The Academic Review Process:

How Can We Make it More Efficient?

Ofer H. Azar*

Ben-Gurion University of the Negev

Abstract

Recently many editors try to reduce the turnaround times of academic journals. Shorter turnaround times, however, will induce many additional submissions of low-quality papers, increasing significantly the workload of editors and referees, and the number of rejections prior to publication. I suggest several ideas how editors can shorten turnaround times and four ideas how they can still avoid frivolous submissions, thus improving the review process efficiency: higher submission fees; requiring authors to review papers in proportion to their submissions; using differential editorial delay – letting low-quality papers wait more; and banning papers from being submitted after a certain number of rejections.

* Ofer H. Azar, Department of Business Administration, School of Management, Ben-Gurion University of the Negev, P.O.B. 653, Beer Sheva 84105, Israel. Tel.: +972-8-6472675; Fax: +972-8-6477691. E-mail address: azar@som.bgu.ac.il. I thank Gadi Barlevy, James Dana, Eddie Dekel, Glenn Ellison, Ricky Lam, Robert Porter, William Rogerson, Asher Wolinsky and an anonymous referee for helpful discussions and comments. Financial support from The Center for the Study of Industrial Organization at Northwestern University is gratefully acknowledged.

I. Introduction

Recent studies suggest that in the last few decades the publication delay has become much longer than it used to be. Both the submit-accept time and the first-response time (from submission to first editorial decision, henceforth FRT, or FRTs in plural) more than doubled in many economics journals over the last twenty to forty years (Ellison, 2002a; Azar, 2004b). Moreover, similar trends seem to exist in many other disciplines as well (Ellison, 2002b).

Many scholars criticize the long FRT, arguing that it is unreasonable that authors have to wait several months to receive the first decision about their manuscript (see for example Szenberg, 1994). This is especially annoying since the reason for the long FRT is not that referees and editors need months to think about the paper, but that the paper waits months just to be read.[1]

Whether as a result of this criticism or as a means to attract authors, many journals try to reduce their FRT and promise the prospective author an efficient review process. At first, it seems that reducing the FRT is welfare increasing, as it allows disseminating research more quickly. This is particularly important because new research builds on previous results, so every delay causes the entire research chain to be delayed. Moreover, the FRT is especially important, because it delays the average article three to six times (it delays the article in every journal to which it was submitted, including the journals that rejected it), whereas other parts of the publication delay usually occur only once in the life cycle of an article (Azar, 2004a).

One way to avoid the repeated delay caused by the FRT is to allow papers to be submitted to several journals simultaneously. This is common in law journals, for example. This idea was discussed in the literature before; the arguments for and against eliminating the exclusivity of journal submissions are summarized well in the intriguing debate between Szenberg (1994) and Pressman (1994). The main reason against simultaneous submissions,

which is probably also the reason why they are still banned, is that allowing them will increase significantly the number of submissions and therefore also the workload of referees and editors.

While reducing the FRT seems at first desirable, it has potential costs. Some argue that it might reduce the quality of the review process. If the reduction in the FRT is achieved by reducing the time the article spends on the referee's desk waiting to be read, however, the quality of the review process will not be hurt. Azar (2003), however, suggests that even if the quality of the review process is not reduced, a shorter FRT raises other problems. He argues that reducing the FRT significantly will encourage many more submissions of mediocre papers to top journals, increasing the workload of editors and referees. Moreover, Azar claims that the increase in the number of submissions will increase the rejection rates of the journals, causing papers to be rejected more times prior to publication. Considering the time from first submission of an article until its publication (potentially in another journal), this will offset some, and possibly all, of the shorter publication delay due to the shorter FRT.

In the current situation, simultaneous submissions are banned, and the FRTs are 4-5 months and even more, resulting in many papers being published years after they have been written. Are we doomed to live with this situation forever? I propose below several ideas how the FRT can be shortened. Szenberg (1994) suggests another good alternative to shortening the FRT, which is probably even easier to implement – allowing simultaneous submissions. The main drawback of both suggestions is similar – the excessive number of submissions they create and the resulting workload for editors and referees.

The main part of this article suggests four ideas how this drawback can be overcome. These ideas can serve as complementary tools to either a shorter FRT or to simultaneous submissions. The advantages and disadvantages of these ideas are discussed shortly. There is a lot to gain if we can reduce the FRT or allow simultaneous submissions without incurring

the problems caused by excessive submissions. Not only authors will be happier, but readers will benefit as well because research will be disseminated faster. My hope is that this article will stimulate discussion and action among academics, as well as additional research, which eventually will make the review process more efficient than it is today.

II. Why is it important to study the review process?

Since my goal is not only to present my ideas but also to encourage others to do research on the review process, let me discuss briefly why this area is important. There are two main reasons why research on the academic review process is important. First, this process affects the productivity of all economists as well as scholars in other disciplines. Even a small improvement in productivity is very important, because the small improvement is multiplied by thousands of scholars whose productivity is improved. Research on the review process and the profession more generally may suggest how to create better incentives for researchers to publish high-quality work, reduce the time researchers spend on refereeing tasks, reduce the time authors spend on polishing papers (thus allowing them more time to work on new ideas; see Ellison, 2002a; 2002b), and so on. If such research improves productivity by 30 percent every year, the present value of it, taking a discount rate of six percent per year, is that of five years of research in all other areas together!

The second reason for the importance of research on the academic review process is that of diminishing returns. The first article on any topic is likely to contribute more to our knowledge than the thousandth article. Obviously, this article is not the first to discuss the review process; several others, often published in top journals, preceded it. The literature addressed several issues in the review process, including the use of single-blind versus double-blind review (Blank, 1991; Laband and Piette, 1994), payment to referees (Engers and

Gans, 1998; Chang and Lai, 2001), the value-added from the review process (Laband, 1990), the information versus selection function of journals (Meyer, 2000), and the slowdown in the review process (Ellison, 2002a; 2002b; Azar, 2004b). Yet, compared to its importance, the review process has received too little attention. There is much to be added in the topics mentioned above, and there are additional issues that have not been discussed in the literature at all.

The importance of the production process is not limited to academic research. In any industry, improving productivity is a very important task. Outside academia (and government), however, firms have large monetary incentives to invest in improving productivity. Indeed, firms invest billions of dollars on activities that are aimed to increase productivity. Shareholders have incentives to increase productivity, and boards of directors and managers are appointed to supervise the activities necessary to increase productivity and maximize profits.

In academia, there are incentives for individuals to increase their efficiency, but hardly any incentives in the aggregate level to improve the efficiency of the system. Moreover, no one has both the power and the incentives to improve the system. While society may benefit from increased productivity in academia, there are no managers with the authority and incentives to make the production process of academic knowledge optimal. It is therefore important that researchers will dedicate time to think seriously about the issues involved in this process.

Prices in academia often do not reflect supply and demand. The amount of donations and governmental financial support that universities receive is not the true value of the research that they produce; neither are the funds given within the university to a particular department or a faculty member the exact value of the research of the department or professor. The private cost of submitting a paper to a journal is not equal to the social cost,

and the private benefits from having a publication are not the same as the social benefits. The amount journal readers pay for the journal does not reflect the true value of the information in the journal. Research papers are available for free on the Internet, and so on. In short, there is no invisible hand in academia that makes the production process of academic research optimal. We need to be the non-invisible hand, and the more this hand is connected with the brain, the better. In other words, we should study seriously how the production process of academic research works today, what can be changed, and what are the possible effects of these changes. In the absence of market forces or central planning, there is no reason to assume that the current processes in academia are optimal.

III. Current first response times and differences between disciplines

Although many economics journals publish data about the dates in which the initial and final versions of the article were received, or the acceptance date, no economics journal I have encountered publishes information about the FRT of each article published. Going over dozens of journals, however, I found several journals that publish aggregate FRT statistics; Ellison (2002a) and websites of various journals provided me some additional data. Table I presents the FRT in various journals, ranked according to the median FRT for new submissions (the mean when the median is not available).

[Table I here]

Table I suggests that the average of the mean FRT (for new submissions when more than one figure is presented) in the top five journals (*Econometrica, AER, JPE, QJE,* and *REStud*) is 126 – 133 days, depending whether the turnaround time in the *QJE* is taken to be 47 days (all papers) or 82 days (only papers sent to referees). In the other six economics

journals in the table, this average is 129 days. As explained in the comments to the table, these numbers probably underestimate the true FRT, although the bias is not large. Assuming that the journals in the table are representative of other economics journals as well, we can conclude that in most economics journals today the FRT is about three to six months, with an average somewhere between four and five months. If a problem of self-selection is present, however (i.e. if journals publish turnaround statistics only when they have a relatively low FRT), then of course the true average FRT is even higher.

Table I includes also FRT in journals in accounting and finance because the short FRTs in finance and accounting journals provide an interesting comparison with economics. While a full discussion of the differences in FRT between disciplines and their reasons justify a separate article and are beyond the scope of this article, I discuss this issue shortly below.

There are various reasons why FRTs differ across disciplines. Some reasons have to do with the nature of the discipline. We would expect that when many researchers work on the same problem at the same time and there is only one correct answer to the problem, it will be very important for them to publish their results very quickly (before a competing scholar obtains the same results and publishes them first). This will create pressure for very short FRT, because slow journals will not attract good papers. Moreover, when the problems under consideration are more objective in the sense that they usually have only one correct answer, it might be easier to evaluate whether a certain paper makes a significant contribution or not, thus helping to attain a short FRT. These reasons might explain why the FRT tends to be much shorter in the hard sciences compared to economics.

Another reason why the FRT might differ among disciplines is the extent to which scholars depend on previous results for their work. In disciplines in which research depends more on previous results, such as the hard sciences, there is more social gain from quick publication of results compared to disciplines in which new research is less dependent on

previous results (the humanities, for example). Even within disciplines we can observe differences – in game theory, for example, papers build more on previous results than in economic history. If actual FRTs are influenced by the socially optimal FRT[2], we would expect, other things being equal, to see shorter FRTs in those disciplines in which new research is heavily dependent on previous results.

The FRT can also differ among disciplines due to other reasons, such as the extent of specialization in the discipline, how busy is the average scholar in the discipline, how many publications are expected for tenure and promotion, whether publishing books is counted seriously for tenure and promotion, the extent of competition between journals in the discipline, and how much time it takes (in hours) to read a paper, think about it, and write a report on it. Generally, we can think about the FRT in a demand and supply framework: the demand for a short FRT comes from the authors, while the supply comes from the editors and referees. The more important it is for authors to have a short FRT, and the less costly (or more beneficial) it is for editors and referees to accommodate a short FRT, the shorter we would expect the FRT to be.

The comparison between economics, finance, and accounting, however, is very illuminating because these disciplines are quite similar to each other. Indeed, the discipline of finance is often referred to as financial economics, and research in accounting can often be categorized as economics of information. All three disciplines have two main methods of inquiry: abstract mathematical models, and empirical tests using econometric methods. A well-trained economist can probably read and understand a paper in finance or accounting in about the same way that he can understand a paper in economics which is not in his specific field. Many of the finance professors were in fact trained as economists, and even the ones that were not, usually took several graduate courses in economics. These observations suggest that there are very little objective differences between economics, finance and

accounting, and we should therefore expect the FRT to be similar in those disciplines. Yet Table I suggests otherwise.

Why is the FRT in finance only about a third of the FRT in economics? Finance articles are not less complex to referee than economics articles, and finance professors are not less busy than economics professors. There are no big differences between the disciplines in their research characteristics. The big difference in the FRT seems to be the result of a combination of editorial efforts, social norms, and higher submission fees in finance.

Editors in finance seem to be much more concerned with how fast referees respond than editors in economics. For example, the *Journal of Financial Economics* publishes on its website tables with the names of its editorial board members and ad-hoc referees, indicating the number of submissions each person reviewed, and his average review time. Obviously, this creates some pressure on referees not to be the culprits who take months to return a referee report. As another example for editorial efforts to shorten the FRT in finance, the *Journal of Finance* announced in the past "We record both the time the reviewer took and the quality of the review. The quality of the author as a reviewer is a strong influence on our decision of to whom we will send the manuscript. Once again, the best way to ensure a good review is to be a good reviewer when called upon" (Elton and Gruber, 1987). Telling the referees explicitly that responding quickly will affect the quality and speediness they will receive when they submit their papers clearly creates pressure on referees to respond quickly and thus shortens the FRT.

How are refereeing practices related to social norms? When a referee decides how quickly to referee a paper, he takes into account how quickly others referee papers (or in other words, the social norm about how quickly a referee should respond), for several reasons. First, feeling uncomfortable for delaying the publication process is affected by the social norm. If everyone else sends the referee report in two weeks, the referee will feel much

more uncomfortable sending a report after three months, than if everyone else sends reports after three months. Second, the editor's expectations from the referee depend on the social norm, and the referee cares about how quickly he is compared to those expectations because he wants to retain the editor's good will (for example because the referee might want to submit to the journal himself in the future, or the editor might be asked to write a letter on him when he is up for tenure or promotion). Third, referees might want to reciprocate (even though the authors of the papers they referee are not exactly those who refereed their papers in the past) and provide similar service (both in terms of the quality of the review and in terms of how quickly they respond) to the service they receive when they submit papers. Thus, if every time I submit a paper I receive a response in one month I will probably feel more obligated to be a quick referee myself than if I wait five months every time I submit a paper.

It is quite obvious that the social norm about how quickly referees should respond is shorter than in economics. The social norm is how much time it takes referees to complete a report on average, and given the different FRTs in economics and finance and the fact that most of the FRT is due to the time it takes referees to send their reports, it is obvious that the social norm is different. Why is it different? One reason is the greater editorial efforts to have a short FRT in finance. Even editorial efforts in the past can affect the current FRT, because they have affected the social norm, and social norms take time to change (for a model describing evolution of social norms with application to refereeing, see Azar 2004c). Another reason why the social norm might be different in finance and economics is discussed below – the different submission fees. We can say that for the social norm to be different a more fundamental difference has to exist, but once the social norm is different, it has its own dynamics and it can continue to affect referees' behavior and the FRT even if the fundamental differences (e.g. submission fees or editorial efforts) have changed.

Consequently, even if economics journals will adopt the same submission fees and make the same efforts to reduce the FRT as in finance journals, it can take many years before the FRT in economics approaches that in finance, because the social norm in economics is different and it might change very slowly.

[Table II here]

Why do submission fees affect the FRT, and how are submission fees different in economics and finance? As Table II indicates, submission fees in finance (and accounting) are much higher than in economics, creating several effects that might lead to a shorter FRT. First, because submission fees are higher, finance journals have more incentives to compete for submissions. A major way to compete with other journals is to provide authors with a shorter FRT. Second, the higher submission fees alleviate the problem of excessive submissions of low-quality papers to top journals, so the need to use a long FRT to prevent those submissions is not as important as in economics. Third, the higher submission fees allow finance journals to pay referees, thus encouraging the referees to respond in a timely manner, as their refereeing payment might depend on responding by a certain deadline. It should be noted, however, that paying referees does not seem to be the major reason for the difference between the disciplines. Some economics journals also pay referees and are far from having FRTs similar to finance. Moreover, a comparison between an economics journal that pays referees and those that do not suggests that the effect of paying referees on shortening the time it takes referees to send their reports is relatively small (see Hamermesh 1994).

Since the difference between economics, finance and accounting seem to come from differences in editorial efforts, submission fees, and social norms, rather than from anything which is inherently different in the nature of the discipline, it seems possible to reduce the FRT in economics significantly. Editorial efforts and submission fees can be changed

quickly; the social norm about how quickly referees should send their reports will follow more slowly, but will change eventually as well. The next section discusses in more detail ideas how the FRT can be shortened.

IV. How can the FRT be reduced?

How can we reduce the FRT? Editors have various ways to shorten the FRT (see also Hamermesh, 1994; Pressman, 1994; Szenberg 1994). First, they can try to minimize the time it takes them to choose referees and to make editorial decisions once the referee reports are received. Second, they determine how quickly they ask the referees to respond and how much pressure they put on referees that do not respond on time. Third, they can provide monetary incentives for the referees to respond quickly, as some journals do, by paying the referees for timely reports. Fourth, editors can use electronic submissions as a way to cut the delays imposed by the mail (this is especially helpful when the editor does not reside in the same country as the author or the referees).

Fifth, when a referee does not respond in a reasonable time, editors can send the paper to another referee or make an editorial decision with the one or two reports they already have, dismissing the tardy referee from the refereeing task. Sixth, the editors can exclude tardy referees from being referees in the future. This idea is not new: the editors of *The Economic Journal* reported in an editorial note in 1973 "The people in the third category [referees who took over 2 months to send their report] were dropped from our list of referees, unless there was a good reason for the delay" (Champernowne, Deane and Reddaway, 1973).

The two last ideas, however, have the potential problem of inducing referees that care neither about their relationship with the editor nor about their colleagues (the article's authors and its potential readers) to be tardy on purpose, to avoid the current or future refereeing tasks.[3] But referees who do not care about the editor and their colleagues can simply refuse to

referee the paper rather than to respond only after a very long time. Refusing seems to dominate being tardy in all dimensions: it takes less effort, does not hurt the authors much (I assume here that the referee is not spiteful, just too busy or lazy, so that he prefers to hurt the authors as little as possible), and probably makes the editor less angry.

Seventh, when a tardy referee submits a paper to the journal, the editor can send his paper to another referee who is known as a procrastinator. If this policy is known, it will induce referees to respond faster, at least if they consider the journal as an outlet for their own papers. The *Journal of Finance* implemented and announced publicly such a system, as was mentioned in the previous section.

Finally, another idea on how to shorten the FRT is that adopted by the electronic journals *The B.E. Journals in Economic Analysis & Policy*.[4] When an author submits a paper he commits to review two other papers (unless he prefers to pay $350). If he does not do so, he is fined $200 for each review not completed. The credit for the review is received only if the review is sent within 21 days, giving a large incentive ($200) to send the review quickly.

V. Why do lower first response times increase the number of submissions?

One of the main reasons for the negative relationship between the FRT and the number of submissions is that promotion and merit-based pay depend on the publication record. Consequently, the publication delay, including the FRT, postpones the stream of monetary rewards, in addition to making them last fewer years. The FRT, however, is more important to the author than other parts of the publication delay because it delays the publication of the paper whether the journal accepts it or rejects it, while other parts of the delay usually affect only accepted papers. Moreover, having the paper accepted is the critical signal regarding the

quality of the paper, not its actual publication, so the acceptance – publication lag is less important from the author's perspective than the submission – acceptance lag (which includes the FRT as part of it).

The FRT can therefore be thought of as a cost of submission in addition to the submission fee. Higher FRT increases the total submission cost and as a result the author submits the paper more selectively, reducing the total number of submissions. Azar (2003) models this issue and using evidence on the costs of submissions and the benefits of publications he obtains the result that shortening the FRT will increase the number of submissions and the workload of referees and editors significantly. Moreover, as a result the number of times papers are rejected prior to publication will increase and offset some, or even all, of the lower publication delay (each submission will take less time, but a paper will be delayed by more submissions prior to publication).

There are two additional reasons, however, why lower FRTs will cause more submissions, reinforcing the conclusion of Azar (2003) about the dangers of shortening the FRT. First, the author often wants the paper to be published (or at least accepted) in a certain amount of time, for example when a tenure decision is about to be made in a few years. If the author has two years until the tenure decision and he evaluates that it will take a year between receiving the first positive decision (i.e. revise and resubmit) and the final acceptance of the paper (after one or more revisions are made), then he wants to receive a positive response in no more than one year. If the FRT is three months, the author can plan to submit to about four journals at most; if it is six months, he only has time to try two journals. When the author has time for more journals, he can afford to submit to more top journals with low acceptance rates at the beginning, hoping to reap the large benefits of publication in a top journal. When there is time for fewer journals, the author has to start submitting to lower-quality journals with higher acceptance rates sooner, to avoid having no publication at all. As a result, the

paper is accepted after fewer submissions on average when the FRT is longer, because of the submission strategy that the author adopts.

The second reason why lower FRTs cause more submissions is that acceptance chances are often decreasing with time. This is the case for a policy-oriented paper that deals with a problem that may not be interesting any loner if too much time elapses. Moreover, for any paper, as time passes the probability that another article on the same topic will be published (and reduce the acceptance chances of the paper) is increased (not the probability per unit of time, but the cumulative probability that such article preempts the author's paper). While the author does not face a rigid time in which he wants the paper to be published, he wants to start submitting to lower-quality journals not too late. This creates a similar effect on submission strategy as was discussed above, leading again to negative relationship between the FRT and the number of submissions.

It is therefore risky to reduce the FRT if nothing else is changed. With zero submission fees in many journals and nominal ones in others (see Table II), if the FRT will be very short, say a month, even authors of papers with a very small probability to be accepted by a top journal will find it optimal to submit the paper to almost all journals in a decreasing order of quality, creating huge burden on referees and editors. In fact, Azar (2003) suggests that even the current FRT may be lower than optimal given the other characteristics of the publishing process and the incentives to publish. The next sections describe how we can solve the potential problem of increased number of submissions, allowing to reduce the FRT without the risks mentioned above.

VI. Idea #1: higher submission fees

A natural idea (at least for economists) is to prevent excessive submissions by increasing submission fees.[5] In fact, increasing submission fees, in addition to deterring low-quality papers from being submitted to high-quality journals, also provides journals with resources that allow them to compensate referees for timely reports, making a reduction in the FRT more easily attainable. Hamermesh (1994) shows that even a small monetary incentive (about $35) to complete the refereeing task within a few weeks changes the behavior of referees, although not of the slowest referees.[6] Higher incentives, however, may affect also those slowest referees. Moreover, compensating referees may also induce better refereeing from two reasons: first, being paid a significant amount may cause referees to feel obligated to do a thorough job. Second, if the payment is high enough, people may become interested in serving as referees. Writing good referee reports is a way to encourage the editor to ask for the referee's services again in the future.

How to determine submission fees optimally is an important topic that deserves treatment in separate articles, but I will raise a few points here. A significant increase in submission fees, while having the benefit of deterring excessive submissions of low-quality papers, also creates inequality between authors with different monetary resources or with different monetary incentives. Since the FRT is the major cost of submission (Azar, 2003), the increase in submission fees has to be much higher (in percentage) than the decrease in the FRT to prevent an increase in submissions. Azar estimates, for example, that to keep the same number and quality of submissions while reducing the FRT from four to two months, we have to increase the submission fee from $50 to about $375.

Authors from certain countries will not be able to afford such fees at all. In other cases, even if the author can afford the submission fee, he will not want to pay such a high fee. In many countries the monetary incentives to publish are not as large as in the United

States. First, the increase in salary as a result of promotion is not everywhere as significant as in the United States. Second, some countries do not have personal contracts but rather a governmental imposed compensation scheme, meaning that professors who reach the top of the scheme do not have any monetary benefits from publications. If they continue to do research (which is not always the case), it is because the satisfaction they derive from doing research (and the status and respect it gives them) exceeds the effort costs. But this does not mean that they will still want to do research and publish if it costs them thousands of dollars as submission fees. Sending an average paper to a top journal implies acceptance chances of about 9 percent. A submission fee of $400 means that the author can expect to pay on average $400/0.09 = $4440 before the paper is accepted in a top journal. Many authors do not have such strong incentives that will justify paying such huge amounts, and they will avoid submitting their work to journals if submission fees are around $400.

Charging differential fee according to the country of residence may solve the geographical discrimination. *The RAND Journal of Economics* (with submission fees of $50 / $85 for subscribers / non-subscribers), for example, states, "The Editor-in-Chief may waive the submission fee for submissions from authors in developing or transforming countries." It is harder to solve the discrimination between different authors in the same country, however. While journals could in principle ask for income tax forms and charge smaller fees from authors with smaller income, this is not likely to happen. As a result, wealthier authors will have an advantage, as they will be able to afford more submissions. Indeed, maybe the reason that the highest submission fees are in finance and accounting journals (see Azar, 2003) is that professors in these fields can afford higher submission fees than their colleagues in economics departments because they earn more (at least in the United States). Another problem is that certain institutions fund the submission fees of their members / faculty; increasing submission fees will have little effect on the submission strategy of these authors.

VII. Idea #2: paying by reviewing others' papers

A related idea is to increase the cost of submission to the author not by means of a monetary payment but rather by obtaining from him a commitment to review papers in a proportion to the number of papers he submits.[7] Journals use their authors as a main source of referees for a long time, but this is done informally and authors do not face any legal sanction if they refuse to referee papers for the journal.[8] Moreover, the number of refereeing tasks someone receives is not proportional to the number of his submissions. In particular, authors of accepted papers are asked to referee much more than authors of rejected papers. While this might seem reasonable because authors of accepted papers are more grateful to the journal and therefore may accept the refereeing assignments with more willingness to do a good job, it is contrary to the correct incentives. To prevent submissions of low-quality papers, a cost should be imposed on authors of rejected papers, not those of accepted papers.

The suggested system involves an explicit commitment to referee papers based on the number of submissions (or alternatively, on the number of submissions that were rejected). If most papers are sent to two referees, then the author should commit to referee two papers for every paper he submits. If the paper has several authors, they can decide how to split this commitment. Since a submission creates a commitment to undertake significant refereeing work, authors will choose to submit more selectively and therefore to fewer journals, other things equal. This system also allows requiring timely response from the referee for him to reduce his "refereeing debt," thus helping to reduce the FRT.

This idea has several potential problems that should be addressed, however. First, we cannot really make someone referee a paper seriously against his will. So there should be sanctions against authors who refuse to referee or who referee not seriously or not in a timely

manner. The simplest sanction, which was adopted by *The B.E. Journals in Economic Analysis & Policy*, is monetary. You do not want to referee to reduce your "refereeing debt"? Pay with money for the refereeing services you received.[9] This, however, raises problems of inequality in a similar manner to the idea to raise submission fees. Wealthy authors will be able to send their papers to many journals, paying with cash, while poorer authors will have to devote their scarce time to referee others' papers.

Moreover, while it is easy to define what a timely report is, it is hard to define criteria for the quality of the review. Obviously, the editors expect a serious report, not just comments of the type "I read the abstract and introduction and do not think the paper should be published because it offers nothing new. Please credit my refereeing debt for this review." So low-quality reports should not credit the referee debt. Naturally, the editor will have to decide whether the report is satisfying, but in the absence of clear criteria some conflicts may arise if the editor is too tough, and if he is not, some people may take advantage of it and devote less effort to the reports they write than the effort others made on their papers.

Another potential problem is that the composition of the optimal referees may be different from the pool of authors. The editor may want to use certain referees but will not be able because they did not submit papers to the journal, or they submitted but chose to pay with cash, or they already paid (with reviews) their refereeing debt. Of course, editors today also face the same problem, because today no one owes them to referee papers. But when refereeing stops to be voluntary and becomes a payment for refereeing services received, the willingness of people to referee papers they do not have to referee may be reduced. An idea that might mitigate this somewhat is to allow journals to exchange debts. Thus, an author who owes to one journal but is essential for refereeing in another journal can be used in the second journal and be credited in the first one.

A smaller problem is that the editor can also get "stuck" with authors who owe reports but cannot serve as appropriate referees. It might not seem as a problem at first, but it is problematic in two ways: first, each "uncollected debt" means that the editor has to ask someone who do not owe a report to referee a paper. Second, authors who know in advance that they will not be suitable referees will take their commitment to referee less seriously, losing the effect of deterrence of low-quality submissions.

VIII. Idea #3: differential delay – let low-quality papers wait more

An intriguing idea that retains the role of the delay in preventing excessive submissions of low-quality papers while not increasing the delay in publication of accepted papers is the following. Referees will send their referee reports shortly after they receive the paper. The editor will reach his editorial decision based on the reports, and if it is positive (accept / revise and resubmit), he will mail it immediately to the author. If it is negative (reject), he will wait some time before sending the decision. Therefore, the delay for papers that are eventually published in the journal will be minimal, and yet the delay for low-quality rejected paper will be high enough to deter excessive submissions. The delay can even be gradual – the lower the quality of the paper, the higher delay the editor will impose on the author. One can argue that in equilibrium any author who does not receive a decision in a short time will infer that the paper is going to be rejected and will withdraw it. But the journal can institute a policy that a paper cannot be withdrawn within a certain period of time after submission (and if it is withdrawn, the author can be punished in future submissions or otherwise, just as authors who are caught submitting simultaneously to different journals face potential punishment).

There are several problems with this idea, however: first, it seems unfair. This is especially true in those cases where the editor and referees make a mistake (for anecdotal evidence of such mistakes, see Gans and Shepherd, 1994; Shepherd, 1995); not only they reject a good paper, but also they delay the editorial decision unnecessarily. But editorial delays for rejected papers can be similar to what they are today for all papers (so rejected papers are not worse off than today), and using a gradual delay makes it unlikely that a good paper, even if not accepted, will be delayed for a long time.

A second problem is that even a paper that is not suitable for publication in a top journal may be good enough to be published in a lower-quality journal, and it seems a waste to delay it. Once again, however, the delay can be roughly what it is today anyway, so there is no disadvantage in the suggested review process compared to the current situation. Moreover, a paper that is generally good will have only a short delay. In addition, the authors of the low-quality papers that wait for relatively long times should have realized that they are wasting scarce resources and that their papers are of much lower quality than that expected in a top journal.

IX. Idea #4: limiting the number of rejections a paper is allowed to receive

An additional way to prevent excessive submissions is to limit the number of rejections a paper is allowed to receive. Suppose that a paper will be able to receive at most four rejections. If it is rejected four times, the author will have the option of posting it on a central non-reviewed electronic database or discarding the paper altogether. Most authors will not spend all their four trials on top journals. After one or two rejections from top journals, they

will try lower-quality journals, because a second- or third-tier journal is still better than no journal at all.

Such a system has several advantages. The number of submissions of low-quality papers will be reduced both because the authors will be more selective in their submissions and because no paper will be submitted more than four times. This will save the scarce time of referees and editors. Yet, even those papers who were rejected four times and are presumably of low-quality will not be lost – upon the authors' consent, they will be available electronically. Readers who are interested in the subject area will easily find the paper in this central database by searching the relevant keywords or even words in the abstract or the full text.

Moreover, implementing such a system is achievable. Every author will have a unique ID, and every paper will get a unique ID. A central institution (let us denote it by ECD for Economics Central Database) will handle the database about submissions in all journals as well as the non-reviewed database of electronic papers that were rejected four times.

Any submission to a journal will require also submission of an electronic version of the paper. The journal will send the electronic version of the paper together with the authors' ID and the paper's ID to the ECD, which will add the information to its database. The ECD will not inform the journal how many times the paper was rejected, unless it was already rejected four times, so no bias in the journal's decision can occur. If the paper was rejected four times, the journal will refuse to consider it for publication. Once the journal makes an editorial decision about a paper it will inform the ECD whether it is a rejection or not, so that the ECD can update its database. As a by-product, this system will also enable to verify that authors do not submit the paper simultaneously to two or more journals.

One of the nice things about this system is that it does not require full cooperation of all economics journals, only cooperation of most of the good journals. Suppose that a low-

quality journal, the *Journal for Moon Economics* (a fictitious journal, of course), does not want to participate. So the ECD does not know about rejections from the *Journal for Moon Economics*. But papers that were rejected from a low-quality journal are not likely to be submitted to high-quality journals anyway, so it does not really hurt the system that the rejection from the *Journal of Moon Economics* was not recorded.

Consider now a paper that was submitted to four high-quality journals participating in the ECD and rejected from all of them. Suppose that the author then submits to a journal that does not participate in the ECD. If all high-quality journals are part of ECD, then it must be a low-quality journal, and there is no harm in low-quality papers being sent to low-quality journals. Any high-quality journal that will not want to be part of the ECD will automatically be the natural submission candidate for papers that received four rejections, which are presumably not of high quality. Moreover, authors will not want to spend one of their four allowed submissions on such a journal (because they can submit to it on their fifth submission), so this journal is not likely to receive many submissions of papers that were not rejected four times before. Consequently, the quality of a non-participating journal will become very low even if initially it was a good journal, so no editor will want to remain outside the ECD if most high-quality journals participate in it. Moreover, a condition for submission to participating journals can be that the authors commit not to submit more than four times even to non-participating journals. Whether authors obey their commitment is easy to observe because published articles in non-participating journals can be compared to the electronic versions of rejected papers.

There are some technical issues, of course. One has to prevent authors from changing the title and submitting a paper that received four rejections as a new one. Here the electronic versions submitted will play a role. A few graduate students employed by the ECD will sample authors (with an emphasis on those with papers who received multiple rejections) and

check that submissions under different titles are of different papers. It is even simpler to catch such cheating if the paper is co-authored, because two papers at about the same time by the same authors raise immediate suspicion. Minor changes will not be considered a new paper. Major changes might be approved but the author will have to ask permission from the ECD in advance. The incentive to cheat is not big, however, since a paper that already received four rejections is not likely to be accepted in a good journal anyway. And scholars, after all, are usually honest people (see also Laband and Piette, 2000). The very rare occurrence of plagiarism and simultaneous submissions of papers shows this. With little resources dedicated to supervision, the occurrence of cheating by changing the title of the manuscript is likely to be very rare.

X. Conclusion

One of the main criticisms of the academic review process in economics (and in some other disciplines as well) is the long time it takes to receive a decision about a submitted manuscript. Since many papers are rejected from a few journals prior to publication, a FRT of a few months can delay publication by a year or two, delaying the dissemination of research significantly. Several journals make efforts to reduce the FRT, but reducing the FRT significantly has the risk of inducing many more submissions of low-quality papers, increasing considerably the workload of referees and editors. Moreover, the increased submissions will increase rejection rates and delay the publication process because papers will be rejected more times prior to publication.

The question how to prevent submissions of low-quality papers to high-quality journals is an important question and addressing it can help reduce the cost of the review process, especially if we want to reduce the FRT significantly, but even without such a

reduction. This article discusses four ideas that might help achieving this goal: increasing submission fees, requiring authors to review papers in proportion to the number of papers they submit, using differential delay conditional on the paper's quality, and banning papers from being submitted after a certain number of rejections. I hope that these ideas will stimulate further discussion and research, and that consequently the review process will be more efficient than it is today, increasing the contribution of academia to society.

Finally, I want to propose a few ideas for future research. First, all the ideas mentioned in the article deserve additional discussion. Another topic that should be discussed is how much time referees should invest in suggesting improvements to a paper they recommend to reject. There is a trade-off between giving the author as much helpful feedback as possible, and saving the referees' time. In addition, too much helpful feedback for rejected papers may induce people to submit premature papers just to receive feedback.

Empirical research of the review process is also important. Knowing more about the review process is essential in order to determine whether the process is optimal and if not what can be done. How many times is a paper rejected on average before being accepted? What portion of the editorial delay is caused by referees and what portion by editors? How much time does it take referees to review a paper and write the referee report? Do more experienced scholars write better reports, or maybe since they are busy they write less helpful reports? To what extent do authors revise a rejected paper according to the referee report before sending it to another journal? How similar are the reports by different referees on the same paper? Compared to the importance of the subject, so little work was done that the opportunities for future research are abundant. Such research can suggest changes that will improve the efficiency of the review process and the productivity of scholars in all areas.

Notes

[1] Pressman (1994) claims that usually the delay is caused by the referees and not by the editors.

[2] The reason that the socially optimal FRT is strictly positive is that a very short FRT will cause excessive frivolous submissions of low-quality papers to top journals and consequently a large waste of editorial and refereeing efforts. The trade-off between this cost and the benefit of having new research disseminated faster is different in different disciplines, leading to different optimal FRT.

[3] Future refereeing tasks are avoided in two ways: first, other things equal, the editor will prefer to send an article to a referee that was efficient in the past rather than to one that was tardy. Second, as long as the referee does not return the report, he is less likely to receive another refereeing task (see Thomson, 2001, p. 116).

[4] See http://www.bepress.com/bejeap/policies.html.

[5] *The Journal of Financial Economics* explicitly makes use of the connection between submission fees and number of submissions. In a managing editor's note about its decision to raise submission fees to $400 (subscribers) and $450 (non-subscribers) (on-line at http://jfe.rochester.edu/submfee.htm), Schwert writes "The continued rise in submissions is putting a lot of pressure on the resources of the journal (particularly the referees and the editor). I hope you understand that increasing submission fees is a necessary step to deal with this situation." It is interesting that the journal that finds it essential to increase its submission fees is the one that had the highest submission fees also before the increase, while some top journals (such as *Econometrica*, *QJE* and *REStud*) do not charge submission fees at all. In July 2004 the journal raised its submission fees again, to $500 and $550 for subscribers and non-subscribers, respectively.

[6] For theoretical work on paying referees see Engers and Gans (1998) and Chang and Lai (2001).

[7] This idea was adopted from the policies of the electronic journals named *The B.E. Journals in Economic Analysis & Policy*, see http://www.bepress.com/bejeap/policies.html.

[8] The *Canadian Journal of Economics* makes this commitment a little more explicit than in other journals, writing in the Submission Guidelines for Authors (see http://economics.ca/cje/en/submission.php): "Perhaps 50% of refereeing is done by authors of recently accepted manuscripts. Taking account of the fact that the CJE has an acceptance rate of around 25% and that the median paper has two authors and two referees, this implies that an author of an accepted paper would, on average, referee two or three papers for the Journal. This is statistical information only, as there is no formal or contractual linkage between acceptance of a paper and refereeing. However, there is an informal relationship. Basically, if you want to get a paper accepted in the CJE you should expect to do your share of refereeing, and should be willing to do it promptly. Submission of a paper should be treated as an indication of willingness to do some refereeing for the CJE if the paper is accepted for publication."

[9] *The B.E. Journals in Economic Analysis & Policy* charge $350 for each submission if the author does not want to owe reviews (in addition to a $75 submission fee), but $400 if the author chooses to commit to review and later does not do so.

References

Azar, Ofer H. (2003) 'The review process in economics: is it too fast?', working paper, Northwestern University.

Azar, Ofer H. (2004a) 'Rejections and the importance of first response times', *International Journal of Social Economics* 31(3): 259−274.

Azar, Ofer H. (2004b) 'The slowdown in first-response times of economics journals: can it be beneficial?', working paper, Northwestern University.

Azar, Ofer H. (2004c) 'Evolution of Social Norms with Heterogeneous Preferences: An Application to Refereeing Times', working paper, Northwestern University.

Blank, Rebecca (1991) 'The effects of double-blind versus single-blind reviewing: experimental evidence from the *American Economic Review*', *American Economic Review* 81(5): 1041−1067.

Champernowne, D. G., Deane, P. M. and Reddaway, W. B. (1973) '*The Economic Journal*: note by the editors', *The Economic Journal* 83(330): 495−504.

Chang, Juin-jen and Lai, Ching-chong (2001) 'Is it worthwhile to pay referees?", *Southern Economic Journal* 68(2): 457−463.

Ellison, Glenn (2002a) 'The slowdown of the economics publishing process', *Journal of Political Economy* 110(5): 947−993.

Ellison, Glenn (2002b) 'Evolving standards for academic publishing: a q-r theory', *Journal of Political Economy* 110(5): 994−1034.

Elton, Edwin J. and Gruber, Martin J. (1987) 'Report of the managing editors of the *Journal of Finance* for 1986', *Journal of Finance* 42(3): 805−808.

Engers, Maxim and Gans, Joshua S. (1998) 'Why referees are not paid (enough)', *American Economic Review* 88(5): 1341−1349.

Hamermesh, Daniel S. (1994) 'Facts and myths about refereeing', *Journal of Economic Perspectives* 8(1): 153−163.

Gans, Joshua S. and Shepherd, George B. (1994) 'How are the mighty fallen: rejected classic articles by leading economists', *Journal of Economic Perspectives* 8(1): 165−180.

Laband, David N. (1990) 'Is there value-added from the review process in economics?: preliminary evidence from authors', *Quarterly Journal of Economics* 105(2): 341–352.

Laband, David N. and Piette, Michael J. (1994) 'Does the "blindness" of peer review influence manuscript selection efficiency?', *Southern Economic Journal* 60(4): 896–906.

Laband, David N. and Piette, Michael J. (2000) 'Perceived conduct and professional ethics among college economics faculty', *American* Economist 44(1): 24–33.

Meyer, Dirk (2000) 'On the information and selection function of economics articles – claims and reality', *Journal of Post Keynesian Economics* 23(1): 49–75.

Pressman, Steven (1994) 'Simultaneous journal submissions: the case against', *American Journal of Economics and Sociology* 53(3): 316–333.

Shepherd, George B. (1995) *Rejected: Leading Economists Ponder the Publication Process*, Sun Lakes, Arizona: Thomas Horton and Daughters.

Szenberg, Michael (1994) 'Disseminating scholarly output: the case for eliminating the exclusivity of journal submissions', *American Journal of Economics and Sociology* 53(3): 303–315.

Thomson, William (2001) A Guide for the Young Economist: Writing and Speaking Effectively about Economics, Cambridge, Massachusetts: The MIT Press.

Table I: First response times in various journals (in days)

	Median FRT	Mean FRT	Period	Source / journal issue	Comments
Economics Journals					
Quarterly Journal of Economics	NA	47	1997	Ellison (2002a).	All papers.
		114			Accepted papers only.
		82			Papers sent to referees.
Canadian Journal of Economics		91	1/1/02-12/1/02	The journal's website.	
Journal of Economic History	103	108	2000/2001	September 2001.	Including re-submissions.
Southern Economic Journal	107	122	2001	October 2002.	New submissions only.
American Economic Review	109	122	7/1/00-6/30/01	May 2002.	Rejected papers only.
Econometrica	110	122	2000	January 2002.	New submissions only.
	98	92			Revisions only.
	108	122			All papers.
Economic Journal	137	137	2000	RES Newsletter (Jan 2003).	All papers.
	137	125			Letters advising rejection.
	168	188			Letters inviting revision.
European Economic Review	143	165	2000	May 2002.	
The RAND Journal of Economics	153	131	2000	Summer 2002.	Simple average of the four quarters of the year.
Economic Inquiry	NA	159	1/1/02-4/15/02	October 2002.	
Journal of Political Economy	NA	167	2000	Ellison (2002a).	
Review of Economic Studies	175	171	9/2000-2/2001	The journal's website.	New submissions only.
	194	198			First revision.
	159	138			Second revision.

	Median FRT	Mean FRT	Period	Source / journal issue	Comments
Accounting Journals					
The Accounting Review	51	52	3/1/01-2/28/02	July 2002.	Including re-submissions.
Journal of Accounting and Economics	52	53	12 months ending 4/2002	August 2002.	
Finance Journals					
Journal of Financial Economics	37	43	10/2001-9/2002	The journal's website.	
The Journal of Finance	39	44	3/1/00-7/31/02	The journal's website.	Including re-submissions.

Details about the data and how it was computed are available from the author upon request.

Table II: Submission fees in various journals

The Journal	Submission fee – members / subscribers	Submission fee – others
Economics Journals		
AER	$75	$150
Econometrica	$0	Must be members to submit
Intl. Econ. Review	$55	$120
J. Econ. Theory	$0	$0
J. Monetary Econ.	$100	$175
JPE	$50	$50
QJE	$0	$0
RAND J. Econ	$50	$85
REStat	$0	$50
REStud	$0	$0
Accounting Journals		
The Accounting Rev.	$75	$100
J. Acc. & Econ.	$250	$300
J. Accounting Res.	$200	$200
Finance Journals		
J. Finance	$70	$140
J. Financial and Quantitative Analysis	$200	$260
J. Financial Econ.	$500	$550
Rev. Financial Stud.	$125	$175