists themselves need to believe in order to appreciate the merits of total sharing.

It is arguably much easier, then, to reach the conclusion that total sharing is a good idea than to negotiate a particular one-off information exchange—for which agreement on the details of densities and information most definitely is required.

A second and perhaps even more important worry is that the qualitative Marxian precept falls short of establishing the quantitative conclusion that a total sharing contract is totally fair. Balanced sharing may unfairly favor higher-power programs, but it does not follow that the optimal amount of sharing will be achieved by total sharing rather than some other loosely Marxian scheme, that is, some other scheme that asks higher-power programs to share more.

To these concerns the Hobbesian should respond as follows. There are only a handful of universally binding, open-ended sharing contracts that are practically feasible, in the sense that they can be implemented at a reasonable cost given the available resources. "Everyone should share everything" is certainly feasible, as is "No one is obliged to share a thing", which being the state of nature implements itself. Balanced sharing—"everyone should share equally"—is more complicated to put into practice, but arguably still within reach.

It is among the feasible alternatives that the scientist of practical necessity chooses. They can do nothing, in effect choosing the state of nature—no socially organized sharing—or they can choose balanced or total sharing. It is difficult to think of other plausible candidates. (Share half of what you have? Insist on receiving twice as much as you dispense?) The question, then, is not whether one of these is optimal, all practical difficulties aside, but whether one is clearly better than the others, with practical difficulties (including enforcement) figured into the cost.

Though I happily concede that there is little reason to think total sharing to be the fairest policy among all logical possibilities, I have shown it to be clearly more attractive than the other practical possibilities. On the one hand, the exchanges resulting from a total sharing contract will be fairer than those resulting from a balanced sharing contract. On the other hand, they will be far more advantageous (given proper enforcement) to individual programs than the state of nature, and they seem to be fair enough—no scientist has reason to expect that they will be systematically short-changed under a total sharing contract. So Hobbesian scientists should sign the damn thing.

A Hobbesian vindication of the communist norm is a model-theoretic tool rather than an end in itself. How to put it to work? I have two purposes in mind.

The first is to provide fortification for communist practice. Giving away information is never rational from a narrow, self-interested point of view. Various aspects of modern-day science accentuate the possibility of loss: the profit motive for scientists producing commercially useful research; the fight for tenure-line positions and then for tenure for scientists in the university system; and the superstar economy in general. But society has a powerful interest in total scientific sharing. It may be helpful to point out to scientists, then, that in the long term, and with adequate encouragement or enforcement, they will be even in the narrowest sense better off under a continuation of the communist regime.

Observe that the argument for the contract is genuinely Hobbesian, not merely Rawlsian. Total sharing need not be seen through a veil of ignorance to seem beneficial to the deliberator. Even scientists who know quite well that they are in a relatively high-power research program, and who care for nothing but their own welfare, will see the case for the contract—as even the best-armed strongman, Hobbes thought, even the king himself, should see that his life is excruciatingly vulnerable in a war of all against all.

The second application of the Hobbesian vindication is less useful but more ambitious: correctly deployed, it will surely explain the origins of the communist norm. In what way that explanation might proceed is a matter I cannot explore here. There are many questions to answer, and these require the knowledge and skills not only of philosophers but also of historians, sociologists, psychologists, and anthropologists among others. How was the line from a contractual to a moral obligation crossed—or was there never a contractual phase? Did scientists, calculating or intuiting the virtues of a total sharing contract, at some point recent or distant deliberately write it into their moral code? Did the norm arise by some process of cultural evolution inside the petri dish of western European science (in which case it is reality's probability distributions, not scientists', to which the assumptions made in this chapter will have to apply)? Or does it go back further? Perhaps science has inherited from humanity an age-old social norm governing the sharing of a broad class of information, exemplified by but not exhausted by scientific knowledge. Perhaps, along with egalitarianism, communism of a limited sort is buried deep in the human mind. Perhaps we have always believed, to give the communist norm its liberal expression, that information wants to be free.

6. Mathematical Development

6.1 Race-Winning and Race-Clinching Probabilities

Consider a two-way race to make a single discovery between research programs **F** and **G**. Let F be the event that **F** would make the discovery at some point, if allowed to go on indefinitely, and G the corresponding event for **G**. Then FG is the event that both programs would make the discovery, and hence the event of there being what you might call a "live race". There are two possible outcomes for a live race: either **F** makes the discovery first and so wins the race, or **G** does. Call the first outcome F^*G and the second outcome FG^*; the asterisk, then, indicates the winner. If **F** makes (or rather would at some point make) its discovery, then one of three mutually exclusive events must occur: F^*G, FG^*, or $F\neg G$, the last being the case where **G** fails to make the discovery at all. (I ignore dead heats.) Thus:

$$P(F) = P(F\neg G) + P(F^*G) + P(FG^*).$$

The probability $P(F^*)$ that **F** wins the race, live or not, and so gains credit for the discovery, is what I have called **F**'s *race-winning probability*. It is the sum of the probabilities of two outcomes corresponding to the two ways that a program can win a discovery race: first, that **F** makes the discovery and **G** does not; second, that **F** and **G** both make the discovery—the race is "live"—but **F** does so first. In symbols:

$$\begin{aligned} P(F^*) &= P(F\neg G) + P(F^*G) \\ &= P(F) - P(FG^*). \end{aligned}$$

The probability that **F** wins a given live race is what I have called **F**'s *race-clinching probability*; it is the conditional probability $P(F^*G|FG)$. Since F^*G occurs only if FG occurs, the race-clinching probability of **F**, which I will write as $C(F)$, is

$$C(F) = P(F^*G)\big/P(FG).$$

6.2 *The Marxian Precept*

The mathematical component of the argument for the Marxian precept, the demonstration that exchanges of equal amounts of information disproportionately benefit the program with the higher race-clinching probability (assumed to be the higher-power program), can be formalized as follows.

The benefit bestowed upon a program by an information exchange is directly proportional (I assume in the main text) to the resulting increase in its probability of winning the discovery race. What must be shown, then, is that equal boosts to discovery probabilities result in unequal boosts to race-winning probabilities, with the higher-power program benefiting more.

The probability that **F** wins its race is (from section 6.1)

$$\begin{aligned} P(F^*) &= P(F) - P(FG^*) \\ &= P(F) - P(FG)C(G). \end{aligned}$$

Likewise, the probability that **G** wins is $P(G) - P(FG)C(F)$. Suppose that the discovery probabilities $P(F)$ and $P(G)$ are boosted by the same amount. Then the win probabilities go up by that amount, less the right-hand terms in the expressions above—for **F**, the term $P(FG)C(G)$. Because $P(FG)$ increases, this "penalty" term increases for both programs (if both race-clinching probabilities are non-zero). But it increases more for the program with the lower race-clinching probability, by assumption the lower-power program.

What if the distance between the race-clinching probabilities is narrowed by the exchange? This may compensate for the unequally weighted increase in $P(FG)$, resulting in no net change in the right-hand-side "penalties", in which case an equal information exchange would be fair.

If there are more than two programs competing in the discovery race the mathematics gets tedious, with many terms representing the many different possible outcomes, but I believe that the same result holds for qualitatively the same reasons.

6.3 *The Waiting-Time Model*

6.3.1 Discovery Densities Consider a two-way race between research programs **F** and **G** with discovery densities $f(t)$ and $g(t)$ respectively.

The probability that **F** makes the designated discovery before time T is by definition the area under the density function to the left of T:

$$\int_0^T f(t)\,dt.$$

The probability that **F** would make the discovery at some point or other, that is, **F**'s discovery probability, is:

$$P(F) = \int_0^\infty f(t)\,dt.$$

6.3.2 Race-Clinching Probabilities Recall from section 6.1 that **F**'s race-clinching probability is defined to be $P(F^*G)/P(FG)$, where FG is the event that both **F** and **G** make the discovery and F^*G is the event that both programs make the discovery and **F** does so first.

Both $P(F^*G)$ and $P(FG)$ are fixed by the discovery densities $f(t)$ and $g(t)$, if it is assumed that F and G are stochastically independent—an assumption I make throughout this development, rationalized in the accompanying footnote.[14] The probability of FG:

$$P(FG) = P(F)P(G)$$

14. The independence assumption can be only approximately correct. Suppose that the two research programs are built on similar presuppositions; the success of one program in making a discovery would then suggest that the presuppositions are true and so that the other is more likely to succeed. Equally, if they disagree on a presupposition, their discovery probabilities will for the same sort of reason be negatively correlated. Another source of correlation is programs' exchanging "information" that has some chance of turning out to be misleading. As best I can see, these failures of independence will not tend to pull race-clinching probabilities systematically in one direction or another: race-clinching probabilities are quotients of the form $P(F^*G)/P(FG)$, and the correlations discussed in this note will (I surmise) tend to have much the same effect on both numerator and denominator, either increasing both or decreasing both in at least rough proportion and so making little difference overall.

$$= \int_0^\infty f(t)\,dt \int_0^\infty g(t)\,dt.$$

And the probability of F^*G:

$$P(F^*G) = \int_0^\infty f(t)G(t,\infty)\,dt, \tag{1}$$

where $G(a,b)$ is the definite integral of $g(t)$ between a and b.

To investigate the effect of information exchange on race-clinching probabilities, then, is to investigate its effect on the ratio of these two expressions.

6.3.3 Inflation A uniform inflation of a discovery density multiplies it by some number greater than one, transforming it from $f(t)$ to $jf(t)$.

Suppose that two programs' discovery densities are uniformly inflated, not necessarily by the same factor. What is the effect on their race-clinching probabilities? The answer is: none.

Proof: Suppose that **F**'s density is inflated by a factor of j and **G**'s by k. Let $\hat{P}(\cdot)$, $\hat{f}(t)$, and $\hat{g}(t)$ be the new probability distribution and the new discovery densities respectively—that is, the distribution and densities that result from inflation. (Thus, $\hat{f}(t) = jf(t)$ and $\hat{g}(t) = kg(t)$.) The aim is to show that, assuming independence of the discovery distributions, **F**'s post-inflation race-clinching probability $\hat{P}(F^*G|FG)$ is equal to its pre-inflation race-clinching probability $P(F^*G|FG)$.

Because integration is linear, the definite integral functions will receive the same boost as the densities, so that for example the post-inflation definite integral $\hat{G}(T,\infty)$ is the pre-inflation definite integral boosted by k, or in symbols $\hat{G}(T,\infty) = kG(T,\infty)$. Linearity also underlies the proof as a whole:

$$\begin{aligned}
\hat{P}(F^*G|FG) &= \frac{\hat{P}(F^*G)}{\hat{P}(F)\hat{P}(G)} \\
&= \frac{\int_0^\infty \hat{f}(t)\hat{G}(t,\infty)\,dt}{\int_0^\infty \hat{f}(t)\,dt \int_0^\infty \hat{g}(t)\,dt} \\
&= \frac{\int_0^\infty jf(t)kG(t,\infty)\,dt}{\int_0^\infty jf(t)\,dt \int_0^\infty kg(t)\,dt}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{jk\int_0^\infty f(t)G(t,\infty)\,dt}{jk\int_0^\infty f(t)\,dt\int_0^\infty g(t)\,dt} \\
&= \frac{jkP(F^*G)}{jkP(F)P(G)} \\
&= P(F^*G|FG).
\end{aligned}
$$

Inflation by any mix of inflating (or deflating) factors therefore leaves race-clinching probabilities unchanged.

6.3.4 Advancement Suppose two research programs receive an equal advancement, transforming their discovery densities from $f(t)$ and $g(t)$ to $f(t+j)$ and $g(t+j)$ respectively. What is the effect on the race-clinching probabilities?

Consider **F**'s race-clinching probability $P(F^*G)/P(FG)$. Because the advancement has no effect on the areas under either density, there is no effect on $P(F)$ or $P(G)$, hence (given independence) on the denominator $P(FG)$.

Intuitively, there is also no effect on the numerator $P(F^*G)$. For a formal proof, integrate equation (1) by substitution with $u = t + j$, then note that for advancement to make sense, the original discovery densities for both programs must be zero for values of t between 0 and j.

The race-clinching probability is therefore unchanged.

6.3.5 Compression In mathematical terms, $f(t)$ is compressed by transforming it to $jf(jt)$ (for $j > 1$; if j is less than one, you have rarefaction rather than compression). Suppose that two programs' discovery densities are compressed or rarified by the same factor j. Then the race-clinching probabilities are left unchanged.

Proof: Since compression, like advancement, preserves the total area under the discovery densities, what must be shown is that the probability of F^*G is unchanged by the transformation. As before, let $\hat{P}(\cdot)$, $\hat{f}(t)$, and $\hat{g}(t)$ be the transformed probability distribution and the transformed discovery densities respectively. Thus

$$\hat{P}(F^*G) = \int_0^\infty \hat{f}(t)\hat{G}(t,\infty)\,dt.$$

Consider $\hat{G}(t,\infty)$ first. Let $u = jt$ and integrate by substitution with $du = j\,dt$:

$$\begin{aligned}
\hat{G}(T,\infty) &= \int_T^\infty \hat{g}(t)\,dt \\
&= \int_T^\infty jg(jt)\,dt \\
&= \int_{jT}^\infty g(u)\,du \\
&= G(jT,\infty).
\end{aligned}$$

Returning to the probability of F^*G, integrate by substitution with the same variable:

$$\begin{aligned}
\hat{P}(F^*G) &= \int_0^\infty \hat{f}(t)\hat{G}(t,\infty)\,dt \\
&= \int_0^\infty jf(jt)G(jt,\infty)\,dt \\
&= \int_0^\infty f(u)G(u,\infty)\,du \\
&= P(F^*G).
\end{aligned}$$

6.3.6 Races with Many Programs The proofs above are straightforwardly generalized to races in which there are three or more programs competing to make the same discovery. To see this, observe that for three programs **F**, **G**, and **H**:

$$P(F^*GH) = \int_0^\infty f(t)G(t,\infty)H(t,\infty)\,dt$$

and so on. The rest is left as an exercise to the reader.

6.4 The Multiple-Stage Model

6.4.1 Two-Stage Model Let the completion densities for a two-stage research program be $f_1(t)$ and $f_2(t)$. The first density gives the probability that the first stage is completed at any particular time; when it is completed, work begins on the second stage, completion of which constitutes discovery. The time taken to discovery is of course the sum of the time taken to complete each stage.

Given stochastic independence of the stages, the program's discovery density is:[15]

$$f(T) = \int_0^T f_1(t) f_2(T-t)\,dt.$$

6.4.2 Inflation of a Single Stage Suppose that the completion probability of a two-stage model's first stage is uniformly inflated by a factor j, yielding a new completion density $jf_1(t)$. Then the new discovery density for the program as a whole is:

$$\begin{aligned}\hat{f}(T) &= \int_0^T jf_1(t) f_2(T-t)\,dt \\ &= jf(T).\end{aligned}$$

The effect of a boost to the first stage is therefore a uniform inflation of the entire discovery density by the same factor.

6.4.3 Advancement of a Single Stage Suppose that the completion density of a two-stage model's first stage is advanced by a factor j, yielding a new completion density $f_1(t+j)$. Then the new discovery density for the program as a whole is:

$$\hat{f}(T) = \int_0^T f_1(t+j) f_2(T-t)\,dt.$$

15. Because the completion densities are equal to zero for all negative values of t, this definite integral is equal to the definite integral over the entire range of t, that is, from $-\infty$ to ∞. It is therefore the convolution of $f_1(t)$ and $f_2(t)$, a fact I will put to work in section 6.4.4.

Define a new variable $u = t + j$ and apply integration by substitution, obtaining

$$\hat{f}(T) = \int_{j}^{T+j} f_1(u) f_2(T + j - u)\, du.$$

Now consider the effect of advancing the original discovery density $f(T)$ by a factor of j, that is, consider $f(T + j)$:

$$f(T + j) = \int_{0}^{T+j} f_1(t) f_2(T + j - t)\, dt.$$

You will see that

$$\hat{f}(T) = f(T + j) - \int_{0}^{j} f_1(t) f_2(T + j - t)\, dt.$$

Advancing $f_1(t)$ by j makes sense only if the density is zero for values of t less than j (otherwise, you are moving some probability mass into t's negative realm); thus, the right-hand term is zero and so the effect of advancing the first stage of a two-stage project is to advance the discovery density as a whole by the same amount.

6.4.4 Compression of a Single Stage Suppose that the completion probability of a two-stage model's first stage is compressed by a factor j, yielding a new discovery density $\hat{f}(T)$ for the program as a whole. Ideally I would show that the compression of the first stage is equivalent to a compression of the whole, that is, that there exists some factor k such that

$$\hat{f}(T) = k f(kT).$$

This does not hold true generally, but it is approximately true for an interesting and large subclass of cases, provided that the compression factor j is fairly close to 1.

Let $\phi(s)$ be the characteristic function of the original discovery density and $\phi_1(s)$ and $\phi_2(s)$ be the characteristic functions of the first and second stages' original completion densities respectively. Then because the discovery density is the convolution of the completion densities (see note 15),

$$\phi(s) = \phi_1(s)\phi_2(s).$$

When the first-stage completion density $f_1(t)$ is compressed by a factor j, it becomes $jf_1(jt)$. What is the effect on the characteristic function? Let $\hat{f}_1(t)$ denote the transformed function and $\hat{\phi}_1(s)$ its characteristic function. Then

$$\begin{aligned}
\hat{\phi}_1(s) &= \int_{-\infty}^{\infty} e^{ist}\hat{f}_1(t)\,dt && \text{(definition of characteristic function)}\\
&= \int_{-\infty}^{\infty} e^{ist} jf_1(jt)\,dt \\
&= \int_{-\infty}^{\infty} e^{isu/j} f_1(u)\,du && \text{(letting } u = jt\text{)}\\
&= \phi_1(s/j).
\end{aligned}$$

Thus, the characteristic function of the discovery density as a whole becomes

$$\hat{\phi}(s) = \phi_1(s/j)\phi_2(s).$$

We want this to equal the characteristic function of $kf(kT)$ for some k, which is, using the same reasoning, $\phi(s/k)$, and so by the convolution theorem is equal to

$$\phi_1(s/k)\phi_2(s/k).$$

The goal, then, is to show that there is a k such that

$$\phi_1(s/j)\phi_2(s) = \phi_1(s/k)\phi_2(s/k).$$

It will be especially nice if the correct value for k depends only on j.

Now we approximate. Suppose that j is sufficiently close to one that the characteristic functions of the completion densities are approximately linear over the intervals $[s/j, s]$, for all relevant values of s. Restrict the search for k to the range $j \geq k \geq 1$, so that s/k will fall into such an interval.

Over such an interval, it follows from approximate linearity that there exist constants a, b, c, and d such that, for any s in the interval,

$$\phi_1(s) \approx as + c \quad \text{and} \quad \phi_2(s) \approx bs + d.$$

Then

$$\begin{gathered}
\phi_1(s/j)\phi_2(s) \approx \phi_1(s/k)\phi_2(s/k) \Rightarrow \\
(a\frac{s}{j} + c)(bs + d) \approx (a\frac{s}{k} + c)(b\frac{s}{k} + d) \Rightarrow \\
ab\frac{s^2}{j} + ad\frac{s}{j} + bcs + cd \approx ab\frac{s^2}{k^2} + ad\frac{s}{k} + bc\frac{s}{k} + cd.
\end{gathered}$$

One way for this approximate equality to hold is for the following two approximate equalities to hold:

$$\frac{ab}{j} \approx \frac{ab}{k^2} \qquad \text{(constraint A)}$$

$$ad\frac{s}{j} + bcs \approx ad\frac{s}{k} + bc\frac{s}{k} \qquad \text{(constraint B)}$$

Constraint A implies that

$$k \approx \sqrt{j}.$$

I will show that setting k to the square root of j is also sufficient for constraint B to hold approximately. Since the value of k does not depend on the interval, it is good for the whole function, which is what we need.

Solving constraint B's approximate equality (or rather, the corresponding exact equation) for k:

$$\begin{gathered}
ad\frac{s}{j} + bcs = ad\frac{s}{k} + bc\frac{s}{k} \Rightarrow \\
\frac{ad}{j} + bc = \frac{ad + bc}{k} \Rightarrow \\
k = \frac{j(ad + bc)}{ad + jbc}.
\end{gathered}$$

Since a, b, c, and d are constants, there exists some z such that $ad = zbc$. The above expression for k can then be simplified:

$$\begin{aligned}\frac{j(ad+bc)}{ad+jbc} &= \frac{j(zbc+bc)}{zbc+jbc} \\ &= \frac{j(z+1)}{z+j}. \end{aligned} \tag{2}$$

For values of z with a large magnitude, equation (2) is approximately equal to j; for values close to zero, it is approximately equal to 1. For intermediate magnitudes it takes values from just less than 1 to j.[16] Since j is close to 1, any value for k between 1 and j will satisfy constraint B at least approximately; the choice of $\sqrt{j}$ in particular, being about halfway between 1 and j, will do nicely. That choice also, of course, satisfies constraint A.

Note that, when z is approximately equal to 1, equation (2) is approximately equal to the square root of j, so that constraint B is satisfied almost exactly.

Proof: taking z as equal to 1 and squaring equation (2) gives you

$$\begin{aligned}\frac{4j^2}{(j+1)^2} &= \frac{4j^2}{j(j+2+1/j)} \\ &\approx \frac{4j^2}{4j} \approx j\end{aligned}$$

since for j close to 1, $j + 1/j$ is almost exactly equal to 2. When is z close to 1, then? That is, when is ad roughly equal to bc? Suppose that the completion densities for the first and second stages have roughly the same form, modulo an inflationary factor y. Then

$$\begin{aligned} f_1(t) = yf_2(t) \Rightarrow \phi_1(t) &= y\phi_2(t) \\ &\Rightarrow a = yb \quad \text{and} \quad c = yd \end{aligned}$$

16. Exception: equation (2) is badly behaved when z is in the vicinity of -1. On intervals where z is close to -1, however, the terms in the expressions for $\phi_1(s/j)\phi_2(s)$ and $\phi_1(s/k)\phi_2(s/k)$ with the coefficients ad and bc will roughly cancel out, so that for the purpose of finding a good value of k for these intervals, constraint B can be ignored. In that case, setting k to $\sqrt{j}$ so as to satisfy constraint A will be sufficient for what is wanted.

for any interval over which the characteristic function is approximately linear. Thus,

$$ad = yb\frac{c}{y} = bc$$

and so $z = 1$.

Taking stock, we did not get an exact formula for k that depends only on j. We did get something close: setting k equal to the square root of j will give us approximately what was wanted, and when the completion densities for the two stages have roughly the same form (modulo an inflationary factor), almost exactly what was wanted—all provided that the compression factor j is close to 1.

It remains to be asked to what degree the approximations made along the way—in particular the assumption of the linearity in the small of the characteristic functions and the fudging in the satisfaction of constraint B—diminish the interest of the result. As it happens, for many relevant probability densities, such as the gamma distributions often used to represent waiting-time probabilities, the match between the discovery density with a compressed first stage and the compressed discovery density (using the square root approximation) is very close, and almost exact when the functions have the same form.

6.4.5 Generalization The results above are demonstrated only for a certain subclass of multi-stage research programs, namely, those that have two stages and in which the transformation in question—inflation, advancement, or compression—is applied to the first of those stages. The results are, however, easily and fully generalized.

First, transforming the first stage of a two-stage program has the same effect as transforming the first stage of a multi-stage program, since any number of stages after the first can be represented by a single completion density representing the probability distribution over the time taken for all stages after the first to complete. Simply interpret the second completion density in the proofs above as representing such a distribution.

Second, transforming the second or later stage of a multi-stage program has the same effect as transforming the first stage, since the completion densities for each of the stages are multiplied together in the expression for the discovery density, and multiplication commutes.

References

Arzberger, P., P. Schroeder, A. Beaulieu, G. Bowker, K. Casey, L. Laaksonen, D. Moorman, P. Uhlir, and P. Wouters. (2004). Promoting access to public research data for scientific, economic, and social development. *Data Science Journal* 3:135–152.

Blumenthal, D., E. G. Campbell, M. Gokhale, R. Yucel, B. Clarridge, S. Hilgartner, and N. A. Holtzman. (2006). Data withholding in genetics and the other life sciences: Prevalences and predictors. *Academic Medicine* 81:136–145.

Boyer, T. (2014). Is a bird in the hand worth two in the bush? Or, whether scientists should publish intermediate results. *Synthese* 191:17–35.

Collins, H. M. (1974). The TEA set: Tacit knowledge and scientific networks. *Science Studies* 4:165–185.

Contreras, J. L. (2011). Bermuda's legacy: Policy, patents, and the design of the genome commons. *Minnesota Journal of Law, Science & Technology* 12:61–125.

Dasgupta, P. and P. A. David. (1987). Information disclosure and the economics of science and technology. In G. Feiwel (ed.), *Arrow and the Ascent of Modern Economic Theory*, pp. 519–542. New York University Press, New York.

———. (1994). Toward a new economics of science. *Research Policy* 23:487–521.

Heesen, R. (2016). *To Expedite the Flow of Knowledge*. Ph.D. thesis, Carnegie Mellon University.

Hereditary Disease Foundation. (2014). Hereditary disease foundation FAQ. Accessed online 1/21/14 at http://www.hdfoundation.org/faq.php.

Jasanoff, S. (2006). Transparency in public science: Purposes, reasons, limits. *Law and Contemporary Problems* 69:21–45.

Kitcher, P. (1990). The division of cognitive labor. *Journal of Philosophy* 87:5–21.

Louis, K. S., L. M. Jones, and E. G. Campbell. (2002). Sharing in science. *American Scientist* 90:304–307.

Macfarlane, B. and M. Cheng. (2008). Communism, universalism and disinterestedness: Re-examining contemporary support among academics for Merton's scientific norms. *Journal of Academic Ethics* 6:67–78.

Merton, R. K. (1942). The normative structure of science. *Journal of Legal and Political Sociology* 1:115–126. Originally titled "Science and Technology in a Democratic Order". Page references are to the reprint in Merton (1973).

———. (1957). Priorities in scientific discovery. *American Sociological Review* 22:635–659.

———. (1973). *The Sociology of Science*. University of Chicago Press, Chicago.

National Institutes of Health. (2003). Final NIH statement on sharing research data. Posted at http://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html.

Resnik, D. B. (2006). Openness versus secrecy in scientific research. *Episteme* 2:135–147.

Savage, C. J. and A. J. Vickers. (2009). Empirical study of data sharing by authors publishing in PLoS journals. *PLoS ONE* 4(9):1–3.

Strevens, M. (2003). The role of the priority rule in science. *Journal of Philosophy* 100:55–79.

Tenopir, C., S. Allard, K. Douglass, A. U. Aydinoglu, L. Wu, E. Read, M. Manoff, and M. Frame. (2011). Data sharing by scientists: Practices and perceptions. *PLoS ONE* 6(6):1–21.

Wade, N. (1981). *The Nobel Duel: Two Scientists' 21-Year Race to Win the World's Most Coveted Research Prize*. Anchor Doubleday, New York.

Watson, J. D. (1968). *The Double Helix*. Atheneum, New York.