Evolving Standards for Academic Publishing: A q-r Theory

Glenn Ellison[1]

Department of Economics
Massachusetts Institute of Technology

September 2001

Abstract

Over the last several decades academic journal articles have become longer and changed in other ways. The review process has also changed – it now often involves extensive revisions of initial submissions. This paper develops a model of evolving standards for academic publishing. Papers are assumed to vary along two quality dimensions: q and r. The former is interpreted as the importance of a paper's main ideas and the latter as other aspects of quality. Observed trends are thought of as reflecting increases in r-quality. A static equilibrium model in which an arbitrary social norm determines how q and r are weighted is used to discuss comparative statics explanations for increases in r. The paper then develops a learning model in which referees (who have a biased view of the importance of their own work) try to learn the social norm from observing how their own papers are treated and the decisions editors make on papers they referee. The learning model predicts that social norms will gradually and steadily evolve to increasingly emphasize r-quality.

Glenn Ellison
Department of Economics
Massachusetts Institute of Technology
50 Memorial Drive
Cambridge, MA 02142-1347

1 Introduction

I encourage readers of this paper to first put it down for two minutes and thumb through a thirty- or forty-year-old issue of an economics journal. This should convey better than I can with words how dramatically economics papers have changed over the last few decades. Papers today are much longer.[1] They have longer introductions.[2] They have more sections discussing extensions of the main results. They have more references.[3] The publication process has also changed dramatically. Around 1960 most papers were accepted or rejected on the initial submission.[4] In the early 1970's successful authors would submit a paper, receive reports, make revisions and get a final acceptance within about nine months. Today extensive revisions are the norm and getting an acceptance takes 20 to 30 months at most top economics journals.[5] The phenomenon I am describing is not unique to economics. Similar trends can be seen in many other academic disciplines. In this paper I develop a model to organize the observed trends and develop potential explanations.

Section 2 presents data on the facts to be explained. It discusses the duration of the review and revision process and the form of published papers in a large number of academic disciplines.

Section 3 formulates a static model of the journal review process. The central premise of the 'q-r theory' is that we can usefully regard academic papers as varying along two quality dimensions: q and r. I think of q as reflecting the importance of a paper's main contribution and r as reflecting other aspects of quality (generality, robustness checks, extensions, discussions of related literature, etc.) that are typically improved in revisions and in the final stage of preparing a paper for submission. My thought is that the various

---

[1] Ariel Rubinstein's 1982 *Econometrica* article, "Perfect Equilibrium in a Bargaining Model," for example, is about as long as the average note in *Econometrica* in 2001.

[2] For example, Amartya Sen's "The Impossibility of a Paretian Liberal" and two other papers in the January-February 1970 issue of the *Journal of Political Economy* had one paragraph introductions. No introduction in the Feburary 2000 *JPE* was shorter than seven paragraphs and two were longer than Sen's entire paper.

[3] An extreme example here is Michael Spence's 1973 *Quarterly Journal of Economics* paper "Job Market Signaling." It contains two references. While my anectodes relate to classic papers, old journals also contain many derivative papers making N^{th} order contributions to long forgotten literatures. These also look very different from today's papers.

[4] For example, the *Quarterly Journal of Economics* and *Econometrica* used the "revise-and-resubmit" option about five times per year.

[5] See Ellison (2001) for much more data on the duration of review process at economics journals.

trends noted above can all be regarded reflections of r-quality having increased.

The model features a continuum of academics. They allocate their time between working on q-quality and r-quality in trying to write one paper to submit to the one journal in the profession. How the profession weights q-quality and r-quality in selecting papers for publication is a commonly understood social norm. Referees evaluate submitted papers using this norm and propose improvements that would bring a paper up to the publication threshold. After revisions are made the editor fills the journal's slots by accepting the fraction τ of papers with the highest quality. A crucial assumption is that initial work on a paper determines its q-quality and subsequent revisions only improve r-quality. In the real world there are obviously many dimensions of quality. One can think of q and r in any way that is consistent with the timing assumption.

Section 4 analyzes the equlibria of the static model. For a range of parameter values the model produces a reasonable reflection of the publication process. Papers with the lowest q are deemed sufficiently unimportant so that no feasible revision could make them acceptable. Authors of papers with intermediate q add as much r-quality as they can in the revision process, but are not always successful. Papers with high q need only achieve a moderate r to be publishable. The authors of these papers never fail to execute the necessary revisions. The result of the writing, review, and revision process is a distribution of paper qualities in $q - r$ space. The distribution depends on a number of factors including the technologies for q and r production, uncertainty in the process and the selectivity of the journal. Differences in any of these factors may account for differences in the publication process across fields or within a field over time.

The other basic observation I make about the static model is that a continuum of social norms are possible. If the community agrees that q-quality is very important, then authors will spend most of their time developing main ideas. If r-quality is very important, then authors spend very little time on ideas and focus on revisions. Nothing in the model prevents either extreme or something in the middle from being part of an equilibrium. Differences in social norms provide another potential explanation for differences across fields or over time.

In the companion empirical paper (Ellison, 2001) I've noted that it is hard to attribute

much of the slowdown of the economics publishing process to observable changes in the profession. Providing any equilibrium explanation may be difficult — the slowdown is a dramatic event and my impression is that the economics profession has not changed all that much over the last thirty years. This motivates investigating whether there are reasons to think that social norms might tend to shift. The second half of this paper develops one model of the evolution of social norms that is broadly consistent with observed trends.

Section 5 formulates this model. The most important actors are the journal's referees. They attempt to learn the prevailing social norm from two sources: seeing what revisions they are asked to make on their own papers and seeing whether editors accept or reject papers they refereed.

Section 6 analyzes a base version of the model. It has a continuum of steady-state social norms corresponding the set of equilibria of the static model. The model has an interesting disequilibrium feature: *if* referees try to to hold authors to an infeasibly high standard, then they will learn (correctly) that overall quality standards must be lower than they thought and (incorrectly) that the social norm places more weight on r-quality than they thought. The latter inference reflects that the papers that are unexpectedly accepted are papers with relatively low q's that were revised to death. This dynamic cannot explain the long gradual trends we see in economics — as in most disequilibrium models beliefs converge quickly to a neighborhood of some equilibrium and there is little further movement.

Section 7 adds the assumption that academics are biased and think their work is slightly better than it really is.[6] The equilibria of the no-bias model are no longer steady-states — referees would try to hold authors to the higher standard they mistakenly feel is being applied to their papers. Beliefs about the social norm cannot stray far from the equlibrium set. Instead, referees end up perpetually trying to hold authors to a standard that is just slighly too high. The observation of the previous section about how referees learn when

---

[6]In the psychology literature this is referred to as an overconfidence bias. An often discussed example is Svenson's (1981) finding that about 90 percent of the U.S. college students in his study estimated themselves to be safer (and more skillful) at driving than the median subject. Psychologists have reported that experts in many fields are overconfident in assessing their own ability to answer questions. Lichtenstein, Fischoff and Phillips (1982) provide a nice (but early) survey. Overconfidence is frequently mentioned in finance to motivate the existence of agency problems and to justify the actions of noise traders as in De Long et al (1990). Odean (1998) is a notable recent contribution with a detailed summary of psychology and finance papers on overconfidence. Rabin (1998) discusses a number of other psychological biases and their relevance to economics.

standards that are too high still applies. The result is a gradual evolution of social norms to increasingly weight r-quality. The dynamics are slow and steady. This makes them a plausible candidate for explaining observed trends.

I make a number of additional observations. For example, the model does not just predict social norms will go on trending forever — the evolution ceases well before papers have no q-quality at all. The story is also not just a simple case of overconfidence causing standards to creep up. If referees were underconfident it would not predict a trend toward less r-quality.

There has been little related theoretical work on the dynamics of standards. The most notable are Sobel's analyses of a models in which candidates produce work of multidimensional quality in an attempt to qualify for membership in a club of elites. (One could think of the set of people publishing in a journal as an example.) Sobel (2000) notes that when overall effort is endogenous and judges compare candidates to existing elites random shocks to the weighting of the quality dimensions result in declining standards — candidates allocate their effort to meet the standard at minimum cost and changes in weights can only make it easier to match the assessed performance of earlier successful candidates. Sobel (2001) considers models in which admission to the club is determined by a vote of elites and elites apply heterogeneous quality weights. It describes how different voting rules can lead to rising, falling, or fluctuating standards.

The fact that I have not to followed the current trend and given this paper an overly general title and a seven page introduction should not be taken to indicate that there are not broader lessons to be learned from it. There are all kinds of social norms, e.g. standards for politeness, standards for language and violence on television, hazing at fraternities, hours worked by young doctors and lawyers, years spent in higher education, distributions of grades, etc. Many of these norms have commonly perceived trends, but other than the literature on fashion cycles, e.g. Karni and Schmeidler (1990) and Pesendorfer (1995), most of the existing literature on social norms doesn't focus on dynamics. I will not try to draw conclusions about other norms from my results. My view is that one would need to analyze each application separately and think about how the norm is learned to assess whether a drift should be expected and if so in what direction.

4

On a theoretical level, the paper's innovation is to note that one can produce a model that explains a long gradual trend by making a slight perturbation to a model with a continuum of equilibria. In an early presentation of this paper Robert Barro asked a penetrating question: 'so are you trying to tell us that you're going to explain a thirty-year trend by saying that we've been out of equilibrium the whole time?' My answer is "Yes!" I hope to convince readers that models of the type I introduce make such arguments possible.

2   Some data from various academic disciplines

Table 1 presents evidence on how the form of an academic article has changed. The table lists the average length in pages and the average number of references for articles in top journals in a number of disciplines. Economics papers are roughly twice as long as they were twenty five years ago and have about twice as many references. In almost all fields papers seem to be longer now than in 1975. The increases are more moderate in the sciences. With the exceptions of law and history, articles now also tend to have more references. While economics has experienced substantial reference growth it has a long way to go to catch many social sciences.

Table 2 provides some evidence on how long it takes to review papers and on how extensively they are revised for publication. In almost every case it takes longer to get a paper accepted now than it did in 1975. While *Econometrica* and the *Review of Economic Studies* have the most drawn out publication processes among the listed journals, similar trends are visible in computer science, psychology, statistics, linguistics and finance. I have been told that many rounds of revisions are also the norm in marketing, political science, and a number of other social sciences. A slowdown is also visible in some sciences, but the time scale is completely different.
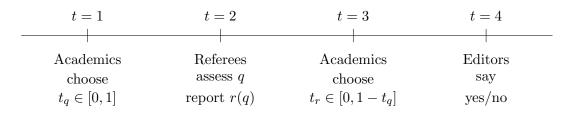
3   The static model

In this section I describe a simple static model of academic publishing. The main actors in the model are a continuum of academics (of unit mass). Each is endowed with one unit of time and may write one paper. There is one journal that publishes a mass τ of papers with

$0 < \tau < 1$. Academics' preferences are lexicographic in publications and leisure time, i.e. they attempt to maximize the probability of publishing an article in the journal; holding the probability of publication fixed they prefer more leisure to less.

Papers can be fully described by two dimensions of quality, q and r. The q dimension is intended to reflect the contribution inherent in the main ideas of the paper. One way I think of q-quality is as a measure of what I would take out of a paper if I were to teach it in a graduate course. The r dimension is intended to reflect additional aspects of quality that may be improved when referees ask authors to generalize theoretical results, to check the robustness of empirical findings, to extend the analysis to consider related questions, to improve and tighten a paper's exposition, to make clear relationships to other papers in the literature, etc.

Social norms for evaluating papers are assumed to be common and commonly known. Under the (α, z)-*social norm*, papers are regarded as worthy of publication if and only if $\alpha q + (1 - \alpha)r \geq z$. The parameter α may reflect two different value judgements. It can reflect what people think makes a paper valuable. It can also reflect what people think authors should be required to do. For example, a referee might argue that while he feels that a particular high q, low r paper is "better" than the marginal paper in a journal, it should still be rejected because the good idea does not excuse the author's failure to make r improvements required of everyone else.

The timeline of the model is illustrated below. While the model is described as a four step process with three groups of players, at the moment the authors are the only ones acting in a nonmechanical way.

$t = 1$	$t = 2$	$t = 3$	$t = 4$
Academics choose $t_q \in [0,1]$	Referees assess q report $r(q)$	Academics choose $t_r \in [0, 1 - t_q]$	Editors say yes/no

In the first stage of the model, authors choose the fraction $t_q \in [0,1]$ of their time to devote to thinking up and developing the main ideas of the paper. The result is a paper of q-quality $q \sim F(q|t_q)$. Assume that F is continuously differentiable in t_q and for each

6

t_q F has an everywhere positive density $f(q|t_q)$ on the interval $[0, m(t_q)]$ (with $m(t_q) > 0$ being possibly infinite). Natural specifications will have the q distribution increasing in t_q. For example, q might be assumed to be uniformly distributed on $[0, t_q]$ or exponentially distributed with mean t_q.

In the second stage of the model authors submit their papers to the journal. The journal's referees correctly assess the quality q of the paper and report that it will be acceptable for publication if and only if authors are able to revise it and achieve an r-quality of at least $r(q)$ as defined by $\alpha q + (1 - \alpha)r(q) = z$. In practice one can think of $r(q)$ as a measure of the number of improvements referees ask for in their reports and the difficulty of the tasks.

In the third stage authors choose the amount of time $t_r \in [0, 1 - t_q]$ to spend on revisions. The production of r-quality is again a random process. Specifically, assume that $r = h(t_r) + \eta$, where η is a random variable uniformly distributed on $[0, \sigma]$ with $\sigma > 0$. Assume that the production of revisions is a decreasing returns activity with $h(0) = 0$, $h' \geq 0$ and $h'' < 0$. To ensure that time will be allocated to both dimensions of quality I assume also that $h'(0) = \infty$, and $h'(1) = 0$.

In the fourth stage, editors accept the fraction τ of papers for which $\alpha q + (1 - \alpha)r$ is highest for publication. Note that editors have a minor role in the model — they do not try to impose a personal view of overall quality and instead only make the minor adjustment of moving the bar up or down to ensure that the proper number of papers are accepted. I view this as a good decriptive model of many busy editors. I think that some such model is necessary to account for why economists continue to submit fifty page papers with myriad extensions despite editors' claims that they abhor this and wish that authors would just concisely explain their ideas.

An *equilibrium* of the model is a quadruple, $(\alpha, z, t_q^*, t_r^*(q))$ such that t_q^* and $t_r^*(q)$ are chosen to maximize the probability that $\alpha q + (1 - \alpha)r \geq z$ (and are as small as possible if there are multiple choices that yield the same probability of publication) and such that the fraction of papers with $\alpha q + (1 - \alpha)r \geq z$ is exactly τ. I will refer to (α, z) as a *consistent social norm* if there exist choices of t_q and $t_r(q)$ for which $(\alpha, z, t_q, t_r(q))$ is an equilibrium.

4 Analysis of the static model

4.1 Characterization of equilibrium

The analysis of the equilibrium is a straightforward backward induction argument. Consider an equilibrium $(\alpha, z, t_q^*, t_r^*(q))$. Because of the lexicographic preference for publications over leisure, at $t = 3$ authors will devote all of their remaining time to improving their paper's r-quality unless the paper is sure to be rejected anyway or is sure to be accepted even if less time is devoted to revisions. As a result we have,

Proposition 1 *In any equilibrium of the model, let $\underline{q} = (z - (1 - \alpha)(h(1 - t_q^*) + \sigma))/\alpha$ and $\overline{q} = (z - (1 - \alpha)h(1 - t_q^*))/\alpha$. Then, $\underline{q} < \overline{q}$ and*

$$
t_r^*(q) = \begin{cases}
0 & \text{if } q < \underline{q}, \\
1 - t_q^* & \text{if } \underline{q} < q \leq \overline{q}, \\
h^{-1}((z - \alpha q)/(1 - \alpha)) & \text{if } \overline{q} \leq q \text{ and } \alpha q < z. \\
0 & \text{if } \overline{q} \leq q \text{ and } \alpha q \geq z.
\end{cases}
$$

Note that $\underline{q}$ may be less than zero and $\overline{q}$ may be greater than the upper bound of the support of q. Hence, the extreme cases may not arise for particular parameterizations of the model.

In the first stage, the time t_q allocated to trying to develop the main ideas for a paper will be chosen to maximize the probability of eventual publication. Write $G(z; t)$ for the probability that $\alpha q + (1 - \alpha)r$ is at most z when $t_q = t$ and t_r is chosen optimally as in the above proposition. Note that this probability is the same as it would be if t_r was simply set equal to $1 - t_q$. Hence,

$$
G(z; t) = \int_{q=0}^{\min\left(m(t), \frac{z - (1-\alpha)h(1-t)}{\alpha}\right)} \min\left(1, \frac{z - \alpha q - (1 - \alpha)h(1 - t)}{(1 - \alpha)\sigma}\right) f(q|t)dq
$$

Note that G is uniformly continuous in z and t. It is strictly increasing in z whenever $G(z; t)$ is strictly between zero and one. Write $G^{-1}(p; t)$ for the inverse of this function for a fixed t. The equilibrium time allocation is easily described by

Proposition 2 *In the first stage of any equilibrium, the time allocated to developing q-quality satisfies*

$$
t_q^* \in Argmax_t G^{-1}(1 - \tau; t).
$$

8

<u>Proof</u>

In equilibrium, each author's paper is accepted with probability τ. Hence, z must satisfy $G(z; t_q^*) = (1 - \tau)$. If t_q^* does not belong to $\text{Argmax}_t G^{-1}(1 - \tau; t)$, then any t which does maximize that expression has $G^{-1}(1 - \tau; t) > G^{-1}(1 - \tau; t_q^*) = z$. Because G is strictly increasing in z whenever $0 < G(z; t) < 1$, this implies that $G(z, t) < G(G^{-1}(1 - \tau; t)) = \tau$, which contradicts the optimality of t_q^*.

QED.

4.2   Some examples

Figure 1 illustrates the distribution of paper qualities in a "typical" equilibrium — in this case generated by assuming that $\tau = 0.3$, q is uniformly distributed on $[0, t_q]$, the technology for r production is $h(t_r) = \sqrt{t_r} - t_r/2$ with $\sigma = 0.2$, and the social norm for judging papers has $\alpha = 0.5$ and $z \approx 0.504$.

For these parameters, the equilibrium effort allocated to q production turns out to be $t_q^* \approx 0.826$. All three possible outcomes of a submission occur. Authors of papers with q-quality less that $\underline{q} \approx 0.48$ realize that their papers have no chance of becoming acceptable and do not attempt to revise them. Authors of papers with $q \in (\underline{q}, \overline{q}) \approx (0.48, 0.68)$ devote as much time as possible to revising (setting $t_r = 1 - t_q^* \approx 0.174$) and have their papers accepted with probability strictly between zero and one. Authors of papers with quality $q \in [\overline{q}, m(t_q) = t_q^*] \approx [0.68, 0.83]$ do the minimal revision necessary to ensure that their papers will be accepted with probability one.

The figure shows the outline of the support of the equilibrium distribution of paper qualities in (q, r)-space. Paper qualities are distributed with a constant density within these regions. Papers in the lower left box are those for which authors set $t_r = 0$. These papers are never accepted. The upper right region is divided into a triangle of papers that are revised then rejected and a trapezoid of papers that are accepted. The mass of papers in this upper region is, of course, τ.

FIGURE 1 HERE

The form of the equilibrium seems to reflect fairly well the functioning of an economics journal. One observation I'd like to make is that the "marginal" rejected papers have

relatively low q-quality compared to the pool of accepted papers — they all have $q \in [\underline{q}, \overline{q}]$. The marginal rejected papers are not relatively low in r-quality. Their authors have spent as much time as possible revising and achieved r-qualities that are on average superior to those of the accepted papers.

While I think that the case illustrated above is the primary one of interest and subsequent arguments focus on it, the equilibrium can take other forms for different parameter values. Most notably, when α is sufficiently small (i.e. q-quality is of little importance) all authors set $t_r = 1 - t_q^*$ — there will always be some chance that any idea, no matter how vacuous, can be developed into a publishable paper and no paper's idea is good enough to make its eventual acceptance a sure thing. The following proposition formalizes this observation. Figure 2 graphs the equilibrium distribution of paper qualities and the acceptance and rejection regions for such a case: $q \sim U[0, t_q]$, $h(t_r) = \sqrt{t_r} - t_r/2$, $\sigma = 0.2$, $\alpha = 0.2$ and $z \approx 0.532$.

FIGURE 2 HERE

Proposition 3 *Suppose that the upper bound of the q-distribution, $m(t_q)$, is finite and uniformly bounded for all $t_q \in [0, 1]$. Then, there exists $\overline{\alpha} > 0$ such that for all $\alpha \in (0, \overline{\alpha})$, $t_r^*(q) = 1 - t_q^*$ for all $q \in [0, m(t_q^*)]$, i.e. all papers are revised to the greatest extent possible and no paper achieves a level of q-quality sufficient to ensure that it will be accepted with probability one.*

Proof

To see that $\underline{q} \le 0$ note that if a paper with $q = 0$ is revised to the greatest extent possible and gets the best possible draw on r-quality its overall quality will be $(1-\alpha)h(1-t_q^*)+\sigma$. In equilibrium,

$$\text{Prob}\{\alpha q + (1-\alpha)r \ge (1-\alpha)(h(1-t_q^*)+\sigma)\} \le \text{Prob}\{r \ge h(1-t_q^*) + \sigma - \frac{\alpha}{1-\alpha}m(t_q^*)\}$$
$$= \min\left(1, \frac{\alpha}{1-\alpha}\frac{m(t_q^*)}{\sigma}\right)$$

For α sufficiently small the expression is less than τ, and hence there is a positive probability of the paper with $q = 0$ being acceptable.

10

Similarly, to see that $\bar{q} > m(t_q^*)$ note that if a paper with q-quality $m(t_q^*)$ is revised to the greatest extent possible but gets the worst possible draw on r-quality its overall quality level is $\alpha m(t_q^*) + (1-\alpha)h(1-t_q^*)$ and

$$\text{Prob}\{\alpha q + (1-\alpha)r \geq \alpha m(t_q^*) + (1-\alpha)h(1-t_q^*)\} \geq \text{Prob}\{r \geq h(1-t_q^*) + \frac{\alpha}{1-\alpha}m(t_q^*)\}$$
$$= \max\left(0, 1 - \frac{\alpha}{1-\alpha}\frac{m(t_q^*)}{\sigma}\right)$$

Hence, for any $\tau < 1$ the probability that a paper with q-quality $m(t_q^*)$ fails to be among the best τ is strictly positive if α is sufficiently small.

QED.

Some of the results I'll give later will depend on the form of the equilibrium. To simplify the statements of these results I will give the forms pictured in the figures (and a couple of other forms) names.

Definition 1 *I will say that an equilibrium is "typical" or has the typical form if $0 < \underline{q} < \bar{q} < m(t_q^*)$ and $r(m(t_q^*)) > 0$. An equilibrium has the "low α" form if $\underline{q} < 0 < m(t_q^*) < \bar{q}$. It has the "somewhat low α" form if $0 < \underline{q} < m(t_q^*) < \bar{q}$. It is of the "high α" form if $0 < \underline{q} < \bar{q} < m(t_q^*)$ and $r(m(t_q^*)) < 0$.*

The somewhat low α form is similar to the low α form, with the only difference being that in the former authors of the lowest q papers do not revise their papers. The high α form is similar to the typical form, but the highest q papers are so good that $r(q)$ is negative, i.e. referees tell the authors that even if they revised the paper to make it worse it would still be publishable. The authors of these papers obviously exert no effort on revisions and have their papers accepted with quality to spare. In the model with $q \sim U[0, t_q]$, $h(t) = \sqrt{t} - t/2$, $\sigma = 0.2$ and $\tau = 0.3$ the equilibrium has the low α form for $\alpha \in (0, 0.2285)$, the somewhat low α form for $\alpha \in (0.2286, 0.3363)$, the typical form for $\alpha \in (0.3364, 0.5869)$ and the high α form for $\alpha \in (0.5870, 1)$. In other specifications for the model the equilibrium can take on other forms. For example, if the distribution of q is unbounded (or τ is large) we can simultaneously see papers of the lowest q-quality being resubmitted and papers of the highest q-quality being accepted with no revisions.

11

4.3 The multiplicity of consistent social norms

In the model described above not all social norms are consistent. If there is only room in the journal for a small fraction of papers, then the quality threshold z must be high. This, however, is really the only constraint. Nothing in the model restricts the weight the community places on q-quality versus r-quality. There are a continuum of consistent social norms with any α being possible. The following proposition gives a formal statement to this effect, and Figure 3 graphs the set of consistent social norms for the model with $q \sim U[0, t_q]$, $h(t) = \sqrt{t} - t/2$, $\sigma = 0.2$ and $\tau = 0.3$.

Proposition 4 *In the model described above, for any $\alpha \in [0, 1]$ there exists an unique $z^*(\alpha)$ such that $(\alpha, z^*(\alpha))$ is a consistent social norm.*

Proof

For any fixed α, let $G(z; t)$ be the CDF of $\alpha q + (1-\alpha)r$ as above. Let $H(z) = \inf_t G(z; t)$. Because G is uniformly continuous, t is chosen from a compact set and $\lim_{z \to \infty} G(z; t) = 1$ for all t, H is continuous with $H(0) = 0$ and $\lim_{z \to \infty} H(z) = 1$. Hence there is a solution $z^*(\alpha)$ to $H(z) = 1 - \tau$. $(\alpha, z^*(\alpha))$ is a consistent social norm.

It is not possible for both (α, z) and (α, z') to be consistent social norms with $z < z'$. In that case, an agent setting t_q equal to the the equilibrium choice under the (α, z') norm would surpass the z threshhold with probability greater than τ.

QED.

FIGURE 3 HERE

4.4 Explanations for observed increases in r-quality

The various trends in economics papers mentioned in the introduction can all be thought of as reflecting an increase in the r-quality of published papers. The length of papers, the space devoted to introductory material and related literature, and the number of extensions a paper develops are direct measures of aspects of r-quality. A longer review and revision process may also be associated with higher levels of r-quality if an efficient way to generate r-quality is to have authors work jointly with other experts in the field.

The model makes clear that a number of explanations for an increase in r-quality are possible. One idea that has often been suggested to me is that r-quality may be increasing because research is adding to a growing stock of knowledge. This might be captured in the model by assuming that the technologies for producing q and r are changing. On a depressing note, one could argue that academics have begun to exhaust the set of ideas within a paradigm and model this as a change in $F(q|t_q)$ that reduces the marginal benefit of t_q. More optimistically, one could argue that accumulated knowledge or advances in technology allow authors to produce more r-quality per unit time. Either way, the straightforward prediction of the model would be that more effort will be devoted to r-quality and published papers will be higher in r-quality.

Another potential explanation is that academic communities may have grown and publishing may have become more competitive.[7] We could think of this as a decrease in τ. This might lead to an increase in paper quality along both dimensions, but might not. Intuitively, reducing τ has two effects: it allows the journal to be more selective when choosing from the pool of resubmitted papers and affects authors' time allocation decisions. If authors react to increased competition by gambling on bold projects that require a lot of t_q, the overall effect of τ on r-quality can be ambiguous.

Why might decreases in τ lead authors to gamble on high values of t_q? An intuitive answer is that t_q^* reflects authors setting α times the effect of t_q on q-quality equal to $(1-\alpha)$ times the effect of t_r on r-quality conditional on their luck being such that their paper is marginal for the journal. When the journal is highly nonselective, this is conditioning on getting a very bad draw from the q distribution. When the journal is extremely selective, the conditioning is on getting a very good draw. With a functional form like $F(q|t_q) \sim U[0, t_q]$, higher percentiles of the q distribution increase more with increases in t_q, and hence the return to t_q is greater when one conditions on the paper being marginal for a more selective journal.

Figure 4 illustrates the effects of changes in τ on time allocation and quality in a model

---

[7]In Ellison (2001) I provide some support for such an explanation in economics. I note that the economics profession has grown moderately over the last 30 years, but that this growth together with decreases in the number of articles published by some top journals and an increase in the relative status of the top journals may have led to a substantial increase in competition.

with $q \sim U[0, t_q]$, $h(t) = \sqrt{t} - t/2$, $\sigma = 0.2$, and $\alpha = 0.5$ The left panel shows that t_q^* is decreasing in τ. The right panel shows the effect of τ on the mean q- and r-quality of the published papers. The relationship beween τ and r-quality is nonmonotonic. At some points the effect of τ on t_q^* dominates, while at others the effect of journal selectivity dominates.

FIGURE 4 HERE

The explanation I focus on in the latter half of this paper is that the r-quality of published papers will increase if α decreases, i.e. if social norms shift to place relatively less emphasis on q-quality and relatively more emphasis on r-quality. This effect is again straightforward: a decrease in α raises the r-quality of published papers for two reasons: academics react by allocating more of their time to producing r-quality; and when choosing from the pool of resubmitted papers the journal places more emphasis on r-quality.

5   A dynamic model of evolving norms

The static model above has a continuum of equilibria corresponding to different social norms. In this section I describe a dynamic model of the evolution of norms. The model involves a population of author-referees who are trying to learn and apply the profession's standards. In contrast to what one might think when one hears the term "learning model", I will not model agents as arriving with different beliefs and examine whether there is convergence to a common belief. Instead, the model will be constructed so that agents will have common beliefs at every point in time and I will focus on whether agents' attempts to learn the prevailing norm lead to a shift in norms.[8]

The model involves a discrete set of time periods $t = 0, 1, 2, \ldots$. At the start of period t all academics believe that that the social norm is (α_t, z_t). They then write a paper and try to publish it as in the static game of Section 2 and serve as a referee. The data they receive via referee reports and editorial decisions will suggest to them that the social norm

---

[8]This focus is not intended to suggest that whether a community will reach a common norm and the effects of belief or preference heterogeneity on the evolution of social norms are not very interesting. Instead, as in Ellison and Fudenberg (2000), the common beliefs learning model is motivated solely by the desire for a tractable model that highlights an important effect.

is in fact, $(\hat{\alpha}_t, \hat{z}_t)$. This leads them to alter their beliefs according to

$$\left(\begin{array}{c} \alpha_{t+1} - \alpha_t \\ z_{t+1} - z_t \end{array} \right) = k \left(\begin{array}{c} \hat{\alpha}_t - \alpha_t \\ \hat{z}_t - z_t \end{array} \right)$$

for some constant $k \in (0, 1)$. To complete the specification I must describe what data academics gather and how they draw the $(\hat{\alpha}_t, \hat{z}_t)$ inferences from the data.

What data do academics get? I assume that academics get two types of data points. First, when an author submits a paper of quality at least $\underline{q}$ I assume that the referee reports he or she receives give him or her a data point of the form $(q, r(q))$.[9] These datapoints should all lie on the line $\alpha_t q + (1 - \alpha_t) r(q) = z_t$. I will allow, however, for the possibility that academics are subject to an overconfidence bias when judging the quality of their own work. In particular, I'll assume that they overestimate the r-quality of their initial submission and this leads them to believe that they have been required to achieve an r-quality that is ϵ higher than what they have actually been required to achieve.[10] In this case, the $(q, r(q))$ datapoints actually lie on the line $\alpha_t q + (1 - \alpha_t) r(q) = z_t + (1 - \alpha_t)\epsilon$.

Second, whenever an academic referees a paper that is of sufficiently high quality to be resubmitted he or she gets a data point of the form $(q, r, \text{Accept}/\text{Reject})$.[11] Academics expect to see all papers lying above the line $\alpha_t q + (1 - \alpha_t) r = z_t$ being accepted and all papers lying below this line being rejected.

If each academic saw a finite number of data points, then their analyses would lead to a divergence in the second period beliefs even if they entered the model with common beliefs. To keep the model tractable I avoid this by invoking word-of-mouth communication. I

---

[9]Note that in a slight departure from the static model I've assumed that journals do not provide the author with a list of revisions sufficient to make the paper publishable if the paper's q-quality is so low as to make it inconceivable that a revision will be publishable. I believe that this is a good description of journal practices. I did not try to incorporate such behavior by referees in the static model, however, because it seemed a needless complication and because it creates a possibility for another type of equilibrium multiplicity that I did not feel was important: authors won't save time for revisions if referees won't ask for large revisions because they don't think they're feasible.

[10]Given that the acceptance fronteir is downward sloping in q-r space this assumption is almost equivalent to assuming that authors believe their papers to be of slightly higher q-quality than they actually are, or to assuming that have biased views of both the q- and r-quality of their work. I've chosen the formulation above because it makes some results a little cleaner (especially those about small α behavior.)

[11]Another important data source for real world academics is reading journals. Given the word-of-mouth assumption below, this would just provide redundant observations on all of the acceptances. For this reason, nothing would be changed if I included this data source in the model. An ambiguity in the model is what happens if the assumed standards is so excessively high that fewer than τ papers are resubmitted. In the simulations in the next section I assume that in this case editors accept some papers that are not resubmitted and that these acceptances are observed by the referees.

assume that each academic talks to every other academic in each period and thereby sees all of the data points that were generated in that period. While this makes each author's own experiences a measure zero subset of his or her dataset, I do not want to lose the possibility of inferences being affected by authors' misperceptions of the quality of their own work. I thus assume that the $(q, r(q))$ observations academics receive by hearing others talk about the referee reports they received are contaminated by the authors' biases (and that the listeners do not realize this).

What do academics do with the data they obtain each period? Fitting the (α, z) model involves estimating the slope and intercept of a line that fits the $(q, r(q))$ data and divides the acceptance and rejection regions. Typically, no line will do both jobs perfectly. I assume that this does not cause academics to lose faith in their model of the world and that they go ahead and try to fit the data as well as possible with the (α, z) model. A justification for not worrying about academics noticing that they are estimating a misspecified model is that in a more realistic model academics would only receive a finite number of datapoints and there would be a random component to each observation, so the form of the misspecification would not be so apparent. The idea that academics struggle to reconcile hard-to-reconcile observations without abandoning biased self images does not seem unrealistic to me.

Formally, I assume that academics' period t analyses take the form

$$(\hat{\alpha}_t, \hat{z}_t) = \arg \min_{\alpha, z} L(\alpha, z; \mu_1, \mu_2),$$

where $L(\alpha, z)$ is a loss function that describes how poorly the the data (a measure μ_1 describing the $(q, r(q))$ points and a measure μ_2 describing the $(q, r, Accept/Reject)$ points) fit the hypothesis that all referees and the journal editor are applying the (α, z) social norm. Specifically, I assume that

$$L(\alpha, z; \mu_1, \mu_2) = L_1(\alpha, z; \mu_1) + L_2(\alpha, z; \mu_2),$$

where

$$L_1(\alpha, z; \mu_1) = \int \left(r(q) - \frac{z - \alpha q}{1 - \alpha} \right)^2 d\mu_1(q),$$

is a standard mean squared deviation measure of the distance (in the r-dimension) between

16

the $(q, r(q))$ data points and the line $\alpha q + (1 - \alpha)r = z$ and

$$L_2(\alpha, z; \mu_2) = \int_{R_{UA}} \left(\frac{z - \alpha q}{1 - \alpha} - r \right) d\mu_2(q, r) + \int_{R_{UR}} \left(r - \frac{z - \alpha q}{1 - \alpha} \right) d\mu_2(q, r),$$

where R_{UA} is the set of (q, r) values for which papers were "unexpectedly accepted" despite failing to meet the (α, z) standard and R_{UR} is the set of "unexpectedly rejected" papers that met the (α, z) standard but were rejected. L_2 can be thought of as the product of the fraction of accept/reject decisions that are inconsistent with the (α, z) model and the average degree of error (in the r-dimension) that appears to be embodied in the "unexpected" decisions.

Obviously other loss functions would be reasonable. The most natural would probably be the negative of the log likelihood of the data under a hypothesis involving referees and the editor trying to apply the (α, z) norm but making idiosyncratic errors in judging the quality of each paper. Analyzing such a specification would require examining integrals of CDF's and PDF's, however, and I felt that the specification above was the best compromise in terms of reflecting a similar goodness-of-fit notion and being tractable.

6 Analysis of the dynamic model with no overconfidence bias $(\epsilon = 0)$

In this section I discuss the behavior of the dynamic model when academics do not have an inflated view of the quality of their own work. The main observations are that consistent social norms are steady states of the model and that when referees are too demanding academics infer both that their standards were too high and that q-quality must be relatively less important than they had thought.

6.1 Steady states

When there is no overconfidence bias it is easy to see that any consistent social norm of the static model is a steady state of the dynamic model.

Proposition 5 *Suppose $\epsilon = 0$ in the dynamic model and (α_0, z_0) is a consistent social norm. Then, $(\alpha_t, z_t) = (\alpha_0, z_0)$ for all t.*

<u>Proof</u>

All points $(q, r(q))$ in the data obtained from referees' reports lie exactly on the line $\alpha_0 q + (1 - \alpha_0) r(q) = z_0$. The editor's decisions are also consistent with imposing the (α_0, z_0) standard, i.e. all rejected papers have $\alpha_0 q + (1 - \alpha_0) r - z_0 \leq 0$ and all accepted papers have $\alpha_0 q + (1 - \alpha_0) r - z_0 \geq 0$. Hence, both L_1 and L_2 are zero for $(\alpha, z) = (\alpha_0, z_0)$. Because the q-distribution is nonatomic on a continuous support, L_1 is strictly positive (and L_2 is always nonnegative) for any other (α, z). The unique minimum of the loss function is thus $(\hat{\alpha}_0, \hat{z}_0) = (\alpha_0, z_0)$.

QED.

6.2 Disequilibrium dynamics

In this section I discuss the disequlibrium behavior of the dynamic model. The results in this section (and in the remainder of the paper) will concern the uniform technology for q-production, $q \sim U[0, t_q]$. The whole of what I want to say in this section can be summarized concisely by saying that the dynamic evolution of academics' beliefs about the social norm (α, z) outside of equilibrium follows the pattern illustrated in Figure 5. The figure was constructed by solving the model numerically for various initial beliefs under the assumption that $q \sim U[0, t_q]$, $\tau = 0.3$, $\sigma = 0.2$ and $h(t) = \sqrt{t} - t/2$. The solid line in the figure is the locus of consistent social norms $(\alpha, z^*(\alpha))$. The vectors in the figure are proportional to the change in beliefs $(\alpha_{t+1} - \alpha_t, z_{t+1} - z_t)$ that occurs for various initial beliefs.[12] To help organize the dynamics I have placed a vertical dashed line at $\alpha = 0.3363$. This is where the equilibrium shifts from the somewhat low α form to the typical form.

FIGURE 5 HERE

The locus of consistent social norms and the dashed line divide (α, z) space into four regions. In three of the four regions the learing process is mostly just a straightforward adjustment of the overall quality threshold: when referees try to impose a standard that would not allow the editor to fill the journal they infer from the unexpected acceptances

---

[12]The speed with which beliefs change in the model is determined by an arbitrary scaling parameter k. The figure uses $k = 0.6$. Note that unlike all other graphs in this paper the x and y axes in this figure do not have the same scale. The y-axis has been magnified by a factor of 1.5. These choices reflect an attempt to maximize the visibility of the directions and minimize clutter.

that they must reduce z; when referees are too soft and the editor has to turn down some papers they recommend they learn to choose a higher z.[13]

What is most important to my main argument is what happens in the fourth region — the upper right part of Figure 5. Suppose that the (α_t, z_t) standard is unreasonably high, i.e. $z > z^*(\alpha)$. Suppose also that the distribution of resubmitted papers has the typical form. Academics will correctly perceive that referees are asking them to meet a very high standard. At the same time, they will see that some papers they thought were submarginal are being accepted. The diagonal arrows in the figure indicate that the conflicting data leads economists to change their beliefs in two ways: they infer that overall quality standards are lower than they had thought and that r-quality is relatively more important than they had thought.

To illustrate why academics make this inference, Figure 6 contains an enlarged view of the data academics get in one such case.[14] The bold line represents the $(q, r(q))$ datapoints they get from referee reports. The outlined area is the support of the (uniform) quality distribution. All papers above and to the right of the bold line are accepted. The journal editor also accepts papers in the shaded region (to the surprise of the referees). Lower quality papers are rejected. The lines below the q-axis and to the left of the r-axis illustrate the support of the q and r distributions among resubmitted papers. The bold portions of these lines are meant to illustrate that the unexpectedly accepted papers are from the low end of the q distribution and the high end of the r distribution. How do academics reconcile the suprise acceptances with the highly demanding referees' reports they've seen? The dashed line graphs the social norm $(\hat{\alpha}, \hat{z})$ that best fits the data. The flatter line allows academics to account for many of the unexpected acceptances while maintaining a good fit to the high q part of the $(q, r(q))$ data. Concluding that the acceptance frontier is flatter than they had thought is equivalent to concluding that r-quality is more important.

FIGURE 6 HERE

Figure 7 contains a similar diagram illustrating academics' inferences when referees'

---

[13]The dynamics in the $z < z^*(\alpha)$ and α large region fit this description only if α is not too large. I discuss what happens in the "high α" case at the end of this section.

[14]The figure graphs the quality distribution and the best fit $(\hat{\alpha}, \hat{z})$ when initial beliefs are that $\alpha = 0.5$ and $z = 0.53417$ with $q \sim U[0, t_q]$, $h(t) = \sqrt{t} - t/2$, $\sigma = 0.2$ and $\tau = 0.3$.

beliefs (α_t, z_t) are unreasonably tough and α_t is sufficently low so that the distribution of resubmitted papers has the low α form — all papers are resubmitted and all authors devote the same maximum effort to their revisions. Here, the unexpectedly accepted papers are uniformly distributed in the q dimension, and the dotted line illustrates that the best fit is obtained by slightly lowering z while leaving the slope of the line unchanged.

FIGURE 7 HERE

Even with the simple loss function I've chosen, getting analytic expressions for the optimal inference from an inconsistent (α_t, z_t) is difficult. Proposition 6 is a characterization of the dynamics that brings out the main observations I've mentioned above. The proposition characterizes the dynamics for initial beliefs that are close to being consistent, i.e. for z_t close to $z^*(\alpha_t)$. Parts (a) and (b) note that in the low α and somewhat low α cases academics do not adjust their estimate of α (at least approximately) and adjust their estimate of z toward $z^*(\alpha_t)$. Part (c) notes that when referees' beliefs correspond to a standard that is too low the dynamics are similar in the "typical" case: the dynamics are approximately vertical when the social norm is approximately consistent. Part (d) relates to my main observation. It notes that when z_t is slightly larger than $z^*(\alpha_t)$ the dynamics involve both a reduction in α and a reduction in the overall quality standard (after accounting for the change that is induced mechanically by the change in α). The shifts in the two parameters are comparable in magnitude. The proof of the proposition is contained in Appendix A.

Proposition 6 *Consider the dynamic model described above with $F(q|t_q) \sim U[0, t_q]$. Let $t_q(\alpha, z)$ be academics' optimal time allocation when they believe that the social norm is (α, z). Write $b(z) \approx a(z - z^*(\alpha))$ as shorthand for $\lim_{z \to z^*(\alpha)} \frac{b(z)}{z - z^*(\alpha)} = a$.*

(a) Suppose that for a given $\alpha_t \in (0, 1)$ the unique equilibrium for the social norm $(\alpha_t, z^*(\alpha_t))$ has the low α form. Then there exists a constant $a > 0$ such that for z_t in a neighborhood of $z^*(\alpha_t)$ the dynamics have

$$\alpha_{t+1} - \alpha_t = 0$$

$$z_{t+1} - z_t = a(z^*(\alpha_t) - z_t)$$

(b) Suppose that for a given $\alpha_t \in (0, 1)$ the unique equilibrium for the social norm $(\alpha_t, z^*(\alpha_t))$ has the somewhat low α form and that $\frac{\partial^2 G}{\partial t^2}$ is strictly positive at $t = t_q(\alpha_t, z^*(\alpha_t))$. Then

20

there exists a constant $a > 0$ such that for z_t close to $z^*(\alpha_t)$ the dynamics have

$$\alpha_{t+1} - \alpha_t \approx 0$$

$$z_{t+1} - z_t \approx a(z^*(\alpha_t) - z_t)$$

For z_t slightly larger than $z^*(\alpha_t)$ the dynamics have $\alpha_{t+1} - \alpha_t = 0$.

(c) Suppose that for a given $\alpha_t \in (0,1)$ the unique equilibrium for the social norm $(\alpha_t, z^*(\alpha_t))$ has the "typical" form and $z_t < z^*(\alpha_t)$. Then there exists a constant $a > 0$ such that for z_t close to $z^*(\alpha_t)$ the dynamics have

$$\alpha_{t+1} - \alpha_t \approx 0$$

$$z_{t+1} - z_t \approx a(z^*(\alpha_t) - z_t),$$

(d) Suppose that for a given $\alpha_t \in (0,1)$ the unique equilibrium for the social norm $(\alpha_t, z^*(\alpha_t))$ has the "typical" form and $z_t > z^*(\alpha_t)$. Let $M_q = (\underline{q}(\alpha_t, z_t) + t_q(\alpha_t, z_t))/2$ and $M_r = \frac{z_t - \alpha_t M_q}{1 - \alpha_t}$. Then there exist constants $a_1 > 0$ and $a_2 > 0$ such that for z_t close to $z^*(\alpha_t)$ the dynamics have

$$\alpha_{t+1} - \alpha_t \approx -a_1(z_t - z^*(\alpha_t))$$

$$z_{t+1} - z_t \approx a_2(z^*(\alpha_t) - z_t) + (\alpha_{t+1} - \alpha_t)(M_q - M_r)$$

The one notable feature of the Figure 5 that I have left out of the discussion so far is what happens when $z < z^*(\alpha)$ and α is sufficiently high so that the equilibrium has the high α form. It is apparent from the figure that in this case academics conclude that q is relatively more important than they had thought. The argument is similar to the argument for why academics conclude that r is more important than they had thought when standard are too high in a typical equilibrium. The unexpectedly rejected papers are relatively low q papers and hence a steeper line allows referees to account for many of the unexpected rejections while maintaining a good fit to the $r(q)$ data for high q papers. I have not discussed this case in more detail because in a couple of ways the argument seems less plausible. First, it involves academics mistakenly not achieving levels of r-quality that they could have achieved. Second, the effect derives from academics trying to fit datapoints with $r(q)$ negative.

Before moving on I wanted to emphasize that I would not regard this model as potentially accounting for the observed trends in academic publishing. The model features a substantial change in norms only if the initial norm is far from the locus of consistent norms. Even with such initial beliefs we would not see a long gradual evolution of norms. Instead, beliefs would initially move quickly to a neighborhood of the equilibrium locus and there would be little subsequent movement.

7 The overconfidence bias and gradual evolution

In this section I show that adding a slight overconfidence bias produces a model in which social norms slowly and steadily evolve over a long period to place ever more emphasis on r-quality.

7.1 ϵ-perturbations of dynamics with a continuum of steady-states

Before discussing the model, it is instructive to discuss its stucture in more generality. The dynamic model is of the form

$$\left(\begin{array}{c} \alpha_{t+1} - \alpha_t \\ z_{t+1} - z_t \end{array} \right) = k \left(\begin{array}{c} \hat{\alpha}_t - \alpha_t \\ \hat{z}_t - z_t \end{array} \right)$$

where $(\hat{\alpha}_t, \hat{z}_t) = \arg\min_{\alpha,z} L(\alpha, z; \alpha_t, z_t)$.[15] A social norm (α_t, z_t) is a steady-state only if it is a solution to

$$\frac{\partial L}{\partial \alpha}(\alpha_t, z_t; \alpha_t, z_t) = 0$$
$$\frac{\partial L}{\partial z}(\alpha_t, z_t; \alpha_t, z_t) = 0$$

This is a system of two equations in two unknowns. Ordinarily, one would expect such a system to have one solution (or zero or a few). The fact that the dynamic model of the previous section has a continuum of equilibria indicates that it is somewhat special.

What happens if we take a dynamic model with a continuum of equilibria and perturb it slightly? The answer depends on how the system is pertubed. To take a simple example,

---

[15] While I earlier defined L as depending on the measures μ_1 and μ_2 describing the two types of data, μ_1 and μ_2 are themselves functions of α_t and z_t. I use α_t and z_t as arguments of L in this section to clarify the nature of the dynamic.

consider a dynamic of the form above with the loss function

$$L(\alpha, z; \alpha_t, z_t) = (\alpha - \alpha_t)^2 + (z - z_t)^2 + (z - \alpha)^2.$$

This model has every point on the line $z_t = \alpha_t$ as a steady state. One thing that can happen with an ϵ-perturbation is an ϵ-order shift in the set of steady states. For example, if we perturb the loss function above to

$$L(\alpha, z; \alpha_t, z_t, \epsilon) = (\alpha - \alpha_t)^2 + (z - z_t)^2 + (z - (\alpha + \epsilon))^2,$$

the system has a continuum of steady states given by $z_t = \alpha_t + \epsilon$.

With a generic perturbation, however, the continuum of steady states will disappear. For example, if we instead perturb the system to

$$L(\alpha, z; \alpha_t, z_t, \epsilon) = (\alpha - \alpha_t)^2 + (z - z_t)^2 + (z - \alpha)^2 + \epsilon z^2,$$

the only remaining steady-state is $\alpha_t = z_t = 0$. What happens to the former equilibria? In this example, the ϵ perturbed dynamics are

$$
\begin{aligned}
\hat{\alpha} - \alpha_t &= \frac{1+\epsilon}{3+2\epsilon}(z_t - \alpha_t) - \frac{\epsilon}{3+2\epsilon}(\alpha_t + z_t) \\
\hat{z} - z_t &= \frac{1+\epsilon}{3+2\epsilon}(\alpha_t - z_t) - \frac{\epsilon}{3+2\epsilon}(\alpha_t + z_t)
\end{aligned}
$$

The first terms on the right hand sides of these equations tell us that from any initial condition the dynamics lead quickly to a neighborhood of the nearly stable locus $z_t = \alpha_t$. The system then evolves at an ϵ rate in a neighborhood of this locus toward the steady state. If the initial condition is very from the steady state a long gradual evolution would be observed.

Obviously, the most natural way to account for a long gradual trend in an economic variable will usually be to view the trend as reflecting a continuous shift in the equilibrium of a model due to a trending exogenous variable. The most general idea the dynamic model is intended to convey is that the disequilibrium dynamics in perturbations of models with a continuum of equilibria may provide an alternate method for explaining some such trends.

7.2 Dynamics in the social norms model: a gradual trend and a stopping point

When referees' beliefs are very far from any consistent social norm, the ϵ-perturbation has little impact on the dynamics of the model — beliefs will evolve quickly to a neighborhood of the set of consistent social norms of the unperturbed model. Once beliefs approach the former equilibrium locus, however, the overconfidence becomes important. The equilibria of the no-bias model are, of course, no longer steady states. If academics' initial beliefs correspond to a consistent social norm, then their misperceptions of the referee reports they receive will lead them to conclude that overall quality standards must be higher than they had thought (i.e. $\hat{z} - z_t > 0$.) For this reason, referees will perpetually try to hold authors to a standard that is slightly higher than is feasible.

What happens when referees are slightly too tough? In the typical case, the addition of the ϵ-overconfidence does not affect the argument of the previous section that academics will conclude that r-quality is more important than they had thought whenever standards are higher than is feasible. As a result, there is no new set of steady states lying just above the set of steady states in the unperturbed model. Instead, what we would see in each period is referees seeing small number of surprise acceptances, making slight downward revisions to their estimates of α and z, but adopting a new standard that is again slightly too high. In each period, these changes might seem minor. With a long run perspective they will be seen as a long, gradual evolution of social norms through the near equilibrium set in the direction of placing an increased emphasis on r-quality.

Must we look forward to a world with no q-quality whatsoever? In my model at least, the answer is no. In the neighborhood of a consistent social norm the effect of a small overconfidence bias is to make academics think that z must be slightly higher than they had thought. In the low α and somewhat low α cases, when the standard is slightly too high academics infer only that z must be slightly lower than they had thought. It turns out that the two effects exactly offset for some z slightly greater than $z^*(\alpha)$. As a result, the ϵ-perturbed model does have a continuum of steady-states lying just above the low α and somewhat low α portions of the set of steady states of the unperturbed model. The dynamics of the model are such that the evolution of social norms comes to a halt as soon

24

as the somewhat low α region is reached. In practical terms, the evolution stops as soon as no paper's idea is good enough to let its author be sure that with enough revisions he or she will eventually be able to get an acceptance. If one believed that economics or another discipline had already reached the point, the model's prediction would be that the trend toward increased emphasis on r-quality should come to a halt.

Figure 8 illustrates the dynamics of the system with a small overconfidence bias. The figure was generated by solving the model numerically for various initial beliefs under the assumption that $q \sim U[0, t_q]$, $\tau = 0.3$, $\sigma = 0.2$, $h(t) = \sqrt{t} - t/2$ and $\epsilon = 0.01$. The thick solid curve on the left side of the figure is the locus of steady states $(\alpha, z^\epsilon(\alpha))$. On the right side of the figure I've graphed eight curves illustrating the evolution of social norms from eight initial conditions.[16] The curves illustrate that regardless of whether referees' initial beliefs are too tough or too soft, there is an initial shift in beliefs toward the near equilibrium locus followed by an evolution through the near equilibrium set in which referees continually decrease their estimates of α.

The arrows on the curves give a feeling for the speed of movement. The arrows mark the beliefs that prevail after one period of evolution and then after every ten additional periods.[17] The fact that the arrows are initially far apart on each curve reflects that the initial movement toward the nearly stable set is rapid — from most starting points the curves get sufficiently close to the nearly stable set to be indistinguishable to the naked eye within ten or twenty periods. The fact that arrows then become closely and regularly spaced as norms evolve along the nearly stable set reflects that this evolution is slow and steady. With the parameters I've chosen the evolution from a near equilibrium with $\alpha = 0.8$ to a near equilibrium with $\alpha = 0.4$ takes about two hundred periods.

FIGURE 8 HERE

Proposition 7 provides a formal description of some properties of the system. Part (a) notes that the model has a continuum of equilibria covering roughly the low α and somewhat low α ranges. Part (b) notes that when α is such that the equilibrium has the

---

[16]While the eight curves appear to join together well before they reach the bold curve on the left side, they actually remain separate and in an extremely magnified figure what now appears to be one central curve would be revealed to be a set of nearly parallel curves.

[17]In constructing the figure I took $k = 0.5$

typical form there is no steady state: at $(\alpha, z^*(\alpha))$ the data is better fit by increasing z; when z is slightly higher than $z^*(\alpha)$ the fit is improved by reducing α.

Proposition 7 *Consider the dynamic model described above with $F(q|t_q) \sim U[0, t_q]$. Let $t_q(\alpha, z)$ be the optimal time allocation when academics believe that the social norm is (α, z).*

(a) Suppose that for a given $\alpha \in (0, 1)$ the unique equilibrium for the social norm $(\alpha, z^*(\alpha))$ with $\epsilon = 0$ has the low α or somewhat low α form. Suppose that $\frac{\partial^2 G}{\partial t^2}$ is strictly positive at $t = t_q(\alpha, z^*(\alpha))$. Then for ϵ sufficiently small there exists a value of z, $z^\epsilon(\alpha)$ such that $(\alpha, z^\epsilon(\alpha))$ is a steady state of the model with ϵ-overconfidence Further, there exists a positive constant a such that, $z^\epsilon(\alpha) - z^*(\alpha) \approx a\epsilon$ when ϵ is small. In the low α case $z^\epsilon(\alpha)$ is exactly equal to $z^*(\alpha) + 2(1 - \alpha)\sigma\epsilon$ when ϵ is sufficiently small.

(b) Suppose that for a given $\alpha \in (0, 1)$ the unique equilibrium for the social norm $(\alpha, z^*(\alpha))$ with $\epsilon = 0$ has the typical form. When $\epsilon > 0$ we have $\frac{\partial L}{\partial z}(\alpha, z^*(\alpha); \alpha, z^*(\alpha)) < 0$. For ϵ sufficiently small there exists a $z > z^*(\alpha)$ for which the distribution of paper qualities and expectedly and unexpectedly accepted papers has the typical form and the loss function has $\frac{\partial L}{\partial z}(\alpha, z; \alpha, z) = 0$. For any such z, $\frac{\partial L}{\partial \alpha}(\alpha, z; \alpha, z) < 0$. Hence, the system does not have a steady state where the distribution of paper qualities has the typical form. Let $z^\epsilon(\alpha)$ be the smallest value of z such that $\frac{\partial L}{\partial z}(\alpha, z; \alpha, z) = 0$. Then there exist positive constants a and b such that $z^\epsilon(\alpha) - z^*(\alpha) \approx a\epsilon$ and $\frac{\partial L}{\partial \alpha}(\alpha, z^\epsilon(\alpha); \alpha, z^\epsilon(\alpha)) \approx -b\epsilon$ for small ϵ.

For those who do not believe that academics suffer from overconfidence bias (and those who are wondering if observed changes in standards should be taken as providing evidence of such a bias), I would like to note that a gradual evolution toward lower α norms can be generated by many small modifications to the basic dynamic model. Essentially, any perturbation that makes referees perpetually try to hold authors to a slightly infeasibly high standard will do the same thing. One example would be an assumption that referees try to make up more extensive lists of revisions than is standard to impress editors with their thoroughness. Another would be an assumption that referees are competitive (or spiteful) and try to hold back others in their field by imposing standards that are slightly higher than the norm (perhaps because making requests that are even more demanding would expose their spite).

One criticism of any model like that presented here is that it is inherently nonrobust — many small changes to the base model will eliminate the continuum of equilibria. In a slightly different model where there is a unique stable equilibrium, a standard analysis would conclude that adding an ϵ-overconfidence bias would not fundamentally alter the dynamics if ϵ is sufficiently small. My view of this criticism is that it is correct and highlights an important point — readers should think about whether there are things I've left out of the model that they feel are more important than the overconfidence bias and if so whether they might have opposing or other effects. At the same time, however, I also feel that the opposite of this argument may be equally important in many models yet is typically ignored. Economists generally like models with a unique equilibrium and often add assumptions about strict concavity or convexity to guarantee uniqueness. If we insisted that models only added ϵ-concavity then their predictions would also be nonrobust to making small changes if ϵ is sufficiently small relative to the changes. In some cases (this one included), I would argue that we might get a better understanding of what we can and cannot predict confidently by not simply making whatever assumptions are needed to ensure uniqueness and instead focusing on models with a dimension of uncertainty and thinking carefully about what we can say confidently beyond that many outcomes are possible.

8 Conclusion

I have proposed that it is helpful to think about changes in the academic publishing process with a two-dimensional q-r model of quality. The slowdown of the process and the increased length of papers may reflect an increase in r-quality. I have discussed a couple of different ways in which we might account for the changes. First, the changes may be attributable to changes in the condition of academic disciplines. Any trend that is common across fields, such as increases in the number of academics or technological progress that makes it easier to add certain types of quality to papers, could potentially explain the common trends in publishing. Second, I've proposed that within any field there may be a tendency for the social norms for weighing different aspects of quality to shift. In the model, academics' struggles to reconcile the high standards being applied to them with the mediocrity they see in journals lead them to place increasing weight on r-quality. Researchers react by

27

spending less and less time developing new insights and more and more time padding and polishing papers. Ultimately, which of the potential explanations are most important is an empirical question. I attempt to begin this sorting out in Ellison (2001).

What are the welfare effects of changes in publication standards? There are no answers within my model, because individual preferences over q and r do not appear. If preferences are time invariant, then welfare changes as standards evolve. Whether changes are welfare increasing or welfare decreasing depends on how the prevailing social norm differs from the optimum given authors' and readers' preferences. The changes of the last few decades have had a substantial impact on economists. Many young economists report spending as much time revising old papers as working on new ones. Guiding larger and more frequent revisions is an additional burden on referees and editors. While the thought that all this effort may be misguided may be horrifying, it can also be seen as a very optimistic message. If many social norms are indeed possible, then academic communities may be able to achieve dramatic welfare improvements by simply discussing what standards members would like to have and agreeing on a change.

The models of this paper may also be useful in thinking about the substantial cross-field differences in the form of academic articles and the structure of the journal review process. The comparative statics explanations discussed in the paper apply to differences in technologies, etc. across fields. An additional factor that seems potentially important in explaining the science-social science divide is the role of science journals in establishing priority. Applying the evolving social norm model to explain cross-field differences is less straightforward. One can not, for example, say that it predicts that fields that have been around longer will have more drawn out review processes. What the model predicts is that once the revise-and-resubmit comes to be common, social norms will start evolving to place more emphasis on r-quality and this trend will continue until it reaches the point that no paper's idea is good enough to ensure that the author can revise it to make it acceptable. How fast the evolution proceeds and where it will stop depend on all of the factors that are involved in a comparative statics explanation.

Among the many extensions to the model that one might pursue, I view the incorporation of randomness in the assessment of quality and in academics' data samples as the

28

most intriguing. Referees who receive signals suggesting that standards are higher than they are may conclude that r-quality is more important than they thought, while referees who receive signals suggesting that standards are lower than they are will not draw the opposite conclusion. Even without an overconfidence bias we may see increasing emphasis placed on r-quality. Sobel's (2001) work suggests that randomness in assessments due to heterogeneity in tastes may have more subtle effects.

How else might one examine the evolutionary model empirically? With ideal data one could directly examine referees' demands and see how they are affected by referees' experiences as authors. Another feature that distinguishes the evolutionary model from equilibrium models is that transitory shocks (such as the appointment of a revision-loving editor for a fixed term) may have permanent effects.

In addition to trying to change how people think about academic publishing, I have tried to make the general point that a long run trend can be a disequilibrium phenomenon. Comparative statics of equilibria will remain the standard for explaining trends, but I hope that models like the one developed here will find other applications.

Appendix A

<u>Proof of Proposition 6</u>
(a) In the low α case for (z, t_q) in a neighborhood of $(z^*(\alpha_t), t_q(\alpha_t, z^*(\alpha_t)))$ we have

$$G(z, t_q) = \frac{1}{\sigma}\left(\frac{z - \alpha_t t_q/2}{1 - \alpha_t} - h(1 - t_q)\right).$$

This function has $\frac{\partial G}{\partial t_q} = -\frac{\alpha_t}{2\sigma(1-\alpha_t)} + \frac{h'(1-t_q)}{\sigma}$, which is independent of z. It also has $\frac{\partial^2 G}{\partial t_q^2} = -h''(1-t_q)/\sigma > 0$, and hence in a neighborhood of $z^*(\alpha_t)$, $t_q(\alpha_t, z_t)$ is independent of z_t. The continuity of the thresholds implies that the ordering $\underline{q} < 0 < m(t_q) < \bar{q}$ holds for all z_t in a neighborhood of $z^*(\alpha_t)$.

Write $r_a(q; \alpha_t, z_t)$ for the level of r-quality that is actually necessary for an editor to accept a paper of q-quality q when the population believes the standard is (α_t, z_t). Because $t_q(\alpha_t, z_t)$ is locally independent of z_t and all authors revise to the greatest extent possible, $r_a(q; \alpha_t, z_t) = r(q; \alpha_t, z^*(\alpha_t))$ for z_t near $z^*(\alpha_t)$. The situation is thus like that pictured in Figure 7. When $z_t < z^*(\alpha_t)$ the situation is similar, but involving a parallelogram of unexpectedly rejected papers lying just above the $(q, r(q))$ line.

To describe academics' inferences, I'll make a change of variables and analyze the minimization of the loss function over m and w where $m = \frac{-\alpha}{1-\alpha}$ is the slope of the line and $w = \frac{z - \alpha M_q}{1-\alpha}$ is the level of r quality required of a paper with the mean q-quality, $M_q \equiv (t_q(\alpha_t, z_t) + \max(0, \underline{q}(\alpha_t, z_t)))/2$. I'll carry out the calculations for $z_t > z^*(\alpha_t)$. The calculations in the opposite case are identical but with some signs reversed.

With the change of variables L_1 takes on a very simple form.[18]

$$
\begin{aligned}
L_1(m, w) &= \int_0^{t_q(\alpha_t, z_t)} (m(q - M_q) + w - r(q; \alpha_t, z_t))^2 \frac{1}{t_q(\alpha_t, z_t)} dq \\
&= \frac{1}{t_q(\alpha_t, z_t)} \int_0^{t_q(\alpha_t, z_t)} (m(q - M_q) + w - (m_t(q - M_q) + w_t))^2 \, dq \\
&= \frac{1}{t_q(\alpha_t, z_t)} \left[(m - m_t)^2 \frac{(q - M_q)^3}{3} + (m - m_t)(w - w_t)(q - M_q)^2 + (w - w_t)^2 q\right]_0^{t_q(\alpha_t, z_t)} \\
&= \frac{1}{12} t_q(\alpha_t, z_t)^2 (m - m_t)^2 + (w - w_t)^2
\end{aligned}
$$

By making a few comparisons it is easy to see that the minimum of $L_1 + L_2$ can not be achieved for $w > w_t$, or for $w < r_a(M_q, \alpha_t, z_t)$, or with a slope m for which $w + m(-M_q)$ is outside the interval $[r_a(0; \alpha_t, z_t), r(0; \alpha_t, z_t)]$. (For example, to show the second note that for any such w, $L(m, w) > L(m_t, r_a(M_q, \alpha_t, z_t))$ because L_1 is smaller at $(m_t, r_a(M_q, \alpha_t, z_t))$ and L_2 has its global minimum there.) In the range containing any potential minumum, L_2 turns out to also be a simple quadratic.

$$L_2(m, w) = \int_0^{t_q(\alpha_t, z_t)} \int_0^{w - r_a(M_q; \alpha_t, z_t) + (m - m_t)(q - M_q)} r \frac{1}{\sigma} dr \frac{1}{t_q(\alpha_t, z_t)} dq$$

[18]To make the formulas more readable I'll often omit the μ_1 and μ_2 arguments of L_1 and L_2. μ_1 and μ_2 are determined by α_t and z_t, so I'll also sometimes substitute α_t and z_t (or m_t and w_t) as arguments. I will also omit arguments of M_q, t_q, $\bar{q}$, and $\underline{q}$ to improve readability.

$$= \frac{1}{2\sigma t_q(\alpha_t, z_t)} \int_0^{t_q(\alpha_t, z_t)} (w - r_a(M_q; \alpha_t, z_t) + (m - m_t)(q - M_q))^2 dq$$

$$= \frac{1}{24\sigma} t_q(\alpha_t, z_t)^2 (m - m_t)^2 + \frac{1}{2\sigma}(w - r_a(M_q; \alpha_t, z_t))^2.$$

Hence,

$$L(m, w; m_t, w_t) = \frac{1}{6} t_q(\alpha_t, z_t)^2 (m - m_t)^2 + (w - w_t)^2 + \frac{1}{2\sigma}(w - r_a(M_q; \alpha_t, z_t))^2.$$

The minimum clearly involves $\hat{m} = m_t$. This gives the first part of the conclusion: $\hat{\alpha} = \alpha_t$.

The minimizing value for w is $\hat{w} = \frac{2w_t + (1/\sigma) r_a(M_q; \alpha_t, z_t)}{2 + 1/\sigma}$. Recall that r_a is independent of z_t in a neighborhood of $z^*(\alpha_t)$. We thus have

$$r_a(M_q; \alpha_t, z_t) = \frac{z^*(\alpha_t) - \alpha_t M_q}{1 - \alpha_t}$$

$$= \frac{z^*(\alpha_t) - z_t}{1 - \alpha_t} + \frac{z_t - \alpha_t M_q}{1 - \alpha_t} = \frac{z^*(\alpha_t) - z_t}{1 - \alpha_t} + w_t.$$

Substituting this into the expression for $\hat{w}$ gives

$$\hat{w} - w_t = \frac{1}{(2\sigma + 1)(1 - \alpha_t)}(z^*(\alpha_t) - z_t).$$

Using the identities $\alpha_t = \hat{\alpha}$ and $z_t = \alpha_t M_q + (1 - \alpha_t)w_t$ allows us to conclude as desired that

$$z_{t+1} - z_t = k(\hat{\alpha}M_q + (1 - \hat{\alpha})\hat{w} - z_t) = k(\alpha_t M_q + (1 - \alpha_t)\hat{w} - \alpha_t M_q + (1 - \alpha_t)w_t)$$

$$= k(1 - \alpha_t)(\hat{w} - w_t) = k\frac{1}{2\sigma + 1}(z^*(\alpha_t) - z_t).$$

(b) The proof for the somewhat low α case is very similar. One difference is that the form of G is different. This makes it possible that the second derivative of G will not be strictly positive at $t_q(\alpha_t, z^*(\alpha_t))$, in which case $t_q(\alpha_t, z_t)$ might not be differentiable at $z_t = z^*(\alpha_t)$. This would cause many complications. To avoid them, I've just assumed in the proposition that $\frac{\partial^2 G}{\partial t_q^2}(\alpha_t, t_q(\alpha_t, z^*(\alpha_t))) > 0$. This ensures that $t_q(\alpha_t, z)$ is differentiable in z at $z^*(\alpha_t)$.

The analysis of the $z_t > z^*(\alpha_t)$ case then proceeds exactly as above to show that

$$\hat{\alpha} - \alpha_t = 0$$

$$\hat{w} - w_t = \frac{1}{2\sigma + 1}(r_a(M_q; \alpha_t, z_t) - w_t).$$

Write $\tau_a(z_t) = 1 - G(z_t; t_q(\alpha_t, z_t))$ for the fraction of papers achieving the z_t standard given the inital beliefs. By the envelope theorem we know that in a neighborhood of $z^*(\alpha_t)$,

$$\tau_a(z_t) \approx \tau - \frac{\partial G}{\partial z}(z^*(\alpha_t); t_q(\alpha_t, z^*(\alpha_t)))(z_t - z^*(\alpha_t))$$

$$= \tau - \frac{t_q(\alpha_t, z^*(\alpha_t)) - \underline{q}(\alpha_t, z^*(\alpha_t))}{(1 - \alpha_t)\sigma t_q(\alpha_t, z^*(\alpha_t))}(z_t - z^*(\alpha_t)).$$

The mass of extra papers the editor accepts when he lowers the standard from $z^*(\alpha_t)$ to $z^*(\alpha_t) - dz$ is $\frac{t_q - q}{t_q(1 - \alpha_t)\sigma}dz$. Hence, in a neighborhood of $z^*(\alpha_t)$ we have

$$
\begin{aligned}
r_a(M_q; \alpha_t, z_t) - w_t &= \frac{(1 - \alpha_t)\sigma t_q}{(1 - \alpha_t)(t_q - \underline{q})}(\tau_a(z_t) - \tau) \\
&\approx -\frac{1}{(1 - \alpha_t)}(z_t - z^*(\alpha_t))
\end{aligned}
$$

This gives

$$
z_{t+1} - z_t \approx a(z^*(\alpha_t) - z_t),
$$

for $a = \frac{k}{2\sigma + 1}$ as desired.

The analysis of the $z_t < z^*(\alpha_t)$ case adds another slight complication — the formula for L_2 is slightly different because there are no papers in the triangle bounded by $(\underline{q}, h(1 - t_q) + \sigma)$, $(\underline{q}, r_a(\underline{q}, \alpha_t, z_t))$, and $(\underline{q} + (1 - \alpha_t)(r_a(\underline{q}, \alpha_t, z_t) - r(\underline{q}, \alpha_t, z_t))/\alpha_t, h(1 - t_q) + \sigma)$ that can be unexpectedly rejected. The loss function thus takes the form

$$
\begin{aligned}
L(m, w; m_t, w_t) &= \frac{2\sigma + 1}{24\sigma}(t_q - \underline{q})^2(m - m_t)^2 + (w - w_t)^2 + \frac{1}{2\sigma}(w - r_a(M_q; \alpha_t, z_t))^2 \\
&\quad - \frac{1}{\sigma(t_q - \underline{q})}\int_T I(r > r(q; m, w))(r - r(q; m, w))dqdr,
\end{aligned}
$$

where T is the triangle bounded by the three points above and I is the indicator function. To show that $\alpha_{t+1} - \alpha_t \approx 0$ it suffices to show that for any $c > 0$ there exists a δ such that $|\hat{m} - m_t| < c(z^*(\alpha_t) - z_t)$ whenever $z_t \in (z^*(\alpha_t) - \delta, z^*(\alpha_t))$. To see this note that for all z_t in some interval below $z^*(\alpha_t)$ we have for any $m < m_t - c(z^*(\alpha_t) - z_t)$

$$
\begin{aligned}
\frac{\partial L}{\partial m}(m, w) &= \frac{2\sigma + 1}{12\sigma}(t_q - \underline{q})^2(m - m_t) - \frac{1}{\sigma(t_q - \underline{q})}\frac{\partial}{\partial m}\int_T I(r > r(q; m, w))(r - r(q; m, w))dqdr \\
&< -\frac{2\sigma + 1}{12\sigma}(t_q - \underline{q})^2 c(z^*(\alpha_t) - z_t) \\
&\quad + \frac{A(T)}{\sigma(t_q - \underline{q})}\sup_{(q, r) \in T}|\frac{\partial}{\partial m}I(r > r(q; m, w))(r - r(q; m, w))|,
\end{aligned}
$$

where $A(T)$ is the area of the triangle T. The area of the triangle is $\frac{1 - \alpha_t}{2\alpha_t}(r_a(q; \alpha_t, z_t) - r(q; \alpha_t, z_t))^2$. The effect of a dm change in m on $r(q; m, w)$ is $q - M_q$, which is largest when q is farthest from M_q. This gives

$$
\frac{\partial L}{\partial m}(m, w) \leq -\frac{2\sigma + 1}{12\sigma}(t_q - \underline{q})^2 c(z^*(\alpha_t) - z_t) + \frac{1 - \alpha_t}{4\alpha_t\sigma}(r_a(q; \alpha_t, z_t) - r(q; \alpha_t, z_t))^2.
$$

The $(r_a(q; \alpha_t, z_t) - r(q; \alpha_t, z_t))^2$ term is a second order effect in $z^*(\alpha_t) - z_t$. Hence, $\frac{\partial L}{\partial m}(m, w) < 0$ for all $m < m_t - c(z^*(\alpha_t) - z_t)$ when z_t sufficiently close to $z^*(\alpha_t)$. Combining this with a similar calculation of the derivative for $m > m_t + c(z^*(\alpha_t) - z_t)$ allows us to conclude that $\hat{m} \in (m_t - c(z^*(\alpha_t) - z_t), m_t + c(z^*(\alpha_t) - z_t))$ as desired.

A similar calculation shows that the result on $z_{t+1} - z_t$ is also unaffected by the second order change in the loss function.

(c) When the equilibrium has the typical form $G(z; t_q)$ is given by

$$G(z; t_q) = \frac{z - (1 - \alpha_t)(h(1 - t_q) + \sigma/2)}{\alpha_t t_q}.$$

An explicit calculation of the first two derivatives of this function shows that $\frac{\partial^2 G}{\partial t_q^2}(z; t_q) > 0$ in a neighborhood of $(z^*(\alpha_t), t_q(\alpha_t, z^*(\alpha_t)))$. This again implies that $\frac{\partial t_q}{\partial z}(\alpha_t, z^*(\alpha_t))$ exists and that the distribution of paper qualities has the typical form for z_t in some neighborhood of $z^*(\alpha_t)$.

The $(q, r(q))$ line and the set of unexpectedly rejected papers have exactly the same form here as in $z_t < z^*(\alpha_t)$ case of part (b) of the proposition. The result is thus identical to the result for that case.

(d) As above t_q is differentiable in z at $z^*(\alpha_t)$. The distribution of paper qualities and outcomes has the typical form pictured in Figure 6 for all z_t in some neighborhood $(z^*(\alpha_t), z^*(\alpha_t) + \delta)$.

Write w_a for $r_a(M_q; \alpha_t, z_t)$. I first note that a number of simple comparisons imply that academics must infer that standards are lower and q is less important than they had thought. Specifically, we must have $\hat{m} > m_t$ and $\hat{w} \in (w_a, w_t)$. To see this, note first that the line given by $(\hat{m}, \hat{w})$ can not be entirely above the (m_t, w_t) line over the whole range $[q(\alpha_t, z_t), t_q(\alpha_t, z_t)]$ of q-qualities of resubmitted papers. From any such estimate both L_1 and L_2 are decreased by moving the line down until there is an intersection. Next, note that it is also impossible for the minimum to have $\hat{w} > w_t$ with the $(\hat{m}, \hat{w})$ line intersecting the (m_t, w_t) line at $q \in [q, t_q]$. In this case, L_1 and L_2 are both reduced by rotating the fitted line about the point were it intersects the line $r = r_a(q; m_t, w_t)$ in the direction that reduces $\hat{w}$. The estimates also can not have $\hat{w} < w_a$. In this case increasing $\hat{w}$ to w_a and setting $\hat{m} = m_t$ reduces L_1 and makes L_2 equal to its global minimum. This gives $\hat{w} \in (r_a(M_q; \alpha_t, z_t), w_t)$. It is then easy to see that $\hat{m} \geq m_t$. Otherwise, slightly reducing $\hat{m}$ would reduce L_1 and also reduce L_2. (The latter is because the gain from to improving the fit in $(q, \min(\bar{q}, M_q))$ is greater than the loss from worsening the fit in the (possibly empty) interval $[M_q, \bar{q}]$.) Finally, all of the extreme values ($\hat{m} = m_t$, $\hat{w} = w_t$ and $\hat{w} = w_a$) can be ruled out by looking at derivatives of the loss function.

To obtain the proposition's further conclusion that the changes in m and w are both first order in $z_t - z^*(\alpha_t)$ I examine the first order conditions of the loss function. Even after the discussion above there remain a few possibilities for exactly where the best fit line may intersect the other lines in the figure: it may be above or below $r_a(q; \alpha_t, z_t)$ at the left edge and may be above or below $r(t_q; \alpha_t, z_t)$ at the right edge. The first order conditions are slightly different in the four cases. I will work out the equations for the simplest case: assuming that the best fit line is strictly between $r_a(q; \alpha_t, z_t)$ and $r(q; \alpha_t, z_t)$ throughout the interval $[q, t_q]$.

In this case, the loss function is very similar to the loss function in the cases above. The only differences are that the set of unexpectedly accepted papers does not have papers with q-qualities above $\bar{q}$ and is also missing a second-order triangle in (q, r)-space below $(\bar{q}, r(\bar{q}; \alpha_t, z_t))$. As above, the second-order triangle can be ignored. The estimates have the same asymptotics as those obtained by minimizing loss functions given by integrals identical to those in part (a) but with different lower and upper bounds. Specifically we

can examine the minimizer of $L = L_1 + L_2$ where

$$L_1(m, w) = \frac{1}{12}(t_q - \underline{q})^2(m - m_t)^2 + (w - w_t)^2$$

and

$$L_2(m, w) = \frac{1}{24\sigma}\frac{(\bar{q} - \underline{q})^3}{t_q - \underline{q}}(m - m_t)^2 + \frac{1}{2\sigma}\frac{\bar{q} - \underline{q}}{t_q - \underline{q}}\left(w - w_a - (m - m_t)\frac{t_q - \bar{q}}{2}\right)^2.$$

(Note that I've omitted the arguments (α_t, z_t) arguments of t_q, $\bar{q}$ and $\underline{q}$ to improve readability.) The first order conditions for this minimization are of the form

$$\begin{aligned}
c_1(\hat{m} - m_t) - c_2(\hat{w} - w_a) &= 0 \\
-c_2(\hat{m} - m_t) + c_3(\hat{w} - w') &= 0,
\end{aligned}$$

where $w' = \frac{2\sigma(t_q - \underline{q})w_t + (\bar{q} - \underline{q})w_a}{2\sigma(t_q - \underline{q}) + (\bar{q} - \underline{q})}$ is a weighted average of w_t and w_a and c_1, c_2 and c_3 are

positive constants: $c_1 = \frac{1}{6}(t_q - \underline{q})^2 + \frac{1}{12\sigma}\left(\frac{\bar{q} - \underline{q}}{t_q - \underline{q}}\right)^3 + \frac{1}{4}(\bar{q} - \underline{q})(t_q - \bar{q})$; $c_2 = \frac{(\bar{q} - \underline{q})(t_q - \bar{q})}{2\sigma(t_q - \underline{q})}$; and

$c_3 = 2 + \frac{\bar{q} - \underline{q}}{\sigma(t_q - \underline{q})}$. Adding c_2/c_3 times the second equation to the first gives

$$\hat{m} - m_t = \frac{c_2 c_3}{c_1 c_3 - c_2^2}\frac{2\sigma(t_q - \underline{q})}{2\sigma(t_q - \underline{q}) + (\bar{q} - \underline{q})}(w_t - w_a).$$

We know that the true minimum has $\hat{m} > m_t$ and that $w_t > w_a$, so the first order conditions of this case can only give the true minimum if the constant in this expression is positive. The equations also give

$$\hat{w} - w_t = \frac{c_2^2 - c_1 c_3 \frac{\bar{q} - \underline{q}}{2\sigma(t_q - \underline{q}) + (\bar{q} - ql)}}{c_1 c_3 - c_2^2}(w_t - w_a),$$

From the discussion above this can be the true minumum only if the leading constant is negative.

The fact that $w_t - w_a \approx (z_t - z^*(\alpha_t))/(1 - \alpha_t)$ implies that $\hat{m} - m_t$ and $\hat{w} - w_t$ are first order in $z_t - z^*(\alpha_t)$. The longer expression for $z_{t+1} - z_t$ follows from the calculation

$$\begin{aligned}
\hat{z} - z_t &= \hat{\alpha} M_q + (1 - \hat{\alpha})\hat{w} - (\alpha_t M_q + (1 - \alpha_t)w_t) \\
&= (1 - \alpha_t)(\hat{w} - w_t) + M_q(\hat{\alpha} - \alpha_t) - \hat{w}(\hat{\alpha} - \alpha_t) \\
&= (1 - \alpha_t)(\hat{w} - w_t) + (M_q - M_r)(\hat{\alpha} - \alpha_t) + (M_r - \hat{w})(\hat{\alpha} - \alpha_t)
\end{aligned}$$

The last term in this expression is second order in $z_t - z^*(\alpha_t)$.

QED.

Proof of Proposition 7

(a) Recall from the proof of Proposition 6 that $t_q(\alpha_t, z_t)$ is differentiable in z at $z^*(\alpha_t)$ and that the distribution of paper qualities will also have the low α or somewhat low α form for z_t in a neighborhood of $z^*(\alpha_t)$. The same computation as in the proof of part (b) of

Proposition 6 shows that for z_t in some neighborhood $(z^*(\alpha_t), z^*(\alpha_t) + \delta)$ the loss function is of the form $L = L_1 + L_2$ with

$$L_1(m, w; m_t, w_t) = \frac{1}{12}(t_q - \underline{q})^2(m - m_t)^2 + (w - (w_t + \epsilon))^2$$

$$L_2(m, w; m_t, w_t) = \frac{1}{24\sigma}(t_q - \underline{q})^2(m - m_t)^2 + \frac{1}{2\sigma}(w - r_a(M_q; \alpha_t, z_t))^2.$$

Differentiating, (m_t, w_t) is a steady state if and only if $w_t - r_a(M_q; \alpha_t, z_t) = 2\sigma\epsilon$. In the previous proof we also saw that $w_t - r_a(M_q; \alpha_t, z_t) \approx \frac{1}{1-\alpha_t}(z_t - z^*(\alpha_t))$. Hence, for ϵ sufficiently small we can find a $z_t > z^*(\alpha_t)$ that satisfies the equation for a steady state.

In the somewhat low α case the equations above give $z^\epsilon(\alpha) - z^*(\alpha) \approx 2(1 - \alpha)\sigma\epsilon$. In the low α case we saw earlier that the expression $w_t - r_a(M_q; \alpha_t, z_t) = \frac{1}{1-\alpha_t}(z_t - z^*(\alpha_t))$ is exact when z_t is sufficiently close to $z^*(\alpha_t)$, and hence the expression for $z^\epsilon(\alpha) - z^*(\alpha)$ is an equality as well.

(b) Again for a given α_t the distribution of papers has the typical form if z_t is sufficiently close to $z^*(\alpha_t)$. For w_t in a neighborhood of $r(M_q; \alpha_t, z_t)$ and (m, w) sufficiently close to (m_t, w_t) with $m \leq m_t$, $w \geq w_t$ the expression for the loss function is again a slight variant of that in the previous proposition:

$$L_1(m, w; m_t, w_t) = \frac{1}{12}(t_q - \underline{q})^2(m - m_t)^2 + (w - (w_t + \epsilon))^2$$

$$L_2(m, w; m_t, w_t) = \frac{1}{24\sigma}\left(\frac{\overline{q} - \underline{q}}{t_q - \underline{q}}\right)^3(m - m_t)^2 + \frac{1}{2\sigma}\frac{\overline{q} - \underline{q}}{t_q - \underline{q}}\left(w - w_a - (m - m_t)\frac{t_q - \underline{q}}{2}\right)^2,$$

where again I've written w_a for $r_a(M_q; \alpha_t, z_t)$ to save space. A direct computation shows that $\frac{\partial L}{\partial m}(m_t, w_t; m_t, w_t) = 0$ and $\frac{\partial L}{\partial w}(m_t, w_t; m_t, w_t) < 0$ when (m_t, w_t) corresponds to a consistent social norm. This is the first result mentioned in part (b).

We will find $\frac{\partial L}{\partial z} = 0$ with L parameterized by (α, z) if and only if $\frac{\partial L}{\partial w} = 0$ when L is parameterized by (m, w). A simple calculation of derivatives shows that $\frac{\partial L}{\partial w}(m_t, w_t; m_t, w_t) = 0$ if and only if

$$w_t - w_a = 2\sigma\frac{t_q - \underline{q}}{\overline{q} - \underline{q}}\epsilon = 0.$$

As above this has a solution z_t with $z_t - z^*(\alpha_t) \approx 2(1-\alpha_t)\sigma\frac{t_q - \underline{q}}{\overline{q} - \underline{q}}\epsilon$. This establishes the next claim in part (b) and the fact that the smallest solution has $z^\epsilon(\alpha_t) - z^*(\alpha_t) \approx a\epsilon$ for some $a > 0$.

If (m_t, w_t) were a steady state of the dynamics in which the distribution of paper qualities had the typical form it would have to satisfy the first-order condition above and the additional constraint that $\frac{\partial L}{\partial m}(m_t, w_t; m_t; w_t) = 0$. We know that $\frac{\partial L_1}{\partial m}(m_t, w_t; m_t; w_t) = 0$ for any m_t and w_t. The w FOC can only be satisfied for $w_t > r(M_q; \alpha_t, z_t)$. For such values of w_t and for $m \leq m_t$ sufficiently close to m_t we have

$$L_2(m, w_t; m_t, w_t) = \frac{1}{\sigma(t_q - \underline{q})}\left(\int_{\max(0, \overline{q}-M_q)}^{(t_q - \underline{q})/2}\int_0^{(m-m_t)q} r\,dr\,dq + \int_{M_q-\overline{q}}^{(t_q-\underline{q})/2}\int_0^{w_t - w_a - (m-m_t)q} r\,dr\,dq\right).$$

The first term is the surprise due to the unexpected acceptance of papers with q-qualities above $\overline{q}$, and the second is the suprise due to the unexpected acceptance of papers with

35

lower q-qualities. Evaluating the integrals and differentiating gives

$$\frac{\partial L_2}{\partial m}(m_t, w_t; m_t, w_t) = -\frac{(t_q - \underline{q})^2 - 4(\bar{q} - M_q)^2}{8\sigma(t_q - \underline{q})}(w_t - w_a).$$

The derivative with respect to α is simply this expression divided by $(1 - \alpha_t)^2$. This is negative, which is the second conclusion of part (b). It is also immediate from the expression that $\frac{\partial L}{\partial \alpha}(\alpha_t, z_t; \alpha_t, z_t) \approx b\epsilon$.

Looking at the numerator in the equation for $\frac{\partial L}{\partial m}$ we can also see that the gain from reducing α vanishes as we approach the somewhat low α region and $\bar{q}$ approaches t_q (or as α approaches one and $\bar{q}$ approaches $\underline{q}$.)

QED.

References

De Long, J. Bradford; Shleifer, Andrei; Summers, Lawrence H.; and Waldmann, Robert J. "Noise Trader Risk in Financeial Markets." *Journal of Political Economy* 98 (August 1990): 703-738.

Ellison, Glenn. "The Slowdown of the Economics Publishing Process." Manuscript. Cambridge, Mass.: MIT, 2001.

Ellison, Glenn and Fudenberg, Drew. "Learning Purified Mixed Equilibria." *Journal of Economic Theory* 90 (January 2000): 84-115.

Karni, Edi, and Schmeidler, David. "Fixed Preferences and Changing Tastes." *American Economic Review* 80 (May 1990): 262-267.

Lichtenstein, Sarah; Fischhoff, Baruch; and Phillips, Lawrence. "Calibration of Probabilities: The State of the Art to 1980." In *Judgement under Uncertainty: Heuristics and Biases*, edited by Daniel Kahneman, Paul Slovic and Amos Tversky. Cambridge: Cambridge University Press, 1982.

Odean, Terrance. "Volume, Volatility, Price, and Profit When All Traders Are Above Average." *Journal of Finance* 53 (December 1998): 1887-1934.

Pesendorfer, Wolfgang. "Design Innovation and Fashion Cycles." *American Economic Review* 85 (September 1995): 771 - 792.

Rabin, Matthew. "Psychology and Economics." *Journal of Economic Literature* 36 (March 1998): 11-46.

Rubinstein, Ariel. "Perfect Equilibrium in a Bargaining Model." *Econometrica* 50 (January 1982): 97-109.

Sen, Amartya. "The Impossibility of a Paretian Liberal." *Journal of Political Economy* 78 (January-February 1970): 152-157.

Sobel, Joel. "A Model of Declining Standards." *International Economic Review* 41 (May 2000): 295-303.

Sobel, Joel. "On the Dynamics of Standards." *Rand Journal of Economics*, forthcoming (2001).

Spence, A. Michael. "Job Market Signaling." *Quarterly Journal of Economics* 87 (August 1973): 355-374.

Svenson, Ola "Are We All Less Risky and More Skillful than Our Fellow Drivers." *Acta Psychologica* 47 (1981): 143-148.
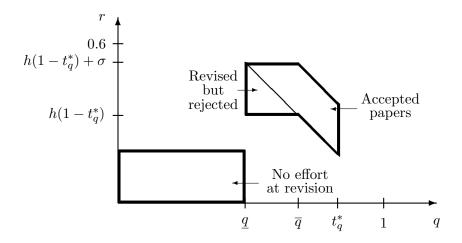
Figure 1: The equilibrium distribution of paper qualities in a "typical" equilbrium: $\tau = 0.3$, $q \sim U[0, t_q]$, $h(t) = \sqrt{t} - t/2$, $\sigma = 0.2$, $\alpha = 0.5$, $z \approx 0.504$.
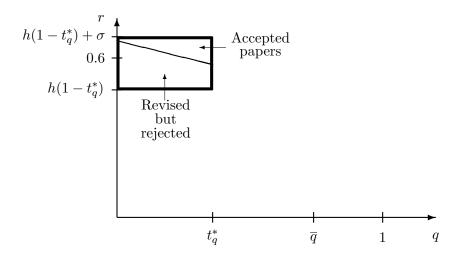

Figure 2: An example of the equilibrium quality distribution in a "low α" equilibrium: $\tau = 0.3$, $q \sim U[0, t_q]$, $h(t) = \sqrt{t} - t/2$, $\sigma = 0.2$, $\alpha = 0.2$, $z \approx 0.532$.
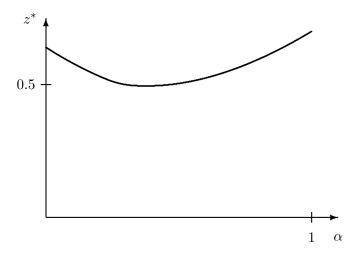
38

Figure 3: Consistent social norms (α, z) in the model with $q \sim U[0, t_q]$, $h(t) = \sqrt{t} - t/2$, $\tau = 0.3$, and $\sigma = 0.2$.

Time allocation
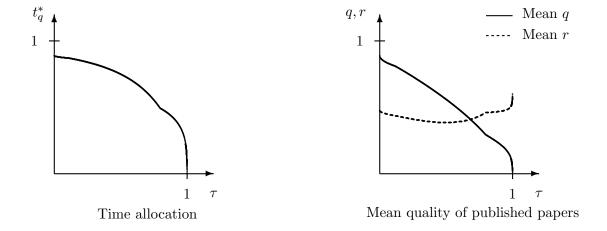
Mean quality of published papers

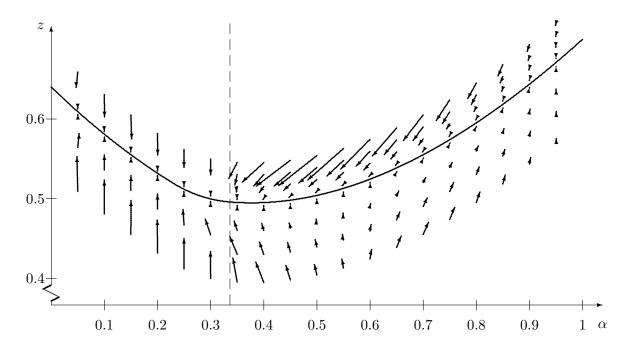Figure 4: Effect of journal selectivity on time allocation and observed paper quality

Figure 5: The disequilibrium evolution of social norms when $\epsilon = 0$.
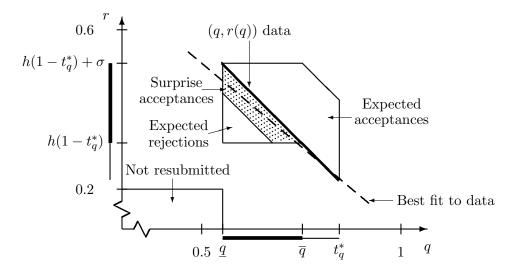

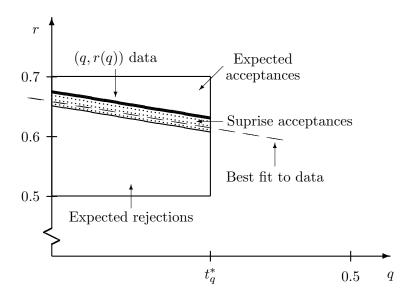Figure 6: Academics' inferences when standards are too high: the "typical" case

40

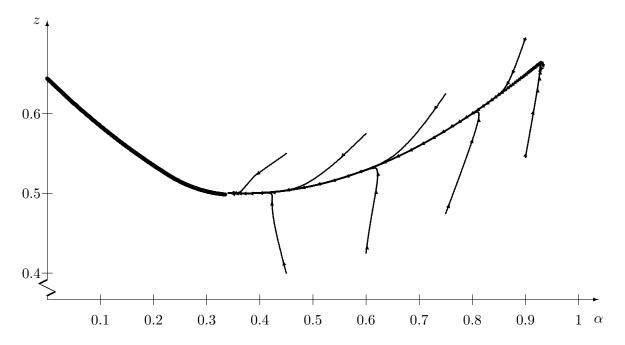Figure 7: Inferences when standards are too high: the "low α" case.

Figure 8: The disequilibrium evolution of social norms when $\epsilon > 0$.

41

Table 1: Mean references per nonreview article and average page lengths

Field	Journal	Pages 1975	Pages 1999	References 1977/8	References 1998
Accounting	Accounting Review	12.1	20.9	12.3	32.4
Anthropology	Current Anthropology	11.2	17.8	61.4	57.3
Anthropology	Amer. Jrn. of Physical Anthropology	8.9	14.0	26.6	47.3
Biology	Cell	9.1	11.5	35.4	49.0
Chemistry	Journal of the Amer. Chemical Society	6.4	7.9	32.8	38.6
Communication	Journal of Communication	9.0	18.6	10.0	52.7
Computer Science	IEEE Transactions on Info. Theory	7.3	13.7	17.5	23.9
Computer Science	Journal of the ACM	12.0	29.0	15.0	37.8
Computer Science	Communications of the ACM	7.6	7.0	14.6	6.9
Demography	Demography	13.7	13.5	16.9	37.5
Ecology	Ecology	9.8	12.7	33.0	56.2
Economics	Econometrica	14.8	31.1	15.1	27.2
Economics	Journal of Political Economy	22.2	29.7	18.0	29.7
Economics	Quarterly Journal of Economics	15.5	36.7	15.9	31.2
Economics	Review of Economic Studies	12.9	23.5	12.2	29.2
Education	American Educational Research Journal	15.0	36.4	19.4	50.7
Education	Harvard Educational Review	23.8	26.8	47.2	43.7
Eng. - Aero.	AIAA Journal	5.8	7.8	11.5	21.4
Eng. - Elec.	IEEE Transactions on Electron Devices	6.6	7.2	13.9	17.7
Eng. - Materials	Journal of Materials Science	7.7	7.8	18.0	21.6
Eng. - Mech.	Journal of Eng. Mechanics and ASCE	15.3	8.5	14.9	19.2
Finance	Journal of Finance	13.9	34.8	15.0	30.8
Geography	Ann. of the Assoc. of Amer. Geographers	12.7	24.4	37.1	67.1
Geology	Geology	4.8	4.0	19.8	25.2
Geology	Journal of Geology	17.8	15.3	35.7	50.7
History	American Historical Review	22.5	31.0	85.9	62.7
History	Journal of American History	22.8	28.4	75.2	65.9
Law	Harvard Law Review	59.1	92.4	156.6	47.7
Law	Yale Law Journal	56.2	85.4	165.9	59.1
Linguistics	Language	18.8	[a]30.1	30.9	61.2
Mathematics	Annals of Mathematics	21.6	35.0	18.0	28.4
Mathematics	Inventiones Mathematicae	22.7	34.6	16.1	25.2
Medicine	New England Journal of Medicine	4.0	7.2	23.0	24.9
Medicine	Journal of the Amer. Medical Assoc.	4.1	6.6	9.5	24.2
Oceanography	Limnology and Oceanography	10.2	11.3	26.5	42.3
Paleontology	Journal of Paleontology	13.8	12.8	19.2	45.4
Philosophy	Philosophy of Science	15.8	20.6	16.0	32.3
Physics	Physical Review Letters	3.3	4.0	15.4	20.8
Physics	Physical Review B	8.9	8.5	25.4	28.6
Physics	Physical Review D	8.9	10.7	22.9	31.5
Political Science	American Journal of Political Science	16.9	24.2	27.6	37.2
Political Science	American Political Science Review	12.7	[a]16.1	37.1	51.0
Psychology	Psychological Review	21.6	[a]30.3	51.4	60.0
Psychology	Journal of Personality and Social Psych.	8.3	14.4	23.8	59.3
Psychology	Developmental Psychology	7.8	11.5	12.6	45.8
Sciences	Nature	4.2	5.7	17.8	25.8
Sciences	Science	7.7	5.5	23.3	28.5
Sociology	American Sociological Review	15.5	17.2	35.8	49.2
Sociology	American Journal of Sociology	22.7	39.3	26.2	60.1
Statistics	Journal of the Amer. Statistical Assoc.	5.9	11.5	13.5	28.3
Women's Studies	Signs	22.1	31.0	22.5	48.5

Notes: a - Data not adjusted for larger page size.

Table 2: Duration of the review process at various journals: 1975 and 1999

Field	Journal	Mean delay 1975	Mean delay 1999
		1975	1999
Mean submission-final resubmission time			
Biology	Cell	1.2	2.2
Computer Science	IEEE Transactions on Info. Theory	6.9	13.6
Computer Science	Journal of the ACM	10.4	21.0
Economics	Econometrica	9.3	26.3
Eng. - Aero.	AIAA Journal	5.3	8.7
Eng. - Elec.	IEEE Transactions on Electron Devices	4.6	6.0
Eng. - Elec.	Proceedings of the IEEE	2.6	6.9
Eng. - Materials	Journal of Materials Science	1.5	5.2
Finance	Journal of Finance	[a]6.5	18.6
Mathematics	Annals of Mathematics	2.0	5.5
Mathematics	Inventiones Mathematicae	1.3	10.8
Philosophy	Philosophy of Science	0.9	5.8
Physics	Physical Review B	0.4	1.3
Political Science	American Journal of Political Science	6.0	7.6
Psychology	Psychological Review	1.7	18.8
Statistics	Annals of Statistics	15.0	18.0
Statistics	Journal of the Amer. Statistical Assoc.	10.1	18.6
Stat. - Bio.	Biometrika	5.3	15.1
Stat. - Psych.	Psychometrika	8.1	16.8
Mean submission-acceptance time			
Ecology	Ecology	8.9	[b]9.2
Economics	Quarterly Journal of Economics	8.5	13.0
Economics	Review of Economic Studies	12.1	28.8
Geology	Geology	2.2	4.6
Oceanography	Limnology and Oceanography	7.0	[c]8.8
Mean submission-publication time			
Astronomy	Astrophysical Journal	7.9	10.7
Chemistry	Journal of the Amer. Chemical Society	7.8	6.2
Geology	Journal of Geology	12.4	10.6
Linguistics	Language	12.8	24.8
Philosophy	Philosophy of the Social Sciences	13.8	22.1
Physics	Physical Review Letters	2.1	5.8
Psychology	Developmental Psychology	9.8	24.1
Sciences	Nature	4.3	4.3

Notes: *a* - Data for 1979. *b* - Does not include time for post-acceptance revisions (which occurred for about 40% of papers.) *c* - Does not include time for post-acceptance revisions (which occurred for about 90% of papers.)