

On the Epistemic Import of Models of Scientific Communities

Franco Menares-Paredes

July 2025

Abstract

This paper develops a framework to understand the epistemic significance of models of scientific communities, proposing that the outcomes of such models should be treated as evidence capable of informing belief dynamics and guiding policy decisions. Unlike much previous literature—which primarily examines the explanatory or methodological roles these models may play—I argue explicitly for their evidential role. I address common concerns about their highly idealized and abstract nature, suggesting that although traditional measures of evidential quality like representational accuracy and predictive success are often challenging to apply, these models nonetheless provide genuine, context-sensitive evidence. The paper emphasizes the need to assess this evidence comparatively, acknowledging the practical limitations and conceptual difficulties inherent in studying complex epistemic systems empirically. I conclude that, despite justified skepticism regarding their evidential strength in absolute terms, models of scientific communities represent a significant improvement over intuitive or speculative approaches, making them valuable epistemic tools, particularly in contexts characterized by uncertainty and limited empirical access.

1 Introduction

While the issue of the epistemic import of models and computer simulations has received considerable attention in the philosophical literature, much less has been said about models of scientific communities in particular. In this paper, I propose a framework to think about the epistemic import of this latter family of models. While some authors have specifically explored the types of explanations these models may provide (see Šešelja, 2023), and others have concentrated on the methodological role models of scientific communities can play, stressing how simulations might serve as tools for philosophers (see Mayo-Wilson & Zollman, 2021), here, I will focus rather on the issue of how the results of simulations run over these models¹ should shape our beliefs and eventually

¹In some contexts, it may be helpful to distinguish more finely between the model itself, the simulations run on it, and the outputs those simulations produce (for discussion, see Mohseni

provide justification for action. So, for example, after learning of some model's results we may conclude that a certain hypothesis is valuable and plausible enough as to justify allocating resources to extend its study through empirical means. Alternatively, based on the insights gained from certain models, we might conclude that a previously favored line of thought is less promising than initially thought, thereby motivating a redistribution of research efforts within a particular field. There are some cases where models' results may even motivate and justify implementing certain policies. I characterize model results as a kind of evidence—something that influence our beliefs upon learning of them. In what follows, I aim to clarify the precise sense in which models of scientific communities can serve this evidential role, and to consider some concerns that might arise from treating them in this way.

2 Models of Scientific Communities as Evidence

When it comes to ask about the epistemic significance of models of scientific communities, the literature has so far largely focused on discussing what type of explanations these models could possibly provide. However, that is not always what people interested in the social dynamics of science have in mind when engaging with this type of inquiry. One of the pioneers in the use of models to study scientific communities, Philip Kitcher, opens the chapter where he presents his model with the following words aiming to frame his project:

The general problem of social epistemology, as I conceive it, is to identify the properties of epistemically well-designed social systems, that is, to specify the conditions under which a group of individuals, operating according to various rules for modifying their individual practices, succeed, through their interactions, in generating a progressive sequence of consensus practices. (Kitcher, 1995, p. 303).

And then he goes on to set his research question in the following terms:

[H]ow will the whole system best work to promote a progressive sequence of consensus practices? Is it possible that the individual ways of responding to nature matter far less than the coordination, cooperation, and competition that go on among the individuals? (Kitcher, 1995, p. 303).

What Kitcher sets for himself is the task of solving an optimization problem, not of providing an explanation of a particular phenomenon. The issue of what kinds of explanations models of scientific communities might offer does not appear to be relevant to his project. And much like Kitcher's, the work of many other modelers in this area does not seem primarily aimed at explaining

(2024) and Durán (2020)). I will make such distinctions where ambiguity or confusion might otherwise arise; elsewhere, for the sake of readability and economy of language, I will use expressions like “the model,” “simulations,” or “model results” more or less interchangeably.

phenomena relevant to their object of study—at least not except in a derivative way. Much of the work done in this program seems to fit the spirit expressed by Kitcher’s following words:

How do we best design social institutions for the advancement of learning? The philosophers have ignored the social structure of science. The point, however, is to change it. (Kitcher, 1990, p.22)

If not explanation, then how should we think about the role that models of scientific communities are meant to play in inquiries like this? If our primary aim is to study the relationships between different elements of complex systems—such as scientific communities—with an eye toward optimizing their functioning along some dimension of interest, what can models do for us? To me, the natural way to make sense of the place of models in this context is to think of them as putative sources of evidence: learning the results of some model simulations should eventually lead to a change in our beliefs about the system, and possibly even become action-guiding. In this sense, models of scientific communities can be understood as evidence with the potential to shape our beliefs and inform our decisions.

Consider other settings in which people face optimization problems. A paradigmatic example is that of a doctor selecting the optimal treatment for a patient suffering from a particular disease. Doctors rely on the results of randomized controlled trials (RCTs) or meta-analyses to inform their decisions. To determine the best course of action, they review the available evidence, update their beliefs accordingly, and act on the basis of those beliefs. Researchers, for their part—similarly to Kitcher—also aim to solve an optimization problem: what treatment is best? The results of their efforts serve as evidence to inform the doctor’s decision.

Results from experiments such as RCTs are paradigmatic instances of evidence produced to inform decisions, but the notion of evidence is far broader than that. Not only RCTs and meta-analyses serve to assist decision-makers; we also often encounter testimony or expert judgment being treated as evidence in contexts such as legal trials or policy-making. It is therefore important to recognize that the notion of ‘evidence’ need not be tightly bound to “observational” or “experimental” input, as salient examples like the one described in the previous paragraph might suggest. There is a broader sense in which what counts as evidence is determined by its role in belief dynamics. From the perspective of belief dynamics, evidence is whatever can make a difference to an agent’s doxastic state. For example, in standard Bayesianism, “evidence” is simply what agents condition on. Many modeling contexts make it natural to interpret such conditioning events as observational inputs of some kind, but this is not a structural requirement of the framework. There are forms of evidence whose connection to data and observation is far more indirect.

Consider when, in the morning, you check the weather app on your phone to inform your decision about whether to take an umbrella with you before leaving home. You take what the app reports as evidence relevant to that decision.

If the app forecasts rain, your confidence in that outcome will likely increase, and this might eventually prompt you to take the umbrella. Interestingly, what underlies the app’s forecast is most likely a simulation model of the weather. As the results of such simulations are conveyed in the form of a forecast, we treat them as evidence—shaping our beliefs and, in some cases, guiding our actions.

Simulation outputs have the potential of changing our beliefs, influence our decisions and justify courses of action. I maintain that there is no principled difference between weather simulations and simulations of scientific communities that would justify treating the outputs of the former as evidence—as we routinely do—while doubting the conceptual appropriateness of regarding the outputs of the latter in evidential terms. Of course, there are salient differences between these cases that might lead some to suspect the analogy is ill-suited. In what follows, I remain attentive to these concerns. For now, I hope to have convinced the reader that thinking of the outputs of simulations of scientific communities in evidential terms—as potentially playing a role in one’s belief dynamics—is a sensible conceptualization of the issue.

3 Assessing the Evidential Import of Models of Scientific Communities

One thing is whether something is pertinent to shaping one’s beliefs; another is how much of a difference it should actually make. To illustrate: suppose it’s your first day in an unfamiliar city, and you know almost nothing about its weather. You’re deciding whether to take an umbrella before heading out. You check the weather app, but it’s broken and, instead of showing today’s forecast, it displays the weather from the same date last year. Suppose you have no other forecast available. Learning this isn’t entirely beside the point—it’s not like being told the name of the city’s founder—but there are likely other considerations that should carry much more weight, such as looking outside through the window. Suppose you do so and see a sunny day. In that case, learning that it rained on this date last year probably won’t change your mind, though it might slightly adjust your confidence. But suppose instead that you see a cloudy sky. In this case, you would naturally think that it raining is an open option, so learning that it rained on this date last year becomes more relevant. So, in this scenario, the information from the app might carry more weight in shaping your beliefs about whether to take an umbrella.

Even though in most scenarios, learning what the weather was like on the same date last year wouldn’t noticeably change one’s beliefs, the previous example suggests that it might still make a difference in an uncertain scenario. My aim was to illustrate that the evidential import is not a binary matter. In a given context, a piece of evidence might not be the most relevant for shaping beliefs about a particular issue—it might be only weakly informative, unreliable, or ambiguous. But that does not mean it ceases to be evidence altogether. A shift in context may render that same piece of information more relevant. It’s all

a matter of degree and context. Consider a case where the weather app is working properly: in such a case, looking outside the window would carry much less weight than in the previous scenario, where the app was broken. However, even if rendered rather irrelevant, glancing outside would still provide information pertinent to your decision.

The most common worries raised regarding models of scientific communities points to their highly idealized character and its rather loose connection to empirical data (see e.g. Martini & Fernández Pinto, 2017). I think these concerns are fair, but without making the proper nuances we risk to throw the baby out with the bathwater. It is true that models of scientific communities might not provide evidence of the highest quality, but that doesn't mean their epistemic significance is null. We should think of the evidential import of models of scientific communities as a matter of degree. The pertinent question is then how to assess the degree of evidential import that a certain model has. Unfortunately, there seems to be no straightforward answer to this question. In what follows, I will discuss a proposal by Thicke (2020) to assess the epistemic import of models of scientific communities, and I will examine some heuristics that might help us to navigate the issue.

Thicke (2020) proposes assessing models of scientific communities based on their representational accuracy and predictive success. While this might be a reasonable way to evaluate many scientific models, it overlooks some of the difficulties characteristic of models of scientific communities. While in other kinds of models it might be easy to identify the relevant elements and relationships of the system, in the case of models of scientific communities, as we will discuss shortly, mapping the model's elements to its intended target often is not as easy. Likewise, when it comes to assess a certain model's predictive success, we must recognize how difficult measuring the outcomes of these systems is, so comparing measures of the system with its model's results is usually not an option. In these situations, to judge how strong a piece of evidence a model is, we might then resort to considerations about the model's assumptions plausibility or the consistency of the model's results with our observations of the system. However, as these touchstones might as well not be immediately accessible, we might need to find alternative standards.

Models need not always be understood in a straightforward representational way, as Thicke's proposal seems to assume. This complicates the task of assessing the epistemic import of social models of scientific inquiry in terms of representational accuracy. Thinking more broadly about models in terms of evidence give us the flexibility we need to make sense of the nuanced role I propose they should play in shaping our beliefs. A model can provide evidence that support certain kind of interventions in specific contexts but not necessarily in others. Some results might be worth considering while others might be better disregarded as meaningless artifacts. This feature might well be a consequence of the intensive methodological use of this type of modeling. Recognizing that the role of models as methodological tools and their role as evidence informing decisions for both modelers and non-modelers are deeply intertwined is crucial for properly interpreting the epistemic import of the outcomes of this type of

inquiry.

Suppose we learn of a model—built and analyzed by a competent modeler—that suggests that a scientific community that allocates resources by prioritizing scientists with a high volume of publications and paying little attention to quality indicators like journal prestige and number of citations performs better than a community that places greater emphasis on such quality indicators. What should we conclude from that? Suppose you serve in the board of a funding body. At the next meeting, should you propose a policy change to allocate grants based on quantity rather than quality when assessing scientists’ trajectories? The answer is that you probably shouldn’t. Why? Well, because you likely hold a strong prior belief that allocating funds to scientists considering the quality of their work is likely to be more beneficial than just throwing money to whoever has more publications. However, having learned the results of such a model should’ve made you less confident in such a belief, even if only slightly.

Maybe you are a modeler yourself, and even though you respect your colleague, you may believe that the type of model they used is not well-suited for capturing the dynamics you consider relevant for drawing their conclusions. So you attribute the counter-intuitive results to this structural limitation of your colleague’s model. Suppose you decide to model the system using your preferred model type and obtain similar results. This suggests that such results are not just the product of the structural assumptions you find problematic. The results are also unlikely just the product of chance, so you increase your confidence in the conclusion, elevating it to a hypothesis worthy of serious consideration. You decide, then, to study it seriously and start to perform robustness analysis on the structural features of your model.

Your inquiry proceeds by scrutinizing some idealizations shared by both models that might be overlooking mechanisms relevant to the process of knowledge production in the actual system. Suppose both models assume that key events—such as research, submission for publication, journal decision-making, revision, and the assignment of credit—occur instantaneously and simultaneously. You hypothesize, however, that the time these events take should influence the system’s behavior. Yet, after reflecting on potential ways to implement this feature in your model, you realize that any plausible approach would likely reinforce the counterintuitive result. You therefore conclude that this idealization is unlikely to be the underlying cause of the outcome.

Imagine you think of an alternative hypothesis that might explain the results: both models assume that scientists can read articles costlessly and instantly determine their reliability. This simplification omits a mechanism through which prestige hierarchies may contribute to scientific learning—by curating studies. You may think that this mechanism allows scientists to allocate their limited reading time more efficiently, so this feature could play a relevant role in the dynamics of the system. There are several existing indexes that intend to gauge the prestige of journals, but these days, we don’t have a straightforward method to estimate the global reliability of the results published in an article and get some estimation of the correlation between prestige and reliability so it wouldn’t be easy to implement this feature in our model. Here is one important advantage

of the modeling methodology. Where we don't count with reliable empirical inputs to measure some feature or aspect of the system we are interested in, we can explore different modeling alternatives to have some insight on the issue. For instance, Ghomi and Stegenga (2023) have used higher-order characteristics of real trial data, such as reported means and standard deviations of measured outcomes, and simulated patient-level data out of this parameters to test some meta-scientific hypothesis like that some features of medical research contribute to overestimate the effectiveness of medical interventions. For many reasons, the access to such patient-level data is extremely restricted. Setting up a model and run simulations of these scenarios allows the authors to explore the outcomes of some scenarios of special interest. For example, an scenario where the are some distortions in the research process and the true effectiveness of a drug is zero.

The strategy employed by Ghomi and Stegenga (2023) offers one way of addressing empirically elusive aspects of a system. Another approach might involve introducing an *ad hoc* parameter—say, a “reliability index”—stipulated to represent the feature in question. This parameter is implemented in the model to play the role it is hypothesized to have in the system, and a range of values is arbitrarily assigned to it. The model is then tested for robustness across those values. If the results of interest remain stable, this may indicate that the modeled aspect is not crucial to the outcomes we care about, and we may ultimately decide to omit it from the model.

Suppose the interesting results from your model are not robust across different values for the *ad hoc* parameter you included. This generates the challenge of operationalizing and estimating its value. It is important to note that this challenge is not unique to the modeling method we are discussing. Scientists often need to come out with constructs to fill the gaps in the theories that they are interested in exploring. This is pretty common in social sciences. Allegedly, many concepts in the natural sciences are also constructs that arise from similar circumstances and fulfill the same role. This situation also presents modelers with an opportunity to build bridges and suggest avenues for further research to more empirically-focused studies. However, it might be a legitimate decision to just leave your model as it stands and draw the general conclusion that in some, partially indeterminate, scenarios an allocation of research funding that observes productivity alone may be more effective in promoting scientific communities' learning than one that considers putative quality measures.

It is usually expected that counter-intuitive results such as the one in our example be presented with some story that could provide an explanation of the results. Suppose you find the more intuitive way to make sense of the results is by appealing to previous results of similar models that highlights the importance of cognitive diversity among scientific communities (as in Zollman, 2010), so when writing your paper, in the conclusion section, you explain the results appealing to this already established piece of theory. This exercise might be seen as providing further support to the hypothesis that the results capture real features of the system, or as suggesting new ways to understand this previously known effect.

What sort of conclusions might we draw in a scenario like this? First, it

seems to me that the most important and well-supported conclusion is that we should be less confident in our intuitions about the relationship between funding allocation policies and the behavior of scientific communities than we might otherwise be inclined to be. We have evidence suggesting that, in some scenarios, our intuitions could lead us stray from our goal of promoting scientific progress by allocating research funding in the way that seems to us is the most efficient way to do it. So this is an invitation to review the evidence we have for holding the default position and weight it against our new evidence.

We could've also opened up new research avenues. Two salient questions plausibly remain open in the story I have told. First, how can we get a sense of where different scenarios fall within the range of possible values for the parameter that remains indeterminate? Perhaps other researchers can find clever ways to operationalize this parameter so that existing data—or data that could feasibly be collected—might offer insight into its plausible values.

Second, at face value, the model suggests that we may be allocating research resources suboptimally, raising the possibility of revisiting current policies. But how bad would it be if we got it wrong? This question invites further modeling work. The original models may lack the features needed to explore the logical space generated by policy changes in sufficient detail. Yet gaining a sense of the risks involved might be important if we are to take that option seriously. So, even if the model results are not yet robust or reliable enough to justify immediate action on its first-order target, they might nonetheless be strong enough to reshape how we study the system. In this way, we also gain insight into where directing new research efforts is likely to be productive.

I have been mainly discussing ways in which we could ameliorate problems concerning assessing the representational accuracy of models of scientific communities. The proposal of assessing the epistemic import of models in terms of their representational accuracy is a quite natural one to make, but relying too much on it might lead us to mistakenly disregard the epistemic import of models whose representational status of its parts is complicated by the complexity of the system they are intended to represent. I propose that one way to ameliorate this difficulty is by resorting to robustness analysis as it might eventually grant us room to keep certain elements of the model indeterminate without our results losing reliability in case the results of interest are shown not to depend on the value of such elements or certain substructures within the model. Furthermore, indeterminacy can be seen as an opportunity to open new research avenues as it invite us to work out ways to operationalize these indeterminate elements or invites us to explore the logical space of the model in more detail.

But what about predictive success? Should we then appeal to this criterion in assessing the evidential import of models? We must note that the similar considerations that make it difficult to assess the representational accuracy of models of scientific communities also complicate the assessment of their predictive success. For starters, how are we supposed to measure the epistemic performance of a scientific community? To determine whether a community has successfully acquired knowledge, we would need something like God's point of view. One of the advantages of the modeling methodology is that it allows

us to adopt such a perspective within the model. But this, in turn, reinforces the disanalogies that motivate skepticism about the method in the first place.

Perhaps the second-best option is to stipulate some proxy for epistemic success—say, sustained consensus around an answer. But this is only partially satisfactory. For one, a community may converge on the wrong answer, and correcting such an error might take a long time. For another, the timescale on which scientific consensus tends to develop makes it highly impractical to use it to evaluate the predictive success of models. The epistemic performance of a scientific community is not something we can track on a timescale that would allow for timely and regular feedback on model outputs.

A possible way to ameliorate this problem could be to identify non-problematically representational elements of the model—if available—and identify patterns in its dynamics that could be observable in the real system and are likely to be generated by independent assumptions. For example, returning to our thought experiment, we might observe that researchers that prioritize productivity over quality tend to publish in a more diverse set of journals than those who focus more on quality. Assume that the assumptions that shapes scientists’ productivity and quality has nothing to do with the assumptions about publication strategies in terms of the diversity of the journals they choose to submit their work to. If this pattern emerges robustly enough, this might well be considered as a prediction of the model—although not the one we are interested in. Conveniently, it shouldn’t be hard to tell apart the two groups of scientists in the real world, so we could check whether the model’s prediction holds in the real system. If it does, this should increase our confidence in the model. If it doesn’t, we should probably be less confident about it.

Taking stock, assessing the evidential import of models of scientific communities is far from straightforward. Representational accuracy and predictive success—although natural criteria—are often of limited use due to the practical difficulties posed by the very systems these models are intended to represent. The complexity and opacity of their targets, the frequent need to leave some parameters partially indeterminate, and the impracticality of obtaining timely empirical feedback all complicates the extent to which we can reliably assess their significance. These factors makes any assessment of the evidential quality of such models highly uncertain. Still, the preceding discussion shows that dismissing their evidential import altogether would be premature. On a more positive note, I have pointed to some ways of mitigating these challenges. And we should be careful not to conflate uncertainty about the quality of a piece of evidence with that evidence being of poor quality. We should adopt a nuanced perspective: even if models of scientific communities cannot deliver conclusive evidence, they do nonetheless offer valuable insights that are pertinent in shaping our beliefs. These insights shouldn’t be disregarded, especially if we consider that the alternatives are either pragmatically unfeasible or even less reliable. In the next section, I turn to this comparative perspective, arguing that despite justified skepticism, these models often remain among our best available epistemic tools to study their subject.

4 Models in Perspective

We ought to evaluate the epistemic merits of a given methodology in light of the alternatives available to us. While the modeling approach to the study of scientific communities is not without limitations, it nonetheless represents a clear improvement over earlier philosophical approaches, which often relied on intuition and armchair speculation. One of the key values of empirical input in theoretical inquiry is that it provides a source of constraint—a way for the world to “push back” when our theorizing goes astray. In this spirit, modeling enhances the rigor of inquiry by requiring the explicit articulation of assumptions, the structuring of speculation, and the use of mathematical and computational tools to trace inferential consequences. These features provide a form of constraint that, while distinct from direct empirical input, plays a structurally analogous role and should be expected to improve the reliability of investigations into systems as complex and elusive as scientific communities.

It is also worth bearing in mind that our intuitions are notoriously unreliable when it comes to anticipating the behavior of complex systems. We are often surprised by the outcomes of economic policies, ecological interventions, or even traffic flow patterns—domains in which many interacting parts produce results that defy our expectations. This should not come as a surprise: evolution has equipped us to reason effectively about small-scale, local problems, not to intuitively grasp the emergent dynamics of large, interconnected systems. So we have little reason to trust our untutored judgments in these domains.

It must also be noted that researchers working within this methodology are typically keen to incorporate all relevant empirical insights available to them—or at the very least, there is nothing inherent to the approach that prevents them from doing so. The difficulty seems to lie more with the subject matter than with the method itself. As it happens, the kinds of questions that tend to fall within the remit of philosophy are often those that are difficult to access empirically and that demand intensive conceptual work. This may be a consequence of the broader division of academic labour. In the case of the epistemology of scientific communities—as with many questions involving social dynamics—pragmatic constraints make experimentation nearly impossible. Moreover, it is difficult to devise measures that adequately capture the key features of the system. In practice, the empirical inputs we do have are limited to a handful of sources: bibliometrics, citation and collaboration networks, funding and patent records, and career trajectory data, which offer partial insight into the production, dissemination, and impact of scientific knowledge.

If we were interested in modeling the trajectory of a mid-sized physical body, we would have at our disposal a reliable theory identifying the relevant features of the system’s components, along with well-established experimental methods for measuring those features. In macroeconomics, to take another example, we also have relatively mature theoretical frameworks that specify the relevant variables, many of which are routinely measured—think of metrics like GDP, unemployment rates, or wages. While economists often disagree about which models best apply to particular situations, there is reasonable agreement about

what variables are in play and how to quantify them. By contrast, one of the central challenges in studying how scientific communities acquire knowledge is that it is far from obvious what the relevant variables are, or how to measure them. Consider the basic problem of measuring the epistemic performance of a scientific community: it is not clear how this should be defined, let alone tracked.

But in regretting the limited empirical access we have to these systems, we should not lose sight of the fact that empirical studies come with their own problems, and are not always synonymous with reliability. The replication crisis across several disciplines serves as a cautionary tale: we may be equally justified in adopting a degree of skepticism toward empirical findings, particularly when meta-scientific circumstances or methodological choices leave room for distortion. It should not be taken for granted that a counter-intuitive result from a small-sample experiment conducted under narrow and artificial conditions is always more trustworthy than a robust pattern emerging from a well-designed set of models.

Some criticisms of this research program argue that models of scientific communities lack sufficient empirical grounding to inform policy—a role often attributed to them by their proponents (see Thicke, 2020). But such critiques often rely on an overly coarse-grained understanding of both the aims of this modeling tradition and what counts as empirical input in the construction of a model. First, we should recognize that, in one relevant sense, there is no sharp categorical difference between the policy of an individual researcher or team deciding to investigate a certain problem, the policy of approving a new drug for clinical use, and the policy of imposing a tax on a particular class of imports. It is this broad sense of ‘policy’ that I take to be relevant here. I do not think modelers seriously expect that a single model should, on its own, justify a wholesale reorganization of science. But it is not too ambitious to think that a single model might play a role in shaping inquiry—for example, by helping us identify questions worth pursuing or assumptions worth scrutinizing.

Furthermore, given that there is nothing categorically different between the various policy contexts I just described, I see no principled reason why it would be appropriate for models to inform policy in some cases but not in others. I find no compelling grounds to exclude evidence from models and simulations—even highly idealized ones—from the space of considerations relevant to policy-making. It may be reasonable to demand stronger empirical support for policy changes in domains where the stakes are particularly high, such as public health or macroeconomic regulation. But as I have argued throughout, evidential support is a matter of degree. In high-stakes contexts, the weight of the evidence from modeling may well be small, but to disregard it entirely seems unlikely to serve the goal of making informed decisions. At the same time, we should remain open to the possibility that our resistance to certain model results may stem not from careful reasoning but from prejudiced assumptions or faulty inference. If we cannot clearly reconstruct the basis for a belief we find intuitive, and we are confident in the assumptions underlying a model that challenges it, it may be more rational to shift our confidence in the direction

of the model. A commitment to informed decision-making requires more than a default deference to prior beliefs—it calls for scrutiny of those beliefs, and a willingness to revise them when better-supported alternatives are available.

Empirical support is not exhausted by the performance of experiments or by fitting parameter values to measurements. Regardless of how idealized a model’s assumptions are, or how its parameters are set, we must acknowledge that the structure of the model is not arbitrary. If we take a model to represent its target system, it is because we believe—rightly or wrongly—that it captures some aspect of how that system behaves. Mere stipulation is not enough. To the extent that we regard a model as a representation of its target, it carries at least some empirical support, insofar as its structure is taken to reflect, however loosely, the dynamics of the system. Even highly abstract models are constructed under informal empirical constraints: our background knowledge, direct observations, and second-hand reports shape the mental models we rely on when theorizing, and those mental models, in turn, constrain the models we build. In this sense, empirical support, like evidential strength, is a matter of degree. A highly idealized model with made-up parameter values may be only weakly supported, but not entirely devoid of empirical content. After all, many of its core assumptions—for example, that scientists learn from the world, communicate with one another, or try to optimize some performance measure (whether epistemic or instrumental)—are grounded in everyday experience. While it makes sense to hold that a model’s evidential force increases with the strength of its empirical support, it is a mistake to claim that abstract models lack empirical support altogether.

All things considered, the epistemic value of models of scientific communities is best appreciated in light of the alternatives available. While these models may fall short of the standards we apply to more empirically tractable domains, they often represent a clear improvement over purely intuitive or speculative approaches. Especially in contexts where empirical access is limited, and reliable measurement difficult, modeling offers a way to discipline our theorizing, make assumptions explicit, and explore the space of possibilities in a structured and transparent way. Their evidential import may be limited in absolute terms, but relative to the alternatives—particularly in philosophical and conceptual work on scientific practice—they remain among our best tools for making informed judgments.

5 Conclusion

In this paper, I have proposed that models of scientific communities can be understood as providing evidence—evidence that may influence our beliefs and, in some cases, support policy decisions. I argued that the weight such evidence should carry depends on the quality of the model, which is often assessed in terms of representational accuracy and predictive success. However, the complex nature of systems like scientific communities can render these criteria difficult to apply. Still, this should not lead us to dismiss the evidential value of such models

altogether. Tools like robustness analysis and the identification of patterns in representational components can help us assess their reliability.

Two considerations are worth keeping in mind. First, our alternatives for studying these phenomena are few and often no better equipped to incorporate the features that would make our conclusions more reliable. Given the challenges of investigating epistemic systems, it is difficult to deny that modeling is a useful—perhaps even the best available—tool for the task. Second, in contexts where access to the system is limited and empirical inputs are imperfect, results from isolated models may not justify radical belief revision or high-stakes policy shifts. But this does not mean such results should be dismissed altogether. They may still be informative in low-stakes contexts—for instance, in decisions about which research avenues to pursue—or when considered alongside other forms of evidence.

There is nothing in principle that prevents this methodology from incorporating richer empirical inputs, and models must necessarily reflect at least some empirical constraints if we are to recognize them as representations of their intended targets. Concerns about their empirical detachment may stem from overly ambitious expectations—expectations that may not be realistic given the nature of the systems under investigation. Taking a more measured perspective should help us better appreciate the relative merits of modeling in this domain. I invite readers open to utilitarian considerations to reflect on the relatively low social cost of maintaining such a research program, especially when weighed against its potential to support better-informed decisions and to improve the overall efficiency of our scientific enterprise.

References

- Durán, J. M. (2020). What is a Simulation Model? *Minds and Machines*, 30(3), 301–323. Retrieved from <https://doi.org/10.1007/s11023-020-09520-z> doi: 10.1007/s11023-020-09520-z
- Ghomi, H. T., & Stegenga, J. (2023). Simulation of trial data to test speculative hypotheses about research methods. In K. Hens & A. de Block (Eds.), *Advances in experimental philosophy of medicine* (pp. 111–128). Bloomsbury Academic.
- Kitcher, P. (1990). The division of cognitive labor. *Journal of Philosophy*, 87(1), 5–22. doi: 10.2307/2026796
- Kitcher, P. (1995). *The advancement of science: Science without legend, objectivity without illusions* (1. issued as an Oxford Univ. Press paperback ed.). New York: Oxford University Press.
- Martini, C., & Fernández Pinto, M. (2017). Modeling the social organization of science. *European Journal for Philosophy of Science*, 7(2), 221–238.
- Mayo-Wilson, C., & Zollman, K. J. S. (2021). The computational philosophy: Simulation as a core philosophical method. *Synthese*, 199(1-2), 3647–3673. doi: 10.1007/s11229-020-02950-3
- Mohseni, A. (2024). Methods for modelers of science. In J. Veigl & A. Currie (Eds.), *Methods in the philosophy of science: A user’s guide*. Cambridge, MA: MIT Press.
- Thicke, M. (2020). Evaluating formal models of science. *Journal for General Philosophy of Science / Zeitschrift für Allgemeine Wissenschaftstheorie*, 51(2), 315–335. doi: 10.1007/s10838-018-9440-1
- Zollman, K. J. S. (2010). The epistemic benefit of transient diversity. *Erkenntnis*, 72(1), 17–35. doi: 10.1007/s10670-009-9194-6
- Šešelja, D. (2023). What kind of explanations do we get from agent-based models of scientific inquiry? In H. Andersen, T. Marvan, H. Chang, B. Löwe, & I. Pezlar (Eds.), *Proceedings of the 16th international congress of logic, methodology and philosophy of science and technology*.