

University of Groningen

## Communism and the Incentive to Share in Science

Heesen, Remco

*Published in:*  
Philosophy of Science

*DOI:*  
[10.1086/693875](https://doi.org/10.1086/693875)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2017

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*  
Heesen, R. (2017). Communism and the Incentive to Share in Science. *Philosophy of Science*, 84(4), 698-716. <https://doi.org/10.1086/693875>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Communism and the Incentive to Share in Science

Remco Heesen\*†

---

The communist norm requires that scientists widely share the results of their work. Where did this norm come from, and how does it persist? I argue on the basis of a game-theoretic model that rational credit-maximizing scientists will in many cases conform to the norm. This means that the origins and persistence of the communist norm can be explained even in the absence of a social contract or enforcement, contrary to recent work by Michael Strevens but adding to previous work emphasizing the benefits of the incentive structure created by the priority rule.

---

**1. Introduction.** The social value of scientific work is highest when it is widely shared. Work that is shared can be built on by other scientists and used in the wider society. Work that is not shared can only be built on or used by the original discoverer and would have to be duplicated by others before they can use it, leading to inefficient double work.<sup>1</sup>

To put the point more strongly, work that is not widely shared is not really scientific work. Insofar as science is essentially a social enterprise, rep-

Received July 2016; revised December 2016.

\*To contact the author, please write to: Faculty of Philosophy, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DA, UK; e-mail: rdh51@cam.ac.uk.

†Thanks to Kevin Zollman, Michael Strevens, Jan Sprenger, Teddy Seidenfeld, Stephan Hartmann, Lee Elkin, Liam Bright, Thomas Boyer-Kassem, Carl Bergstrom, Arif Ahmed, two anonymous referees, and audiences at the Bristol-Groningen Conference in Formal Epistemology, the Logic Colloquium in Helsinki, and EPSA15 in Düsseldorf for valuable comments and discussion. This work was partially supported by the National Science Foundation under grant SES 1254291 and by an Early Career Fellowship from the Leverhulme Trust and the Isaac Newton Trust.

1. Of course scientific work is often duplicated by others even when it is shared (so-called replications). But this is not inefficient in the same way, as after the replication is shared the work is known by all to be more certainly established than if only one or the other instance was shared.

Philosophy of Science, 84 (October 2017) pp. 698–716. 0031-8248/2017/8404-0005\$10.00  
Copyright 2017 by the Philosophy of Science Association. All rights reserved.

representing the cumulative stock of human knowledge, work that other scientists do not know about and cannot build on is not science (cf. the distinction between Science and Technology in Dasgupta and David [1994]). The sharing of scientific work is thus a necessary condition not merely for the success of science but in an important sense for its very existence.

The sociologist Robert Merton first noticed that there exists an institutional norm in science that mandates widely sharing. He called this the *communist norm*, according to which “the substantive findings of science . . . are assigned to the community. . . . The scientist’s claim to ‘his’ intellectual ‘property’ is limited to that of recognition and esteem” (Merton 1942, 121). Subsequent empirical work by Louis, Jones, and Campbell (2002) and Macfarlane and Cheng (2008) confirms that over 90% of scientists recognize this norm of sharing. Moreover, most scientists (if not as many as 90%) consistently conform to the communist norm.

The existence of this norm raises two questions. Where did it come from? And how does it persist? In light of what I said above, these are important questions. A good understanding of what makes the communist norm persist tells us which aspects of the institutional structure of science can be changed without affecting the communist norm. Understanding its origins might allow us to reinstate the communist norm if it disappeared for whatever reason. Insofar as we value the existence and success of science, these are things we should want to know.

Strevens (2017) gives what he calls a “Hobbesian vindication” of the communist norm by showing that scientists should be willing to sign a contract that enforces sharing. The claim is that, from a credit-maximizing perspective, it is not rational for an individual scientist to share her work (which would help other scientists more than her), but every scientist is better off if everyone shares than if no one shares.

In contrast, I argue that in many circumstances *sharing is rational from a credit-maximizing perspective for an individual scientist*. If my argument is successful, it provides a more detailed account of the origins and the persistence of the communist norm than Strevens’s Hobbesian vindication. It also adds to a tradition of work in philosophy and economics that has emphasized how individual scientists’ “selfish” desire to receive credit for their work furthers the aims of science (e.g., Kitcher 1990; Dasgupta and David 1994; Strevens 2003).

Because the existence of a norm can itself change what is in scientists’ interests to do, the sense of “rational” in the statement above needs to be clarified. For this purpose, I rely on the terminology for social norms developed by Bicchieri (2006). I explain this terminology in section 2 and use it to state Strevens’s position more precisely.

Section 3 sets out my own position by explaining how the idea that scientists can publish and claim credit for intermediate results can be used to

establish the rationality of sharing. Sections 4 and 5 make this more precise by describing a game-theoretic model of scientists needing to decide whether to share their intermediate results and establishing conditions under which rational credit-maximizing scientists should be expected to share.<sup>2</sup>

Section 6 fleshes out my explanation of the persistence of the communist norm and considers some objections. I extend my explanation to include the origins of the norm in section 7, which involves considering boundedly rational scientists and some historical evidence. A brief conclusion wraps up the article.

**2. Social Norms and Communism.** The question that this article focuses on is whether it is in a scientist's interest to behave in accordance with the communist norm. More specifically, would it be in scientists' interest to share their work even in the absence of a norm telling them to do so? To clarify the question, I use some terminology defined by Bicchieri (2006). She defines a *social norm* as follows:

Let  $R$  be a *behavioral rule* for situations of type  $S$ , where  $S$  can be represented as a mixed-motive game. We say that  $R$  is a social norm in a population  $P$  if there exists a sufficiently large subset  $P_{cf} \subseteq P$  such that, for each individual  $i \in P_{cf}$ :

*Contingency*:  $i$  knows that a rule  $R$  exists and applies to situations of type  $S$ ;

*Conditional preference*:  $i$  prefers to conform to  $R$  in situations of type  $S$  on the condition that:

(a) *Empirical expectations*:  $i$  believes that a sufficiently large subset of  $P$  conforms to  $R$  in situations of type  $S$ ;

and either

(b) *Normative expectations*:  $i$  believes that a sufficiently large subset of  $P$  expects  $i$  to conform to  $R$  in situations of type  $S$ ;

or

(b') *Normative expectations with sanctions*:  $i$  believes that a sufficiently large subset of  $P$  expects  $i$  to conform to  $R$  in situations of type  $S$ , prefers  $i$  to conform, and may sanction behavior. (Bicchieri 2006, 11)

2. The idea of using game theory to get a better understanding of norms in science goes back at least to Bicchieri (1988).

The crucial feature of this definition is the requirement of normative expectations. This says that an individual's preference to conform to the norm is conditional on others' expectations (possibly enforced by sanctions). For example, norms surrounding the sharing of food are plausibly social norms: in the absence of others expecting them to share, many people might prefer not to share even if they knew most other people shared. In contrast, if an individual knows that in a particular country most people drive on the right side of the road, she would probably prefer to do the same even if others had no expectations about her behavior.

The language of game theory is useful to sharpen these ideas. Recall that conforming to a behavioral rule  $R$  constitutes a (*Nash*) equilibrium if no individual has an incentive to deviate unilaterally; that is, everyone prefers to conform given that everyone else does.

If knowledge of  $R$  and empirical expectations (that others will conform to  $R$ ) are sufficient to make an individual prefer to conform to  $R$ , then  $R$  is an equilibrium of the underlying game  $S$ . But if normative expectations are required, that is, if individuals only prefer to conform to  $R$  if others expect them to conform (and, possibly, are willing to back this up with sanctions), then  $R$  is not an equilibrium of the "original" game: it is only made into an equilibrium by the existence of the norm itself. So the existence of a social norm transforms the underlying game by changing people's preferences, thus creating a new equilibrium (Bicchieri 2006, 25–27).

Is the communist norm a social norm in this sense; that is, are normative expectations a necessary ingredient to make it in scientists' interest to share their work? In order to answer this question, an account of scientists' interests is needed that is independent of the communist norm, so that the question can be asked whether a self-interested scientist would share her work in the absence of a normative expectation.

A scientist's achievements create for her a stock of *credit*. This credit is the means by which she advances her career, which determines both her income and her status in the profession. Insofar as a scientist is someone who is interested in building a career in science, it is then in her interest to maximize credit.<sup>3</sup>

This is not to deny that a scientist may have other interests, either as a scientist (e.g., to advance human knowledge) or apart from being a scientist (e.g., to have time for other pursuits). But these are idiosyncratic, while credit maximization is an interest that all scientists share. This makes it a particularly powerful tool to explain scientists' behavior.

3. This claim has been defended by various philosophers and sociologists of science, including Merton (1957, 1969), Latour and Woolgar (1986, chap. 5), Hull (1988, chap. 8), Kitcher (1990), and Strevens (2003).

The institutions of science put a premium on originality. Credit is awarded to the first scientist to publish some particular result or discovery. This feature of science is known as the *priority rule*, and the extent to which it shapes scientists' behavior is well documented (Merton 1957, 1969; Kitcher 1990; Dasgupta and David 1994; Strevens 2003).

By rewarding only the first scientist, the priority rule encourages scientists to work and publish quickly (Dasgupta and David 1994). In this way, it seems that the priority rule creates an incentive for scientists to share their work. However, "the same considerations give you a powerful incentive not to share your results before you have extracted every last publication from them" (Strevens 2017, 3). If results were shared before publication, this would improve other scientists' chances of scooping important discoveries for which those results are relevant. So, Strevens argues, there is a split in the motivations provided by the priority rule: "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" (3–4). Thus, at the crucial stage at which scientific progress can be sped up by sharing, the priority rule provides no incentive to do so, according to Strevens.

Strevens then goes on to show that a social contract, in which all scientists agree to widely share their work (even before discovery), would be beneficial to all scientists. Putting this all together, Strevens has effectively claimed that the problem of sharing has the structure of a Prisoner's Dilemma: every scientist would be better off if every scientist shared, but each individual scientist has an incentive not to share.<sup>4</sup> The communist norm is thus a social norm on Strevens's view: without normative expectations to transform the game (into something that looks more like a Stag Hunt), widely sharing scientific work is not an equilibrium.

**3. Communism and Intermediate Results.** In this article I argue that, given the priority rule, it is often in a scientist's own interest to share her work widely. In other words, in many realistic situations sharing widely is an equi-

4. Strevens is not the only one to make this claim. For example, Dasgupta and David (1994, 500) maintain that "[the priority rule] sets up an immediate tension between cooperative compliance with the norm of full disclosure (to assist oneself and colleagues in the communal search for knowledge), and the individualistic competitive urge to win priority races." See also Arzberger et al. (2004, 146), Resnik (2006, 135), Borgman (2012, 1072), and Soranno et al. (2015, 70).

librium of the relevant game even in the absence of normative expectations. The problem of sharing is thus not like a Prisoner's Dilemma: the role of the communist norm is not to change scientists' preferences to make sharing attractive (at least not primarily).

An important part of my argument is the insight that major discoveries can often be split into multiple smaller discoveries. Boyer (2014, 18 and 21) gives some examples: the construction of the first laser can be split into a theoretical development and the actual construction based on that theory, and the experimental test of the Einstein-Podolsky-Rosen thought experiment by Aspect, Dalibard, and Roger (1982) was preceded by a number of papers defining and refining the experiment.

In these cases each of the smaller discoveries was published as soon as it was done, rather than after the major discovery was completed. It is not obvious that the scientists involved were acting in their own best interest. While credit can be claimed when a smaller discovery is published, the advantage that the smaller discovery gives on the way toward the major discovery is thereby lost. In fact, Schawlow and Townes seem to have lost the race to build the first working laser at least partially because their publication of the theoretical idea spurred on other teams.

Boyer (2014) provides a model to analyze this trade-off. In his model the benefits of sharing intermediate results outweigh the costs, with costs and benefits both measured in credit assigned via the priority rule. Although Boyer does not specifically discuss the communist norm, his result could be used to argue that normative expectations are not necessary to explain it: the priority rule encourages wide sharing of scientific work even before the potential of future discoveries based on this work has been exhausted.

One may worry that Boyer's result is not general enough to support claims about the origins or persistence of the communist norm. By his own admission, he only shows that "there exist simple and plausible research situations for which the [credit] incentive to publish intermediate steps is sufficient" (Boyer 2014, 29). I aim to show that in fact many if not most research situations are such that there is a credit incentive to publish intermediate results. This requires a more general model, which I call the Intermediate Results Game. I relax Boyer's assumptions that there are only two scientists, that the scientists are equally productive, that different intermediate results are equally hard to achieve, and that scientists share either all or no intermediate results.

The key claim is that, contra Strevens (2017), a social contract may not be needed to enforce sharing. The reason for this is the possibility to claim credit for intermediate results.

**4. The Intermediate Results Game.** The Intermediate Results Game is intended to investigate scientists' incentives when they are working on a proj-



ect that can be divided into a number of intermediate stages.<sup>5</sup> An *intermediate stage* is a part of the project that, when completed successfully, yields a publishable intermediate result in the sense of Boyer (2014, sec. 2). I assume that stages can only be completed in one order.<sup>6</sup> The number of intermediate stages of the project is denoted  $k$ .

A number of scientists  $n \geq 2$  compete to complete this research project.<sup>7</sup> Note that “scientist” may refer to someone working in the natural sciences, the social sciences, the humanities, or any other field in which the priority rule applies. Moreover, teams of collaborating scientists are represented by one scientist in the model.

Whenever a scientist completes an intermediate stage, she has to make a choice: she can either publish the result or keep it to herself. Publishing benefits the scientist, because she thereby claims credit for completing that intermediate stage as well as any preceding stages that remain unpublished, in accordance with the priority rule. The amount of credit is given by the parameter  $c_j > 0$  for each stage  $j$ , with  $C = \sum_{j=1}^k c_j$  denoting the total credit available. Publishing also benefits the scientific community: other scientists no longer need to work independently on the stages that have been published. Publishing thus “expedites the flow of knowledge.” I use  $E$  to denote this strategy.

If the scientist keeps her result secret instead, she can start working on the next stage before anyone else can. This improves her chance of being the first to successfully complete the next stage, thus allowing her to claim credit for more stages later. Holding onto a discovery until a more expedient time might thus be beneficial to the scientist. Call this strategy  $H$ . When a scientist completes the last stage she always publishes, claiming credit for all unpublished stages.

An interesting feature of the priority rule is its uncompromising nature: there are no second prizes, even if the time interval between two discoveries is very small. This feature was noted by Merton (1957, 658), who quotes the French scientist François Arago as saying: “‘about the same time’ proves nothing; questions as to priority may depend on weeks, on days, on hours, on minutes.”

5. Although it was developed independently, the game turns out to be essentially identical to the one studied by Banerjee, Goel, and Krishnaswamy (2014). In sec. 5 I discuss their main theorem, which is roughly speaking a weaker result in a more general model (see Heesen 2017, secs. 4 and 5, for more discussion; this content is also available as an online-only appendix). However, Banerjee et al. do not give a detailed defense of their assumptions, nor do they apply their theorems to explaining the communist norm.

6. This assumption can be relaxed. See Heesen (2017, sec. 5).

7. Compare Merton (1961), who observed that different scientists frequently work on the same research problem, often unbeknownst to each other.



To incorporate this feature into the Intermediate Results Game, it needs to be able to distinguish arbitrarily small time intervals. This suggests that a continuous-time probability distribution is needed to model the *waiting time* (the time it takes a given scientist to complete an intermediate stage): using discrete time units might place two discoveries in the same time unit even though in reality one of them happened (slightly) earlier than the other. For this purpose I use the exponential distribution.

The assumption that waiting times are exponential is equivalent to the assumption that scientists' productivity is a (nonstationary) Poisson process. Empirical work has shown that scientists' productivity fits a Poisson distribution quite well. Huber (1998a, 1998b) has established this for the rate at which patents are produced by inventors, Huber and Wagner-Döbler (2001a) for publications in mathematical logic, Huber and Wagner-Döbler (2001b) for publications in nineteenth-century physics, and Huber (2001) for publications in modern physics, biology, and psychology.

On this basis, I assume that the time scientist  $i$  takes to complete stage  $j$  follows an exponential distribution with parameter  $\lambda_{ij} > 0$ . That is, the probability that it will take scientist  $i$  more than  $t$  time units to complete stage  $j$  is  $e^{-t\lambda_{ij}}$ .<sup>8</sup> The parameter can be interpreted as the speed at which the scientist works. In particular,  $1/\lambda_{ij}$  is the expected time scientist  $i$  needs to complete stage  $j$ . The speed parameter may vary by scientist and by stage, allowing for differences in difficulty between stages and differences in talent, skill, resources, or specialization between scientists.

The exponential distribution has some formal features that I will make use of (Norris 1998, sec. 2.3). First, it is "memoryless." This means that after a certain amount of time has passed and the waiting time has not ended yet, the distribution of the remaining waiting time is just the original exponential distribution. Second, if scientist  $i$  is working on stage  $j_i$  then the waiting time until any one of the scientists finishes the stage she is working on is exponentially distributed with parameter

$$\sigma_{j_1, j_2, \dots, j_n} = \sum_{i=1}^n \lambda_{j_i}.$$

8. Compare this with Boyer's assumption that there is a fixed probability  $\lambda$  that a given scientist will solve a stage in a time unit. As noted above, by using discrete time units this model provides no way of applying the priority rule when two scientists finish the same stage in the same time unit. To address this, suppose each time unit is divided into  $x$  equal parts, and in each part the scientist completes the stage with probability  $\lambda/x$ . The probability that the scientist has not completed the stage at time  $t$  (where  $t$  is measured in the original time units) is  $(1 - \lambda/x)^{xt}$ . A continuous-time model is obtained by taking the limit as  $x$  goes to infinity. Then the probability that the scientist has not completed the stage at time  $t$  is  $\lim_{x \rightarrow \infty} (1 - \lambda/x)^{xt} = e^{-\lambda t}$ . So, in addition to being independently empirically justified, exponential waiting times naturally arise as the limiting case of Boyer's model with continuous time.

In the special case in which all scientists are working on the same stage  $j$ , I write  $\sigma_j = \sum_{i=1}^n \lambda_{ij}$ . Third, the probability that scientist  $i$  is the first one to finish the stage she is working on is  $\lambda_{ij} / \sigma_{j,j_2,\dots,j_n}$ .

In general, whether there is an incentive to share in this game depends on the amount of credit given for each stage and the speed with which the scientists can solve the stages. The results presented in the next section show that sharing is incentivized whenever the following assumption is satisfied.

**Assumption 1** (Proportional Credit). The speed parameters and the credit rewards stand in the following relation: for every scientist  $i$  and for each pair of stages  $j < j'$ ,

$$c_j \lambda_{ij} \geq c_{j'} \lambda_{ij'}.$$

This assumption states that either the credit given for each stage is proportional to its difficulty or earlier stages are awarded more credit than later ones relative to their difficulty.

Is Proportional Credit likely to hold in practice? It may seem reasonable to reward scientists proportional to the difficulty of their contributions. But in practice, rewards for scientific contributions tend to be based on their social value (Merton 1957; Strevens 2003), which may not always correlate with difficulty. Additionally, it may happen that the scientist who finishes the last stage (“puts it all together”) gets a relatively large share of the credit.

From a descriptive perspective, these might be the kinds of cases in which scientists do not share their intermediate results, and the game suggests why. From a normative perspective, perhaps the appropriate conclusion is that Proportional Credit should be enforced. If scientists are rewarded proportionally to difficulty, without extra credit for completing the last stage of a research project, then sharing is incentivized.

**5. The Incentive to Share in the Intermediate Results Game.** The Intermediate Results Game consists of a sequence of (probabilistic) events, in which the scientists can intervene at specific points through their choice of strategy by publishing their work ( $E$ ) or keeping it secret ( $H$ ). In its simplest instantiation there are two scientists ( $n = 2$ ) and the research project has two stages ( $k = 2$ ). The extensive form of the game is given in figure 1.

At the root node Nature decides which of the two scientists is the first one to complete the first stage of the project with the indicated probabilities. This leads to one of two decision nodes marked with a number indicating which scientist makes a decision at this node.

The scientist can choose one of two strategies ( $E$  or  $H$ ), then Nature decides who is the next scientist to complete the stage she is working on, and so on, until one of the scientists completes the second stage. At this point the game

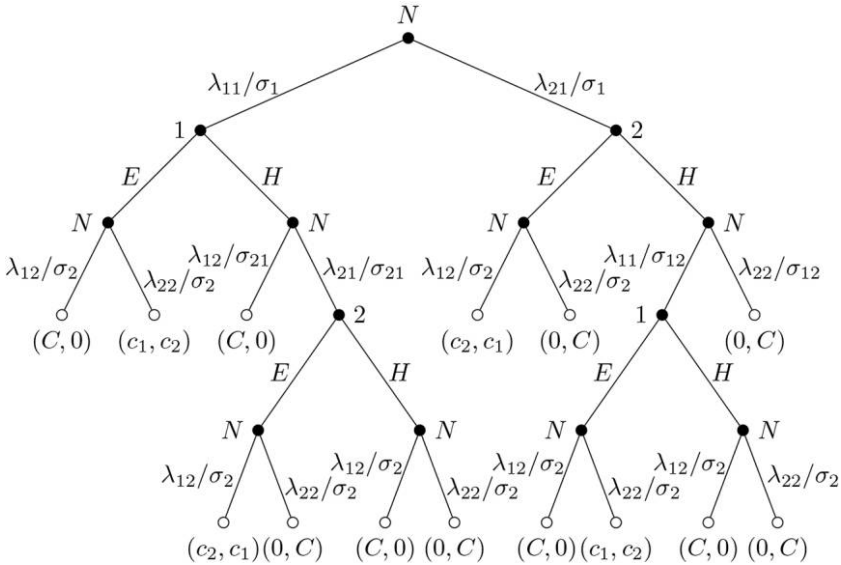


Figure 1. Extensive form of the Intermediate Results Game with perfect information when  $n = 2$  and  $k = 2$ .

ends, with payoff pairs indicating credit awarded to each scientist (see Heesen [2017], sec. 3, for a more detailed explanation of fig. 1).

It is implicitly assumed in figure 1 that each scientist knows when another scientist completes a stage, even when she keeps the result secret. Is it realistic to assume that scientists have this kind of information? It depends. In small fields in which everyone knows what everyone else is working on, word gets around when one of the labs has solved a particular problem, even when they manage to keep the details to themselves. Or with preregistration of clinical trials becoming more common, scientists might know that a particular trial has finished without knowing its outcome.

But in other fields this kind of information might not be available. If this assumption is dropped, scientists are unable to distinguish between certain decision nodes, indicated by *information sets* (see fig. 2). This yields a *game of imperfect information*. In contrast, the version of the game in which scientists can make these distinctions (as in fig. 1) is a *game of perfect information*. I analyze both versions of the game.

One way to find an equilibrium in a game of perfect information is by *backward induction*. This involves identifying what a rational scientist will do at a terminal decision node and then going backward through the tree, identifying rational actions for the scientists by assuming other scientists will play rationally downstream.

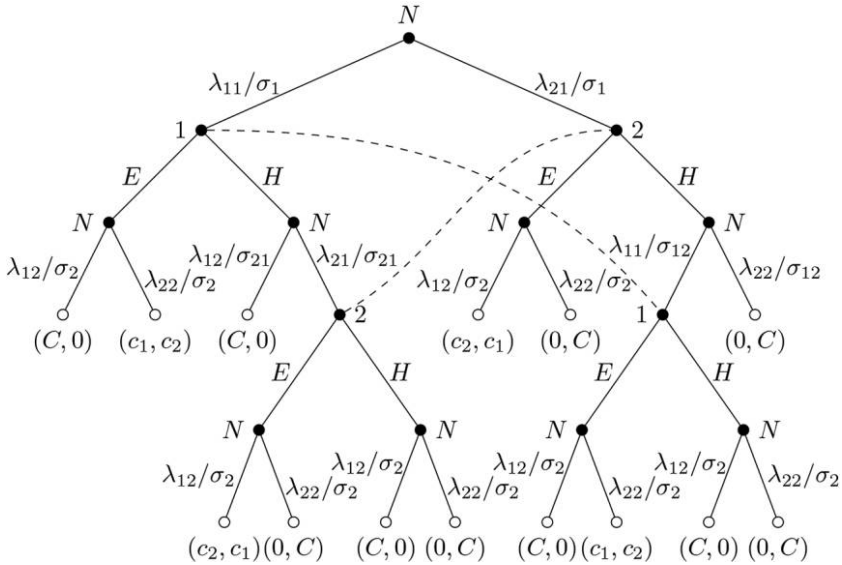


Figure 2. Extensive form of the Intermediate Results Game with imperfect information when  $n = 2$  and  $k = 2$ . Dashed lines indicate information sets.

In figure 1 it is rational for the scientists at the two lower decision nodes to play strategy  $E$ : this yields either the same payoff or a higher payoff than playing strategy  $H$ . Assuming that the scientists play  $E$  at the lower nodes, and assuming Proportional Credit, it is also rational for the scientists at the two higher nodes to play strategy  $E$ . Thus, under Proportional Credit the backward induction solution of this game is for both scientists to play  $E$  at both of their decision nodes.

The following theorem shows that this backward induction analysis also goes through when there are more than two scientists or more than two stages (for proofs of all theorems, see Heesen [2017]). Moreover, any other equilibrium of the game is behaviorally indistinguishable from the backward induction solution. That is, while there may be other equilibria, these differ only in that some scientists make different decisions at decision nodes that will not actually be reached in the game.

**Theorem 1.** Consider the Intermediate Results Game with perfect information with  $n \geq 2$  scientists and  $k \geq 1$  stages, and assume Proportional Credit.

- a) The game has a (unique) backward induction solution in which all scientists play strategy  $E$  at every decision node.
- b) There are no equilibria (in pure or mixed strategies) that are behaviorally distinct from the backward induction solution.

An equilibrium analysis thus yields a unique prediction for the game with perfect information. How about the game with imperfect information? Equilibria can be identified by analyzing the normal form of the game. Table 1 gives the expected credit for each scientist in two examples, one in which the first scientist is thrice as fast as the second and one in which the second stage can be completed thrice as quickly as the first stage (cf. Heesen 2017, table 4.1). Note that because the scientists cannot distinguish between their two decision nodes, only two (pure) strategies are available to them.

Since the credit given for each stage is equal in both cases, example 1 satisfies Proportional Credit while example 2 does not. In example 1, the only equilibrium is the one in which both scientists play strategy *E*, and this is a *strict equilibrium* (a scientist who deviates is strictly worse off). In example 2, both scientists play strategy *H* in the unique and strict equilibrium.

The features of example 1 generalize for different numbers of scientists and stages.

**Theorem 2.** Consider the Intermediate Results Game with imperfect information with  $n \geq 2$  scientists and  $k \geq 1$  stages, and assume Proportional Credit.

- a) The game has an equilibrium in which all scientists play strategy *E* at every information set.
- b) There are no other equilibria (in pure or mixed strategies).
- c) The equilibrium is strict.

Note that theorem 2a was first proved by Banerjee et al. (2014, theorem 2.1). In fact they show that the Intermediate Results Game with imperfect information has an equilibrium in which all scientists share under the following somewhat more general condition: for every scientist  $i$  and for each pair of stages  $j < j'$ ,

$$\frac{c_i(\sigma_j - \lambda_{ij})}{c_j(\sigma_j - \lambda_{ij})} \geq \frac{\lambda_{ij'}}{\sigma_j}.$$

Because Banerjee et al. show neither uniqueness nor strictness of the equilibrium, the following interpretation of theorems 1 and 2 (on which I base my explanation of the communist norm in the next two sections) may not be valid under their more general condition.<sup>9</sup>

Theorems 1 and 2 say that if not every scientist immediately shares any stage that she completes, there is at least one scientist who is irrational in the sense that she would have had a higher expected credit if she had played a

9. Banerjee et al. only prove uniqueness in case a significant proportion of the scientists commits to sharing before the game starts.

TABLE 1. NORMAL FORM OF THE INTERMEDIATE RESULTS GAME WITH IMPERFECT INFORMATION

	Example 1		Example 2	
	<i>E</i>	<i>H</i>	<i>E</i>	<i>H</i>
<i>E</i>	(24, 8)	(26 1/4, 5 3/4)	(16, 16)	(15, 17)
<i>H</i>	(23 1/4, 8 3/4)	(27, 5)	(17, 15)	(16, 16)

Note.—Scientist 1's strategy as the rows, and scientist 2's strategy as the columns. Example 1,  $\lambda_{11} = \lambda_{12} = 3$ ,  $\lambda_{21} = \lambda_{22} = 1$ , and  $c_1 = c_2 = 16$ . Example 2,  $\lambda_{11} = \lambda_{21} = 1$ ,  $\lambda_{12} = \lambda_{22} = 3$ , and  $c_1 = c_2 = 16$ .

different strategy. In other words, if all scientists are rational expected credit maximizers they must all share every stage they complete.

**6. Explaining the Persistence of the Communist Norm.** I take the results from section 5 to give an explanation for the *persistence* of the communist norm, that is, the fact that real scientists publish their intermediate results in a large range of cases. The explanation runs as follows.

Suppose scientists are generally sharing their intermediate results. If a given scientist withholds an intermediate result, she thereby lowers her expected credit (this is just what it means for sharing to be a strict equilibrium). Hence, the scientist has a credit incentive to return to sharing. So credit incentives can correct (small) deviations from the communist norm.

Note that I do not claim that real scientists are rational credit maximizers. All that follows for real scientists is that they have a credit incentive to conform to the norm (even when they fail to act on it). This fact, combined with the fact that real scientists are at least somewhat sensitive to credit incentives, constitutes my explanation of the persistence of the norm.

My explanation relies on three basic principles: scientists' sensitivity to credit incentives, intermediate results being given sufficient credit as specified by Proportional Credit, and the priority rule as the mechanism for assigning credit. These ingredients are sufficient to explain the persistence of the norm. In particular, there is no need for a social contract, normative expectations, or altruism.

This leads to a potential objection. On my construal, the communist norm is not a social norm in Bicchieri's sense, as normative expectations have no role in the explanation. But the available evidence seems to refute this: scientists (normatively) expect other scientists to conform to the communist norm (Louis et al. 2002; Macfarlane and Cheng 2008). This appears to be at odds with an explanation based on the Intermediate Results Game: since the game is zero-sum, other scientists actually benefit when a given scientist fails to share, so from a credit-maximizing perspective they should be encouraging each other to keep secrets.

But the game considers only those scientists who are directly competing on a given research project. While those scientists may stand to gain if their competitors fail to share, the wider scientific community stands to lose, as it will take longer to complete the research project. I claim that this wider community is the source of any normative expectations regarding sharing behavior. The normative expectations can then also be explained from self-interest, as the completion of the research project may benefit other scientists' research.<sup>10</sup>

This yields an empirical prediction that can help decide between Strevens's explanation and mine. On Strevens's explanation withholding an intermediate result is a breach of a social contract that most directly affects the immediate competitors of the scientist within the research project, who may legitimately regard it as unfair. On my explanation withholding actually benefits the immediate competitors; the most direct negative impact is on those scientists who work on nearby projects. An examination of which scientists (direct competitors or those working on nearby projects) tend to object most vocally when other scientists fail to share may thus help decide whether sharing happens out of self-interested credit maximization or as the result of a social contract.

The scope of my explanation is restricted to the sharing of "intermediate results," that is, results that are significant enough to be publishable in their own right. Strevens points out a limitation of this view: "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 shareables [e.g., details of experimental methods or raw data] will remain hidden" (Strevens 2017, 5). This constitutes an objection to my explanation, as according to Strevens the communist norm requires that any and all results should be shared, regardless of their credit worthiness.

I reply that it is not clear that the communist norm makes such strong requirements. When the material under consideration is too little or too detailed to be considered publishable, scientists' actual compliance with a putative norm of sharing drops off steeply (Louis et al. 2002; Tenopir et al. 2011). If Strevens's aim is to explain a norm of sharing for these cases, he may be trying to explain something that does not exist.

This leaves the question of what to do if one wants to encourage sharing work below publishable size, especially in the kinds of cases in which sharing is currently not standard practice. Strevens (2017) shows that scientists have a common interest in establishing a norm of sharing for such cases.

10. Alternatively, normative expectations may arise simply because everyone in the community is behaving in a certain way. Bicchieri (2006, 40) points out that "some conventions may not involve externalities, at least initially, but they may become so well entrenched that people start attaching value to them."



My contribution, in contrast, is in providing a suggestion for how such a norm could be established. If getting scientists to share these minor results or crucial details is a goal that scientists and policy makers consider important, the game gives clear directions on how to get there: give credit for smaller publications and for sharing crucial details (Tenopir et al. 2011; Goring et al. 2014). Modern information technology readily suggests ways in which this can be done without overburdening existing scientific journals (Piwowar 2013).

Strevens (2017) is able to show scientists' common interest in sharing because his model is not zero-sum. This is because Strevens's model allows for the possibility that the research project is never completed by anyone.<sup>11</sup> By sharing their progress, Strevens assumes, the scientists improve each other's chances of completing the research project, thus increasing the total expected credit. As long as this "extra" credit is divided in such a way that everyone benefits at least a little, it is clear that everyone will be better off if everyone shares.

On this point, Strevens's model is arguably more realistic: research projects sometimes fail to reach their goal. It would be interesting to study a model that incorporates both a positive probability of failure and credit for intermediate results. Whether there would be an incentive to share intermediate results under conditions similar to those I have found here is a question I leave for future research.

There are other ways to change the game that would make it no longer zero-sum. For example, Boyer-Kassem and Imbert (2015, sec. 4) argue that one should consider credit per unit of time (rather than "total credit," which I use). Then sharing benefits all scientists to some extent by decreasing the expected completion time of the research project. For present purposes it makes no difference: my theorems still hold if credit is measured per unit of time (see Heesen 2017, sec. 7).

**7. Explaining the Origins of the Communist Norm.** Above I argued that the results from section 5 explain the persistence of the communist norm. It could be argued that they also explain the *origins* of the norm: the uniqueness clauses in theorems 1 and 2 guarantee that behavior in accordance with the communist norm is the only pattern that rational credit-maximizing scientists could settle on (in cases in which their assumptions are satisfied).

But such an argument would make stringent demands on the scientists' rationality that real scientists are unlikely to satisfy. This section investigates the question whether less than perfectly rational scientists would also learn

11. In contrast, in the Intermediate Results Game the scientists complete all  $k$  stages in finite time with probability 1. The models of Banerjee et al. (2014), Boyer (2014), and Boyer-Kassem and Imbert (2015) have the same feature.

to share their intermediate results, thus giving a more robust account of the origins of the communist norm.

To answer this question I consider a boundedly rational learning rule that makes only minimal assumptions on the cognitive abilities of the scientists. In particular, it requires only that the scientists know which strategies are available to them and that they can compare the credit earned under different strategies.

The rule I consider is *probe and adjust*. Suppose the game with imperfect information is played repeatedly. A scientist using probe and adjust follows a simple procedure: on each round (one instance of the game), play the same strategy as the round before with probability  $1 - \epsilon$  or “probe” a new strategy with probability  $\epsilon$  (with  $0 < \epsilon < 1$ ;  $\epsilon$  is usually “small”). In the case of a probe, pick a new strategy uniformly at random from all possible strategies. After playing this strategy for one round, the probe is evaluated: if the payoff in the probing round is higher than the payoff in the previous round, keep the probed strategy; if the payoff is lower, return to the old strategy; if payoffs are equal, return to the old strategy with probability  $q$  and retain the probe with probability  $1 - q$  (with  $0 < q < 1$ ).

Consider a population of  $n \geq 2$  scientists using probe and adjust to determine their strategy in the Intermediate Results Game with imperfect information. Assume the number of stages  $k \geq 1$  is fixed and Proportional Credit is satisfied. Assume all scientists use the same values of  $\epsilon$  and  $q$  (this assumption can be relaxed; see Huttegger, Skyrms, and Zollman 2014, 837–38). Then the following result can be proven.

**Theorem 3.** For any probability  $p < 1$ , if the probe probability  $\epsilon > 0$  is small enough there exists a  $T$  such that, on an arbitrary round  $t$  with  $t > T$ , all scientists play strategy  $E$  at every information set with probability at least  $p$ .

If, on a given round, all scientists play strategy  $E$  at every information set, they may be said to have learned to share their intermediate results. The theorem says that the probability of this happening can be made arbitrarily high by choosing a small enough probe probability. Moreover, the theorem says that once the scientists learn to share their intermediate results they continue to do so on most subsequent rounds. So even on this cognitively simple learning rule both the origins and the persistence of the communist norm can be explained on the basis of credit incentives if Proportional Credit holds.<sup>12</sup>

12. Because the equilibrium in the game with imperfect information is both strict and unique, various other learning rules and evolutionary dynamics can be shown to converge to it. Examples include fictitious play, the best-response dynamics, and the replicator dynamics.

How historically plausible is my claim that credit incentives are responsible for the origins of the communist norm? It is not entirely clear how one should evaluate this question. But a necessary condition for my explanation to be correct is that credit for scientific work, and in particular credit awarded in accordance with the priority rule, predates the communist norm.

As Merton (1957) points out, scientists' concern for priority goes back at least as far as Galileo. In 1610, he used an anagram to report seeing Saturn as a "triple star" (the first sighting of the rings of Saturn). The device of the anagram served "the double purpose of establishing priority of conception and of yet not putting rivals on to one's original ideas, until they had been further worked out" (Merton 1957, 654).

The communist norm, however, was not established as a norm of science until around 1665. At the time, "many men of science still set a premium upon secrecy" (Zuckerman and Merton 1971, 69). The first scientific journals—the *Journal des Sçavans* and the *Philosophical Transactions*, both founded in 1665—were instrumental "for the emergence of that component of the ethos of science which has been described as 'communism': the norm which prescribes the open communication of findings to other scientists" (69).

**8. Conclusion.** In the introduction I argued that the sharing of scientific results (mandated by the communist norm) is important to the success of science and indeed to the existence of science as we know it. My theorems show that the priority rule gives scientists an incentive to share intermediate results whenever these are awarded credit proportional to their difficulty. This can be used to explain both the origins and the persistence of the communist norm, answering the questions I raised in the introduction.

If my explanation is accepted, the crucial features of the social structure of science that maintain the communist norm are the fact that scientists respond to credit incentives, the priority rule, and intermediate results being awarded sufficient credit. Tinkering with these features thus risks undercutting one of the most central aspects of science as a social enterprise.

By emphasizing credit incentives moderated by the priority rule, this article falls in the tradition of Kitcher (1990), Dasgupta and David (1994), and Strevens (2003). As in those papers, I have picked one aspect of the social structure of science and shown how the priority rule has the power to shape that aspect to science's benefit.

I take my results to show that no special explanation (using, e.g., normative expectations or a social contract) is required for the communist norm, contra Strevens (2017). However, this only applies to whatever is sufficiently rewarded with credit. Sharing scientific work that is too insignificant to be published is not incentivized in the same way. But insofar as this is a problem it suggests its own solution: give sufficient credit for whatever one would like to see shared, and scientists will indeed start sharing it.

## REFERENCES

- Arzberger, Peter, Peter Schroeder, Anne Beaulieu, Geof Bowker, Kathleen Casey, Leif Laaksonen, David Moorman, Paul Uhler, and Paul Wouters. 2004. "Promoting Access to Public Research Data for Scientific, Economic, and Social Development." *Data Science Journal* 3:135–52.
- Aspect, Alain, Jean Dalibard, and Gérard Roger. 1982. "Experimental Test of Bell's Inequalities Using Time-Varying Analyzers." *Physical Review Letters* 49 (25): 1804–7.
- Banerjee, Siddhartha, Ashish Goel, and Anilesh Kollagunta Krishnaswamy. 2014. "Re-incentivizing Discovery: Mechanisms for Partial-Progress Sharing in Research." In *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, 149–66. New York: ACM.
- Bicchieri, Cristina. 1988. "Methodological Rules as Conventions." *Philosophy of the Social Sciences* 18 (4): 477–95.
- . 2006. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge: Cambridge University Press.
- Borgman, Christine L. 2012. "The Conundrum of Sharing Research Data." *Journal of the American Society for Information Science and Technology* 63 (6): 1059–78.
- Boyer, Thomas. 2014. "Is a Bird in the Hand Worth Two in the Bush? Or, Whether Scientists Should Publish Intermediate Results." *Synthese* 191 (1): 17–35.
- Boyer-Kassem, Thomas, and Cyrille Imbert. 2015. "Scientific Collaboration: Do Two Heads Need to Be More than Twice Better than One?" *Philosophy of Science* 82 (4): 667–88.
- Dasgupta, Partha, and Paul A. David. 1994. "Toward a New Economics of Science." *Research Policy* 23 (5): 487–521.
- Goring, Simon J., Kathleen C. Weathers, Walter K. Dodds, Patricia A. Soranno, Lynn C. Sweet, Kendra S. Cheruvelil, John S. Kominoski, Janine Rüegg, Alexandra M. Thorn, and Ryan M. Utz. 2014. "Improving the Culture of Interdisciplinary Collaboration in Ecology by Expanding Measures of Success." *Frontiers in Ecology and the Environment* 12 (1): 39–47.
- Heesen, Remco. 2017. "The Incentive to Share in the Intermediate Results Game." Technical report, PhilSci Archive, <http://philsci-archive.pitt.edu/13321/>.
- Huber, John C. 1998a. "Invention and Inventivity as a Special Kind of Creativity, with Implications for General Creativity." *Journal of Creative Behavior* 32 (1): 58–72.
- . 1998b. "Invention and Inventivity Is a Random, Poisson Process: A Potential Guide to Analysis of General Creativity." *Creativity Research Journal* 11 (3): 231–41.
- . 2001. "A New Method for Analyzing Scientific Productivity." *Journal of the American Society for Information Science and Technology* 52 (13): 1089–99.
- Huber, John C., and Roland Wagner-Döbler. 2001a. "Scientific Production: A Statistical Analysis of Authors in Mathematical Logic." *Scientometrics* 50 (2): 323–37.
- . 2001b. "Scientific Production: A Statistical Analysis of Authors in Physics, 1800–1900." *Scientometrics* 50 (3): 437–53.
- Hull, David L. 1988. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: University of Chicago Press.
- Huttegger, Simon M., Brian Skyrms, and Kevin J. S. Zollman. 2014. "Probe and Adjust in Information Transfer Games." *Erkenntnis* 79 (4): 835–53.
- Kitcher, Philip. 1990. "The Division of Cognitive Labor." *Journal of Philosophy* 87 (1): 5–22.
- Latour, Bruno, and Steve Woolgar. 1986. *Laboratory Life: The Construction of Scientific Facts*. 2nd ed. Princeton, NJ: Princeton University Press.
- Louis, Karen Seashore, Lisa M. Jones, and Eric G. Campbell. 2002. "Macroscopic: Sharing in Science." *American Scientist* 90 (4): 304–7.
- Macfarlane, Bruce, and Ming Cheng. 2008. "Communism, Universalism and Disinterestedness: Re-examining Contemporary Support among Academics for Merton's Scientific Norms." *Journal of Academic Ethics* 6 (1): 67–78.
- Merton, Robert K. 1942. "A Note on Science and Democracy." *Journal of Legal and Political Sociology* 1 (1–2): 115–26.
- . 1957. "Priorities in Scientific Discovery: A Chapter in the Sociology of Science." *American Sociological Review* 22 (6): 635–59.
- . 1961. "Singletons and Multiples in Scientific Discovery: A Chapter in the Sociology of Science." *Proceedings of the American Philosophical Society* 105 (5): 470–86.
- . 1969. "Behavior Patterns of Scientists." *American Scholar* 38 (2): 197–225.

- Norris, James R. 1998. *Markov Chains*. Cambridge: Cambridge University Press.
- Piwowar, Heather. 2013. "Altmetrics: Value All Research Products." *Nature* 493 (7431): 159.
- Resnik, David B. 2006. "Openness versus Secrecy in Scientific Research." *Episteme* 2 (3): 135–47.
- Soranno, Patricia A., Kendra S. Cheruvilil, Kevin C. Elliott, and Georgina M. Montgomery. 2015. "It's Good to Share: Why Environmental Scientists' Ethics Are out of Date." *BioScience* 65 (1): 69–73.
- Strevens, Michael. 2003. "The Role of the Priority Rule in Science." *Journal of Philosophy* 100 (2): 55–79.
- . 2017. "Scientific Sharing: Communism and the Social Contract." In *Scientific Collaboration and Collective Knowledge*, ed. Thomas Boyer-Kassem, Conor Mayo-Wilson, and Michael Weisberg, chap. 1. Oxford: Oxford University Press.
- Tenopir, Carol, Suzie Allard, Kimberly Douglass, Arsev Umur Aydinoglu, Lei Wu, Eleanor Read, Maribeth Manoff, and Mike Frame. 2011. "Data Sharing by Scientists: Practices and Perceptions." *PLoS ONE* 6 (6): e21101.
- Zuckerman, Harriet, and Robert K. Merton. 1971. "Patterns of Evaluation in Science: Institutionalisation, Structure and Functions of the Referee System." *Minerva* 9 (1): 66–100.