Abstract

The Constitution of the United States empowers the Congress to pass copyright laws to promote knowledge creation in the society and more specifically scientific knowledge. Many interesting economic studies have been conducted on copyright law, but very little research has been done to study the impact of the law on knowledge creation. In this paper we develop and analyze an agent-based model to investigate the impact of copyright on the creation and discovery of new knowledge. The model suggests that, for the most part, the extension of the copyright term hinders scholars in producing new knowledge. Furthermore, extending the copyright term tends to harm everyone, including scholars who have access to all published articles in the research field. However, we also identify situations where extending copyright term promotes rather than hinders knowledge creation. Additionally, scholars that publish copyrighted materials tend to out-perform those who do not creating a potential tension between individual incentives and the public good.

Keywords:
Knowledge Creation, Copyright Law, Copyright Extension, Division of Labor, Complex Systems

Introduction

1.1 The number of scientific publications has been growing. Larsen and Ins (2010) estimated 4.7% of overall growth in scientific publications between 1997 and 2006. Policy makers are interested in promoting a higher growth of knowledge in the society as it contributes to the advancement and development of the country. As an important policy tool, the Constitution empowered the Congress to pass copyright laws "to promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries." Where this is currently understood as pertaining to "the advancement of useful knowledge or discoveries" (Walterschied 2002).

1.2 Lawmakers have extended the copyright term many times since its inception in 1790. There are many interesting economic studies on copyright law. However, to our knowledge, there has been no research directly focused on how the extension of the copyright term contributes to the production of knowledge, the very goal for which this policy tool was designed. In this paper, we are planning to take a step towards this goal by implementing and analyzing an agent-based model of discovery and publication. The model suggests that the extension of copyright term generally hurts knowledge creation. More importantly, the extension does not only hurt those with lower access to copyrighted materials, but also hinders scholars with access to all copyrighted papers. However, we also find scenarios where extending the copyright term moderately helps knowledge creation. In both cases, scholars that publish copyrighted material tend to out-perform those that do not creating a potential tension between individual incentives and the public good.

1.3 Before explaining the model, a few introductory remarks about historical copyright extensions and the motivation of scholars will be helpful. The first copyright law secured fourteen years of rights for authors. The author could also renew the copyright for another fourteen years. If she did not renew the copyright, the work was passed to public domain. The renewal mechanism in the first copyright policy insured only works that were worth renewing would remain protected. Others would return to public domain. Since 1790, the copyright term has been expanded multiple times, but more importantly in 1976, the system of renewal of the copyright was removed from the law, and every book and other creative works received the maximum term. In 1998, in the Copyright Term Extension Act (CTEA), the Congress extended the terms to life of the author plus seventy years, and for works of corporate authorship to hundred and twenty years after creation or ninety five years after publication, whichever endpoint is
earlier. The constitutionality of the act was challenged in the court and ultimately, the United States Supreme Court held the act constitutional.[3] In "Free Culture", Lessig (2004) elaborates on the consequence of extensive increase in copyright term on public domain.[4]

1.4 Accounting for the motivation of scholars and scientists is central to understanding how to promote knowledge creation. Aristotle believed that the desire for knowledge was part of human nature. While there is not a strong consensus on a specific account of motivating factors for knowledge creation, there is a general sentiment that pursuit of truth is probably an important component. Social epistemologists argue that credit, respect and honor can also motivate scholars and scientists to create new knowledge and publish papers (Goldman 1999, 2004). Such factors also encourage openness in science (Scotchmer 2004). Bonilla (2001) takes this idea even further and argues that scientists' main motivation is not the pursuit of truth but the pursuit of recognition. Accepting the role of credit seeking in the expansion of knowledge, Hull (1988) discusses the types of credits considered the most important and the effects that striving for credit creates on science.[5] Scholars may prefer minimal copyright protection because it increases the level of access to their scholarly works and results in academic promotions and improves their academic prestige (Landes et al. 2009, 48). For these reasons, we will focus on the impact of copyright law on the production of publications and citations rather than on, say, economic factors.

1.5 Even with a restricted focus, attempting to study the effect of copyright law on knowledge creation is wrought with complexities. The environments that scholars find themselves in are far from simple; they may involve different legal systems over different periods of time while experiencing rapid technological advancement. Furthermore, given the nature of the issue, a direct comparison between different possible copyright laws is not feasible as the current law is the only binding one. Historical comparison is also not helpful, because the rapid advancement of information technology in the 21st century has significantly impacted knowledge creation. Thus, real systems may be too complex and difficult to analyze if we are interested in assessing the specific effect of copyright law on knowledge creation. Payette (2011a, 2011b) has argued that agent based modeling provides a particularly useful method of approaching the study of science. We will follow this suggestion here. Agent based modeling provides a method of abstracting away from many real-world complexities and isolating the particular interdependent dynamics we are interested in studying. Such an approach will allow comparisons that would not otherwise be possible given the legal and historical constraints.

The Model[6]

2.1 The aim of the model is to measure the impact of the term length of the copyright law on knowledge creation, specifically in the sciences. This requires a representation of a given field or area of study, scholars doing research in the field, and a way to represent copyright and access to copyrighted material. The model consists of a fixed set of agents (scholars) that explore a network of connected research topics (an "epistemic plane"), attempting to publish new research that builds upon what is already known. Published research may be copyrighted, limiting access to that research for some researchers. We will examine the effects of varying copyright terms on the efficiency by which scholars publish new research as well as how copyright terms might impact citation tendencies.[7]

Epistemic Planes

2.2 We will use a simple spatial network, a lattice, to represent a field of research (see figure 1).[8] The nodes on this network represent potential research topics and the distance between nodes corresponds to the similarity of topics. Each node is in one of two states: an established (published) finding or a currently unknown (unpublished) fact. At a given point in time, each scholar is located at a particular node—i.e., her current focus in research. The links between nodes are bidirectional and constrain both the movement of scholars and the potential citations for new publication, move on this below. The lattice is arranged on a torus, so the space is enclosed but without borders. Furthermore, there are no intrinsic differences in the value or importance of any research topic, they are all considered equally valuable. We refer to this representation as an "Epistemic Plane."[9]

2.3 The notion of an epistemic plane abstracts away from certain complexities of real research fields. Real research topics differ in importance and value, are not connected in a simple uniform way, and are not connected in a bidirectional fashion, as some may be necessary to form a foundation for future topics. While such idealizations distance the model from the target phenomenon, they are helpful in facilitating a detailed analysis and allow for a visual representation of the model.
The number of steps between nodes corresponds to the similarity of topics: the farther you get from one point in epistemic plane, the larger the difference in research questions and concepts.[10] We include an adjustable parameter, the Neighborhood Size, that specifies how closely topics must be related in order to have potential citation connections and/or allow for movement of scholars between topics. The neighborhood of a given point consists of all other points on the epistemic plane within the specified (Euclidian) distance.

Each point in the epistemic landscape is designated as "discovered" (published) or "undiscovered" (unpublished). Each publication is also assigned a copyright status that determines accessibility. "Copyrighted" works are available only to high-access scholars. "Open-access" works are available to everyone. At the beginning of each simulation all points within a radius of 10 an arbitrary center point are designated as previously discovered. Each of the initially known points is randomly assigned a copyright status with a 50% chance of each status and scholars are located at radius 2 of the center on knowledge points in which they have access.

Scholars are responsible for the discovery of new points and the publication of new knowledge. At any point in time, each scholar is located at a specific point in the epistemic plane. As time progresses, scholars will move through the network seeking to read published work and publish new research when reaching undiscovered nodes. Scholars are initially located in the center of the previously known topics and will begin by reading articles, provided they have access, and moving to neighboring points. Scholars can publish new research if they (i) are in the neighborhood of an undiscovered point and (ii) have read a sufficient number of articles within the neighborhood of the undiscovered point—we refer to this as the "citation requirement" for publication. Scholars are allowed to cite their own work. Publishing is the scholar's highest priority and each will do so whenever possible, citing all articles she has read within the neighborhood of any new publication. If publishing is not an option, they will choose an accessible article within their current neighborhood that they have not yet read, move to that location and read the article.[11] When selecting an article to read, scholars give preference to the most recently published articles. The idea behind this strategy is to facilitate reaching the frontier of knowledge discovery as soon as possible. Only if there is no option to read or publish within her current neighborhood will a scholar relocate to a more distant topic on the epistemic plane.

The algorithm for scholar behavior is as follows. In all the scenarios if the number of available knowledge patches is more than one, the scholar selects one randomly.

1. If there are undiscovered points in the neighborhood of your current location, and you meet the citation requirements for publication, move to that point and publish it. The point then becomes discovered.
2. If there are discovered points in the neighborhood of your current location that you have not yet read and that you can access, choose the most recently published article (break ties at random), move to that point and read the article.
3. If there are no reading or publication opportunities in the neighborhood of your current location, choose a random discovered node that you can access and have not read, move to that location and read the article.
4. Otherwise, do nothing.

Research Cycles and Copyright Term

There are two types of scholars with respect to the publications they can access: those that can only read open-access articles (AL0 scholars) and those that can read any article (AL1 scholars). Scholars are initially located on the epistemic plane in the center of "previously known" results. At the beginning of a simulation, each scholar is randomly assigned an access level, with equal weights assigned to AL0 and AL1. This access level remains constant for each scholar throughout the duration of the simulation.[12] Published papers are randomly assigned a copyright status with an equal chance of being an open access publication or a copyrighted publication.[13] We assume that all scholars have read their own publications regardless of access-level and copyright status. We will also consider a case where each scholar is randomly assigned a publishing habit and does not vary form publication to publication.

Discovery on the epistemic plane occurs in discrete time steps. Each time step, every scholar will attempt to either publish new
research or read existing work. We call a complete time step a "research cycle." All copyrighted articles remain restricted-access until they have been in existence for a predetermined number of research cycles—this is the **copyright term**. After the copyright term for an article expires, the article passes into the public domain and is treated as being open-access. Newly published articles, if copyrighted, will remain so for the length of the copyright term beginning from the time of publication.[14]

2.10 Figure 2 provides a visual illustration of the model at a fixed research cycle. Discovered points are green and undiscovered points are gray. The two shades of green for published articles represent the current access levels for those articles: open-access in light green and copyrighted in dark green. Figure 3 provides a closer look with agents occupying the epistemic plane. AL0 scholars are represented in blue, AL1 scholars in red.

![Figure 2. Epistemic Plane](http://jasss.soc.surrey.ac.uk/18/2/23.html)
Figure 3: Epistemic Plane with Scholars. The figure shows scholars exploring epistemic plane.

Table 1: Model Parameters

| Parameter, Variable | Description | Value or Range of Values tested in the simulation |
|---------------------|-------------|--------------------------------------------------|
| Neighborhood Size (NS) | The neighborhood size represents the range of nearby concepts or topics that a scholar would perceive. The neighborhood of a given point consists of all other points on the epistemic plane within the specified (Euclidian) distance. Neighborhood size of 3 consists of 28 patches and neighborhood size of 2 makes of 12 patches. | 2 and 3 |
| Required Citation (RC) | Scholars are required to cite minimum number of published papers in the neighborhood, before publishing new knowledge. | 1 to 11 |
| Copyright Term | Copyright term is the number of cycles during the simulation that copyrighted papers remain as AL1 with limited access only to AL1 scholars. | 1, 2, 5, 10, 25, 50, 100, 150, 200 |
| Number of Scholars | Number of scholars determines how many scholars explore knowledge on the epistemic plain. The number of scholars remain unchanged during the simulation. | 10, 50, 100, 200, 500 |

Results

3.1 We will measure the overall rate of publications and citations under various copyright term lengths to understand if changes in copyright term length may have any significant impacts on number of publications and citations. We will also look at both access-
level-one and access-level-zero scholars separately in the model under various copyright term lengths to analyze how well each group of scholars perform in publishing new knowledge and receiving citations to their published articles. We ran 180,000 independent experiments with copyright term ranging between 0 and 200 cycles, with varying neighborhood sizes and citation requirements. The number of scholars in these experiments also varied ranging from 10 to 500 and scholars were randomly assigned an access level of one or zero. Three important themes emerged from running the model. 1) There was a general tendency for longer copyright terms to hinder knowledge creation. 2) This tendency, however, was not universal and there are scenarios where some extension of the copyright term can (moderately) help knowledge creation. 3) When scholars are assigned fixed publishing habits, those scholars that publish only copyrighted papers systematically outperform those that publish only open access papers.

Does copyright hurt knowledge creation?

3.2 The prevailing trend in most simulations was that increasing the copyright term resulted decreased knowledge production. The declining trend is intuitive, because under shorter copyright protection regimes access to published articles goes up and scholars can more easily read and cite discovered knowledge. Consequently, scholars can publish new papers more rapidly. The impact of increase in copyright cycles is more dramatic when the required citations are higher. For example, when eleven citations are needed, the number of published papers declines about seventy percent when we extend copyright protection from zero to two hundred cycles, while the drop in publications is only about twenty percent where required citations are five (see graph 1). Although the general trend is intuitive, there are several more specific features of the results that are particularly interesting.

![Graph 1](http://jasss.soc.surrey.ac.uk/18/2/23.html)

Graph 1. Total publications under different copyright cycles. The set of graphs compares three required citations of five, seven, nine and eleven. In all four graphs the number of scholars is 200 and the size of the neighborhood is 3. We ran the model for 250 cycles.
3.3 First, the decline in knowledge creation was steep with respect to increases in shorter copyright terms and relatively flat with respect to increases in longer copyright terms. For example, in the scenario where we require eleven citations, 63% of the 70 percent total decline in knowledge production happens when copyright is increased from 1 to 25 cycles. The curve begins to flatten with respect to increases after 25-cycle point. However, when the required citation is lower, we see a more gradual decrease in knowledge creation. The increase from 150 to 200 cycles has very little negative impact regardless of citation requirements.[16] The same trends also occur with respect to the impact of copyright term on citation rates, where the results are qualitatively similar to those regarding publications.

3.4 Second, increasing the copyright term tends to hurt all scholars. However, the negative impact is significantly larger for scholars with restricted access, leading to large inequalities between the groups. The decline in the publication of AL0 scholars is intuitive because increasing the copyright term means these scholars will have access to fewer articles to read and as the result fewer opportunity to publish. It is, however, interesting to note that most of this inequality typically emerges in the initial increase of copyright term from 0 to 25. And, increasing the copyright term beyond this point tends to, very moderately, reduce both inequality and overall productivity (see graph 2). The same trends hold for citations received. Additionally, the resulting inequalities are most striking in scenarios where more citations are required. For example, with required citations set at eleven and copyright term sets at 50, AL1 scholars are responsible for contributing 92% of total publications.

Graph 2. Average publications made by AL0 and AL1 scholars. These graphs measure knowledge production under different copyright cycles. In all graphs, the neighborhood size is set at three and we included two hundred scholars to the epistemic plane.

3.5 To further investigate the impact of increased copyright term on AL0 scholars, we tracked the number of times scholars were forced to look for accessible articles outside the current neighborhood where they located due to the lack of accessible articles. In other words, AL0 scholars had to change their research subjects to be able to read and find new ideas to publish. In reality, scholars may show more perseverance in exploring the subject they are currently investigating and may decide not to switch.
subject if they could not find articles to read too quickly. However, measuring number of times AL0 scholars switch research subject gives us some indications on how copyright extension may cause challenges for scholars with limited access to copyrighted materials. Naturally, by tightening copyright protection, AL0 scholars had to shift neighborhood more frequently (see graph 3).

Graph 3. Average Number of shifts under different copyright cycles. The set of graphs illustrates the impact of copyright extension on number of times AL0 scholars had to switch research subject.

3.6 Finally, it is important to note that all parties are negatively impacted by an increase in copyright term, including the AL1 scholars. The lack of activity among AL0 scholars reduces knowledge discovery in the epistemic plane. Therefore, fewer published articles are available to AL1 scholars to read and cite which culminates in lower rate of publication among AL1 scholars. In the scenarios examined in graph 2, although AL1 scholars do much better than AL0 scholars under the presence of copyright, everyone is worse off than they would be without any copyright. In other words, copyright creates a relative advantage for AL1 scholars, but simultaneously decreases the absolute productivity of those same scholars. [17]

3.7 Increasing copyright term has slightly different effects depending on the number of scholars working to produce knowledge. Naturally, increasing the number of scholars in epistemic plane causes knowledge production to increase. However, in scenarios where we have many scholars competing to publish new articles, the copyright protection slows down knowledge production more significantly than the scenarios where the field is less crowded. Fewer scholars typically means more opportunities for each to publish, reducing the impact of extending the copyright term. Graph 4 illustrates how extension in copyright protection impacts knowledge in research fields with different number of scholars. A similar trend holds for number of citations received.
Can copyright help knowledge creation?

3.8 The prevalent message of the model, holding in most scenarios, seems to be that increasing copyright term reduces the rate of publication and citation. There are certain circumstances, however, where an increase in copyright term can help knowledge creation. This occurs when very few citations are required for publication and only a small number of scholars in the field. Graph 5 illustrates the effect of increasing copyright for different numbers of required citations in a field with 10 scholars. In this case, the effect of increasing copyright term on knowledge production shifts from moderately positive to moderately detrimental as we vary required citations for new papers from one to five.
Graph 5. Impact of copyright term on knowledge production under different citation requirements. These graphs illustrate knowledge production with various required citations (RC's) in a three-radius neighborhood with 10 scholars in the epistemic plane. Under required citation of two, the number of publications goes up when the copyright cycles changes from one to ten, then it flattens and stays fairly unchanged. It is important to point out that the increase in publications is a modest 2% in this graph.

3.9 In addition to required citations, the number of scholars in the model is another influencing factor in the impact of copyright on knowledge creation. Graph 6 shows knowledge creation as a function of copyright term with various numbers of scholars in the epistemic plane. Longer copyright protection hinders knowledge creation when there are many scholars in the field. But with fewer scholars in the field, an increase in short copyright terms modestly improves knowledge creation with little difference in for increases in already longer copyright terms.
3.10 The explanation for this modest increase in research productivity is due to the division of labor in science. This topic has been explored by many scholars in a number of different settings; see e.g. Kitcher (1990), Stevens (2003), Weisberg and Muldoon (2009) and Zollman (2010). With respect to our model, an increase in copyright term can cause scholars to spread out across the epistemic plane, each exploring separate local areas. The division of labor occurs due to the behavior of AL0 scholars. Because they are unable to access many recent articles, publishing copyrighted papers makes it slightly more difficult for AL0 scholars to publish papers nearby. This partially prevents AL0 scholars from simultaneously working in the current area of another scholar. And, when there are very few other scholars, often there will be other nearby unoccupied areas in which they can publish new research. This causes AL0 scholars to spread out in the epistemic plane and results in a moderate increase in knowledge production. In this setting (low citation requirement, few scholars), we also see that there is very little difference in overall production between AL1 and AL0 scholars.

Is it Better to Publish Copyrighted or Open-Access Papers?

3.11 So far, we have examined the effects of extending the copyright term on overall knowledge production. Typically, copyright hurts knowledge production on the whole. However, we may also want to know whether scholars themselves would prefer to publish open-access or copyrighted articles. To answer this we need to know whether publishing copyrighted material increases or decreases success for individual scholars.

3.12 To address this question, we modified the model to give each scholar a randomly chosen publishing habit for the duration of the simulation. This creates four types of scholars: S00, S01, S10 and S11. The first number designating access level, the second designating publishing habit. This allows us to track the relative success of scholars (in publications) in terms of both publishing habits and access levels. At the beginning of a simulation, each scholar is assigned randomly a publishing habit (50% for each) and an access level (50% for each). In almost all cases examined, scholars that published copyrighted papers outperformed those that published open access papers. Graph 7 shows the results for 200 scholars with a neighborhood size of 3 and a citation requirement of 9 for various copyright terms. Controlling for access-level, Scholars that publish copyrighted papers systematically outperform those that publish open-access papers. In these settings, copyright hinders overall knowledge production, but...
individual scholars that publish copyrighted materials tend to do better than those that do not.

Graph 7. Success of scholars according to access level and publishing habit. 200 scholars with fixed publishing habits. NS = 3, RC = 9, each scholar assigned access level and publishing habit with equal chance.

Discussion

4.1 The founding fathers of the United States intentionally defined the purpose of the copyright law very narrowly as a policy tool to promote knowledge creation and learning. In this paper, we studied the impact of copyright term length on knowledge creation using agent based modeling. Needless to say, copyright law has a broad scope and applies to many subject matters from literary works to sound recording and architectural works. The scope of this research is limited to literary works and more specifically scientific publications. So the findings of this research should not be generalized to all types of copyrighted subject matters.

4.2 Generally speaking, the extension of copyright term hurts knowledge creation in most scenarios. The adverse effect of the extension is more prominent when scholars have to read and cite more articles before publishing new knowledge. Changes in short copyright terms, especially from open-access to limited copyright, have a large negative impact on knowledge creation. Additionally, the effect is more significant the more scholars there are in the field. When copyright does hinder knowledge creation, the extension of the copyright term hurts all scholars, including those that have access to copyrighted materials.

4.3 To make matters worse, it also seems that scholars will be incentivized to publish copyrighted papers, as those scholars that publish copyrighted papers systematically outperform those that do not. Thus, it seems that, in most circumstances, copyright law simultaneously promotes the publication of copyrighted articles while decreasing the overall knowledge production. Within the context of scientific publications, copyright law may be creating an academic version of the tragedy of the commons, where individual incentives and the public good are misaligned in a way that harms the public good (Hardin 1968).

4.4 The increase in copyright term does not always hurt knowledge creation, however. In some scenarios the extension of copyright term moderately improves the rate of publications and citations. In scenarios where a low number of citations are required to publish in the field, a limited copyright protection policy helps scholars generate more knowledge. We believe copyright protection creates a division of labor in the epistemic plane allowing scholars more effectively explore their field. This scenario may be a representative of emerging research fields with few scholars in which researchers do not need to borrow from literature significantly and the creation of the new knowledge is mostly the result of lab research and is not attributed to the previous works.

4.5 In so far as the sole goal of copyright is knowledge production, we should (ideally) regulate narrowly defined research fields and allow copyright protection only in those fields that the copyright protection helps creation of new knowledge—though other policy options may have other advantages or be far easier to implement. If such a nuanced policy option is not feasible, we recommend cutting copyright protection across academic disciplines to increase access to published works. The reason for this is that, even in cases where copyright helps, the effect is relatively minimal and in other cases the hindrance can be substantial. For most parameter values considered, the access to past research enables scholars to receive more credit, in terms of citations and research impact, for their published works. Receiving more credit will arguably motivate scholars to explore the research field
Historically, copyright law has helped publishers to remain profitable and stay in business, thereby providing an infrastructure for authors to publish new knowledge. It seems plausible, however, that with the advancements in information technology and social media, there are opportunities for authors to become less dependent on publication companies. In any case, assuming that publication companies still play a vital role in generating new knowledge by providing the publishing infrastructure, a complete removal of copyright protection may lead to their bankruptcy and thus would be detrimental to knowledge production. Consequently, some copyright protection may be important for keeping these companies afloat. The results of our model, however, suggest that such protection also hinders knowledge creation in these areas. Therefore, if the ultimate goal of copyright law is knowledge creation, it seems that the length of copyright term should not be longer than is necessary to keep publication companies marginally profitable.

In this paper, we took the first step towards modeling the impact of copyright term on knowledge creation. Our model simulates some essential aspects of knowledge creation in scientific fields to keep the model simple and robust. However, the model also has some important limitations and idealizations. For instance, we did not include any stakeholders beyond scholars in the model, and has represented scholars as relatively simple decision-makers that aim to produce publications in a relatively shortsighted manner. Expanding the model to include entities like publishing companies and universities and track financial incentives for publishing companies is a natural step for future research. Secondly, we assumed that copyright does prevent access from AL0 scholars. In reality, however, low-access scholars may have varying degrees of access to the work. Additionally, the copyright process may cause a time-delay in publication not present for other avenues of publication. We have not included these more nuanced and realistic aspects of copyright and access limitations. Models that included such features would be interesting extensions of the current work.

Notes

1 Using an economic model, Landes and Posner (2009, 71–85) argue that the copyright law should balance financial incentive for creators and the cost of expression. In another economic analysis of copyright law, it is shown that the supply of work and economic incentive are positively correlated for big screen movies. However, Copyright Term Extension Act seems to have insignificant impact on new creative works (Hui & Png 2002). This finding confirms another empirical research in which Tor and Dostan (2002) show the extension of the copyright from lifetime plus fifty years to lifetime plus seventy years provides little additional incentive to create. Landes and Posner (2009, 234–253) compare the current length of the copyright law to a short fixed term with the possibility of indefinite renewal right and conclude that the expected economic life of most copyrighted works is short. They argue that the size of the public domain expands under the latter system, and a system of indefinite renewals will separate valuable works from works in which the cost of continuing that protection exceeds the sum of administrative and access costs.

2 In the field of social epistemology, there have been a number of studies on the general role of information availability and research dissemination for making scientific progress. See, for instance, Zollman (2007, 2010, 2012), Grim (2009), Grim et al. (2013). These studies, however, do not directly examine copyright.

3 See Eldred v. Ashcroft, 537 U.S. 186 (2003) case.

4 On the de facto length of copyright term, Lessig states: “In 1973, more than 85 percent of copyright owners failed to renew their copyright. That meant the average term of copyright in 1973 was just 32.2 years. Because of the elimination of the renewal requirement, the average term of copyright is now the maximum term. In thirty years, then the average term has tripled, from 32.2 years to 95 years.” (Lessig 2004, p. 135)

5 Credit comes in a variety of forms from prestigious prizes to citations. Of these, one sort of credit is most fundamental—the use that one scientist makes of the work of another. The success that is central to science is not career advancement but mutual use. Science has the cumulative character it has in part because of this sort of credit. Because scientists must use the work of other scientists, they are forced to cooperate in a metaphorical sense with even their closest competitors, i.e., use their work.” (Hull 1988, p. 514)

6 The simulation code of the model is available via the CoMSES Computational Model Library. https://www.openabm.org/model/4400/version/1/view
The model and simulations were implemented using NetLogo.

We use a lattice to represent the epistemic plane for simplicity and illustration; any regular uniform connected network with bidirectional links could serve as an epistemic plane.

The concept of an epistemic plane is related to, but importantly distinct from, that of an epistemic landscape (Weisberg & Muldoon 2009; Grim 2009). Epistemic landscapes represent an area of possible attitudes toward a given topic, some of which are more valuable than others. The boundaries of epistemic landscape are determined by the topic, the coordinates of the landscape correspond to approaches, and the height represents the epistemic significance of the particular approach. An epistemic plane, on the other hand, treats all locations as having equal epistemic worth. Its boundaries delineate a field and its points denote particular topics.

In one of the earlier studies in simulating knowledge discovery, Gilbert (1997) designed a fairly simple model to simulate the creation of scientific papers. He introduced the concept of “kene”, analogous to the concept of “gene”, which carries the information of the paper, and each paper represents a new quantum of knowledge. In his model, the kene of a newly produced paper is a function of the cited papers. Coordinates in an epistemic plane are similar to quantum of knowledge in Gilbert's model. Similarities of the kenes create knowledge topics that are closer to each other. The same concept applies to epistemic plane.

The concept of reading articles in this model is similar to acquiring food in models of foraging behavior. Sellers et al. (2007) developed an agent-based model to study the key activities of chacma baboons in South Africa. Baboons, agents in this model, move to other parts of the map to find food or water to fulfill their needs. Roberts and Goldstone (2006) developed an agent-based model to study the group foraging behavior in human. In this model, each agent makes probabilistic movement decisions to find food.

We also examined different distributions of access levels including 25% AL1, 75% AL0 and 25% AL0 and 75% AL1. The results in these cases were qualitatively similar to the results of the 50% distribution presented below.

It is estimated that about 50% of papers are currently available in one of the open access forms in several countries including the United States. (Archambault et al. 2013)

Note that the copyright term is expressed in terms of numbers of cycles in the model, not in terms of years. Cycles represent the publication or research rate of the scholars and the way in which this translates to years will vary from discipline to discipline.

Here are copyright terms we explored in this agent based molding: \{0, 1, 2, 5, 10, 20, 50, 100, 150, 200\}. We ran the experiments with a series of neighborhood sizes and required citations: \{(1,1), (1,2), (2,1), (2,2), (2,3), (2,4), (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (3,7), (3,8), (3,9), (3,10), (3,11)\}.

Such a result is to be expected in the model. Given the nature of exploration on the epistemic plane, the number of potential new publications will increase over time, as the frontier between the known and unknown topics expands. Since our model assumes a fixed number of scholars their density in the will decrease over time. Because scholars spread out over time, the effect of copyright will be reduced as there will also be fewer non-accessible articles nearby.

Under required citations of 5, 7 and 9 AL1 scholars perform slightly better by initial increase of copyright term from 1–5. We think this happens because of division of labor effect. We discuss the division of labor effect in the following section.

Appendix 1

Epistemic Plane – Copyright Term Agent Based Model Overview, Design concepts, and Details (ODD) Protocol

http://jasss.soc.surrey.ac.uk/18/2/23.html
Elements of the ODD protocol

| Overview | 1. Purpose |
|----------|------------|
|          | This model is designed to study the impact of copyright term on knowledge creation. We would like to measure whether increasing the length of the copyright term promotes or stifles the publication and citation of papers and how it impacts scholars with different levels of access to published knowledge. |

2. Entities, state variables, and scales
The model has two entities: scholars (agents) and square patches of knowledge i.e. quanta of knowledge. The quanta of knowledge (patches) are placed on a 100*100 torus. The torus was extended to 200*200 when we had 500 scholars in the model. There are two types of scholars: Access Level One (AL1) who can access all published articles and Access Level Zero, who can access level zero published articles. Quanta of knowledge are either discovered (green) or undiscovered (gray) and discovered knowledge is either published with access level zero (light green) or with access level one (dark green). For each scholar, we have state variables that keep the set of published and cited papers. For each patch we have state variables that keep the cycle in which they get published, access level, and the author.

3. Process overview and scheduling
In each cycle, if possible, scholars publish papers by moving to an undiscovered knowledge point i.e. gray patch within the neighborhood they are located. The neighborhood size is set at the beginning of the simulation. Scholars publish papers only if they have read at least minimum number of papers in the neighborhood, which is also set at the beginning of simulations. If a scholar is unable to publish, she tries to read a published paper that is available based on her access level in the neighborhood. If no paper is accessible in the neighborhood, she will randomly select an accessible paper outside neighborhood to read by moving to the green patch.

In each cycle all the copyrighted papers are checked and if the copyright has been expired, the access level of the paper changes and the article becomes level 0.
4. Design Concepts

The basic principle of the model is the concept that scholars actively aim to create new knowledge. We explore the impact of copyright term length on total number of citations and publications in the model.

Agents decide between publishing and reading. The priority is given to publishing and when there is no publishing option available i.e. when the scholar does not meet citation requirement, she will choose to read.

Scholars can sense published articles (green patches) and whether they have access to them. They can also sense undiscovered knowledge points (gray patches) in the neighborhood they are located, if they have read minimum requirement in the set neighborhood.

Scholars' interaction is indirectly through publications. If a scholar publishes an undiscovered knowledge point, she becomes the author of the paper and receives publishing credit. The authors of the cited articles receive citation credit.

Stochasticity is used in the following processes:
- The access level of the newly published article is randomly selected.
- If a scholar has more than one option to publish or read, she randomly selects the patch.
- Scholars are randomly located in radius 5 from the center of the grid on a published article in which they have access.

To be able to study the impact of copyright term on credits generated through citation and publication, the following outputs are measured:
- Total publications and total citations for each access level of scholars in the set number of cycles
- Average number of shifts that scholars with lower access have to make in order to find articles to read
- Total publication and citation for each access level
- Number of scholars for each access level
5. **Initialization**

- The variables of the patches are reset.
- Scholars are created based on number set on the scroll bar on the interface.
- Scholars are randomly given level one or level zero access to published knowledge.
- Patches in radius 10 are marked as published with unknown author. They receive a random publication cycle between zero and copyright term. The rest of the grid will remain unpublished.
- Scholars are randomly located on published articles in which they have access within radius 2.

6. **Input Data**

The following environmental variables are set before the simulation begins:

- Copyright term: the number of cycles that access level one articles remain under copyright protection.
- Neighborhood size: The size of the neighborhood in which scholars explore to read and publish.
- Minimum required citations: the number of papers that scholars need to read in the neighborhood before publishing new papers.
- Percentage published open access: this number controls what portion of the newly published articles is created open access. The rest will be protected under copyright.

7. **Submodels**

"Explore" sub-model has three components: Publish, Read in the neighborhood, Read beyond neighborhood. Each scholar executes one of these functions in every cycle in the same order as follow. The publish function allows scholars to move to an unpublished patch and publish it as long as they have already read or published (self citations) minimum required papers in the neighborhood. If the scholar cannot publish, she reads preferably in the neighborhood by moving to a published paper as long as she has access to it, otherwise she randomly moves (reads) to a published article outside the neighborhood.

---

**References**

ARCHAMBAULT, E., Amyot, D., Deschamps, P., Nicol, A., Rebout, L., & Roberge, G. (2013). Proportion of Open Access Peer-Reviewed Papers at the European and World Levels—2004–2011. *Rapport, Commission Européenne DG Recherche & Innovation*.

BONILLA, J. P. Z. (2002). Scientific Inference And The Pursuit Of Fame: A Contractarian Approach. *Philosophy of Science, 69*(2), 300–323. [doi:10.1086/341055]

GILBERT, N. (1997). A simulation of the structure of academic science. *Sociological Research Online, 2* (2): http://www.socresonline.org.uk/2/2/3.html.

GOLDMAN, A. I. (1999). *Knowledge in a social world*. Oxford: Oxford University Press. [doi:10.1093/0198238207.001.0001]

GOLDMAN, A. I. (2004). Group knowledge versus group rationality: Two approaches to social epistemology. *Episteme, 1*(01),

http://jasss.soc.surrey.ac.uk/18/2/23.html 20/10/2015
GRIM, P. (2009, November). Threshold Phenomena in Epistemic Networks. In AAAI Fall Symposium: Complex Adaptive Systems and the Threshold Effect. 53–60.

GRIM, P., Singer D. J., Fisher, S., Bramson A., Berger W. J., Reade, C., Flocken, C., and Sales, A. (2013). Scientific Networks on Data Landscapes: Question Difficulty, Epistemic Success, and Convergence. Episteme, 10(4), 441–464. [doi:10.1017/epi.2013.36]

HARDIN, G. (1968). The tragedy of the commons. Science, 162(3859), 1243–1248. [doi:10.1126/science.162.3859.1243]

HUI, K. L., & Png, I. P. (2002). On the supply of creative work: evidence from the movies. American Economic Review, 92(2), 217–220. [doi:10.1257/000282802320189285]

HULL, D. L. (1988). Science as a process: an evolutionary account of the social and conceptual development of science. Chicago: University of Chicago Press. [doi:10.7208/chicago/9780226360492.001.0001]

KITCHER, P. (1990). The division of cognitive labor. The Journal of Philosophy, 87(1), 5-22. [doi:10.2307/2026796]

LANDES, W. M., Posner, R. A., & Landes, W. M. (2009). The economic structure of intellectual property law. Harvard University Press.

LARSEN, P. O., & von Ins, M. (2010). The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. Scientometrics, 84(3), 575–603. [doi:10.1007/s11192-010-0202-z]

LESSIG, L. (2004). Free culture: How big media uses technology and the law to lock down culture and control creativity. New York: Penguin.

PAYETTE, N. (2011a). For an Integrated Approach to Agent-Based Modeling of Science. Journal of Artificial Societies and Social Simulation, 14(4), 9: http://jasss.soc.surrey.ac.uk/14/4/9.html.

PAYETTE, N. (2011b). Agent-Based Models of Science. In Scharhorst A., Börner, K. & Van den Besselaar (Eds.), Models of Science Dynamics. Encounters Between Complexity Theory and Information Sciences., pp. 127–157. New York: Springer Verlag.

ROBERTS, M. E., & Goldstone, R. L. (2006). EPICURE: An Agent-Based Foraging Model. In Artificial Life X: proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems (Vol. 10, p. 379). Harvard: The MIT Press.

SELLERS, W. I., Hill, R. A., & Logan, B. S. (2007). An agent-based model of group decision making in baboons. Philosophical Transactions of the Royal Society B: Biological Sciences, 362(1485), 1699–1710. [doi:10.1098/rstb.2007.2064]

SCOTCHMER, S. (2004). Innovation and incentives. Harvard: MIT press.

STREVENS, M. (2003). The role of the priority rule in science. The Journal of Philosophy, 100(2), 55–79. [doi:10.5840/jphil2003100224]

TOR, A., & Oliar, D. (2002). Incentives to Create Under a Lifetime-Plus-Years Copyright Duration: Lessons from a Behavioral Economic Analysis for Eldered v. Ashcroft. Loy. LAL Rev., 36, 437.

WALTERSCHEID, E. C. (2002). The nature of the intellectual property clause: A study in historical perspective William S. Hein & Co., Inc.

WEISBERG, M., & Muldoon, R. (2009). Epistemic Landscapes and the Division of Cognitive Labor. Philosophy of Science, 76(2), 225–252. [doi:10.1086/644786]

ZOLLMAN, K. J. S. (2007). The communication structure of epistemic communities. Philosophy of Science, 74(5), 574–587. [doi:10.1086/525605]

ZOLLMAN, K. J. S. (2010). The epistemic benefit of transient diversity. Erkenntnis, 72(1), 17–35. [doi:10.1007/s10670-009-9194-6]

ZOLLMAN, K. J. S. (2012). Social network structure and the achievement of consensus. Politics, Philosophy & Economics, 11(1), 26–44. [doi:10.1177/1470594X11416766]

Law References

U.S. Const., Art. 1, Sec. 8, Cl. 8

Copyright Act of 1790, 1 Stat. 124 (1790)

Copyright Act of 1831, 4 Stat. 436 (1831). §16

Copyright Act of 1909, 35 Stat. 1075 (1909). §24

http://jasss.soc.surrey.ac.uk/18/2/23.html
Copyright Act of 1976, 90 Stat. 2541 (1976).

Copyright Term Extension Act of 1998, 112 Stat. 2827 (1998). §102

Eldred v. Ashcroft, 537 U.S. 186 (2003)