Standards and the distribution of cognitive labour

A model of the dynamics of scientific activity

Rogier De Langhe & Matthias Greiff

Rogier De Langhe
Centre for Logic and Philosophy of Science
Ghent University, Belgium
Rogier.DeLanghe@UGent.be

Matthias Greiff
Department of Economics
University of Bremen, Germany
Tel: ++49 421 218 3595
greiff@uni-bremen.de

Abstract: We present a model of the distribution of labour in science. Such models tend to rely on the mechanism of the invisible hand (e.g. Hull 1988, Goldman & Shaked 1991 and Kitcher 1990). Our analysis starts from the necessity of standards in distributed processes and the possibility of multiple standards in science. Invisible hand models turn out to have only limited scope because they are restricted to describing the atypical single-standard case. Our model is a generalisation of these models to J standards; single-standard models such as Kitcher (1990) are a limiting case. We introduce and formalise this model, demonstrate its dynamics and conclude that the conclusions commonly derived from invisible hand models about the distribution of labour in science are not robust against changes in the number of standards.

The authors are grateful for helpful comments from the participant of the TILPS Workshop "Formal Modeling in Social Epistemology", October 2008, Tilburg University (The Netherlands), as well as comments from Simon Deichsel, Wolfram Elsner, Torsten Heinrich and Ulrich Krause.

1. Introduction

This paper presents a model of the dynamics of scientific activity. An understanding of the division of labour in science might contribute to more effective research policy and better institutional design. Moreover it can clarify a number of more general questions concerning scientific knowledge, the product cognitive labour. Why does knowledge tend to cluster? Is dissent irrational, and if so, why is disagreement a persistent feature of science? Why do scientists sometimes refuse to update their beliefs after being confronted with conflicting evidence?

More specifically, scientific activity exhibits a number of puzzling features which the model will need to explain. On the one hand, dissent and discussion seems to be omnipresent in science. But it has been argued that there is an ever growing body of scientific results on which a consensus is formed; and for some it seems only a matter of time until all dissent will have disappeared. "The positive argument for [convergent] realism is that it is the only philosophy that doesn't make the success of science a miracle" (Putnam 1975, 73). But then again, Larry Laudan (1981) put forward the pessimistic meta-induction argument: he compiled a long list of once successful theories which are now ridiculed by the scientific community. In sum, a powerful model of the dynamics of scientific activity has to provide an account of three aspects of science that seem difficult to reconcile: the existence of dissent, the emergence of consensus and the dissolution of that consensus. This is a tough challenge, as Larry Laudan himself noted: "[S]tudents of the development of science, whether sociologists or philosophers, have alternately been preoccupied with explaining consensus in science or with highlighting disagreement and divergence. […] neither approach has shown itself to have the explanatory resources to deal with both." (Laudan 1984, 2)

This paper discusses a number of models that have taken up this challenge and presents a new model. More specifically, in section 2 invisible hand models of the distribution of labour are discussed and shown to provide a satisfactory explanation of a community characterised by 1 standard. Section 3 lays out a model which generalizes this approach to J standards. Invisible hand models are a limiting case of this more general model. Section 4 formalises this model and presents several simulations to illustrate its dynamics. Finally, section 5 argues that the conclusions commonly derived from invisible hand models about the distribution of labour in science are not robust against changes in the number of standards considered.

2. Invisible hand models and the distribution of labour in science

Since cooperation is a necessary condition for distribution of labour, no model of the division of labour in science can ignore the benefits of cooperation in science, lest it leave the division of labour itself unexplained. The benefits of cooperation imply that there are increasing returns to adoption: scientists prefer rather than eschew more adopters to their views because more adopters means more opportunities for cooperation. This would lead a community to full specialisation, which is an outcome commonly considered to be epistemically undesirable and moreover conflicting with the actual state of science. The benefits of cooperation present a basic problem for those who attempt to understand the distribution of labour in science: why does the presence of these benefits not lead to fully specialised scientific communities, as one would expect? Invisible hand models of the distribution of labour in science offer a solution for this problem.

Petri Ylikoski (1995) offers a general characterization of the essential characteristics of the invisible hand mechanism which is at work in e.g. Kitcher (1990), Goldman & Shaked (1991) and Hull (1988):

1) It is a **decentralised** process: *"There are no explicit agreements or centralised decisions by the participating agents (Brennan & Pettit 1993: 195-196)."* (Ylikoski 1995, p.33)

2) The process is **non-intentional**: *"The agents do not intend to produce the result. They are promoting their own objectives and the result to be explained is a by-product of this promoting. The idea is that the process should work even if the participating agents have no knowledge of the process. This is why the mechanism is called invisible (Ulmann-Margalit 1978: 271)."* (Ylikoski 1995, p.33)

3) Although the process is non-intentional, it *"**needs not be unknown** to the agents participating in its production."* (Ylikoski 1995, p.33)

4) *"The result should be a pattern or a structure that seems to be made or designed intentionally; it should be somebody's handiwork (Ulmann-Margalit 1978: 268-270). This means that the product in question should be somewhat complex and it should not seem to be accidental. To be non-accidental, **the result should be somewhat stable and recurring** (Brennan & Pettit 1993: 191-192)."* (Ylikoski 1995, p.33)

5) In invisible hand explanations, **the result of the mechanism is valued positively**. This contrasts with what has been called the 'invisible backhand': *"The only difference is that the product of the invisible hand is valued positive and the product of the invisible backhand negative (Brennan & Pettit 1993: 192, 204-205)."* (Ylikoski 1995, p.33)

The crux of the invisible hand solution to the basic problem is to offset the scientist's benefit from more adopters by introducing a second factor, competition for credit[1]. Competition brings in decreasing returns to adoption. As more scientists adopt, there is more competition for newness, originality, to be the first to come up with the solution to an important problem,... More generally, decreasing returns are introduced to offset the increasing returns that cause the basic problem. The interplay of cooperation and competition will push and pull a community to a distribution of labour somewhere between full specialisation and full diversity. Ideally, a laissez-faire policy produces an optimal distribution.

One such formulation of the basic problem and the subsequent use of the invisible hand solution is found in Philip Kitcher's "The division of cognitive labour" (1990). Its starting point is the basic problem sketched above, which he calls the "CO-IR-discrepancy": the mismatch between a scientist's individual rationality (IR) and the ideal balance between specialisation and diversity, viz. the community optimum (CO). If scientists were all to pursue the same path, namely that which is best supported by the available evidence, then there is no diversity and the community optimum is unlikely to be reached, provided that, as Kitcher assumes, full specialisation is undesirable. Kitcher solves the discrepancy by de-idealizing the scientist: they are motivated by personal factors such as social and other factors, such as greed, stubbornness and honour rather than high-minded virtues that reflect the community optimum. Scientists freely compete with each other for the reward of being the first to find the solution. As a consequence, they do not just follow the path which is best supported by the available evidence, but discount it with the number of people already pursuing that path. As a result, as by an invisible hand scholarly attention is scattered and yields the community desideratum, viz. more diversity. The introduction of competition offsets the increasing returns that come with cooperation: individual returns *decrease* as the number of scientists following a certain path rises.

[1] The idea of a cycle of credibility stems from Latour and Woolgar (1986), but variants of this are found in Kitcher (1990), Goldman & Shaked (1991) and Hull (1988).

3. A general model of the distribution of labour in science for multiple standards

3.1 The importance of standards

We set out to show that apart from cooperation and competition, there is a third factor which is essential to a distribution of labour, namely standards. Kitcher's solution is a limiting case of this point of view, namely the description of the distribution of labour with 1 standard. The model presented in this paper generalizes Kitcher's solution to *J* standards. Any act of cooperation and any distributed activity requires a *standard* to ensure the compatibility of individual contributions and the coordination of individual efforts. A minimal consensus is required from the individual contributors concerning the goals of the distributed activity and the acceptable procedures to attain these goals. A standard is necessary for the aggregation of individual contributions at a certain time and the cumulation of the aggregated results over time. A model without them would simply fail to explain the distribution of labour in science itself, let alone account for its dynamics.

Standards are essential for a model of the distribution of labour in science and as a good model, Kitcher does indeed incorporate it, although implicitly. The idea of scientists competing for a prize and ending up in a nicely distributed scientific community only makes sense if it is assumed that all these scientists adopt to the same standard. It is after all this standard which determines what the problems are, which problems are important, how they should be solved, what solution is sufficient to claim the prize and how big this prize is. In Kitcher, however, the case of multiple standards is not considered. This is in line with his view of science as embedded in 'consensus practice' but it is a limitation of his model of the distribution of labour in science, because it unnecessarily commits its followers to a single-standard view. To overcome this limitation, a generalization of Kitcher's model is proposed to the case of *J* standards, with Kitcher's own model as a limiting case where there is only 1 standard. This generalization is important for two reasons. Firstly, it allows the model to be used by scholars who are uncomfortable with this single-standard view, such as the large branch of philosophy building on Kuhn (1962) for whom the dynamics of science involves existence of multiple 'paradigms'.[2] Secondly, this generalization will show that Kitcher's conclusions about the distribution of labour in science are not robust against changes in his single-standard view.

3.2. What are standards?

Whereas Kitcher could afford to leave the characterization of standards implicit, a generalization of his model to n standards requires a clear understanding of what standards are and what they do. An interesting way to gain some leverage on this is to characterize the concept of 'standard' in analogy to its use in the economics of network industries. It will be argued that the presence of multiple standards in science produces the same dynamics as that of multiple technological standards in a market.

Standards feature prominently in the literature on network industries. A network is a distributed system constituted of nodes and their interconnections. Its boundaries are defined by a standard. These standards are necessary conditions for inclusion in the network. For example, to run Macintosh software you need an Apple computer. However, you might also run it on a PC, but then you'll need a 'gateway' between networks (an adaptor). In other words, people adopting to one network will incur transaction costs when changing networks. So for agents in a distributed system, standards constitute a barrier to entry. Sometimes these

[2] A recent illustration of discomfort with Kitcher's single-standard view is the 2002 discussion between Philip Kitcher and Helen Longino in Philosophy of Science; Kitcher (2002a, b) and Longino (2002a, b)

barriers are relatively minor and easily overcome, sometimes they are high and lead to significant extra costs. In some systems, barriers to entry are constructed artificially, for example through patents or industrial secrets (e.g. the recipe of the Coca Cola syrup). In other cases barriers to entry arise naturally. One especially significant case of naturally arising barriers to entry is not a feature of the product itself: its rate of adoption. For example, say a company designs an innovative new operating platform. The barrier to entry it is confronted with is that existing operating platforms (most notably Windows) have already been widely adopted to. This not only means that most consumers will already have devices specifically designed to run the pre-existing operating platform, but also, and most importantly, that new software will be written specifically for that platform and not for the newly developed one. Even when exhibiting a very high intrinsic quality, the new operating platform will have great difficulty in conquering market share (it is, however, not impossible). Barriers to entry entail that producers cannot freely compete in a market because of significant costs associated with entering a new market and significant differences among these markets (e.g. different programming languages, different adapters, different consumption patterns,...). The same goes for scientists and the market of ideas: insights, methods, solutions, discussions, conferences, etc. adopt to a certain standard. Adopting to this standard requires a non-trivial investment: barriers to entry include mastering standard-specific expertise, learning standard-specific techniques, getting to know a specific community and identifying the standard-specific puzzles and trends.

Because standards divide a market into different parts separated by transaction costs, free competition is no longer possible. In addition to barriers to entry, free competition is further constrained by the fact that network industries tend to exhibit large economies of scale. These are especially prominent in information-intensive industries such as newspapers, consulting, publishing,... The reason for this is that information is characterized by decreasing marginal costs: once a unit of information is produced (an idea, a book, a score,...) it can be distributed at virtually no cost, unlike for example the car industry or the service sector, where an additional unit of the product keeps on costing significant capital and labour. As such, as more people use it, the cost of production stays the same but marginal costs can keep on falling indefinitely; it's only limit is the total extent of the market. In other words, network industries have a natural tendency toward monopoly. A similar argument can be set up for science: scientists are producers and consumers of information and as it happens, information-intensive industries are typically characterized by falling marginal costs. Just as with barriers to entry, falling marginal costs entail increasing returns to adoption.

So in network industries there are barriers to entry and large economies of scale and these two characteristics also apply to science. Both factors give rise to strong increasing returns to adoption and create enough market disruption to prevent decreasing returns of competition from offsetting these increasing returns. Indeed, the disruptive nature of these kinds of industries is widely known among policymakers and has prompted governments to implement antitrust regulation.

> 'The long-standing public policy concerns over network industries are not accidental, because those industries often embody two major and widely recognized forms of potential market failure: significant economies of scale -with the potential for monopoly- and externalities.' (White 1999, p. 1)

Of course the analogy between science and network industries is not complete. For example, science is often not commercialized and its output not monetized. But analogies are never perfect and this doesn't stop them from being fruitful. One way of arguing for the analogy is to point out that, when turning to economics for a model of distribution of labour in science, it does make more sense to use ideas used to explain the dynamics of information intensive

industries such as Microsoft, rather than to make an analogy with more traditional sectors usually characterized by decreasing returns to adoption such as car manufacturers.

But perhaps the most important reason to resist this analogy is to reach back to factors such as credit, newness, originality,... viz. the usual factors which invisible hand models brought in precisely to avoid those increasing returns characterizing network industries. Surely these have a role to play in science. The next section discusses their position within our framework.

3.3. Newness in a multi-standard view

An important consequence of introducing standards in considering the distribution of labour in science is that scientific contributions will tend to cluster. The presence of these clusters within a field makes it necessary to make the distinction between the dynamics of science *within* a cluster and *between* clusters. This entails that the problem of the distribution of labour can be described at two levels of analysis. Models for the dynamics of science under 1 standard, such as Kitcher's, can additionally be used as a model of what happens within a cluster in models that describe the dynamics between clusters. This distinction between a model of scientific activity within a cluster and between clusters allows us to further clarify our claim about increasing returns to adoption. The basic claim of our model is that there are increasing returns to adoption *between* clusters. This does not preclude decreasing returns to adoption within a cluster; with scientists competing e.g. to win the 'prize' for being the first to find the solution for a certain problem. In the case of 1 standard, no more is needed. However, from a multi-standard point of view, the importance of the problem depends on how many agents find this problem important; a problem from the point of view of one cluster might be irrelevant for someone in another cluster; or what counts as a satisfactory solution for one cluster might not be satisfactory for the other. So whereas these agents within the same cluster are competitors for the 'prize' associated with solving a certain problem, they have a common interest in the number of adopters to the cluster because the importance of the problem (the size of the price; e.g. in terms of recognition, funding, position, etc.) varies with the number of adopters. The characterization of the problem and the amount of 'prize money' are things that Kitcher takes as given but which vary in the case of multiple clusters: the characterization of the problem is relative to the cluster and the number of adopters determines the size of the prize.

Since our model is concerned with the overall dynamics of science, we focus on the distribution of labour across clusters rather than on what happens inside a specific cluster. Standards create barriers that divide a discipline into different parts. The fundamental divisions between scientist's contributions will be the same as the divisions between the standards on which each of these contributions is based. Standards add additional structure to the field. Since all individual contribution must adopt a standard, the change of the crucial boundaries in the field can be modelled by representing the changes in the rate of adoption to the different standards in the field. In other words, the distribution between core research programmes in the field is representative for the distribution of all scientific activity in that field.

At this level of analysis, 'newness' loses its importance for the distribution of labour in science (which is now seen as the distribution of scientific labour across clusters instead of the distribution of scientific labour across scientists). Because of the different levels of analysis that are now distinguished, newness can mean two different things: newness within a cluster and newness as the creation of a new cluster. Both actions are allowed for in our model, but they fail to offset the overall dynamics of increasing returns which governs the competition between multiple clusters competing for adoption. In the first case, newness takes place

within the shared consensus of a particular standard; as such it is simply modelled as a contribution to a cluster. The second case, where a new standard is created, does register as real newness at our level of analysis. However, the success rate of new clusters in network industries is very low (however, it is not impossible). In short, 'newness' does not alter the fundamental dynamics of our model because we are at the level of analysis across clusters.

4. Formalization of the model

Our model addresses the relations across clusters since we believe that this is the most relevant aspect to get a grip on the problem of the distribution of labour in science. Hence we will model the dynamics of clusters competing for adoption. Interestingly, the analytical tools for modeling systems exhibiting increasing returns to adoption have only recently been developed in a series of papers by Arthur (1989) and Arthur et. al. (1983, 1984, 1987). His models were initially designed for problems of technology adoption in network industries, e.g. to model the competition between VHS and Betamax to become the standard video format.

The model will thus focus on standards and their adoption by agents. We model these standards as clusters of contributions. Each turn, all agents in the game (scientists) make a contribution to one of the clusters in the game; there is also a small probability that they start a new cluster. Making a contribution means adopting to the cluster. Adopting to a cluster requires compliance to a basic set of concepts and assumptions; this basic set coordinates individual contributions. They form the core of the research programme implicit in all the contributions made to the cluster. So no matter how diverse the different contributions to the cluster, the cluster itself is a homogenous entity. Its size depends on the number of contributions.

As is customary in these models, the model as a whole is agnostic about the value of these clusters. The choice for one cluster rather than another is left for the agents in the model to decide. For a model about standards in science, this means that the model needs to be 'agnostic' in its conception of scientific value. By committing to a specific conception of scientific value, this part of the problem of the division of labour would be put beyond the model's explanatory scope and a generalization to n standards would be impossible. This contrasts with the invisible hand models described above, where the aims of science could be specified because there is only one standard. Our only claim about these clusters is that it is possible that there are multiple competing conceptions. The 'value' for the agent is then whatever it is that the cluster aims to produce. To indicate that this product can take different forms in different clusters, we leave the specific product of a cluster unspecified and refer to it using the generic term 'output' in the agent's decision function.

4.1 A formal model of formation and dissolution of consensus

Let us consider a population of N epistemic agents. There are J competing clusters. Denote agent n's preferences over clusters by the vector $p_n = (p_{1n}, p_{2n}, ..., p_{Jn})'$ and assume that there is a vector $E = (E_1, E_2, ..., E_J)'$, its j-th element being the available output for cluster j. The simplest way to think about output is to think of it as the number of contributions made to a cluster. We normalize E and denote the vector of relative output by $\hat{E} = (\hat{E}_1, \hat{E}_2, ..., \hat{E}_J)'$.

The most important parameter in our model is c, which we call the strength of increasing returns. We assume that $c \geq 0$. With standards becoming more important c will be higher. In general, c is the weight agents assign to the size of the cluster as measured by relative output.

Let the *likelihoods of pursuit* for each agent n be given by equation (1). The likelihoods of pursuit depends on an individual component (preferences) and a social component (the size of the cluster as measured by output weighted by the strength of increasing returns).

$$\pi_n(t) = p_n + c\hat{E}(t) \quad (1)$$

We define the following decision rule: Make a contribution to the cluster with the highest likelihood of pursuit, i.e. the largest element of the vector $\pi_n(t) = (\pi_{1n}(t), \pi_{2n}(t), ..., \pi_{Jn}(t))'$.

The model evolves by all agents making a contribution each period. At the end of each periods output is updated. By making contributions to a cluster output increases. The process is self-reinforcing. Contributions to a cluster increase output which in turn makes it more likely that agents will contribute to the same cluster next period.[3]

In making his decision to which cluster to contribute a scientist looks at the available output. By output we understand the accumulated knowledge in the form of journals, textbooks and the like. Output is produced by scientists making contributions. We assume that all contributions are contributions to just one single cluster. The quality of contributions is assumed to be homogeneous. Using these assumptions we avoid the task of having to judge the quality of contributions. We can measure output as the weighted sum of past contributions, where output produced within a period equals the number of contributions within this particular period.[4] Output for cluster *j* at time *t* is given by eqn. (2) where $K_j(t)$ denotes the number of contributions to cluster *j* in period *t* and $d \in [0,1)$.

$$E_j(t) = K_j(t-1) + dE_j(t-1) \quad (2)$$

Since eqn. (2) holds for all periods $t \geq 1$ we can use substitution and see that output is the weighted sum of past contributions. We assume that initial contributions $K_j(0)$ are given.

$$E_j(t) = K_j(t-1) + dK_j(t-2) + ... + d^{t-1}K_j(0) = \sum_{s=1}^{t} d^{s-1}K_j(t-s) \quad (2')$$

The number of contributions to cluster *j* in period *t* are given by $K_j(t) = \sum_{n=1}^{N} a_{jn}$ where $a_{jn}=1$ if $j \in \arg\max_{k \in \{1,2,...,J\}} \pi_{kn}(t)$ and 0 else.

The basic model is a nonlinear Polya process with the probability that a new contribution is made to a specific cluster being a function of the contributions already made to that cluster. As previous choices matter and increase the probability that a contribution will be made to a cluster, this process is path-dependent and exhibits positive feedback. We are interested in the structure that emerges during this process, where by structure we understand the proportion of agents working within each cluster. As has been shown by Arthur, Ermoliev, and Kaniovski (1983, 1984, 1987), the structure, which in our model is a vector of proportions, tends to a limit random vector. Our model reaches a stable pattern if $E(t+1)=E(t)$. Since our agents face the same output each period, they make the same choice each period and the distribution of agents across clusters stays constant.

Knowing that a stable pattern emerges our next question is concerned with the size of the clusters. Do all clusters have roughly the same size or does one cluster become dominant? Assuming that preferences are drawn from a [0,1] uniform distribution, it is clear that for c=0.0 all clusters are roughly equal in size. In this situation agents only care about their preferences, there is no premium on compatibility and hence there is no positive feedback. Since preferences are uniformly distributed all clusters are of roughly equal size. However, as

[3] Note that the likelihood of pursuit for any cluster is independent from the size of all other clusters. This implies that the standards that define clusters are completely incompatible. The model could be modified to allow for gateways. Then, the likelihood of pursuit would depend on the size of all clusters, with less weight given to the clusters for which there are gateways.

[4] By using the sum of contributions as a proxy for output it would also be possible to relate the model to scientometric data.

soon as $c>0$ increasing returns kick in. Clusters with high output attract more contributions and grow up to a certain point. This is visualized in the first line in figure 1, showing three runs of a simulation with N=1000 agents and J=5 clusters. On the vertical axis we see the size of each cluster, measured as the share of agents contributing to the cluster (formally: $\frac{K_j(t)}{N}$). The horizontal axes measures time.

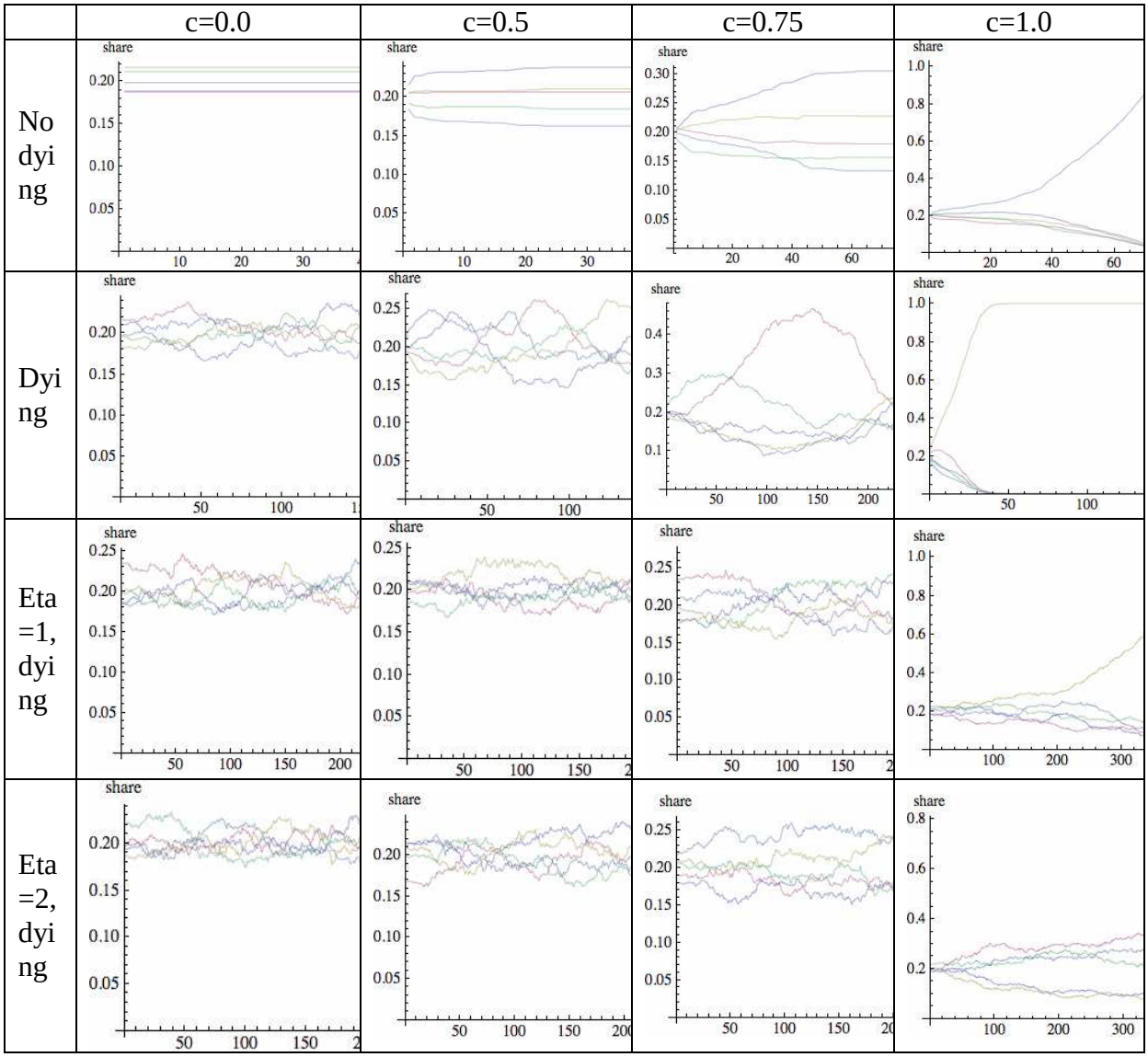

Figure 1: Typical simulation runs with varying parameters for institutional strength (N=1000 agents, J=5 competing clusters).

In the first simulation agents' choices solely depend on preferences. Each period they choose the cluster that is most preferred. Since preferences stay constant and agents are immortal the sizes of the clusters do not change.[5] In the second and third simulations we observe sensitivity to initial conditions. At t=0 the largest cluster is determined by the distribution of preferences. Due to the increasing returns the largest cluster grows faster than all other clusters. After

[5] We will relax both assumptions later and see what happens when agents are not immortal and preferences are allowed to change.

some periods a stable structure emerges at which the size of the dominant cluster (and all other clusters) stays constant. A large cluster can be understood as the existence of high consensus and low disagreement, or much specialisation and low diversity. The maximum size of the dominant cluster increases with *c*, where if $c \geq 1$ the dominant cluster gets 100%, i.e. there is absolutely no disagreement (no diversity) within the particular school of thought.

The resulting process exhibits several features of Arthur's increasing returns model (Arthur 1989). We cannot predict in advance which cluster will get dominant, but we know that one single cluster will get dominant. In Arthur's terms the process is *non-predictable*. The process is *nonergodic*, meaning that small differences at the beginning (the distribution of preferences and initial evidence) are not averaged out over time. Stochastic fluctuations are responsible for selecting the dominant cluster. Having reached a stable state the size of all clusters stays constant. The process is *inflexible*; there is no change from within the system.

4.2 Model refinements

In order to make the model more interesting and realistic we add some refinements. First, our agents do not live forever. Each period their age increases by one unit. Once they have reached a certain age, randomly drawn from a uniform [50,100] distribution they die and get replaced by a new agent with age drawn from a uniform [20,50] distribution. The new agent makes her first contribution to a randomly chosen cluster. The probability that any cluster is chosen is proportional to the size of the cluster. This could be interpreted as the agent makes her first contribution in the same cluster as her teachers worked in. The second line in figure 1 shows typical simulation runs for varying parameters of institutional strength. The only difference to the first line is that agents die and get replaced. For *c*=0.0 the size of the clusters do not affect agents' decision and cluster sizes follow a random walk. As c increases we see emergence and dissolution of consensus. This is clearly visible for *c*=0.75 where consensus reaches its peak around period 100 and more than 40 % of all agents contribute to the dominant cluster. Eventually the dominant cluster ceases to be dominant and we observe the dissolution of consensus. With increasing *c* the size of the dominant cluster gets bigger and dominant clusters are dominant for a longer period of time. This means that with standards being more important consensus exists longer among a bigger share of agents in our epistemic community.

Since agents die and get replaced the process is no longer inflexible for $c < 1$, i.e. there is no stable state at which the size of each clusters stays constant. For $c \geq 1$, however, the process is inflexible. Once a cluster has reached 100 % it stays there forever because agents' preferences do not matter for their decisions.[6]

As a second modification we introduce endogenous preferences. The idea is that when agents make their contributions they invest in learning the methods of the cluster. The resulting skills are specific to the cluster and cannot be transferred to another cluster. The result of the agent's investments are skills which are specific to a cluster, hence the benefits can only be reaped if the agent contributes to the same cluster. It is not possible to appropriate the benefits from that investment if they switch to another cluster. Another reason for endogenous preferences are

[6] A simple example might illustrate this inflexibility for *c*=1. Assume two competing clusters, A and B. Cluster A has reached 100 % which means that $\hat{E}_A = 1$. Likelihoods of pursuit are given by $\pi_A = p_A + 1$ and $\pi_B = p_B$ respectively. Since preferences are drawn from the uniform distribution $U[0,1]$, we must have $\pi_A \geq \pi_B$, so agents will always contribute to cluster A.

that the longer an agent has worked within a cluster, the less likely she is to change since her standing, reputation and accomplishments all depend on the correctness of the cluster.[7]
The change in preferences is modeled as follows. Let there be a vector of intrinsic preferences $\bar{p}_n = (\bar{p}_{1n}, \bar{p}_{2n}, ..., \bar{p}_{Jn})'$ for each agent and denote the number of each agent's contribution to cluster *j* up to time *t* by $k_{jn}(t)$. Assume that at time t the agent makes a contribution to cluster j. Then, her preferences the next period are given by a convex combination of her old preferences and some parameter $\eta \geq 1$ (eqn. 3).

$$p_{jn}(t+1) = \frac{k_{jn}(t)}{k_{jn}(t)+1} p_{jn}(t) + \frac{1}{k_{jn}(t)+1} \eta \qquad (3)$$

For all other clusters $j' \neq j$ to which the agent did not contribute in period *t* preferences do not change, i.e. $p_{j'n}(t+1) = p_{j'n}(t)$. For clusters to which the agent has never made a contribution her preferences are given by $p_{jn}(t) = \bar{p}_{jn}$ for all *t*. The parameter η determines the speed of the preference change and acts as an upper bound on preferences. By increasing η more weight is put on agents' preferences.
The likelihoods of pursuit are now given by $\pi_n(t) = p_n(t) + c\hat{E}(t)$. As can be seen in the third and fourth line of figure 1, allowing for endogenous preferences results in slower change and, if $c < 1$, lower variance in cluster sizes. As agents put more weight on their evidence as a result of past choices they are less likely to switch to another clusters, even if the other cluster is large. For $c \geq 1$ the process is inflexible. At some point all agents contribute to the same cluster. However, with higher η it takes longer until one cluster reaches 100 %.

4.3 Main results and possible extensions

The main results of the model can be summarized as follows.
(1) The resulting division of labour depends on the strength of increasing returns. With stronger increasing returns the size of the largest cluster increases and the community tends to more specialisation. For $c \geq 1$ the dynamics result in a lock-in. All agents contribute to one cluster and the community is completely specialised. The opposite, complete diversification, is achieved for low values of increasing returns.
This can be seen in figure 2 plotting the variance of cluster sizes for different values of *c*. The variance is computed as $Var = \frac{1}{J}\sum_{j=1}^{J}(x_j - \bar{x})^2$ where $x_j(t) = \frac{K_j(t)}{N}$ is cluster size, measured as the number of agents contributing to cluster *j* at time *t*, divided by the total number of agents. Since we assume *J*=5 clusters we know that the mean is $\bar{x} = 0.2$. The variance is a natural way to measure diversification and specialisation. If the community is completely diversified all clusters have equal size and variance is zero. At the other extreme, complete specialisation, variance is given by $\frac{J-1}{J}\left(\frac{1}{J}\right)^2 + \frac{1}{J}\left(\frac{J-1}{J}\right)^2$ which equals 0.16 for *J*=5. Figure 2 shows with weak increasing returns the community is completely diversified. As *c* increases above 0.4 we observe increasing specialisation, and for *c*=1 there is complete specialisation after some periods.

[7] In the literature this is known as „hardening of positions" where as time passes agents put more weight on their own opinion and less weight on the opinion of others (e.g. Hegselmann & Krause 2002, 4). It can also be understood as a process of dissonance reduction (Festinger 1957) where agents adjust their preferences in order to reduce the discrepancy between their preferences and choices.

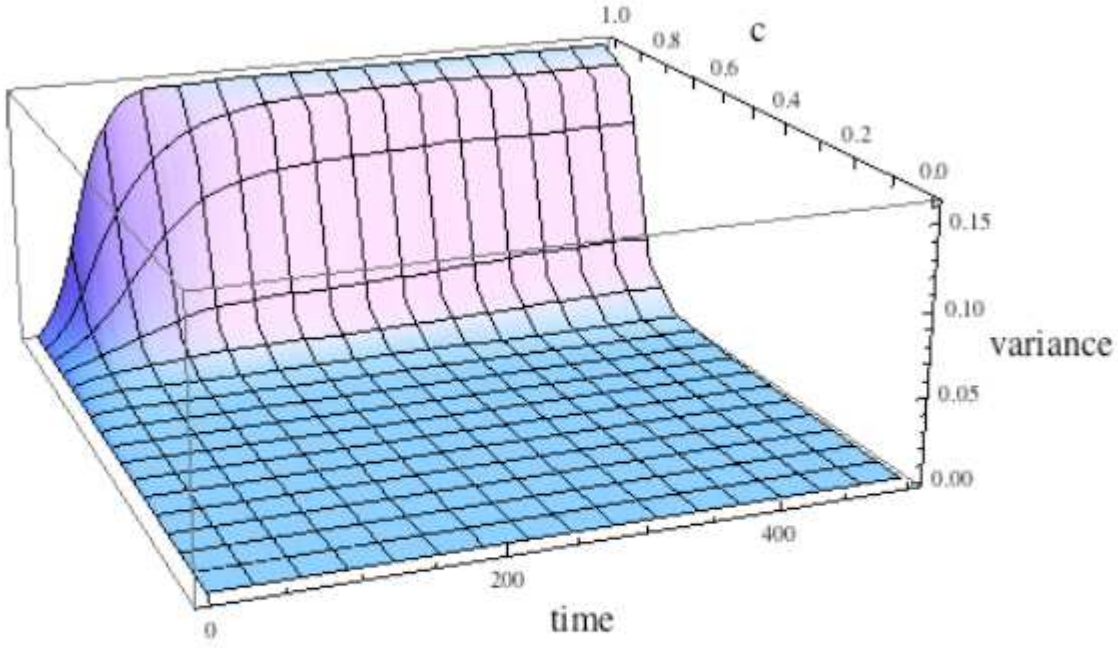

Figure 2: Variances of cluster sizes for varying values of c. Variances are average values obtained by running 100 simulations for each value of c (c=0.0, c=0.1, ..., c=1.0). At t=0 all clusters start with the same size, hence variance is 0. All simulations were run with 500 agents, 5 clusters, and 500 periods.

(2) By introducing endogenous preferences the change in the division of labour between clusters becomes slower because agents are more likely to stick to their choices. With strong increasing returns ($c \geq 1$) the distribution of labour will still reach a lock-in where the community is fully specialised, although the time it takes to get to the lock-in will be longer. This can be seen from table 1, showing the time it takes to get to the lock-in for c=1.0 and varying η.

η	time to lock-in
0	66.34
1.0	391.64
1.2	503.18
1.4	517.04
1.6	495.54
1.8	512.10
2.0	516.30

Table 1: Time to lock-in for varying strength of preference change. Times to lock-in are average values from 100 simulations with 500 agents, 5 clusters, and c=1.0. Each simulation run for 1000 periods. The case η=0 corresponds to no preference change.

(3) For $0 \leq c < 1$ the distribution of labour is flexible if agents are not immortal. Altough the dominant cluster can be quite large, the community never reaches full specialisation. As a consequence of agents dying and getting replaced the largest cluster eventually looses its dominant position and a new cluster gets dominant (paradigm change). For $c \geq 1$ we have a

lock-in, meaning that once a cluster reaches 100 % the community will stay at full specialisation. This could change by endogenizing the number of clusters and allowing agents to create a new clusters, or by introducing exogenous shocks (anomalies) that solve the lock-in by lowering the weight agents put on the cluster's size. These are, however, subjects for further research.

5. Conclusion

Invisible hand models rely on competition to solve the CO-IR discrepancy. This solution requires that there is full competition and that all agents are after the same (e.g. credit). We have argued that the necessary presence of standards in science entails that the scope of this solution is limited to the special case where there is only 1 standard because standards cause fragmentation of the market, undermining competition and hence Kitcher's solution to the CO-IR discrepancy. It turns out that Kitcher's model only describes the features of an atypical case of a system which is in most of its possible states very different from this single-standard case. More specifically, Arthur (1989) lists five features of increasing returns models: there are *multiple* possible *equilibria*, which equilibrium will be selected is *unpredictable* in advance, the equilibrium is *not necessarily optimal*, the system exhibits *inflexibility* (it can 'lock in') and *path-dependence*. These stand in sharp contrast to characteristics of invisible hand models, which typically have only one optimal and predictable equilibrium that is not path-dependent. Because of the atypical character of the single-standard case, the conclusions derived from invisible hand models are not robust against changes in the number of standards. The main conclusion from such models is that there need not be a conflict between individual rationality and the community optimum. Private vices become public virtues:

> *"The very factors that are frequently thought of as interfering with the rational pursuit of science –the thirst for fame and fortune, for example- might actually play a constructive role in our community epistemic projects, enabling us, as a group, to do far better than we would have done had we behaved like independent epistemically rational individuals."* (Kitcher 1990, 16)

As a result, individual responsibility of scientists is downplayed and the task of the institutions of science is to accommodate scientist's cravings rather then direct them toward higher epistemic ends:

> *"social institutions within science might take advantage of our personal foibles to channel our efforts toward community goals rather than toward the epistemic ends that we might set for ourselves as individuals."* (ibid.)

> *"the really neat thing about the reward system in science is that it is so organized that, by and large, more self-serving motivations tend to have the same effect as more altruistic motivations."* (Hull 1997, p.123)

From the perspective of our generalized model, these views about individuals, institutions and their interrelation are no longer tenable. Individual scientists cannot escape their responsibility because the dynamics of the model imply that small changes can have large consequences. A laissez-faire institutional design is bound to miss the community optimum because of the monopolistic tendencies the system exhibits.[8] As a consequence, institutional design must play a more active role to attain the community optimum. The general direction that it should take is that of softening the market disruption produced by the presence of multiple standards. This could involve the implementation of an active pluralist policy which aims to reduce transaction costs between clusters. A big step in this direction would be a pluralist education or at least a historical overview of the development of the discipline a scientist will work in,

[8] Hence we postulate an 'invisible backhand' rather than an 'invisible hand' (see section 2).

such that the transaction costs of not adopting to the mainstream cluster are not already gigantic from the outset of the scientist's career.

A final consequence of generalizing to *J* standards is that whereas the single-standard case can be individualist and a-historical, the shift to multiple standards makes social and historical aspects relevant.[9] In multi-standard versions of our model social aspects of science are important because of the occurrence of network externalities that exert causal influence on scientific activity but are irreducible to the individual level; the connection among the nodes is more important than the nodes themselves. Historical aspects become relevant in multi-standard versions because the dynamics is path-dependent, viz. previous states of the system exert causal influence on future states of the system.

We have presented this model as a generalization of Kitcher's model. While we think our model nicely captures most cases (2 to J standards), Kitcher's model is still better suited for the 1 standard case while our model has nothing informative to say on this, only that everyone will always adopt. Because Kitcher only treats the single-standard case, he can afford to present a model at a lower level of analysis, a level which is better suited to highlight the salient features of the atypical single-standard case. We could also change our model's level of analysis. The clusters would then become paths and increasing returns would be absent because we're inside a cluster (c=0). The model then predicts an equal distribution across paths and by introducing epistemic and non-epistemic motives we arrive at Kitcher's model. So our model is indeed a generalization, and (as is so often the case with general models; cf. Weisberg & Matthewson 2008) this generality goes at the cost of describing certain specific cases. The single-standard case is such a case for our model.

References

Arthur, W. Brian (1989), "Competing technologies, increasing returns, and lock-in by historical events", *The Economic Journal*, 99, pp. 116-131.

Arthur, W. Brian, Yu. M. Ermoliev & Yu. M. Kaniovski (1983), "A generalized urn problem and ist applications", *Cybernetics*, 16, pp. 61-71.

Arthur, W. Brian, Yu. M. Ermoliev & Yu. M. Kaniovski (1984), "Strong laws for a class of path-dependent stochastic processes, with applications", in: Arkin, Shiryayev and Wets (eds), *Proceedings of the Conference on Stochastic Optimization, Kiev 1984*, Lecture Notes in Control and Information Sciences, Springer.

Arthur, W. Brian, Yu. M. Ermoliev & Yu. M. Kaniovski (1987), "Path-dependent processes and the emergence of macro-structure", *European Journal of Operational Research*, 30, pp. 294-303.

Brennan, Geoffrey & Pettit, Philip (1993), "Hands Invisible and Intangible", *Synthese*, 94, pp. 191-225.

Festinger, Leon (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.

Goldman, Alvin & Shaked, Moshe (1991), "An economic model of scientific activity and truth acquisition", *Philosophical Studies*, 63, pp. 31-55.

[9] This enables us to meet i.a. Philip Mirowski's concern that *"A relevant congenital tic of the American philosophy profession (although, it must be conceded, not its alone) is a demonstrated unwillingness to regard science as an historically changing entity, not just in the realm of epistemic ''values'' but also in terms of actual social structures."* (Mirowski 2004 , 285)

Hegselmann, Rainer & Ulrich Krause (2002), "Opinion dynamics and bounded confidence: models, analysis and Simulation", *Journal of Artificial Societies and Social Simulation*. Vol. 5, <http://jasss.soc.surrey.ac.uk/5/3/2.html>.

Hull, David (1988). *Science as a Process*. Chicago: The University of Chicago Press.

Hull, David (1997), "What's wrong with invisible hand explanations?", *Philosophy of Science*, 64(S), S117-S126

Kitcher, Philip (1990), "The division of cognitive labour", *Journal of Philosophy*, Vol. 87(1), pp.5-22.

Kitcher, Philip (2002a), "The Third Way: Reflections on Helen Longino's The Fate of Knowledge", *Philosophy of Science*, 69(4), pp. 549-559.

Kitcher, Philip (2002b), "Reply to Helen Longino", *Philosophy of Science*, 69(4), pp. 569-572.

Kuhn, Thomas (1962). The structure of scientific revolutions. Chicago, Chicago University Press.

Latour, Bruno & Woolgar, Steve (1986). *Labouratory Life. The Construction of Scientific Facts*. Princeton: Princeton University Press.

Laudan, Larry (1981), "A confutation of convergent realism", *Philosophy of Science*, Vol. 48(1), pp. 19-49.

Laudan, Larry (1984). *Science and values*, London: University of California Press.

Longino, Helen (2002a), "Science and the Common Good: Thoughts on Philip Kitcher's Science, Truth, and Democracy", *Philosophy of Science*, 69(4), pp. 560-568.

Longino, Helen (2002b), "Reply to Philip Kitcher", *Philosophy of Science*, 69(4), pp. 573-577.

Matthewson, John & Weisberg Michael (2008), 'The structure of tradeoffs in model building', *Synthese*, published online July 16th 2008, http://www.springerlink.com/content/j127376grp53g314/

Mirowski, Philip (2004), "The scientific dimensions of social knowledge and their distant echoes in 20th-century American philosophy of science", *Studies in History and Philosophy of Science*, 35(2), pp. 283-326.

Putnam, Hilary (1975). *Mathematics, matter and method*. Cambridge: Cambridge University Press.

Ullmann-Margalit, Edna (1978), "Invisible-Hand Explanations", *Synthese*, 39, pp. 263-291.

White, Lawrence (1999), "U.S. Public Policy toward Network Industries", *AEI-Brookings Joint Center for Regulatory Studies Books and Monographs*.

Ylikoski, Petri (1995), "The invisible hand and science", *Science Studies*, 8, pp. 32-43.