Scientific Originality
and the Economics of Publishing

Hugo Mialon and Sue Mialon[1]

Abstract

This paper formulates a simple model of the interaction between authors and reviewers of scientific research. The model assumes authors want to publish their research, reviewers want high quality research to be published, and derivative research, which refines or extends previously published research, is less risky (in terms of the ultimate quality of the resulting paper) than original research, which attempts to break new ground. The model predicts that both the average quality and the publication rate of original papers are greater than those of derivative papers, which may explain why originality is highly valued in the sciences.

(JEL B41 A14 J22)

When a workshop speaker spent the first half hour of his presentation reviewing someone else's model, George [Stigler] interrupted him to say, "If you haven't made any mistakes, have you done anything *original* yet?"

Quoted in Friedland (1993, p. 781)

[1] Department of Economics, Emory University, Atlanta, GA 30322 (e-mail: hmialon@emory.edu and eco_sjh@shsu.edu). We thank Mark Frank, Don Freeman, Dan Hamermesh, Preston McAfee, Max Stinchcombe, and Tom Wiseman for helpful comments.

1. Introduction

In a masterfully written article, Nobel Laureate economist George J. Stigler (1955) denounced what appeared to him to be the excessive value that the economics profession attaches to originality. He asserted that this over-valuation had led the economists of his time to neglect alot of important work by lesser known economists of the previous generations, much to the profession's detriment. And he warned that "Originality means difference, not improvement, and one may invent new errors as well as new truths" (p. 145).

The present paper attempts to explain why the scientific community might attach a great value to originality. The setting is a strategic model of the interaction between authors and reviewers of scientific research. The model assumes that authors want to publish their research and referees want editors to publish high quality research. Authors choose research efforts, while referees choose evaluation efforts. When authors choose research efforts, they do not know with certainty that their efforts will be successful. However, the likelihood of success varies with the type of research they are doing. Types of research are classified into two broad categories: (1) derivative research, which refines, extends, or tests previously published research, and (2) original research, which attempts to break new ground.

The model assumes that original research is more risky, in terms of the ultimate quality of the paper, than derivative research. Original research, while more likely than derivative research to produce huge leaps in scientific knowledge, is also more likely to be barren or misleading. With this assumption, which is in the spirit of Stigler's warning, the model implies that authors working on original papers invest greater research effort than authors working on derivative papers, while referees evaluate original papers with less intensity than they evaluate derivative papers, assuming weakly increasing returns to quality. The resulting equilibrium average quality and publication rate of original papers are greater than those of derivative papers. These results help explain why the scientific community attaches a greater value to original work.

The model developed in this paper is rooted in a growing literature that attempts to formalize the processes by which scientists develop and evaluate new ideas, products, and processes. Evenson and Kislev (1976) model the experimentation choices of individual scientists in a stochastic research environment, Engers *et al* (1999) model the strategic effort choices of coauthors, and Ellison (2002) models the publishing process, but with only authors acting non-mechanically. In our model, the choices of authors and those of reviewers are both endogenous, and we explicitly solve for the mixed-strategy Nash-equilibrium of their interaction.

Section 2 describes the risks and rewards of original empirical and theoretical research. Section 3 develops the model of the strategic interaction between authors and reviewers of scientific research, solves the model for its equilibrium average quality and publication rate of papers, formulates the distinction between original and derivative research, and compares the equilibrium average qualities and publication rates of original versus derivative papers. Section 4 summarizes the results and proposes avenues for further investigation.

2. The Risks of Original Research

Empirical authors often use data to test predictions that derive from theories developed by other authors. Once they have found the appropriate data, the odds that they will discover relationships in the data that tend either to prove or to reject the theories, if they invest the effort, are usually favorable. Moreover, many empirical authors begin their research with available data: if data relevant to a particular theory are not available, they will not begin to work on it.

However, some empirical authors are highly original in finding the right data, instrument, or natural experiment to test theories. Consider the problem of estimating the returns to education. Because individuals who chose to pursue their education may be a selected sample of individuals with higher unobservable ability, the estimated returns to education may be biased upwards. To solve this problem, Angrist and Krueger (1991) used natural variation in birth dates in conjunction with a birth-date cutoff for school entry as an instrument for completed schooling, while Ashenfelter and Krueger (1994) used data on twins that they collected directly in the field at a festival twins attend annually in Ohio.

Although empirical authors can be highly original in finding natural experiments, the natural world does not offer a framework that is flexible enough to test all of the complex theories that have been developed in the social sciences. To overcome this problem, several empirical authors have devised original methods for experimenting with human subjects in controlled environments, even though controlled experimentation has long been reserved for the physical and biological sciences. Vernon Smith, for example, is regarded as the originator of experimental economics, which seeks to test economic theory by observing the reactions of human subjects to economic incentives that are controlled in a laboratory.

However, much experimental research in economics is based on samples consisting of no more than a class-full of college students motivated by modest financial rewards to play economic games on computer terminals. Findings based on such research may offer only limited insight into behavior in the real world. Given the small samples, minor mistakes made by a research assistant or minor differences in the wording of the experiment may have a crucial effect on the results.

Moreover, authors of original empirical research cannot always resist the temptation of eliminating conflicting responses from the sample, especially if they feel strongly about the results they wish to reach. Replications would prevent the emergence of misleading results. However, controlled experiments involve large setup costs, which are barriers to entry that reduce the potential for replication (see Rubinstein, 2001, for a critique of experimental economics from the perspective of a theorist, and Davis and Holt, 2003, for an assessment of several common objections to experimental economics from the practitioner's perspective).

Research based on data that are created by original methods is therefore more likely to mislead the scientific community than research based on publicly available data, which are often compiled by large government agencies that specialize in gathering information on a particular topic. However, the larger risks associated with original empirical research also create the opportunity for larger rewards. Through their work with natural experiments, Angrist, Ashenfelter, Krueger and others have revolutionized the study of labor markets (see Rosenzweig and Wolpin (2003) for a critical assessment of the contribution of the natural experiment approach to empirical economics). And Vernon Smith was awarded the Nobel prize in economic science in 2002 "for having established laboratory experiments as a tool in mainstream analysis."[2]

Like most empirical research, most theoretical research is derivative. Some theorists vary assumptions in existing theories to test whether the results still hold, others generalize existing theories by relaxing assumptions until the results no longer hold, and still others apply existing theories to new environments by appropriately reinterpreting assumptions. The elaboration of theories is sometimes a mechanistic process. Nevertheless, this work is essential to the progress of science. Stigler (1955) notes that,

> At its inception, a theory contains logical ambiguities or errors and its domain of applicability is exaggerated in certain directions and overlooked in others. These

[2] See http://www.nobel.se/economics/laureates/index.html

> deficiencies are gradually diminished as the theory is "worked over" from many directions by many scientists.

However, a science could not progress indefinitely through only this type of work. Not all theorists work on refining or extending the theories of their predecessors. Some theorists begin their work with seemingly heaven-inspired ideas. Based on their general intuition and broad experience, they make novel assumptions, and devise novel algebras, to arrive at novel predictions, or previously discovered facts. For example, Einstein's theory of general relativity deduced the observed discrepancy in Mercury's orbit, and predicted light bending.

The payoffs to deducing novel predictions (and, to a lesser extent, known facts) tend to be more non-linear in effort than the payoffs to extending and refining previous predictions. Even more than original empirical research, original theoretical research has an all-or-nothing nature. The light bending prediction of Einstein's theory of general relativity would later be confirmed empirically, earning Einstein the status of supreme intellectual icon. However, after developing the theory of general relativity (a theory of the very large) early in his career, Einstein worked for the rest of his life on unifying it with the theory of quantum mechanics (a theory of the very small), but to little avail (Denis, 1993, provides an account of Einstein's unsuccessful struggle to develop a "Unified Field Theory").

Proving Fermat's Last Theorem is an other example of an original theoretical endeavor for which effort did not easily translate into success. Several of the best mathematicians of the last three centuries attempted and failed. Andrew Wiles, a professor of mathematics at Princeton, who had dreamed of proving Fermat's last Theorem from childhood, set out to fulfill his life-long ambition in 1990. For seven years, Wiles devoted his efforts exclusively toward this end:

> Nova: During those seven years, could you ever be sure of achieving a complete

proof?

Wiles: I really believed that I was on the right track, but that did not mean that I would necessarily reach my goal. It could be that the methods needed to take the next step may simply be beyond present day mathematics. Perhaps the methods I needed to complete the proof would not be invented for a hundred years. So even if I was on the right track, I could be living in the wrong century.[3]

By working on the project for seven years, Wiles gave up the opportunity to pursue many other projects for which his efforts would have been more likely to yield fruit sooner. Had he abandoned the project after working on it for six years, he would have had little to show his colleagues for his six years of work. Other mathematicians, perhaps as talented as Wiles, were not willing to bear this risk.

Original research, while more likely than derivative research to be barren or misleading, is also more likely to produce huge leaps in scientific knowledge. The next section of the paper formalizes this observation, and analyzes its implications for the strategic interaction between authors and reviewers of scientific research.

3. Originality in a Publishing Game

Three types of players participate in the game of science: authors, referees, and editors. Authors write papers, referees evaluate them, and editors accept or reject them. We model the interaction between authors and referees, leaving editorial choices exogenous. At time 1, each author chooses an effort level $e^A \in \{e_1^A, e_2^A\}$ in writing a single given paper, where $e_1^A < e_2^A$.

Authors cannot foresee how their papers will turn out at the time they choose their

[3] See http://www.pbs.org/wgbh/nova/proof/wiles.html. For a detailed overview of Wiles' proof of Fermat's Last Theorem, see Cornell (1997).

effort levels. At time 2, authors complete their papers, and Nature chooses their quality. Let $\{Q_1, Q_2, Q_3\}$ denote the set of possible qualities of a paper, where $Q_1 < Q_2 < Q_3$, and $Q_3 - Q_2 = Q_2 - Q_1$. Let $q(e^A) = (q_1(e^A), q_2(e^A), q_3(e^A))$ denote the probability distribution over possible qualities $\{Q_1, Q_2, Q_3\}$ for any given effort level e^A by the author, where $\sum_{i=1}^{3} q_i(e^A) = 1$. According to this formulation, a paper's quality depends only on its author's effort level.

The paper's quality is always low if its author exerts low effort, that is, $q(e_1^A) = (1, 0, 0)$. If the author exerts high effort, the probability that the paper's quality is medium or high strictly increases, while the probability that its quality is low strictly decreases. More precisely, $q(e_2^A) = (q_1, q_2, q_3)$, where $q_i \in (0, 1)$ $\forall i \in \{1, 2, 3\}$.

At time 3, the referee chooses an effort level $e^R \in \{e_1^R, e_2^R\}$, where $e_1^R < e_2^R$, in evaluating the author's paper, not knowing the paper's quality, nor the author's effort level. A referee who exerts high effort is more likely to learn the quality of the paper; a referee who exerts low effort is more likely to remain in the dark.

At time 4, the editor makes a decision based on the referee's evaluation: either the paper is accepted (Y) or rejected (N). Let $p_{ij} \in [0, 1]$ be the ultimate probability that the editor accepts the paper for publication, where $p_{ij} = p(Q_i, e_j^R)$. The probability that a paper is published depends on its quality and the effort level that the referee spent on evaluating it.

Assumption 1.

$$p_{11} < p_{21} < p_{31} \tag{1}$$

$$p_{12} < p_{22} < p_{32} \tag{2}$$

$$p_{11} > p_{12} > 0 \tag{3}$$

$$p_{31} < p_{32} < 1 \tag{4}$$

A higher quality paper is more likely to be published than a lower quality paper, regardless of the referee's evaluation effort (parts (1) and (2) of the assumption). As the referee's evaluation effort rises, the editor is better informed about the quality of the paper. Therefore, if the paper's quality is low, then the paper is less likely to be published as the referee's evaluation effort rises (part (3) of the assumption). Indeed, if the paper's quality is at its lowest and the referee's effort level is at its highest, then the editor is most likely to reject it (p_{12} is the smallest of the p_{ij}'s). On the other hand, if the paper's quality is high, then the paper is more likely to be published as the referee's evaluation effort rises (part (4) of the assumption). Indeed, if the paper's quality is at its highest and the referee's effort level is at its highest, then the editor is most likely to accept it (p_{32} is the largest of the p_{ij}'s).

Assumption 2.

$$p_{11} + p_{31} \geq 2p_{21} \tag{5}$$

$$p_{12} + p_{32} \geq 2p_{22} \tag{6}$$

$$p_{12} + p_{32} - 2p_{22} \geq p_{11} + p_{31} - 2p_{21} \tag{7}$$

Improving a paper from low to medium quality increases its publication probability by some extent, but improving it further from medium to high quality increases its publication probability by an even greater extent. In other words, authors enjoy (weakly) increasing returns to quality (parts (5) and (6) of the assumption). This is especially true for publication in good journals. Moreover, the increasing returns to quality are (weakly) greater the greater is the evaluation effort by referees (part (7) of the assumption).

At time 5, authors and referees receive their payoffs. Authors are mainly interested in publishing. Two observations support this assumption. First, authors' ability is often measured by the quantity of their publications in prestigious journals: authors either publish or perish. Second, authors themselves view publication as a measure of their progress, concrete proof that they have not been lazy. For this reason, authors are assumed to receive a utility level of $u > 0$ if their papers are accepted for publication, and a utility level of 0 if their papers are rejected.[4]

Referees do not want low quality papers to be published, and do not want to delay the publication of high quality papers, partly because they want to impress editors, and partly because they care about the progress of science. In particular, referees experience a utility level of m^Y if a high quality paper is accepted, a utility level of m^N if a low quality paper is rejected, a utility level of $-m^Y$ if a low quality paper is accepted, a utility level of $-m^N$ if a high quality paper is rejected, and a utility level normalized to 0 if a medium quality paper is accepted, or if it is rejected.[5] Low effort in writing and evaluating papers is costless. Let c^A denote authors' cost of high effort in writing, and let c^R denote referees' cost of high effort in evaluating.

The extensive form of the Scientific Publishing Game is presented in Figure 1, in the shape of a champagne glass. The initial node is marked by a double circle at the center of the glass's handle. The dotted line is the referee's information set. The vectors attached to terminal arrows have two components: the upper component is the author's utility, the lower component is the referee's utility.

The normal form of the game is presented in Table 1, where for all $i \in \{1, 2, 3\}$ and $j \in \{1, 2\}$,

$$X_{ij} := p_{ij}(-m^Y) + (1 - p_{ij})(m^N) \tag{8}$$

[4] Authors may also care about the quality of their publications, especially if their long term ambition is to maximize their impact on the profession. To the extent that this is true, authors' preferences would be closer to those of referees, and several of the results derived in the paper would be attenuated.

[5] Even though referees care about the progress of science, they can nevertheless make mistakes. Gans and Shepherd (1994) asked leading economists, including recipients of the Nobel Prize, to describe instances where journals rejected their papers. For example, George Akerlof's seminal contribution to the economics of information, "The Market for Lemons: Quality, Uncertainty and the Market Mechanism" (1970), was rejected by the *American Economic Review* and the *Journal of Political Economy* before finally being accepted by the *Quarterly Journal of Economics* four years after Akerlof first submitted it for publication.

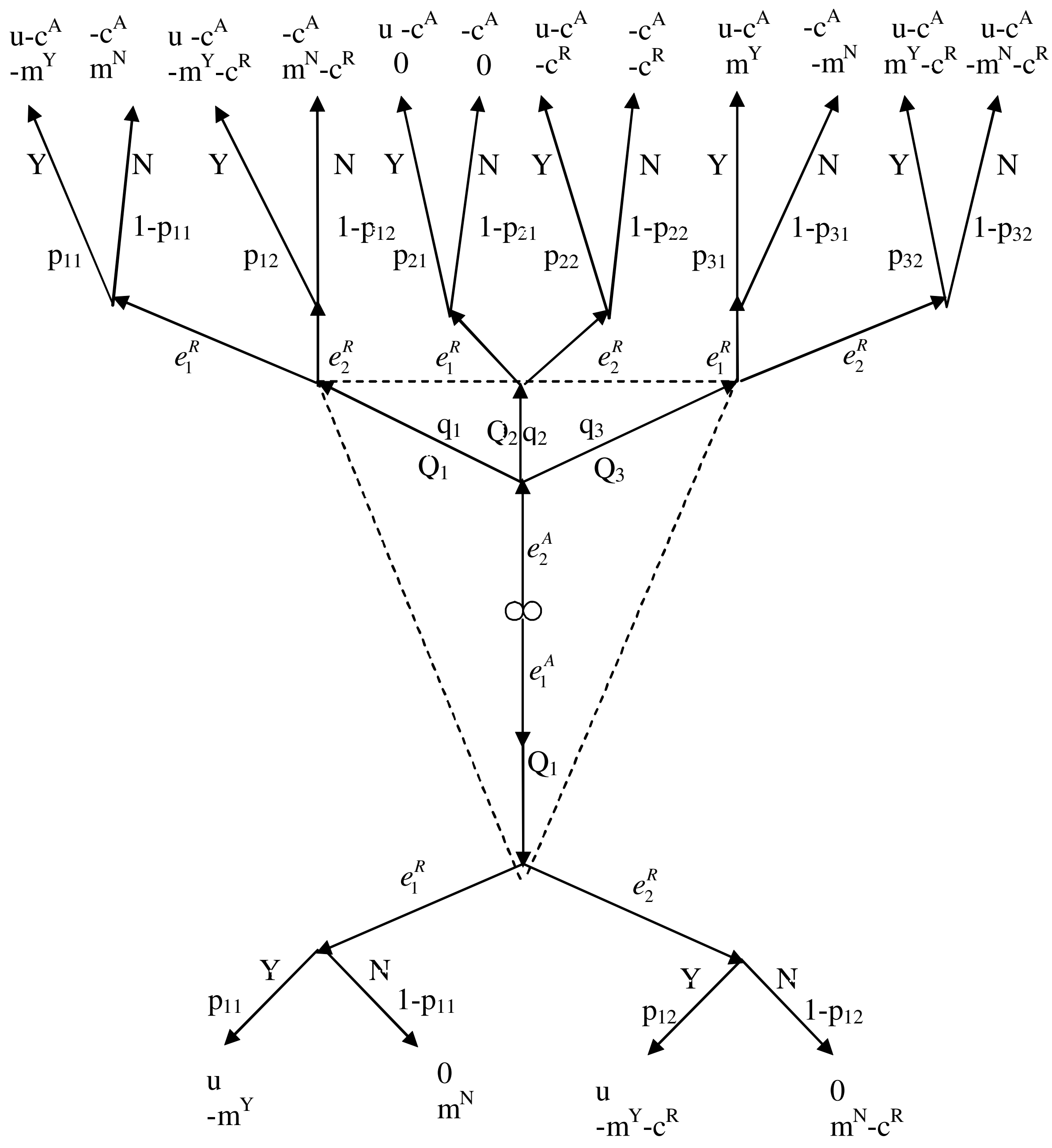

Figure 1: The Scientific Publishing Game, Extensive Form

Table 1: The Scientific Publishing Game, Normal Form

	e_1^R	e_2^R
e_1^A	up_{11}, X_{11}	up_{12}, $X_{12} - c^R$
e_2^A	$u(\sum_{i=1}^3 q_i p_{i1}) - c^A$, $q_1 X_{11} - q_3 X_{31}$	$u(\sum_{i=1}^3 q_i p_{i2}) - c^A$, $q_1 X_{12} - q_3 X_{32} - c^R$

If writing and evaluating papers are neither too costly nor too cheap, the Scientific Publishing Game has the following unique mixed strategy Nash equilibrium:

Proposition 1 *If* $u(\sum_{i=1}^3 q_i p_{i1} - p_{11}) < c^A < u(\sum_{i=1}^3 q_i p_{i2} - p_{12})$ *and* $q_1(X_{12} - X_{11}) + q_3(X_{31} - X_{32}) < c^R < X_{12} - X_{11}$*, then the unique equilibrium is* $\{(\alpha e_1^A, (1-\alpha) e_2^A), (\beta e_1^R, (1-\beta) e_2^R)\}$*, where*

$$\begin{aligned}
\alpha &= \frac{c^R - q_1(X_{12} - X_{11}) - q_3(X_{31} - X_{32})}{(X_{12} - X_{11})(1 - q_1) - q_3(X_{31} - X_{32})} \\
\beta &= \frac{u[\sum_{i=1}^3 q_i p_{i2} - p_{12}] - c^A}{u[\sum_{i=1}^3 q_i p_{i2} - p_{12} - (\sum_{i=1}^3 q_i p_{i1} - p_{11})]}
\end{aligned}$$

Proof. See the Mathematical Appendix for proofs of all propositions stated in the text. ■

Therefore, if writing and evaluating papers are neither too costly nor too cheap, the equilibrium publication rate is

$$\begin{aligned}
P[A] &= \alpha\beta p_{11} + \alpha(1-\beta) p_{12} \\
&\quad + (1-\alpha)\beta(\sum_{i=1}^3 q_i p_{i1}) + (1-\alpha)(1-\beta)(\sum_{i=1}^3 q_i p_{i2}) \\
&= (\beta p_{11} + (1-\beta) p_{12}) \{\alpha + (1-\alpha) q_1\} \\
&\quad + (\beta p_{21} + (1-\beta) p_{22}) \{(1-\alpha) q_2\} \\
&\quad + (\beta p_{31} + (1-\beta) p_{32}) \{(1-\alpha) q_3\}
\end{aligned} \tag{9}$$

As argued before, original research is inherently riskier than derivative research. Let $q^c = (q_1^c, q_2^c, q_3^c)$, where $q_i^c \in [0, 1]$ and $\sum_{i=1}^3 q_i^c = 1$, be the probability distribution over quality Q if authors doing derivative research spend the high level of effort e_1^A, and let

$q^o = (q_1^o, q_2^o, q_3^o)$, where $q_i^o \in [0,1]$ and $\sum_{i=1}^{3} q_i^o = 1$, be the probability distribution over quality if authors doing original research spend the same high level of effort e_1^A.

Assumption 3. The probability distribution over Q if original authors spend e_2^A on research is a mean preserving spread of the distribution over Q if derivative authors spend e_2^A on research: $q_2^o = q_2^c - 2\sigma$, $q_1^o = q_1^c + \sigma$, and $q_3^o = q_3^c + \sigma$, where $\sigma > 0$. Note that the spread preserves mean quality, since $Q_1\sigma - 2Q_2\sigma + Q_3\sigma = 0 \Longleftrightarrow Q_2 - Q_1 = Q_3 - Q_2$, and $Q_2 - Q_1 = Q_3 - Q_2$ is true by assumption.

How do research efforts, evaluation efforts, qualities, and publication rates vary between derivative and original research?

Proposition 2 *The effort intensity of original authors is greater than that of derivative authors, while referees evaluate original papers with lesser intensity than they evaluate derivative papers.*

In the region of parameter space under consideration, if authors put in low research effort, referees would invest high evaluation effort. Then authors' papers would be rightfully rejected with high probability. If authors invested high research effort, referees would put in low evaluation effort. Then authors' papers would end up being wrongfully rejected by editors with high probability. No matter what authors do, referees would best respond by doing something that would increase the rejection rate for the authors. For this reason, authors want to minimize referees' ability to recognize and exploit systematic patterns of behavior in their choices. To do this, they randomize between low and high research effort, and choose their research effort intensity to make referees indifferent between low and high evaluation effort.

But for an given effort intensity by authors, original research is more likely to generate a high quality paper, and more likely to generate a low quality paper, than derivative research. Therefore, with more original research, referees face an increased risk of low quality papers being published or high quality papers being rejected, for any given effort in evaluation. Therefore, original authors need to invest a greater effort intensity in research than do derivative authors in order to make referees indifferent between low and high evaluation effort. For this reason, in equilibrium, the effort intensity of original authors is greater than that of derivative authors.

The extent to which the research effort of original authors is greater is proportional to

$$(X_{12} - X_{11}) - c^R = (m^Y + m^N)(p_{11} - p_{12}) - c^R > 0 \tag{10}$$

which is greater if (1) referees are more averse to editors publishing low quality papers (m^Y is higher), (2) referees are more averse to editors rejecting high quality papers (m^N is higher), (3) higher effort in evaluating papers reduces the publication rate of low quality papers by a larger measure ($p_{11} - p_{12}$ is higher), and (4) evaluating papers is less costly (c^R is lower). These are all incentives for referees to put in high evaluation effort. Referees already have incentive to put in higher evaluation effort for original papers than for derivative papers, since the the latter are more likely to be low or high quality than the former. Therefore, with the additional incentives for referees to put in high evaluation effort, original authors must increase their effort intensity in research by an even greater measure relative to derivative authors in order to make referees indifferent between low and high evaluation effort.

On the other hand, if referees supplied low evaluation effort, authors would supply low research effort. Then low quality papers would end up being accepted by editors with high probability. If referees supplied high evaluation effort, authors would supply high research effort. Then referees would end up needlessly incurring evaluation costs. No matter what referees do, authors would best respond by doing something that would lead referees to want to switch their choice. Thus, referees want to minimize authors' ability to exploit patterns of behavior in their choices too. To do this, they choose their evaluation intensity to make authors indifferent between low and high research effort.

But for any given evaluation intensity by referees, high effort in original research is more likely to generate a low quality paper than high effort in derivative research, which is more likely to generate a medium quality paper. This creates greater incentives for original authors to supply low effort relative to derivative authors. But high effort in original research is also more likely to generate a high quality paper than high effort in derivative research. This creates greater incentives for original authors to supply high effort. But the increase in publication probability between low and medium quality papers is smaller than that between medium and high quality papers. Thus, on the whole, with increasing returns to quality, original authors have greater incentives to supply high effort than do derivative authors. Therefore, referees must reduce their evaluation effort for original papers relative to that of derivative papers in order to make original authors indifferent between low and high research effort. For this reason, in equilibrium, referees evaluate original papers with lesser intensity than they do derivative papers.

Given Proposition 2, the following two results can be proved.

Proposition 3 *The average quality and publication rate of original papers are greater than those of derivative papers.*

Since original authors invest greater effort than do derivative authors on any given paper (the first part of Proposition 2), the average quality of original papers is greater than that of derivative papers. And not only are derivative papers of lower expected quality, but they are also evaluated with greater intensity, than original papers (the second part of Proposition 2), so that referees are more likely to know that they are of lower quality. Therefore, the publication rate of derivative papers must be lower than that of original papers.

While the above results were derived by comparing the choices of a representative original author to those of a representative derivative author, extending the model to allow authors to differ in their effort costs, and giving each of them a choice of whether or not to be original, strengthens the results: low cost authors select into more original research because it allows them to distinguish themselves from high cost authors since it offers a higher variance in returns. For this reason too, more original papers should be of higher quality.

At least one other potential difference between original and critical research would yield the same conclusion. One might argue that most of the "best" ideas have already been pursued. Therefore, the search costs for finding original ideas are relatively large. The author must search those especially unfathomable regions of the intellectual landscape that have yet to be tamed. On the other hand, since derivative research consists of combining and extending previous research, the search costs for finding new derivative ideas are relatively small. The author only has to search those especially visible regions of the intellectual landscape that have already been conquered.

In terms of the above theory, an increase in search costs might correspond to an increase in the author's cost c^A of writing a paper. From Proposition 1, we see that an increase in c^A does not affect the probability α that authors exert low effort, but decreases the probability β that referees exert low effort. Moreover, a decrease in the probability that referees exert low effort results in an increase in the average quality of published papers. Hence, the relative search cost theory also predicts that the average quality of original papers is greater than that of derivative papers.

4. Conclusion

We have formulated a simple game theory of the publication process for scientific research, and performed a comparative static with respect to the risk of original versus derivative research. Assuming only constant or increasing returns to quality, we showed that the equilibrium effort intensity of original authors is greater than that of derivative authors, and the publication rate of original papers is greater than that of derivative papers. These results may help explain why originality seems to be so prized in the sciences.

It would be interesting to test the prediction that the publication rate is greater for original papers. One possible empirical proxy for a paper's originality is the number of times the paper cites other papers. One might think that derivative papers have longer bibliographies than original papers. However, some authors are not aware that their research is not original, that is, that other authors have published similar research. Papers may then have shorter bibliographies, not because they are more original, but because their authors are not abreast of the literature. Moreover, and most importantly, some authors may draw heavily from a small number of papers, while other authors may draw lightly from a large number of papers. The former would then appear to be more original than the latter. This is misleading; in fact, the latter are more original. Thus, derivative papers may actually have shorter bibliographies than original ones.

Originality would be measured more accurately by an index of dispersion in the subfield classifications of the bibliographic elements of the paper. According to this measure, a paper would be more original the broader are the intellectual roots of its underlying research. Research drawing lightly on a broad range of highly divergent ideas would be deemed highly original. On the other hand, research drawing heavily on a small set of very similar ideas would be deemed very derivative. Using a sample of papers submitted to top journals, one might be able to test whether indeed the publication rate is higher for more original papers, controlling for the other characteristics of papers, as well as those of the authors.

A Mathematical Appendix

Proof of Proposition 1. If $u(\sum_{i=1}^{3} q_i p_{i1} - p_{11}) < c^A < u(\sum_{i=1}^{3} q_i p_{i2} - p_{12})$ and $q_1(X_{12} - X_{11}) + q_3(X_{31} - X_{32}) < c^R < X_{12} - X_{11}$ no pure strategy equilibrium exists (because the best responses cycle). Let α denote the probability that an author invests effort level e_1^A, and let β denote the probability that a referee invests effort level e_1^R. Then, an author's expected utility from investing effort level e_1^A is

$$\pi(e_1^A, \beta) = \beta(up_{11}) + (1-\beta)(up_{12}) \tag{11}$$

An author's expected utility from investing effort level e_2^A is

$$\pi(e_2^A, \beta) = \beta[u(\sum_{i=1}^{3} q_i p_{i1}) - c^A] + (1-\beta)[u(\sum_{i=1}^{3} q_i p_{i2}) - c^A] \tag{12}$$

Equating these expected utilities and solving for β yields the unique solution

$$\beta = \frac{u[\sum_{i=1}^{3} q_i p_{i2} - p_{12}] - c^A}{u[\sum_{i=1}^{3} q_i p_{i2} - p_{12} - (\sum_{i=1}^{3} q_i p_{i1} - p_{11})]} \tag{13}$$

Since $u(\sum_{i=1}^{3} q_i p_{i1} - p_{11}) < c^A < u(\sum_{i=1}^{3} q_i p_{i2} - p_{12})$, $\beta \in (0,1)$. On the other hand, a referee's expected utility from investing effort level e_1^R is

$$\pi(e_1^R, \alpha) = \alpha(X_{11}) + (1-\alpha)(q_1 X_{11} - q_3 X_{31}) \tag{14}$$

A referee's expected utility from investing effort level e_2^R is

$$\pi(e_2^R, \alpha) = \alpha(X_{12} - c^R) + (1-\alpha)(q_1 X_{12} - q_3 X_{32} - c^R) \tag{15}$$

Equating these expected utilities and solving for α, yields the unique solution

$$\alpha = \frac{c^R - q_1(X_{12} - X_{11}) - q_3(X_{31} - X_{32})}{(X_{12} - X_{11})(1 - q_1) - q_3(X_{31} - X_{32})} \tag{16}$$

Since $q_1(X_{12} - X_{11}) + q_3(X_{31} - X_{32}) < c^R < X_{12} - X_{11}$, $\alpha \in (0,1)$.

Proof of Proposition 2. Let α^c denote the equilibrium probability that an author exerts high effort in writing a derivative paper, and let α^o denote the equilibrium probability that an author exerts high effort in writing an original paper. Then

$$\alpha^c = \frac{c^R - q_1(X_{12} - X_{11}) - q_3(X_{31} - X_{32})}{(X_{12} - X_{11})(1 - q_1) - q_3(X_{31} - X_{32})} = \frac{c^R - q_1 x_1 - q_3 x_3}{x_1(1 - q_1) - q_3 x_3} \tag{17}$$

and

$$\begin{aligned} \alpha^o &= \frac{c^R - q_1(X_{12} - X_{11}) - q_3(X_{31} - X_{32}) - \sigma(X_{12} - X_{11} + X_{31} - X_{32})}{(X_{12} - X_{11})(1 - q_1) - q_3(X_{31} - X_{32}) - \sigma(X_{12} - X_{11} + X_{31} - X_{32})} \qquad \text{(A.18)} \\ &= \frac{c^R - q_1 x_1 - q_3 x_3 - \sigma(x_1 + x_3)}{x_1(1 - q_1) - q_3 x_3 - \sigma(x_1 + x_3)} \end{aligned}$$

where $x_1 \equiv X_{12} - X_{11}$, $x_3 \equiv X_{31} - X_{32}$.

$$\alpha^c - \alpha^o = \frac{\sigma\left(x_1 + x_3\right)\left(x_1 - c^R\right)}{\left(x_1(1 - q_1) - q_3 x_3\right)\left[\left(x_1(1 - q_1) - q_3 x_3 - \sigma(x_1 + x_3)\right)\right]} > 0 \tag{19}$$

Since $x_1 \equiv X_{12} - X_{11} = (m^Y + m^N)[p_{11} - p_{12}] > 0$ by Assumption 1, and $c^R < x_1 \equiv (X_{12} - X_{11})$ by the mixing condition, $\alpha^c > \alpha^o$. Therefore, any given author exerts lower effort in derivative research than in original research. Moreover, the author's intensity of effort diminishes by a factor proportional to $\sigma\left(x_1 + x_3\right)\left(x_1 - c^R\right) > 0$.

Now, let β^c denote the equilibrium probability that a referee exerts high effort in evaluating an derivative paper, and let β^o denote the equilibrium probability that a referee exerts high effort in evaluating a original paper. Then,

$$\beta^c = \frac{u[\sum_{i=1}^{3} q_i p_{i2} - p_{12}] - c^A}{u[\sum_{i=1}^{3} q_i p_{i2} - p_{12} - (\sum_{i=1}^{3} q_i p_{i1} - p_{11})]} = \frac{u y_2 - c^A}{u(y_2 - y_1)} \tag{20}$$

and

$$\begin{aligned} \beta^o &= \frac{u[\sum_{i=1}^{3} q_i p_{i2} - p_{12}] - c^A + u\sigma[p_{12} + p_{32} - 2p_{22}]}{u[\sum_{i=1}^{3} q_i p_{i2} - p_{12} - (\sum_{i=1}^{3} q_i p_{i1} - p_{11})] + u\sigma[p_{12} + p_{32} - 2p_{22} - (p_{11} + p_{31} - 2p_{21})]} \qquad \text{(A.21)} \\ &= \frac{u y_2 - c^A + u\sigma P_2}{u(y_2 - y_1) + u\sigma(P_2 - P_1)} \end{aligned}$$

where $y_1 \equiv \sum_{i=1}^{3} q_i p_{i1} - p_{11}$, $y_2 \equiv \sum_{i=1}^{3} q_i p_{i2} - p_{12}$, $P_1 \equiv p_{11} + p_{31} - 2p_{21}$, and $P_2 \equiv p_{12} + p_{32} - 2p_{22}$. By Assumption 2, $P_1 \geq 0$, $P_2 \geq 0$, and $P_2 - P_1 \geq 0$.

$$\beta^o - \beta^c = \frac{u\sigma[P_2(c^A - u\,\mathrm{y}_1) + P_1(uy_2 - c^A)]}{[u(y_2 - y_1)]\,[u(y_2 - y_1) + u\sigma(P_2 - P_1)]} \geq 0 \tag{22}$$

Thus, a referee exerts a lower or equal effort in evaluating original papers than in evaluating derivative papers.

Proof of Proposition 3. Let $E[Q]^o$ and $E[Q]^c$ denote the equilibrium average qualities of original and derivative research, respectively. Then

$$\begin{aligned} E[Q]^o &> E[Q]^c \qquad \text{(A.23)} \\ \Longleftrightarrow\ & \alpha^o q_1 Q_1 + (1-\alpha^o)[(q_1+\sigma)Q_1 + (q_2 - 2\sigma)Q_2 + (q_3+\sigma)Q_3] \\ >\ & \alpha^c q_1 Q_1 + (1-\alpha^c)[q_1Q_1 + q_2Q_2 + q_3Q_3] \\ \Longleftrightarrow\ & (\alpha^c - \alpha^o)(q_1Q_1 + q_2Q_2 + q_3Q_3) > (\alpha^c - \alpha^o)q_1Q_1 \end{aligned}$$

Since $\alpha^c - \alpha^o > 0$ by Proposition 2 and $q_1Q_1 + q_2Q_2 + q_3Q_3 > q_1Q_1$, we see that $E[Q]^o > E[Q]^c$. Let $P[A]^o$ and $P[A]^c$ denote the equilibrium publication rates for original and derivative papers, respectively. Then

$$\begin{aligned} P[A]^c &= \alpha^c[\beta^c p_{11} + (1-\beta^c)p_{12}] + (1-\alpha^c)[\beta^c(\sum_{i=1}^{3} q_i p_{i1}) + (1-\beta^c)(\sum_{i=1}^{3} q_i p_{i2})] \qquad \text{(A.24)} \\ &= \beta^c p_{11} + (1-\beta^c)p_{12} + (1-\alpha^c)[\beta^c y_1 + (1-\beta^c)y_2] \end{aligned}$$

and

$$\begin{aligned} P[A]^o &= \beta^o p_{11} + (1-\beta^o)p_{12} + (1-\alpha^o)[\beta^o y_1 + (1-\beta^o)y_2] \\ &\quad + (1-\alpha^o)[\beta^o \sigma P_1 + (1-\beta^o)\sigma P_2] \qquad \text{(A.25)} \end{aligned}$$

Let γ and δ denote the change in effort level by the authors and the change in effort level by the referees, respectively. That is,

$$\begin{aligned}\gamma &\equiv \alpha^c - \alpha^o = \frac{\sigma\left(x_1 + x_3\right)\left(x_1 - c^R\right)}{\left(x_1(1-q_1) - q_3 x_3\right)\left[\left(x_1(1-q_1) - q_3 x_3 - \sigma(x_1+x_3)\right)\right]} \\ &= \frac{\sigma\left(x_1 + x_3\right)}{\left(x_1(1-q_1) - q_3 x_3\right)}(1-\alpha^o) > 0 \qquad \text{(A.26)}\end{aligned}$$

and

$$\delta \equiv \beta^o - \beta^c = \frac{u\sigma[P_2(c^A - u\,\mathrm{y}_1) + P_1(uy_2 - c^A)]}{[u(y_2-y_1)]\,[u(y_2-y_1) + u\sigma(P_2-P_1)]} \geq 0 \qquad (27)$$

Then,

$$\begin{aligned} & P[A]^o - P[A]^c \qquad \text{(A.28)} \\ = \;& \delta(p_{11} - p_{12}) + y_1\left[(1-\alpha^o)\beta^o - (1-\alpha^c)\beta^c\right] \\ & + y_2\left[(1-\alpha^o)(1-\beta^o) - (1-\alpha^c)(1-\beta^c)\right] \\ & + (1-\alpha^o)[\beta^o \sigma P_1 + (1-\beta^o)\sigma P_2] \\ = \;& \delta(p_{11} - p_{12}) + \gamma y_2 - (y_2 - y_1)\left[\gamma\beta^c + \delta(1-\alpha^o)\right] \\ & + (1-\alpha^o)[\beta^o \sigma P_1 + (1-\beta^o)\sigma P_2] \\ = \;& \delta(p_{11} - p_{12}) \\ & + (1-\alpha^o)\left\{ y_2 \frac{\sigma\left(x_1 + x_3\right)}{\left(x_1(1-q_1) - q_3 x_3\right)} - (y_2 - y_1)\frac{\sigma\left(x_1 + x_3\right)}{\left(x_1(1-q_1) - q_3 x_3\right)}\beta^c \right\} \\ > \;& 0 \end{aligned}$$

By Assumption 1, $p_{11} > p_{12}$, and $y_2 \frac{\sigma(x_1+x_3)}{(x_1(1-q_1)-q_3x_3)} > (y_2 - y_1)\frac{\sigma(x_1+x_3)}{(x_1(1-q_1)-q_3x_3)}\beta^c$ also because $y_2 > (y_2 - y_1)$. Therefore, the publication rate of original papers is greater than that of derivative papers.

REFERENCES

Akerlof, George A. (1970), "The Market for Lemons: Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, Vol. 89, 488-500

Angrist, Joshua and Krueger, Alan B. (1991), "Does Compulsory School Attendance Affect Schooling and Earnings," *Quarterly Journal of Economics*, Vol. 106, No. 4, 979-1014

Ashenfelter, Orley and Krueger, Alan B. (1994), "Estimates of the Economic Return to Schooling from a New Sample of Twins, *American Economics Review*, Vol. 84, No. 5, 1157-1174

Ault, Richard W. and Ekelund Jr., Robert B. (1987), "The Problem of Unnecessary Originality in Economics," *Southern Economic Journal*, Vol. 53, No. 3, 651-661

Brian, Denis (1996), *Einstein: A Life*, John Wiley & Sons

Cornell, Gary, Silverman Joseph H, and Stevens, Glenn (1997), *Modular Forms and Fermat's Last Theorem*, Springer-Verlag

Davis, Douglas D. and Holt, Charles A. (1992), "Experimental Economics: Methods, Problems, and Promise," *Working Paper*

Ellison, Glenn (2002), "Evolving Standards for Academic Publishing: A q-r Theory," *Journal of Political Economy*, Vol. 110, No. 5, 994-1034

Engers, Maxim, Gans, Joshua S, Grant, Simon, and King, Stephen P. (1999), "First-Author Conditions," *Journal of Political Economy*, Vol. 107, No. 4, 859-883

Evenson, Robert E. and Kislev, Yoav (1976), "A Stochastic Model of Applied Research," *Journal of Political Economy*, Vol. 84, No. 2, 265-282

Friedland, Claire (1993), "On Stigler and Stiglerisms," *Journal of Political Economy*, Vol. 101, No. 5, 780-783

Freeman, Richard B. (1997), "In Honor of David Card: Winner of the John Bates Clark Medal," *Journal of Economic Perspectives*, Vol. 11, No. 2, 161-178

Gans, Joshua S., Shepherd, George, B. (1994), "How Are the Mighty Fallen: Rejected Classic Articles By Leading Economists," *Journal of Economic Perspectives*, Vol. 8, No. 1, 165-179

Rubinstein, Ariel (2001), "A Theorist's View of Experiments," *European Economic Review*, Vol. 45, 615-628

Rosenzwieg, Mark R. and Wolpin Kenneth I. (2000), "Natural "Natural Experiments," in Economics," *Journal of Economic Literature*, Vol. 38, No. 4, 827-874

Stigler, George J. (1955), "The Nature and Role of Originality in Scientific Progress," *Economica*, Vol. 22, No. 88, 293-302