The Inside Scoop: Acceptance and Rejection at the JIE[1]

Ivan Cherkashin[2]
The Pennsylvania State University

Svetlana Demidova[3]
University of Georgia

Susumu Imai[4]
Queens University

Kala Krishna[5]
The Pennsylvania State University, and NBER

This Version: October 2007

[1] We are grateful to Abdullah Yavas and Bee Roberts for comments on an earlier draft, to Anshul Kumar, Su-Jen Roberts and Supriya Mishra for able research assistance, and to Sergey Lychagin for help with the PERL code for citation extraction. This is a very preliminary draft. Please do not cite without permission. Comments are very welcome. The views expressed here are those of the authors and do not necessarily reflect those of the institutions to which they are affiliated.

[2] Department of Economics, Pennsylvania State University, University Park, PA, 16802. E-mail: ivt101@psu.edu. Phone: (814) 865 2745.

[3] Department of Economics, University of Georgia, Athens, GA 30602. E-mail: demidova@terry.uga.edu. Phone: (706) 542 3689. Fax: (706) 542 3376.

[4] Department of Economics, Queen's University Kingston, Ontario, K7L 3N6, Canada. E-mail: imais@econ.queensu.ca. Tel: (613) 533-2258. Fax: (613) 533-6668.

[5] Department of Economics, Pennsylvania State University, University Park, PA, 16802. E-mail: kmk4@psu.edu. Tel: (814) 865 1106. Fax: (814) 863 4775.

Abstract

There is little work on the inner workings of journals. What factors seem to affect the ability to publish in a journal? Could simple rules (which are already used by some journals) like the immediate rejection of a significant minority of papers, help to streamline the process? At what cost? How well do journals seem to do in choosing papers? What can we say about the extent of type 1 and type 2 errors? Do editors seem to have uniform standards or are some harsher than others? We use data on submissions to the Journal of International Economics to help answer such questions.

Keywords: Publishing in Economics, Performance evaluation, Probit model, Selection bias.

JEL: A14.

1 Introduction

Tenure, promotion, and salaries in academia are based on the quality and quantity of publications. Such factors also affect an individual's market value both in and out of academia. Despite this, there is little to be found on the inner workings of journals and on how well they perform. In part this is because data that matches papers with authors and their characteristics, as well as with the editor handling the submission, is sensitive and rarely available. Most previous work in this area, with a few rare exceptions mentioned below, has used data on published articles from one or more journals or small random samples obtained from the editors of these journals. Rarely has data been available on both published and rejected papers: even when such data is available, information on authors' characteristics has been quite limited. The matched editor/author/paper/outcome data we have constructed is, thus, quite unique. We have compiled data that combines information on paper submissions and editor assignments to the Journal of International Economics, with in-depth data on authors and citations.[1]

We analyze the data with a view to describing, evaluating, and improving the process. The paper proceeds as follows. Section 1.1 provides a selective survey of the literature to date so as to place our work relative to this literature. Section 2 describes the data as well as some interesting patterns that occur. Sections 3 and 4 make up the heart of the paper. Section 3 takes an ex-ante approach. The main question asked is whether a two stage procedure, where a fraction of papers is rejected without going through the refereeing process, might significantly reduce costs without compromising on selection quality. We answer this question by looking at the determinants of rejection and our ability to correctly predict it on the basis on author characteristics alone. If a significant fraction of the papers can be correctly rejected on this basis, then surely an editor can do better by taking a quick look at the paper! We find that there is much to recommend this procedure.

Section 4 tries to evaluate the performance of the JIE along various fronts. We look for evidence on the extent of type 1 (convicting an innocent man, or in our case, rejecting a good paper) and type 2 (letting a guilty man go free, or in our case, accepting a bad paper) errors. We argue that type 1 error is quite low, while type 2 error is high. While there are certainly flaws in our approach, this is the first time that this has been attempted in this area, at least, to our knowledge.

We also incorporate some ex-post data, namely citations, and ask whether citations are a good measure of quality and whether there is any evidence that higher quality papers have a higher probability of acceptance. If citations are closely related to quality, and the acceptance decision is based only on paper quality, then once citations are included, nothing else should matter! Even if citations are not a perfect proxy for quality, including them should reduce the size of the coefficients estimates of the remaining explanatory variables, or

[1] Our data covers ten years, from 1995-2004, with about 3,032 observations. Author data was manually collected from the author's curriculum vitaes whenever feasible. For a sub-sample of 2031 papers we have citation data collected from Google Scholar.

make them less significant, which is what we find!

We compare the distributions of citations of groups of papers. For example, we find that the distribution of citations for papers accepted by the JIE first order stochastically dominates that of papers rejected by the JIE as well as those rejected by the JIE but accepted at higher ranked journals! This is consistent with the JIE doing a good job of not rejecting good papers, or type 1 error being low. However, about 10% of published papers are never cited suggesting that type 2 error could be large.

Finally, we look at co-editor specific effects. We ask whether co-editors seem to differ in their acceptance rates in ways that cannot be accounted for by author characteristics or paper quality. We find significant differences in co-editor acceptance behavior, as well as evidence supporting the hypothesis of differences in acceptance criteria.

Section 6 outlines our policy conclusions as well as directions for future research.

1.1 Existing Work

There are two main groups of papers in terms of the questions asked. The first group deals with questions related to the determinants of the time it takes to publish. The second group deals with whether there is evidence of bias in acceptances.

Despite a proliferation of journals, there seems to have been a significant slowdown in publication process. It seems to have become the norm for papers to undergo multiple revisions, with each round easily taking six to nine months. Even when accepted, papers can take a year or more to come out. Coe and Weinstock (1967) found this process took about 250 days in the 60's, while Yohe (1980) found that between 1966 and 1979 the delay in the publication process increased substantially with the average time between submission and publication being 15.3 months for specialized journals and 23.3 month for major general interest journals. Trivedi (1993) used data on 1134 submissions from 7 econometrics journals from 1986 to 1990. He found that delays were large and increasing over time with the average lag exceeding 31 months in 1986 and 34 months in 1990. Supplementing his data with survey data from authors (34 complete answers from 135 questionnaires) Trivedi constructed delay distributions, but only had data for published papers. He argues that: "It is also important to find out how long the rejected papers stay in the processing line. That statistic reflects the efficiency with which the profession deals with research submitted for publication in journals. These data are also available to the journal editors, but rarely published." He then suggests that the processing delays for rejected papers should be similar to those of accepted papers. However, we find that the delay for accepted papers exceeds that of rejected ones, and that the conditional survival probability increases over time.

Bowen and Sundem (1982) obtained data directly from the editors of leading accounting and finance journals on the durations of all the steps that articles go through between submission and publication or rejection. With data on a random sample of 40 accepted and 40 rejected papers from 9 journals, they compare the duration of different stages in submission among journals. They find that a lion's share of accepted papers (219 out 281) went through one or more revisions. At the same time only 14 out of 326 rejected papers were not rejected in the first round. We find a similar pattern in our data.

Ellison (2002a) describes the changes in submit-accept times for different journals and then proposes and examines possible causes for these patterns. He looks at several possible explanations for the increase in submit-accept times, including democratization of the profession away from an old boy's network, increased complexity of articles, and growth of the profession. To test the first hypothesis, Ellison collects data on authors of the *published* papers only and then regresses submit-accept times on variables that proxy for the authors' position in economic hierarchy such as publications in top journals and contributions to the AEA papers and proceeding or Brookings papers, which are invited but prestigious. He finds no statistically significant relationship between the submit-accept time and authors standing in the profession. It should be noted, however, that this does not account for the possibility that the authors standing could affect the probability of acceptance rather than the time to publication. Moreover, the use of published papers only creates selection bias.

The complexity of papers is somewhat difficult to measure. To test this explanation, Ellison (2002a) uses proxies such as the length of the paper, the number of coauthors, and the degree of specialization as reflected in the JEL index. Since the 1950's, the average paper gained approximately 75% in size while the share of coauthored papers doubled from 30% to 60% from 1970 to 1990 in Econometrica and REStud. Ellison finds that each extra page seems to add 5 days to the time to the first decision. The overall increase in the average number of coauthors from 1.4 to 1.7 accounts for about 10 days of delay. At the same time, he finds no support for the increased specialization hypotheses.

He also argues that there is not much evidence of growth in the profession. Comparing the number of submissions to the best journals such as Econometrica, JPE, AER, and QJE, Ellison fails to record any dramatic trend. However, higher standards for acceptance could reduce submissions and keep acceptance rates constant. Other measures, such as connections with the editor or NBER membership, also failed to have any explanation power or had the "wrong" signs. He argues that while the first response time grew somewhat, the number of revisions and the time spent on them increased more severely.

Ellison (2002b) makes the case that the balance between the importance of the main idea, (q), and other aspects of quality, (r), has changed as referees, who have an upwardly biased view of their own work, update their priors on the social norm regarding the importance of the two. He finds that papers with better ideas (as measured by position in the volume and citations) on average have a shorter reviewing time.[2] However, this theory explains only about a quarter of the increase in the delay. Nevertheless, Ellison's work remains the most extensive and up to date research on publication lags and their possible causes.

[2] In our data, however, the correlation between citations and time to the first decision is not significantly different from zero, and this is so for both accepted and rejected papers separately. However, the correlation between the time to the final decision for accepted papers and citations is slightly negative. This could be because lower quality papers require more polish to be acceptable.

The second direction taken in this literature has been to test for bias in acceptance/rejection. A number of authors look for evidence of biases according to gender, as well as closeness to editors or co-editors of the journal. For example, Laband and Piette (1994) use citation data to test for favoritism. They find, if anything, the opposite: articles published by people in an editors' network tend to have a higher, not lower citation index![3] They speculate that editors seem to use their personal ties to obtain better papers for their journals.[4] Blank (1991) looks at the outcome of an experiment carried out by the AER as an indicator of bias. During the experiment papers were randomly selected into two groups. Those in the first group were sent for a single-blind review, i.e., the referee had information about the author's identity, though the author did not know the identity of the referee. Those in the second group were sent for a double-blind peer review. Under the double-blind review system, referees had no information about the author. One of her key findings was that under double-blind review, rejection rates were higher and referees were more critical about the papers. At the same time Blank did not witness any discrimination by gender, but outlined some differences in acceptance rates on the basis of university ratings: applicants, who worked at near-top universities or from non-academic institutions, had lower acceptance rates under the double-blind review system.

Hamermesh and Oster (1998) look at how productivity and the probability of acceptance vary with age. Their results are based on data about 208 faculty members of the leading 17 economic departments who obtained their degrees between 1959 and 1983. They find that researchers are most productive in the first decade after graduate school, and slow down over time. However, early high productivity seems to be a characteristic of those who remain productive many years later.[5] They also obtained a random sample of submissions to a top general interest journal. This data suggests that the probability of acceptance does not seem to vary with the author's age, though highly cited scholars have a significantly higher probability of acceptance. We find a slow increase in the probability of acceptance with professional age, but do not wish to make too much of this.

How does our work relate to that in the literature? It is complementary to the literature in that it validates some previous findings, and questions others (like the constancy of the acceptance rate as a function of age) using a new data set. It differs from it in a number of ways. First and foremost, it is the only paper that evaluates the performance of a journal and its co-editors directly. We can do so because our unique data set.

[3] Citations may be a bad indicator of quality. A paper with serious flaws may have a high citation index because others cite its defects. Also, insular networks may deliberately cite each other's work making citation numbers suspect.

[4] For obvious reasons the authors use the data on published articles, not on submissions.

[5] However, it is not clear if this is due to talent or the fact that talented academics tend to have a higher initial job placement where (due to lower teaching loads and a more active research environment) it is easier for them to stay at the forefront of research.

2 The Model and the Data

We assume that each article has a quality, q, which cannot be observed directly. The purpose of the editorial process is to identify q and accept the article i if q_i is higher than a threshold level, Q.[6] We distinguish between factors that can be observed at the time of submission, such as the author's identity, and those that cannot, such as the citation index of the paper. While the latter can be used to evaluate the process of selection and outcomes, only the former can be used to help to guide it.

2.1 The Model

We assume that the quality of a paper depends on the author's abilities (a) and efforts (e), as well as an element of luck:

$$q_i = g(a_i, e_i) + \varepsilon_i,$$

and that the article is published if $q_i > Q$.

Ability and effort could be proxied for by the author's education, experience, and performance to date as reflected in his/her publication record. Professional age could also be related to effort as Assistant Professors, who have just completed their Ph.D.s and are working to obtain tenured position, might put in more effort and so be more likely to submit high quality papers, other things being constant. They might also be closer to the frontier, especially if they come from good programmes, than faculty who are not research oriented and whose human capital has depreciated since graduate school.

By making assumptions on the distribution of ε_i, we obtain either the probit or logit model from this setup. However, such simple binary choice models could have significant problems applied directly to our data as there is a selection problem we need to deal with. The selection problem comes from the fact that while we have data on all submissions, we only have curriculum vitaes for authors with a web presence, and as a result, the submissions we can use are restricted. As the acceptance rate for the latter exceeds that of the entire population, there may be selection bias that we need to correct for. Table 1 below shows that that vitaes are more commonly available as time goes on, and that the two groups, those with a web presence and those without, seem to become more different in terms of their acceptance rates: those with a vita were roughly 2.8 times as likely to be accepted in 1995 relative to those without a vita, and 3.75 times as likely to be accepted in 2004, suggesting that only the most marginal authors did not have vita's by 2004. Note also that the acceptance rates themselves have halved over this period from 28% to 14%. While we have CV data for 72% of the authors of accepted articles, we only have 50% of CVs for the authors with unsuccessful submissions. In other words, authors of accepted papers are over-represented in our sample.

[6] Ellison (2002b) argues that overall quality is produced by importance of the main idea and other aspects of quality (quality of math, econometrics, robustness checks and etc.).

The intuition for the expected bias is evident from Figure 1. In Figure 1, think of X as the explanatory variable in the model estimated. The object is to find the coefficient on X that maximizes the likelihood (or minimizes the sum of squared residuals) of the observed data. This gives the solid line shown in Figure 1 which depicts the estimated value of βX when the entire sample is used. If $\beta X + \varepsilon$ exceeds Q, then $y = 1$. Thus, the higher is X, the higher is the probability of acceptance, i.e., of the dependent variable being unity. Hence, most of the high X data points have $y = 1$, and most of the low X points have $y = 0$, though some high X data points are rejected and some low X ones are accepted.

Table 1: Acceptance Rates

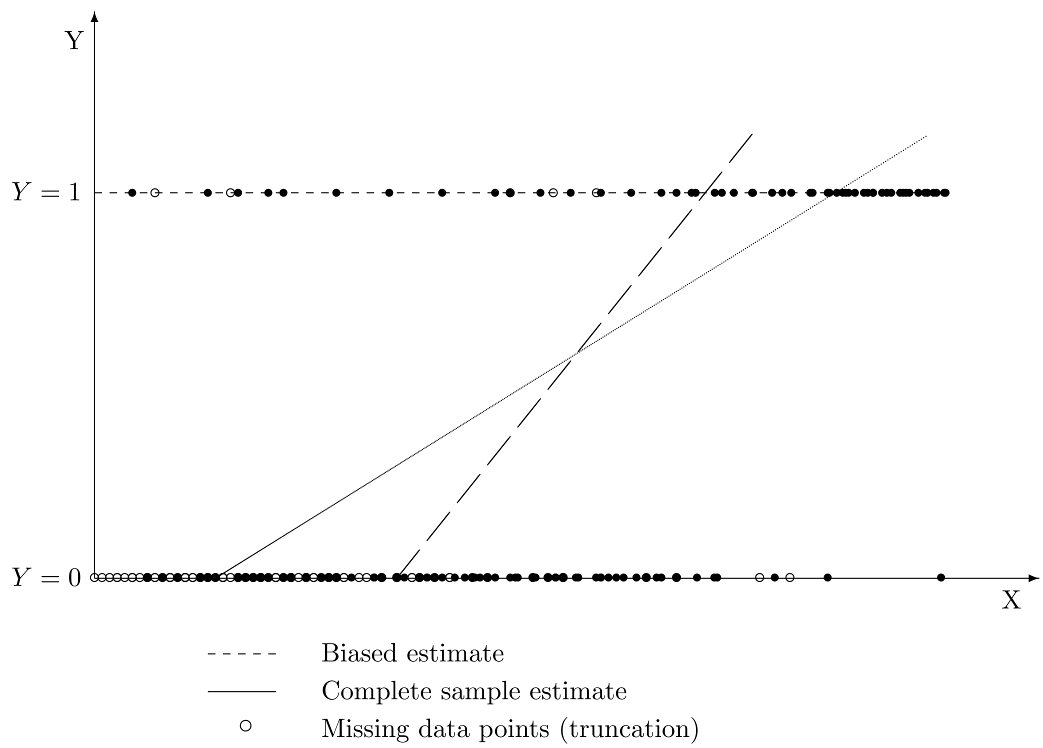
	Total Submissions	Vita Available % Accepted	Vita Not Available % Accepted	All Submissions % Accepted	Fraction with Vita
1995	274	39	14	27	54
1996	286	36	15	25	49
1997	254	31	11	20	53
1998	234	30	5	19	60
1999	353	30	14	21	56
2000	268	28	12	20	55
2001	291	27	8	20	61
2002	313	28	7	19	61
2003	360	23	6	17	67
2004	399	19	4	14	63
Total	3032	28	12	21	59

However, if agents with low values of X are less likely to be in our sample (as they are less likely to have a web presence), then such points are going to be under-represented in the sample due to truncation. To depict this, we remove the points that are not filled in. Low X points are removed more often than high X ones. This in itself is not going to result in bias.[7] However, if given X, submissions without a CV are less likely to be accepted, then we will see more points not filled in at $Y = 0$ than at $Y = 1$, and this is going to bias the estimated slope parameter upward. In this case, if we estimated the model using this truncated sample, we would get the dashed line, which is steeper than the one for the case when the whole sample is used, so that the estimated β is going to be biased upward. Of course, if the selection was random, so all points were equally likely to be removed, then there would be no bias.

Thus, we need to account for this selection bias in our estimation by incorporating the selection equation into the likelihood function as done below. Let X be the set of authors' characteristics that affect the

[7] Analogously, in a standard linear regression, removing low X points more often than high X ones will not bias the fitted line, though it would raise the standard error.

Figure 1: Selection Bias



likelihood of acceptance and Z be the set of authors who have a web presence. These sets could overlap to a greater or lesser extent. We assume that the article i is published in the journal if its latent quality q_i exceeds a threshold level:

$$Y_i = \begin{cases} 1, & \text{if } q_i = X_i\beta + \varepsilon_{1i} > 0, \\ 0, & \text{if } q_i = X_i\beta + \varepsilon_{1i} < 0, \end{cases} \tag{1}$$

where Y_i is an indicator for the paper being published ($Y_i = 1$) or not ($Y_i = 0$). Note that under such a specification we have to include a constant term in $X_i\beta$, which provides an estimate for the threshold level Q.

However, we observe the author's characteristics (the CV) only if $Z_i\gamma + \varepsilon_{2i} > 0$:[8]

$$(Y_i, X_i, Z_i) = \begin{cases} (Y_i, X_i, Z_i) & \text{if } Z_i\gamma + \varepsilon_{2i} > 0 \\ (Y_i, \textit{Not observed}) & \text{if } Z_i\gamma + \varepsilon_{2i} < 0 \end{cases} \tag{2}$$

It is natural to expect that ε_{1i} and ε_{2i} are positively correlated: authors who have a better chance of being published are also more likely to have an established name in profession and have CVs that are easily accessible on the web.[9] To allow for this, we assume that ε_{1i} and ε_{2i} come from the joint normal distribution. This problem is usually referred to as an incidental truncation problem. Versions of such models can be found in a number of applied articles, see, for example, Weiss (1993) and Jenkins et. al. (2004). However, due to specific data structure and the type of truncation, none of these models could be directly applied to our data. We also cannot use the standard Heckman correction for selection as we have no data on the submissions of authors without a web presence. We use maximum likelihood techniques in a way similar to Weiss (1993) to correct for this selection bias. Our problem is in essence a simpler version of his. The likelihood function for our problem is derived in Appendix. Details on its maximization are also provided.

2.2 The Data

We have several sources of data.

2.2.1 Journal Based Data

The JIE displayed steady growth through 1995-2004. Its size doubled as it went from 700 to 1400 pages per year during this period. Its publication pattern changed discretely in 1998: instead of 4 issues per year the JIE started publishing 6 issues. Despite a temporary drop in 2001, due to the publication of two special issues with a slightly larger number of articles, the number of pages per article also increased by the end of the period.

[8] This needs to be qualified as citation data is observed even for some submissions that lack C.V.'s and are not observed for some that have C.V.'s.

[9] If the errors above are uncorrelated, then there is no bias in estimation.

Submission data

We have data on ten subsequent years of article submissions from 1995 to 2004. For each submission, we observe the name of the author, the title of the paper, the date of submission, the name of the co-editor who handled the article, the date of the first decision and subsequent decisions if any, as well as the decisions themselves.

The decision making process at the JIE is as follows. When the JIE receives an article, the editor decides who handles the paper, the editor or a co-editor. After that, whoever is handling the paper sends it to two referees of his/her choice. The referees observe the name of the author as the JIE follows the single-blind review practice. Once the referee reports are in, there are three possible outcomes: accept, decline, or revise and resubmit. In case revisions are requested, additional rounds occur. We observe at most four such rounds in the data. Once the paper is accepted, it joins the queue for publication.[10]

Overall, the JIE received about 3032 submissions of which almost 600 articles were accepted for publication. Thus, about 20% of submissions were finally accepted. As is evident from Table 1, acceptance rates fell by about a half over the period, despite a doubling in the size! The increase in size did result in a blip upward in acceptance rates from 18.8% to 21% in 1999, but the downward trend continued.[11]

Co-editor Information

We include dummy variables for the co-editors who handled the papers. Co-editors have quite different raw acceptance rates. This could occur if co-editors have different views on the minimum acceptable quality of a paper. However, this is not the only possible interpretation. Papers need not be distributed randomly across co-editors. On the contrary, articles would likely be sent to the co-editors whose expertise is closest to the paper, and if some areas are hotter than others, this could result in some editors having higher acceptance rates. Another possibility is that more interesting articles are retained by those assigning papers to co-editors, i.e., there could be a cherry picking effect. This could again lead to differences in raw acceptance rates that have nothing to do with differences in standards. However, by controlling for author characteristics, we control for such composition biases, at least to the best of our ability. Co-editors are only identified by number to preserve confidentiality. There are 21 co-editors who worked with the JIE at some time in this period and who handled a non trivial number of articles.

Backlog

Like many journals, the JIE has a stock of articles that have been accepted but are awaiting publication. We construct a backlog variable to see if this has any effect on the probability of acceptance. The backlog could affect the decisions of the co-editors if information on the backlog is conveyed to the co-editors, who,

[10] There are other possibilities. For instance, an article may be withdrawn. We do have a few such observations in our sample, but far too few to carry out any analysis.

[11] We see slightly higher submission rates in June and July, perhaps as academics finish off leftover projects, and after the summer, in October. We also see slightly higher rates in February which might be due to submissions that occur after being rejected at a general interest journal or after working during the winter break.

in turn, raise standards and reduce the acceptance rate. It could affect submissions if the increase in the backlog was known to authors and this reduced submissions, which, in turn, could raise the probability of acceptance.

We construct the backlog variable by adding up the number of articles accepted during a period less the number of published articles, which gives the change in the backlog. We know that as of September 1, 2002, the backlog was 73 articles. From this we can estimate the backlog for each month of our sample. In other words, if X_t is the number of articles in queue at time t, P_t is the number of articles published at time t, and a_t is the number of articles accepted, then X_{t+1}:

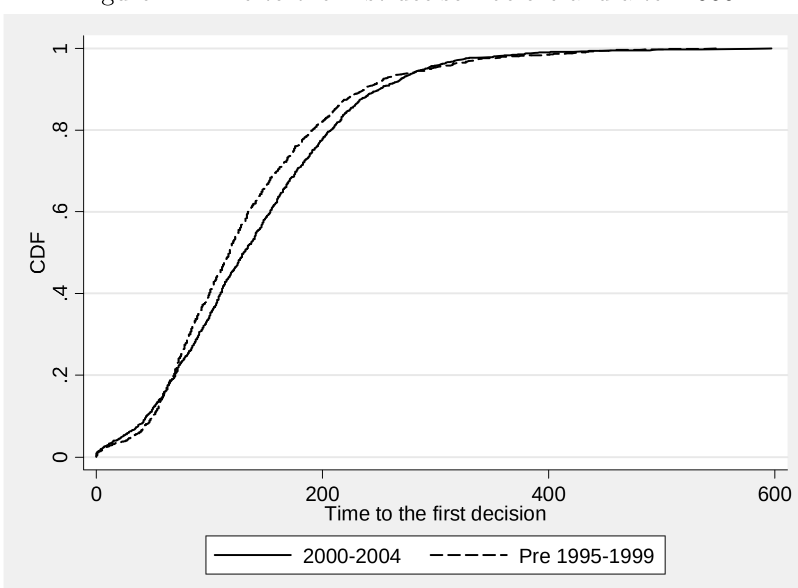
$$X_{t+1} = X_t + a_t - P_t.$$

In our regressions we use the backlog in the previous month as an explanatory variable.[12]

Time to First Decision

We have the date the paper was first submitted and the date of the first decision. The length of the sample we have allows us to test whether the JIE demonstrated any increase in the processing time over the decade. We split our sample into two parts: articles submitted in 1995-1999 and in 2000-2004:

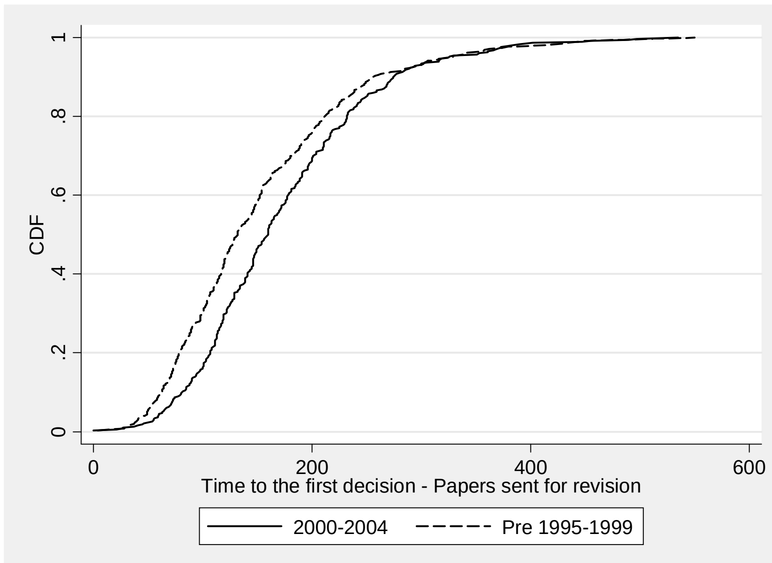
Figure 2: Time to the first decison before and after 2000



[12] The backlog variable we calculate is accurate for later years of the sample (1999-2004), but is less so for 1995-1998. This is a result of not observing all acceptances in the earlier years of the sample. This biases our imputation for the number of papers accepted, and thus, for the backlog variable downward. However, by 1999 this error should be close to zero as it is unlikely that articles submitted in 1994 or earlier are still under revision in 1999.

There is an increase in the time to the first decision[13]. The natural question is where these delays are coming from. Do we observe an increase in waiting time for both rejected papers and papers sent for revision? Plotting cumulative distributions for both categories of articles for earlier and later years gives some insight into the reasons for the increase in time to the first decision. (See Figures 3 and 4.)

Figure 3: Time to the first decisons for papers sent for revision



For papers that were rejected, the time to the first decision remains about the same for the whole period[14]. For articles that were not rejected during the first round, reviewing time increased quite noticeably[15]. Note that this is consistent with the hypothesis of Ellison (2002a) that there has been an increase in the polishing component of quality, r, rather than, q component.

2.2.2 Vita Based Data

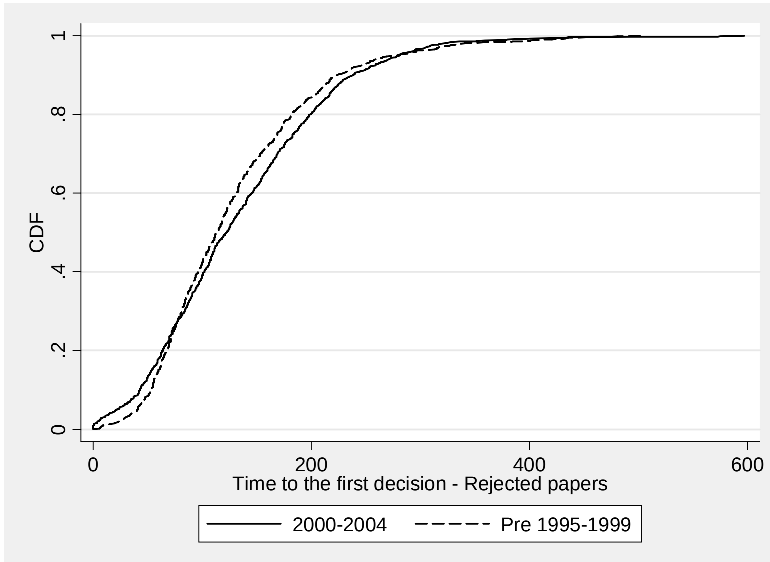
In addition to information on the decisions regarding each paper, the timing of each of the stages, the co-editor assignment, and the applicants' names and titles of the papers, we collected detailed data on authors' background. We obtained this information from their curriculum vitae. We coded data on a number of variables.

Ph.D. Vintage

This variable is not the year the Ph.D. was awarded, but the number of years since the Ph.D. was awarded

[13] The Anderson (1996) test allows us to reject the FOSD hypothesis, while the Kolmogorov-Smirnov test rejects the null hypthesis that the distributions are the same. See Appendix, Table 9, Column 3.
[14] The Anderson (1996) test allows us to reject FOSD hypothesis. See Appendix, Table 9, Column 5.
[15] The Anderson (1996) test does not allow us to reject FOSD hypothesis. See Appendix, Table 9, Column 4.

Figure 4: Time to the first decision for rejected papers



at the time of submission. It helps capture how human capital and incentives vary across the lifecycle. On the one hand, young Ph.D.'s could have more current human capital, higher ambitions, and be willing to invest more in their research to get tenure and because they have a long time to recoup their investments. As a result, they may be more likely to write high-quality papers. They may also be particularly keen on getting an acceptance before tenure and so willing to submit a high quality paper to the journal, where an acceptance is more likely, rather than take their chances elsewhere. On the other hand, with age comes experience: for instance, they might be better able to choose where best to submit a paper or how to sell it, thereby raising their probability of acceptance. Which of these effects is going to dominate is not obvious ex-ante. For this reason, we include a set of dummy variables to proxy for such effects in a flexible manner. In total we have 6 dummies that indicate that an applicant obtained his Ph.D. from zero to two years ago, from two to four, from four to six, from six to 10, from 10 to 20 or that the Ph.D. is not completed at the time of submission. Scholars who received their degrees more than 20 years ago are used as the reference group.[16]

University rank

This gives the ranking of the university that awarded the author's Ph.D. We use the ranking in Kalaitzidakis et. al. (2003), which provides world-wide rankings of the best 200 economic departments and separate rankings of the 120 European economic schools. We record world-wide school rank. If the school is not on this list, we check for the European rank and record it separately. Schools on neither list are labelled

[16] We tried but failed to collect good data on tenure status at the time of submission.

"non-rated". These dummies help proxy for the "quality" of the author, and hence, the article.

Employer type

We specify whether the author was employed at a US, Canada, United Kingdom, or European University[17] or at a university anywhere else on the date of submission. For authors employed outside of academia we code whether they worked in research or international organizations, for instance, such as the Fed, the IMF, or World Bank, or in business. We also distinguish between organizations based in and outside of the United States. Table 2 summarizes the affiliation of authors in our sample.

Table 2: Author's Affiliations (shares)

	US Univ.	*CAN. Univ.*	*UK Univ.*	*EU Univ : Excl. UK*	*Other Univ.*	*Organizations*	*Business*
1995	55	6	8	10	10	11	0
1996	52	9	8	15	11	6	0
1997	46	7	10	18	11	7	1
1998	44	9	8	16	14	8	1
1999	47	6	6	16	10	14	0
2000	45	7	9	25	4	11	0
2001	51	4	7	25	9	5	0
2002	44	5	11	18	11	12	0
2003	39	4	7	22	13	15	0
2004	39	5	7	28	11	8	0

An obvious trend is a decrease in the share of submissions from authors affiliated with the US universities and a corresponding increase in the number of submissions from researchers associated with European Universities. This suggests that at least in International Economics, the US may well be losing ground. The share of submissions from various organizations is stable at about 10%. Very few submissions come from business employees but this could be partly due to their not having a web presence.

Number of previous publications:

These are broken down into those in the leading general interest journals[18], the number of publications in the second tier general interest journals[19], and the number of publications in the JIE prior to submission. We also have the total number of papers in economics journals and the number of publications in books.

Native language

[17] We treated Norwegian and Swiss universities as EU universities though neither Norway or Switzerland is the member of the European Union.

[18] These include: The American Economic Review, Econometrica, The Journal of Political Economy, The Quarterly Journal of Economics, and The Review of Economic Studies.

[19] These include The European Economic Review, The Economic Journal, and The International Economic Review.

Quite a few papers submitted to the JIE are written by non-native speakers who might have a harder time getting their article published. We include a language dummy to allow for this. Unfortunately, many economists do not explicitly state in their CVs if English is their native language. Due to this we define language proficiency based on the other characteristics of the person. We treat a person as proficient in English if he obtained his bachelors and subsequent degrees from a university located in an English speaking country or if English is stated to be the native language.

2.2.3 Publication and Citation Data

We also collected data on the final outcomes with each submission. Here we looked for information of its ultimate fate as well as its reception by the profession.

The fate of the article

For those papers rejected by JIE, and for which we have data on at least one of authors, we record whether the paper was finally published or not. If published, we code where in terms of the ranking of the journal, in deciles (top 10, 10-20,...).

Citation data

The number of citations could be an indicator of the quality of the paper. Of course, there are problems here as well. For example, citations could be negative rather than positive! Moreover, published papers are more likely to be cited just because they are published. However, since many published papers have zero or close to zero citations, while other unpublished ones are highly cited, this is less of an issue today than before the internet.

There are several sources of citation data. The Social Sciences Citation Index® is one possible source. However, it contains citation data only for published papers and only for a subset of journals. The only source that provides citation data on both published and working papers is Google Scholar®. Using it, we collected citation data for 2031 articles. For the rest of articles, Google Scholar® either failed to find any information on the paper or we were not able to identify the match. Articles written in earlier years are likely to have more citations. To provide comparability, we calculate the number of citations per year.

3 A Two Stage Procedure?

Here we first look at the determinants of acceptance. Then we see how well the model can predict acceptance. We run a probit model both with and without controlling for selection bias. Table 10, given in the Appendix, summarizes the main regression findings.

3.1 The Determinants of Acceptance

We have a number of variables in our regression. All of these are ex-ante variables as they are observable at the time of submission.

Vintage

The first block of coefficients in Table 10 corresponds to the results for Ph.D. vintage. Submission of a paper to the JIE is a choice made by the authors. If all authors, irrespective of their vintage, submit the same quality papers to the JIE, there should be no significant coefficients here. However, if a looming tenure decision makes an early acceptance at the JIE more valuable than a slower acceptance at a higher ranked journal, we may see a positive coefficient for the close to tenure years: i.e., a tenure effect in submission choice. Also, if tenure is a way off, even if the chances of acceptance at the JIE are small, low quality submissions may be worth making. Both of these are consistent with acceptance rates increasing with vintage in the early years as seen in the regressions. Moreover, controlling for quality, at least partially, in the full model by using the citation data (column 6 of Table 10) reduces the absolute value of the coefficient as might be expected. Our findings differ somewhat from the results of Hamermesh and Oster (1998). Using a random sample of submissions to one of major economic journals in 1991, they argued that "on average there is no decline with age in acceptance rate of papers submitted", after controlling for the author's quality and experience.

Experience

Good prior publications reveal the ability to write good papers. Thus, it is not surprising that having publications in journals as good as or better than the JIE tends to raise the probability of acceptance at the JIE. The number of publications in other journals (not group 1, 2, or the JIE) has a negative impact on the probability of being published. If this is because writing bad papers makes one more likely to keep on doing so, then this effect should be stronger for people who have been out for a while. To check this, we look for the evidence of a differential impact on the probability of an article's acceptance depending on vintage. The interaction of Ph.D. vintage and the number of publications not in top journals turns out to be significantly negative. The size of the coefficients increases with age, consistent with the above argument.[20]

Language

The language dummy is significant and positive in all specifications we have examined. This highlights the importance of good writing for publication.

Co-editor fixed effects

Co-editors vary substantially in terms of the raw acceptance rates from 10 to 35% (15-54% for the sample). However, this should not be taken as an indicator of bias as it could easily be the case that the quality composition of papers varies across editors. For this reason, we allow for co-editor fixed effects in

[20] Since the JIE has a single blind system of refereeing, it could be that referees take publications in lower ranked outlets as a signal of poor quality, or it could be that such submissions actually tend to be of lower quality.

our regression, and when considering whether there are differences in standards as done in Section 4 below, look at these rather than the raw probabilities. As shown in Table 3, these dummies are significant for a number of co-editors.

Table 3: Co-editor fixed effects

co-editor	% Accepted (Sample)	Probit marginal effect[21]	% Accepted - marginal effect	Citations per year Accepted	Citations per year Rejected	Time to first decision
1	0.33	0.060	0.27	4.6	1.9	187
2	0.28	-0.073	0.35	6.8	1.5	115
3	0.48	0.148**	0.34	6.1	1.4	124
4	0.27	-0.072	0.34	2.9	1.4	127
5	0.40	0.221***	0.18	4.7	1.1	156
6	0.37	—	0.37	9.8	2.5	80
7	0.39	0.081	0.31	3.1	1.6	166
8	0.54	0.299***	0.24	5.7	0.5	218
9	0.22	-0.027	0.25	8.3	2.2	103
10	0.27	0.043	0.23	3.1	0.8	191
11	0.37	0.082	0.29	6.6	0.5	101
12	0.31	0.024	0.29	11.4	1.2	192
13	0.22	0.038	0.18	2.7	2.0	128
14	0.20	0.037	0.16	2.8	1.3	123
15	0.36	0.085	0.28	8.6	1.7	117
16	0.29	0.06	0.23	5.5	2.6	136
17	0.24	0.055	0.18	3.1	0.6	107
18	0.15	-0.094	0.25	9.0	3.6	187
19	0.18	-0.089	0.26	13.6	4.2	180
20	0.41	0.229*	0.18	5.7	1.2	122
21	0.50	0.078	0.42	7.8	1.3	128

Marginal effects are reported for regression estimates. It measures a change in a probability if dummy variable changes from zero to one. Co-editor 6 dummy is omitted to avoid collinearity. *,**,*** denote significance at 10, 5, and 1 percent level, respectively.

[21] The discrete change in the probability for change of dummy variable from 0 to 1 is reported. Evaluated at the means.

University rank

The distribution of submissions by the world-university ranking according to the author's Ph.D. is given below. Note that the share of the top 20 universities constitutes a lion's share of submissions and the share tapers off quite rapidly.[22] Table 4 summarizes the main facts.

Table 4: Submissions and Acceptances by Graduate School Quality Cohorts

	Submissions #	Submissions %	Accepted	Accepted / submissions
Top 20	1186	57%	374	31.5%
Top 50	1436	70%	444	30.9%
Top 200	1851	90%	513	27.7%
Sample Available	2051	100%	519	25.3%
"Population"	3032	—	600	19.8%

Our estimates (in all columns of table 10) show that people who graduated from ranked places (top 30) are more likely to have their papers accepted. At the same time, coefficients for the top 10 - 30 places are about the same, while graduates from the universities that are unranked have a lower probability of having their papers accepted, other things constant.

What might account for the relative stability of acceptance rates across schools? A simple explanation comes from the noting that submission decisions are endogenous. Authors choose where to submit on the basis of expected payoffs. An increase in the payoff from a JIE publication, or a higher subjective probability of acceptance, given quality, would tend to make an author more willing to try his luck with a lower quality paper and so end up with a lower probability of acceptance. It is likely that the probability of acceptance is overestimated at lower ranked schools (as the acceptance rate at the JIE has been falling which may be less well known at lower ranked school), while the value of a JIE publication is higher for them (at lower ranked schools a JIE article would count towards tenure while it would probably not make much of a difference at a highly ranked one). This may well explain the slightly lower acceptance rates at lower ranked institutions. Another explanation could be the desire to get feedback on a paper, even if it has little chance of acceptance. This factor could be important for faculty at lower ranked institutions.

Interactions in university rank and Ph.D. vintage

To see whether university quality might have different effects with Ph.D. vintage, we include interaction terms between university rankings and Ph.D. vintage. Table 5 summarizes these coefficient estimates. These estimates clearly show that graduates from better places are more likely to have a better start. The probability of acceptance falls with Ph.D. vintage. It falls at about the same rate for everyone so that initial

[22] Acceptance rates at lower ranked departments are very volatile due to the small number of submissions. In fact, for some institutions the acceptance rate is 100% due to a single paper being submitted.

differences seem to persist. Though the number of submissions falls with Ph.D. vintage, Table 6 clearly shows that the structure of submissions remains roughly the same as we vary the Ph.D. vintage.

Table 5: Divergence versus Convergence

Rank of the university	Constant term	Slope of intersection with Ph.D. vintage
Top 10 universities	0.816	−0.046
	(0.139)***	(0.012)***
Top 20, excluding Top 10	0.663	−0.016
	(0.143)***	(0.012)
Top 30, excluding Top 20	0.304	0.007
	(0.208)	(0.018)
Top 40, excluding Top 30	0.517	−0.048
	(0.239)**	(0.020)**
Top 50, excluding Top 40	0.492	−0.038
	(0.289)*	(0.029)
Top 100, excluding Top 50	0.489	−0.034
	(0.167)**	(0.016)**
Not one of the 200 best universities	−0.053	−0.005
	(0.119)	(0.010)

In other words, even though those who are 20 years out submit far fewer papers than those who are 5 years out, the distribution across the universities they graduated from remains stable.

Table 6: Shares of Papers' submissions with Ph.D. vintage for various education quality cohorts

University	Ph.D. vintage				
Rank	0 to 5 years	5 to 10 years	10 to 15 years	15 to 20 years	more than 20
Top 10	33%	34%	33%	32%	33%
Top 10 - 20	24%	20%	21%	21%	16%
Top 20-30	7%	8%	7%	7%	9%
Top 30-40	4%	8%	10%	9%	10%
Top 40-50	6%	5%	7%	7%	9%
Top 50-100	14%	14%	12%	15%	17%
Top 100-200	12%	11%	10%	9%	7%
# of Papers	610	563	363	219	197

Affiliation matters

Geography and employer type matter when citations are not accounted for. However, once citations are included, they become insignificant. In other words, the probability of acceptance is higher for scholars employed in research organizations, but this seems to be due to the higher quality of papers submitted by them. This could be explained by a JIE publication being more valuable to them in their careers.

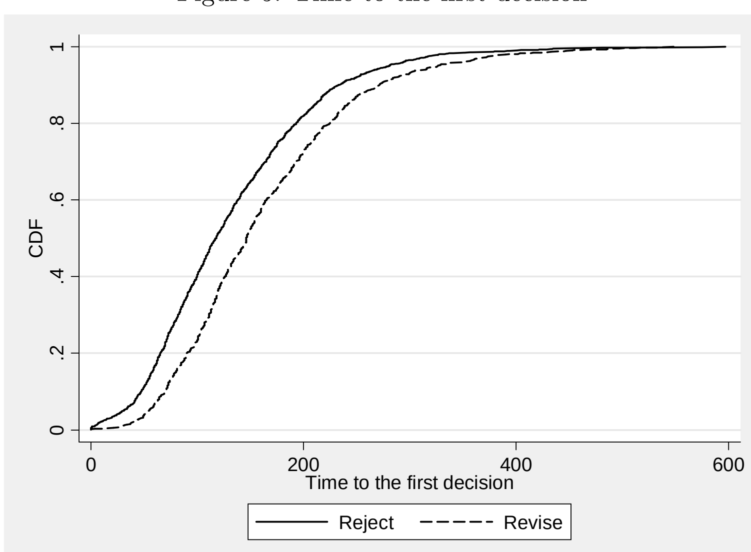
Backlog

The backlog variable turns out to be significant at the 10% level and has a negative coefficient[23].

3.2 Time to First Decision and Survival Probabilities

What does the time to the first decision say about the probability of rejection? Most accepted papers go through at least one revision. Out of about 3,000 submissions, only 17 (or about 0.6%) were accepted with no revision. About 770 (23%) articles were sent for revision and about 600 (78%) of them were finally accepted. Trivedi (1993) hypothesizes that the processing delays for rejected and accepted papers should be of the same order. However, this is not so in our data. The plot below shows the cumulative distribution functions of waiting times for these two groups. The cumulative distribution for rejected papers clearly lies above that for accepted ones, so that the latter FOSD the former.

Figure 5: Time to the first decision



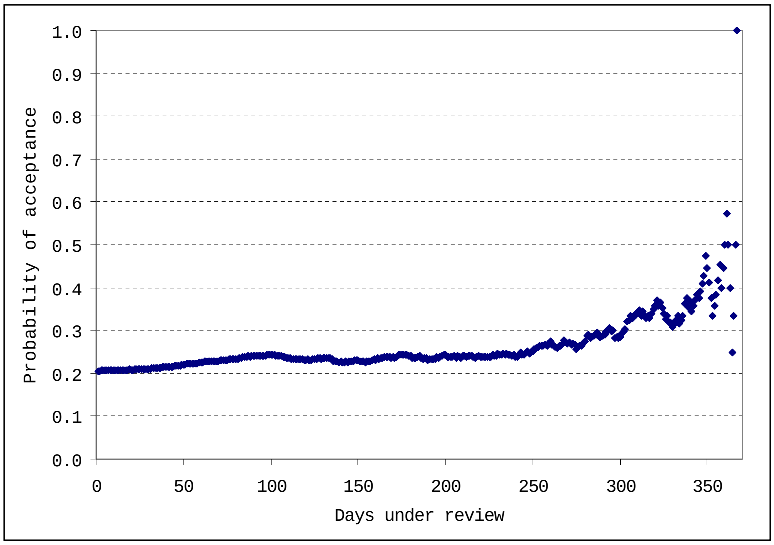
The diagram[24] also suggests that "no news is good news"! This makes sense as it is likely to take more

[23] Detailed estimation results are available on the request.

[24] Both the Anderson (1996) First Order Stochastic Dominance and the Kolmogorov-Smirnov tests support our findings, see

time to review a paper that might be acceptable than to reject a clearly bad one. On average, it takes about 132 days to process a paper that will be rejected and 162 days to handle an article that has to be revised. Papers that were accepted during the first round, i.e., without revision, on average spent 130 days under reviewing, but with a very high standard deviation of 115 days. Below (Figure 6) we plot the probability of acceptance, conditional on survival. That is, the probability that an article will not be rejected given that it has survived for X days from submission. This is clearly increasing.

Figure 6: Acceptance Conditional on Survival



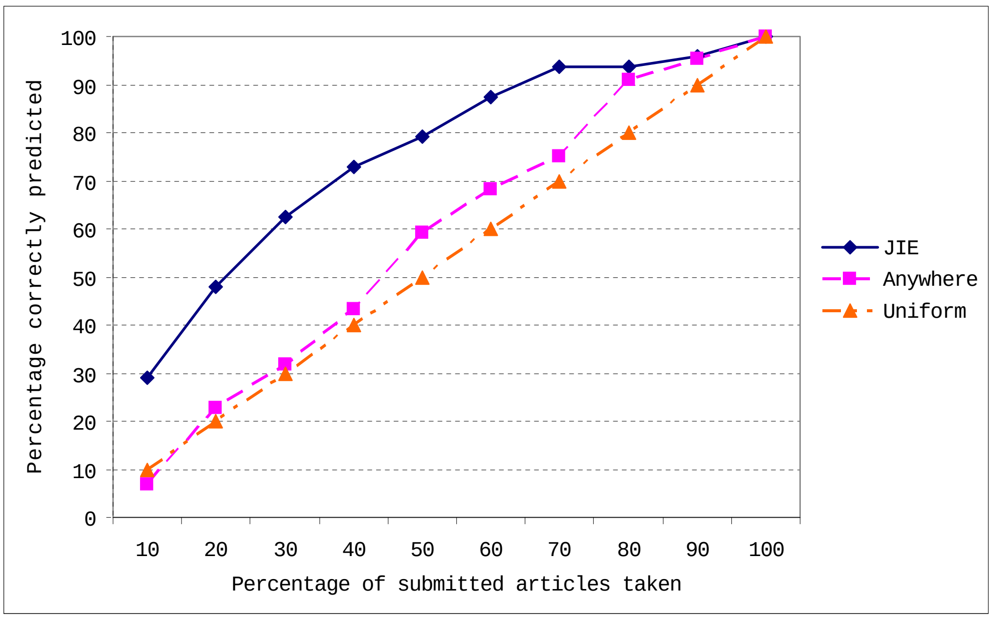
3.3 Streamlining and its Costs

Finally, we ask, how well does the model do in predicting final acceptance? Suppose the journal rejected the papers with the lowest probability of acceptance according to the model *without looking at the paper itself.* How badly would it err? Of course, editors could do better by taking a quick look at the contents, but even without this, how well does our regression perform? Using only those variables that are available ex-ante, and omitting co-editor dummies, we estimate a version of our model. Then we take the predicted probability of being published for each paper submitted to the JIE in 2000 and sort the articles in descending order. In other words, the articles with the highest probability of acceptance are at the top. We ask, if we had taken the top k% of papers, after ordering papers in terms of their predicted probability of acceptance, and only sent these out for refereeing, rejecting the others, what fraction of papers would be wrongfully rejected? This is depicted in Figure 7 below. When k=80%, this number is 5%. Hence, without reading the papers,

Appendix, Table 9, Column 2.

eliminating 20% of the submitted manuscripts will result in at most a 5% wrongful rejection rate relative to the current procedure. Note that this is without any information about the paper itself! Without reading the papers, eliminating 25% of the submitted manuscripts will result in at most a 7.5% wrongful rejection rate relative to the current procedure. At the same time, if we only take the articles not accepted in the JIE, and we take 75% of the papers in this set with the highest predicted probability of being accepted, we will predict only 79.5% of publications other than at the JIE correctly.

Figure 7: Prediction Quality



Information on the co-editor handling the paper improves the model fit. Including this in the regression helps to exactly predict the outcome at $k = 90\%$. When we look at single authored papers only, we do even better. At $k = 70\%$, the error is zero! In other words, none of the worst 30% of papers as classified by our model were actually accepted and half of the best papers account for approximately 87% of articles accepted for publication. The quality of prediction on single-authored papers is far better than that for the whole sample, suggesting that coauthorship is befuddling the model.[25]

[25]Using the best author's characteristics does not improve the fit.

4 Evaluating Performance

4.1 Type 1 versus Type 2 Errors

Out of 3032 papers submitted to JIE, 600 were accepted. Of the 2432 remaining articles, 564 were published elsewhere. About 14% of these were published in journals ranked by Kalaitzidakis et. al.(2003) above the JIE.

Table 7: Final publications for papers rejected by the JIE

Rank of journal	Share (%)
Top 1-10 journal	1.6
Top 11-20 journal	8.3
Top 21-30 journal (Excluding JIE)	4.1
Top 31-40 journal	6.4
Top 41-50 journal	9.0
Top 50-100 journal	9.2
Other ranked journals	4.4
Non-ranked journals	56.9

If we take publication in a journal ranked above the JIE as evidence of a mistake, then type 1 error is about 14%. However, this would be an over estimate if most such papers were rejected by the JIE for not being a good fit, not because they were below the quality threshold. It would be an under-estimate if rejection by the JIE discourages authors from submitting elsewhere.

Table 8: Average citations per year for different groups of papers

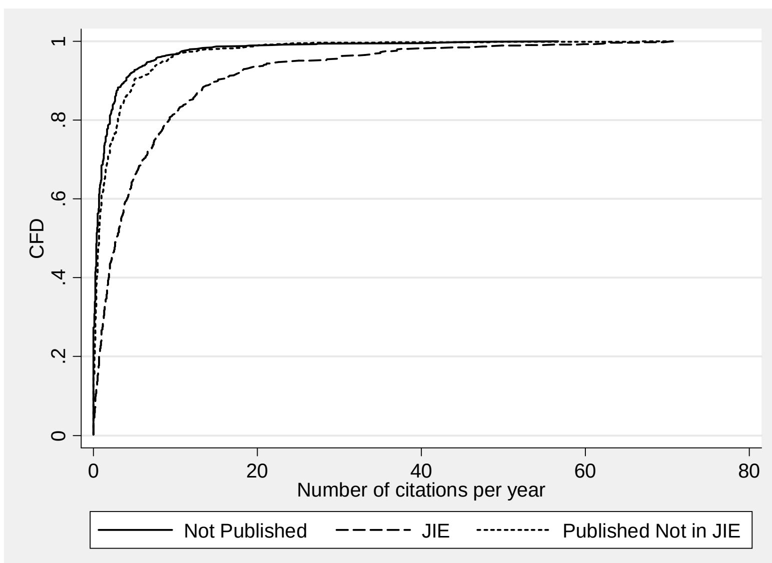
	Citations per year	*Max.citations per year*	*Citations per year*	*Max.citations per year*
	1995 - 2004	1995 - 2004	1995 - 06/2001	1995 - 06/2001
Published in JIE	6.33	70.75	6.30	69.54
Rejected by JIE, but published in a higher ranked journal[26]	3.71	71.00	3.20	38.83
Rejected by JIE, but published anywhere else	2.17	71.00	1.78	21.00
Rejected by JIE, and **not** published anywhere else	1.67	56.67	0.96	11.71

[26] As ranked by Kalaitzidakis et al.(2003)

Table 8 shows that the average number of citations per year varies significantly across groups of articles. For papers published in the JIE, the average number of citations per year is about 6.33. For papers rejected by the JIE, but published elsewhere, it is 2.17. For papers published in better journals, the average number of citations per year is half that of those published in the JIE. It could be argued that it takes more time for articles to be accepted in another journal after being rejected by the JIE and this explains why their citations are lower. However, most papers are available on-line until published, and thus, accumulate citations before publication. Moreover, even for the pre 2001 period, papers published in journals better than the JIE are cited less often, suggesting that type 1 error is not so large.

Figure 8 plots cumulative distribution function of the number of citations per year for three groups of papers. The citations for the papers published in the JIE FOSD both the others, which cannot be distinguished from each other.[27]

Figure 8: Number of Citations per Year (CDF)



4.2 Acceptance, Rejection and Quality

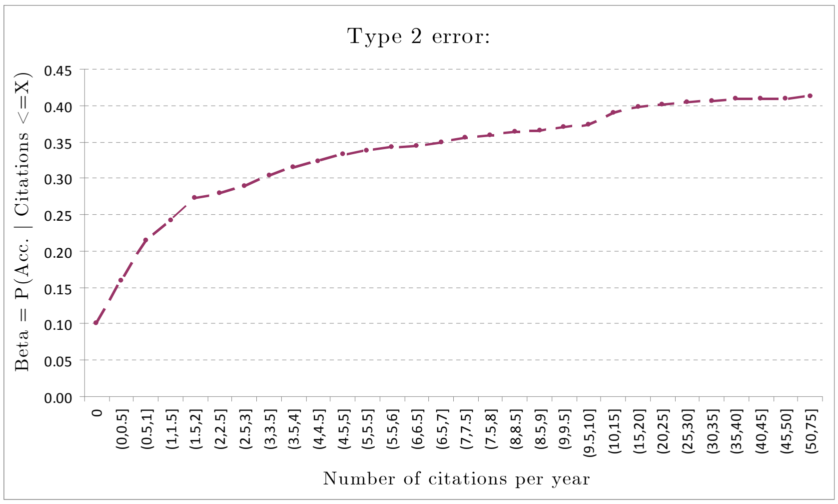
Next, we look at the relation between quality and outcomes. We do so by considering the distribution of citations for accepted papers versus rejected ones. The probability of acceptance equals that of rejection when citations are about 2.25. Thus, articles with 2.25 cites are equally likely to have been accepted or rejected. Similarly, articles with 0.5-1 cites per year – are 4 times more likely to have been rejected than accepted. In other words, they have an odds ratio (A/R) of .25. Hence, the probability that an accepted

[27]Results of the Anderson (1996) test and of the Kolmogorov-Smirnov test are reported in Appendix, Table 10.

paper with this number of cites has a $\frac{.25}{.75}$ or 1/3 chance of being wrongfully accepted. If, as might be expected, published papers are cited more than unpublished ones, this number would tend to under-estimate type 2 error.

From these distributions we derive the conditional probability of acceptance and rejection (which add to unity) given that a paper has x citations or less. Figure 9 depicts the fraction of papers that are accepted but which have x citations or less. If we say that articles with a citation below x were wrongfully accepted, then even with $x = 0$, we have roughly 10% wrongfully accepted articles. This suggests that type 2 error is quite high.

Figure 9: Probability of accepting paper with X or less citations



4.3 Editorial Heterogeneity

By adding dummies for co-editors, in conjunction with the author characteristics, we can capture fixed effects associated with co-editors.[28] Suppose, for example, that more interesting articles are retained by those assigning papers to co-editors. This would lead to differences in raw acceptance rates that have nothing to do with differences in standards. However, by controlling for author characteristics, we could control for such composition biases, at least to the best of our ability, so that the fixed effects capture differences in acceptance rates that are not due to quality differences.

We find, as depicted in Table 3, column 3, that a number of co-editor dummies included in our regressions turns out to be significant in all the specifications. The most generous co-editor is 29 raw percentage points more likely to accept a paper than the least generous one! Moreover, the most lenient editor has a marginal

[28] Such effects can exist if co-editors have different views on the minimum acceptable quality of a paper and would make the outcome more random and decisions less uniform.

effect of .299- with raw acceptance rate of 54%. Thus, his acceptance rate is 30 percentage points higher than it would have been if co-editor 6 had handled it. However, this need not reflect differences in cutoff quality. If such co-editors put in a lot more work in improving a paper than do those with a low acceptance rate, their cutoffs could be similar, while their marginal effects could be positive. If this were the explanation, then the average citations for these co-editors should not differ too much from the average. However, as can be seen in Table 3, the average citation for papers accepted *and* rejected by each of these co-editors is below the average for the JIE! This suggests that their cutoffs are lower and points to differences in standards between co-editors.[29]

5 Conclusion

A better understanding of the way journals operate can help authors, editors, referees and the journal all together. It can help authors understand how well or badly a journal performs, identify strengths and weaknesses (where there may be room for improvement at low cost) to those operating the journal, and provide editors and, in the future maybe referees, feedback on their performance. Such evaluations could also shed light on the biases, if any, inherent in the existing system.[30]

Our main conclusions are that the JIE seems to be doing a good job in identifying quality. However, there is room for improvement. First, a two tier evaluation procedure would likely reduce the burden on all concerned at little or no cost in terms of performance. Second, the preliminary evidence suggests there is a difference in standards and performance across editors. This might be reduced by providing feedback to editors on their relative performance. In the future, with electronic data bases being kept by journals, such feedback could also be provided to individual referees. Editors could also be provided with more information on the performance of referees. Third, the evidence suggests that while type 1 error is relatively small (rejected papers are clearly less well cited no matter what their final fate), type 2 error is large (many accepted papers are poorly cited).

[29] However, editors do handle papers from very different sub fields and we cannot control for this in our data.

[30] For example, the existence of significant unexplained differences between editors in terms of acceptance rates that are negatively correlated with citation rates might serve as a flag for the presence of heterogenous standards.

References

[1] Anderson, Gordon, 1996. "Nonparametric Tests of Stochastic Dominance in Income Distributions," *Econometrica* 64(5), pp: 1183-1193.

[2] Blank, Rebecca M., 1991. "The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from the American Economic Review," *The American Economic Review* 81(5), pp: 1041-1067.

[3] Bowen, Robert M. and Gary L. Sundem, 1982. "Editorial and Publication Lags in the Accounting and Finance Literature," *The Accounting Review* 57(4), pp: 778-784.

[4] Coe, Robert K., and Irwin Weinstock, 1967. "Editorial Policies of Major Economic Journals", *Quarterly Review of Economics and Business* 7(4), pp: 37-43.

[5] Ellison, Glenn, 2002a. "The Slowdown of the Economics Publishing Process," *The Journal of Political Economy* 110(5), pp: 947-993.

[6] Ellison, Glenn, 2002b. "Evolving Standards for Academic Publishing: A q-r Theory," *The Journal of Political Economy* 110(5), pp: 994-1034.

[7] Hamermesh, Daniel S. and Sharon S. Oster, 1998. "Aging and Productivity among Economists," *The Review of Economics and Statistics* 80(1), pp: 154-156.

[8] Jenkins, Stephen P., Lorenzo Cappellari, Peter Lynn, Annette Jackle and Emanuela Sala, 2004. "Patterns of Consent: Evidence From a General Household Survey," ISER Working Paper #2004-24.

[9] Kalaitzidakis, Pantelis, Theofanis P. Mamuneas and Thanasis Stengos, 2003. "Ranking of Academic Journals and Institutions in Economics," *Journal of the European Economic Association* 1(3), pp: 1346-1366.

[10] Laband, David N. and Michael J. Piette, 1994. "Favoritism versus Search for Good Papers: Empirical Evidence Regarding the Behavior of Journal Editors," *The Journal of Political Economy* 102(1), pp: 193-204.

[11] Trivedi, Pravin K., 1993. "An Analysis of Publication Lags in Econometrics," *Journal of Applied Econometrics* 8(1), pp: 93-100.

[12] Weiss, Andrew A., 1993. "A Bivariate Ordered Probit Model With Truncation: Helmet Use and Motorcycle Injuries," *Applied Statistics* 42(3), pp: 487-499.

[13] Yohe, Cary W., 1980. "Current Publication Lags in Economics Journals," *Journal of Economic Literature* 18(3), pp: 1050-1055.

6 Appendix

6.1 The Likelihood Function

The errors in the choice and truncation equations are assumed to be jointly normally distributed with a variance-covariance matrix:

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim N\left(\left(\begin{array}{c} 0 \\ 0 \end{array}\right), \left(\begin{array}{cc} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{array}\right)\right), \tag{3}$$

with $\rho \neq 0$. If $\rho = 0$, the estimates of the choice equation would coincide with those of the probit model and the truncation equation is non-identifiable.

Denote the joint normal distribution of $(\varepsilon_{1i}, \varepsilon_{2i})$ by $G(\varepsilon_{1i}, \varepsilon_{2i})$, its density function by $g(\varepsilon_{1i}, \varepsilon_{2i})$, and marginal density functions by $g_1(\varepsilon_{1i})$, $g_2(\varepsilon_{2i})$ and $G_1(\varepsilon_{1i})$, $G_2(\varepsilon_{2i})$, correspondingly. One could also note that as in the usual probit model, σ_1 and σ_2 cannot be identified separately from β and γ and without loss of generality can be normalized to 1. Therefore, $G(\varepsilon_{1i}, \varepsilon_{2i})$ has a variance-covariance matrix:

$$\Sigma = \left(\begin{array}{cc} 1 & \rho \\ \rho & 1 \end{array}\right).$$

Beyond this point we use ϕ and Φ to refer to the density and cumulative distribution of the standard normal distributions. The probability that the article is accepted for publication given that we observe the author's vita is:

$$Prob[X_i\beta + \varepsilon_{1i} > 0 | Z_i\gamma + \varepsilon_{2i} > 0] = \frac{Prob[\varepsilon_{1i} > -X_i\beta \quad \cap \quad \varepsilon_{2i} > -Z_i\gamma]}{Prob[\varepsilon_{2i} > -Z_i\gamma]} \tag{4}$$

$$= \frac{\int_{-Z_i\gamma}^{\infty}\int_{-X_i\beta}^{\infty} g(E_1, E_2)dE_1dE_2}{\int_{-Z_i\gamma}^{\infty} g_2(E_2)dE_2}. \tag{5}$$

The probability that the article is rejected given that we observe the vita is unity less the above expression. In addition, as

$$\frac{\int_{-Z_i\gamma}^{\infty}\int_{-X_i\beta}^{\infty} g(E_1, E_2)dE_1dE_2}{\int_{-Z_i\gamma}^{\infty} g_2(E_2)dE_2} = \frac{1 - \Phi(-X_i\beta) - \Phi(-Z_i\gamma) + G(-X_i\beta, -Z_i\gamma)}{1 - \Phi(-Z_i\gamma)},$$

the likelihood function, $L(Y_1, .., Y_N,\ X_1, .., X_N, Z_1, .., Z_N)$, can be written as:

$$\prod_{i=1}^{N}\left[\frac{1 - \Phi(-X_i\beta) - \Phi(-Z_i\gamma) + G(-X_i\beta, -Z_i\gamma)}{1 - \Phi(-Z_i\gamma)}\right]^{Y_i}\left[\frac{\Phi(-X_i\beta) - G(-X_i\beta, -Z_i\gamma)}{1 - \Phi(-Z_i^*\gamma)}\right]^{1-Y_i} \tag{6}$$

$$= \prod_{i=1}^{N} \left[1 - \frac{\Phi(-X\beta) - G(-X\beta, -Z\gamma)}{1 - \Phi(-Z\gamma)}\right]^{Y_i} \left[\frac{\Phi(-X\beta) - G(-X\beta, -Z\gamma)}{1 - \Phi(-Z\gamma)}\right]^{1-Y_i}$$

$$= \prod_{i=1}^{N} [1 - g(\beta, \gamma, \rho)]^{Y_i} [g(\beta, \gamma, \rho)]^{1-Y_i},$$

where

$$g(\beta, \gamma, \rho) = \frac{\Phi(-X\beta) - G(-X\beta, -Z\gamma)}{\Phi(Z\gamma)}. \tag{7}$$

Taking logarithms, we obtain the following expression:

$$log(L) = \sum_{i=1}^{N} Y_i \log(1 - g(\beta, \gamma, \rho)) + \sum_{i=1}^{N} (1 - Y_i) \log(g(\beta, \gamma, \rho)). \tag{8}$$

Jacobian for this problem could be written as:

$$\frac{dlog(L)}{dT} = \sum_{i=1}^{N} \frac{dg}{dT} \left[Y_i \frac{-1}{1 - g(\beta, \gamma, \rho)} + (1 - Y_i) \frac{1}{g(\beta, \gamma, \rho)}\right], \tag{9}$$

where $T = \{\beta, \gamma, \rho\}$. That implies that Hessian is:

$$\begin{aligned} \frac{d^2 \log(L)}{dTdT'} \quad &= \quad \sum_{i}^{N} \frac{d^2 g}{dTdT'} \left[Y_i \frac{-1}{1 - g(\beta, \gamma, \rho)} + (1 - Y_i) \frac{1}{g(\beta, \gamma, \rho)}\right] \\ &\quad - \left(\frac{dg}{dT}\right)' \left[Y_i \frac{1}{(1 - g(\beta, \gamma, \rho))^2} + (1 - Y_i) \frac{1}{(g(\beta, \gamma, \rho))^2}\right] \left(\frac{dg}{dT}\right). \end{aligned}$$

Thus, we have to find $\frac{dg}{dT}$ and $\frac{d^2 \log(L)}{dTdT'}$. It is not mandatory to calculate these derivatives, however, providing them analytically significantly improves the performance of optimization routines.

6.2 Kolmogorov-Smirnov and Anderson tests: Detailed results

Table 9: FOSD (Anderson) test of time to the first decision

First Order Stochastic dominance test (**Anderson test**)				
Distribution A	Papers to be Revised	All papers Before 2000	Papers to be Revised Before 2000	Rejected Papers Before 2000
Distribution B	Rejected papers	All papers 2000 and after	Papers to be Revised 2000 and after	Rejected Papers 2000 and after
Decile 1	-8.04***	-1.70*	6.68***	-4.29***
Decile 2	-12.82***	0.69	12.13***	-2.12
Decile 3	-15.00***	5.15***	14.17***	1.61
Decile 4	-16.58***	6.30***	15.01***	3.29*
Decile 5	-14.52***	7.38***	13.47***	4.48**
Decile 6	-12.89***	8.08***	12.12***	7.34***
Decile 7	-11.37***	6.07***	7.65**	5.07***
Decile 8	-8.85***	4.10***	6.70**	4.75***
Decile 9	-5.80***	1.76*	2.56	1.58
Decile 10	0.00	0.00	0.00	0.00
PAT($\chi^2_{(9)}$)	85.74***	34.45***	23.12***	34.42***
Conclusion:	FOSD	Mixed Result	FOSD	Mixed Result
Note *, **,*** denote significance at 10, 5, and 1 percent level, respectively.				

Kolmogorov - Smirnov Test

H_0 : The two samples come from a common distribution				
P-value for K-S	0.000	0.000	0.000	0.005
H_0 : $F_A(X) > F_B(X)$, where F stands for CDF				
P-value	0.000	0.480	0.964	0.072
H_0 : $F_A(X) < F_B(X)$, where F stands for CDF				
P-value	0.998	0.000	0.000	0.000

6.3 Estimation Results

Table 10: Model estimation results

Estimated Model	Probit model	Probit model	Probit model	Probit model	2-equat. specif-n	2-equat. specif-n
Statistics reported	*Marginal effect*[31]	*Marginal effect*	*Marginal effect*	*Coef. estim.*	*Coef. estim.*	*Coef. estim.*
Column #	1	2	3	4	5	6
	Ph.D. vintage variables (years)					
Not Graduated Yet	0.25	0.212	0.041	0.57	-1.307	-1.289
	[3.29]***	[2.51]**	[0.45]	[2.51]**	[3.27]***	[2.48]**
Ph.D. vintage: (0, 2]	0.317	0.247	0.089	0.66	-0.960	-0.956
	[4.26]***	[3.01]***	[1.01]	[3.01]***	[2.67]***	[2.04]**
Ph.D. vintage: (2, 4]	0.403	0.35	0.211	0.927	-0.622	-0.558
	[5.46]***	[4.29]***	[2.36]**	[4.29]***	[1.87]*	[1.29]
Ph.D. vintage: (4, 6]	0.278	0.231	0.108	0.619	-0.672	-0.632
	[3.88]***	[2.92]***	[1.29]	[2.92]***	[2.20]**	[1.59]
Ph.D. vintage: (6, 10]	0.251	0.187	0.089	0.51	-0.524	-0.480
	[3.83]***	[2.57]**	[1.19]	[2.57]**	[1.95]*	[1.37]
Ph.D. vintage: (10, 20]	0.144	0.065	0.004	0.186	-0.248	-0.271
	[2.57]**	[1.05]	[0.07]	[1.05]	[1.21]	[1.01]
	Experience: number of publications in various journals					
# of articles: Group 1	0.022	0.016	0.022	0.048	0.071	0.055
	[3.46]***	[2.17]**	[2.83]***	[2.17]**	[5.06]***	[2.95]***
# of articles: Group 2	0.011	0.01	0.016	0.029	0.028	0.040
	[1.20]	[0.99]	[1.53]	[0.99]	[1.50]	[1.69]
# of articles in JIE	0.022	0.014	0.01	0.041	0.061	0.053
	[2.46]**	[1.35]	[0.91]	[1.35]	[3.09]***	[2.22]**
# of other publications	—	—	-0.005	—	—	—
	—	—	[3.14]***	—	—	—
	Language Effect					
Native speaker dummy	0.052	0.077	0.078	0.224	0.197	0.265
	[1.76]*	[2.22]**	[2.18]**	[2.22]**	[3.05]***	[3.27]***

[31] Marginal effect is the change in the probability for an infinitesimal change in each independent, continuous variable. For

Table 10: Model estimation results (Continued)

Estimated Model	Probit model	Probit model	Probit model	Probit model	2-equat. specif-n	2-equat. specif-n
Statistics reported	*Marginal effect*	*Marginal effect*	*Marginal effect*	*Coef. estim.*	*Coef. estim.*	*Coef. estim.*
Column #	1	2	3	4	5	6
	University ranking variables - Graduation Place (Continued)					
Grad. from top 10	0.158	0.156	0.143	0.445	0.333	0.349
	[4.95]***	[4.10]***	[3.76]***	[4.10]***	[5.02]***	[4.24]***
Grad. from top 10 - 20	0.175	0.208	0.194	0.574	0.375	0.450
	[5.25]***	[5.35]***	[4.96]***	[5.35]***	[5.45]***	[5.37]***
Grad. from top 20 - 30	0.144	0.189	0.16	0.506	0.287	0.398
	[3.03]***	[3.35]***	[2.82]***	[3.35]***	[3.10]***	[3.17]***
Grad. from top 30 - 40	0.034	0.069	0.056	0.193	0.117	0.192
	[0.71]	[1.18]	[0.96]	[1.18]	[1.48]	[1.46]
Grad. from top 40 - 50	0.07	0.097	0.079	0.268	0.158	0.235
	[1.42]	[1.71]*	[1.41]	[1.71]*	[1.64]	[1.92]*
Grad. from top 50 -100	0.073	0.101	0.083	0.281	0.123	0.159
	[1.99]**	[2.30]**	[1.89]*	[2.30]**	[1.64]	[1.71]
Not ranked university	-0.031	-0.021	-0.009	-0.06	-0.052	-0.026
	[1.22]	[0.68]	[0.28]	[0.68]	[0.95]	[0.38]
	Institution affiliation variables					
Affiliated with US univ.	0.087	0	0.001	0	—	—
	[2.69]***	[0.00]	[0.02]	[0.00]	—	—
Affiliated with CA univ.	0.098	0.062	0.074	0.175	—	—
	[1.87]*	[1.04]	[1.22]	[1.04]	—	—
Affiliated with UK univ.	0.057	-0.013	-0.008	-0.039	—	—
	[1.15]	[0.24]	[0.15]	[0.24]	—	—
Affiliated with EU univ	0.006	-0.065	-0.071	-0.195	—	—
	[0.17]	[1.63]	[1.79]*	[1.63]	—	—
Affiliated with organiz.	0.125	0.054	0.051	0.153	—	—
	[2.99]***	[1.16]	[1.12]	[1.16]	—	—

dummy variables marginal effect is the discrete change in the probability for change of dummy variable from 0 to 1. All marginal effects are evaluated at the means.

Table 10: Model estimation results (Continued)

Estimated Model	Probit model	Probit model	Probit model	Probit model	2-equat. specif-n	2-equat. specif-n
Statistics reported	*Marginal effect*	*Marginal effect*	*Marginal effect*	*Coef. estim.*	*Coef. estim.*	*Coef. estim.*
Column #	1	2	3	4	5	6
	Co-editors fixed effects (Continued)					
Co-editor 3	0.102	0.132	0.128	0.362	0.243	0.363
	[2.46]**	[2.65]***	[2.52]**	[2.65]***	[3.09]***	[3.48]***
Co-editor 4	—	—	—	—	0.005	—
	—	—	—	—	[0.04]	—
Co-editor 5	0.165	0.215	0.219	0.576	0.261	0.360
	[4.03]***	[4.60]***	[4.63]***	[4.60]***	[3.38]***	[3.74]***
Co-editor 8	0.316	0.297	0.245	0.776	0.563	0.661
	[3.68]***	[3.19]***	[2.63]***	[3.19]***	[3.68]***	[3.46]***
Co-editor 16	0.124	0.069	0.074	0.193	—	—
	[1.77]*	[0.88]	[0.93]	[0.88]	—	—
Co-editor 20	0.22	0.225	0.213	0.592	0.330	0.357
	[2.43]**	[2.19]**	[2.07]**	[2.19]**	[1.93]*	[1.62]
	Proxies for article quality					
Citations per year	—	0.023	0.022	0.068	—	0.049
	—	[9.56]***	[9.28]***	[9.56]***	—	[14.63]***
Time to first decision	—	—	0.0004	—	—	—
	—	—	[2.40]**	—	—	—
	Constant term					
Constant	—	—	—	-1.517	-2.232	-2.177
	—	—	—	[5.06]***	[8.06]***	[3.29]***

Table 10: Model estimation results (Continued)

Estimated Model	Probit model	Probit model	Probit model	Probit model	2-equat. specif-n	2-equat. specif-n
Statistics reported	*Marginal effect*	*Marginal effect*	*Marginal effect*	*Coef. estim.*	*Coef. estim.*	*Coef. estim.*
Column #	1	2	3	4	5	6
	First stage (truncation) equation					
PhD vintage	—	—	—	—	0.124	0.157
	—	—	—	—	[4.38]***	[3.74]***
Organiz. dummy	—	—	—	—	-0.199	-0.166
	—	—	—	—	[1.95]*	[1.13]
Constant term	—	—	—	—	-2.786	-2.734
	—	—	—	—	[4.48]***	[3.29]***
Correlation (ρ)	—	—	—	—	0.79	0.70
	—	—	—	—	[2046]***	[196]***
Pseudo R^2	0.13	0.20	0.21	0.2	—	—
Number of obs.	1661	1382	1367	1382	1641	1382

Estimates of year, months specific effects, and constant term in probit equation are omitted to save space and could be provided on request.

Robust Z statistics in parentheses for simple probit estimates, t-ratios for the full model with truncation equation. ***, **, * -significant at 1%, 5% and 10% correspondingly

Table 11: FOSD (Anderson) test of number of citations per year

First Order Stochastic dominance test (**Anderson test)**			
Distribution A	Papers published in JIE	Papers published in JIE	Papers published anywhere else
Distribution B	Papers Never Published	Papers Published anywhere else	Papers Never Published
Decile 1	-23.00***	-11.26***	-12.60***
Decile 2	-25.16***	-20.60***	-11.17***
Decile 3	-37.29***	-30.93***	-10.91***
Decile 4	-39.57***	-34.18***	-10.25***
Decile 5	-41.91***	-32.96***	-7.71***
Decile 6	-39.99***	-28.95***	-7.42***
Decile 7	-37.00***	-27.78***	-4.94**
Decile 8	-28.88***	-21.16***	0
Decile 9	-15.80***	-13.28***	—
Decile 10	0.00	0.00	—
PAT($\chi^2_{(9)}$)	306.23***	162.33***	32.18*** ($\chi^2_{(7)}$)
Conclusion:	FOSD	FOSD	FOSD
*, **, *** denote significance at 10, 5, and 1 percent level, respectively.			

Kolmogorov - Smirnov Test

H_0 : The two samples come from a common distribution			
P-value for K-S	0.000	0.000	0.621
H_0 : $F_A(X) > F_B(X)$, where F stands for CDF			
P-value	0.000	0.000	0.519
H_0 : $F_A(X) < F_B(X)$, where F stands for CDF			
P-value	0.999	1.000	0.346