The Inside Scoop: Acceptance and Rejection at the Journal of International Economics[1]

Ivan Cherkashin[2]
The Pennsylvania State University

Svetlana Demidova[3]
University of Georgia

Susumu Imai[4]
Queen's University

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

April 2008

[1]We are grateful to Abdullah Yavas and Bee Roberts for comments on an earlier draft, to Eva Drago, Anshul Kumar, Su-Jen Roberts and Supriya Mishra for able research assistance, to Sergey Lychagin for help with the PERL code for citation extraction and to Sara Herpolsheimer for journal data preparation. Krishna is grateful to Princeton University for support as a Kenen Research Fellow during 2006-2007. 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 these questions.

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

JEL: A14

I Introduction

Tenure, promotion, and salaries in academia are based on publications. These factors also affect an individual's market value 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 I.A provides a selective survey of the literature to date so as to place our work relative to this literature. Section II describes the data as well as some interesting patterns that occur. Sections III and IV make up the heart of the paper. Section III takes an ex-ante approach. The main question 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 of 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.

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

Section IV evaluates the performance of the JIE along various fronts. First, 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. Evidence on the extent of such errors comes from two sources: first, from looking at the ultimate fate of rejected papers; and second, from comparing the distributions of citations of different groups of papers. Roughly 14% of rejected papers end up in journals ranked above the JIE. However, 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 15% of published papers are cited less than once in two years, and 7% are never cited, suggesting that type 2 error could be large. 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 look for evidence on how good a measure of quality citations seem to be. 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 coefficient estimates of the remaining explanatory variables or make them less significant, which is exactly what we find!

Finally, we look at co-editor specific effects. We ask whether co-editors differ in their acceptance rates in ways that cannot be accounted for by the 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 V performs some robustness checks. We correct for a selection bias that arises in the simple probit model due to our ability use only data points, for which whom we can find a CV on the web. We show that even when we correct for this using maximum likelihood techniques, our results are by and large unaffected. Section VI outlines our policy conclusions as well as directions for future research.

A 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 an evidence of bias in acceptances.

Despite a proliferation of journals, there seems to have been a significant slowdown in the 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) find that this process took about 250 days in the 1960's. Yohe (1980) reports 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) uses data on 1134 submissions from 7 econometrics journals from 1986 to 1990. He finds 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 constructs delay distributions, but only has 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) obtain 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 each of 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) looks at the changes in submit-accept times for different journals and then examines possible causes for these patterns. He looks at several possible explanations for this increase, including democratization of the profession away from an "old boys" network, increased complexity of articles, and growth of the profession. To test the first hypothesis, Ellison collects data on the authors of the *published* papers only and 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 AER Papers and Proceeding or Brookings Papers, which are invited but prestigious. He finds no statistically significant relationship between the submit-accept time and the authors' standing in the profession. Note, 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 co-authors, and the degree of specialization as reflected in the JEL index. Since the 1970's, the average paper gained approximately 75% in size, while the share of co-authored papers doubled from 30% to 60% from 1970 to 1999 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 co-authors 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[2], (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.[3] However, this theory explains only about a quarter of the increase in the delay. Ellison's work remains the most extensive and up to date research on publication lags and their possible causes.

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![4] They speculate that editors seem to use their personal ties to obtain better papers for their journals.[5] Blank (1991) looks at the outcome of an experiment carried out by the AER as an indicator of

[2]The other aspects of quality include quality of math, econometrics, robustness checks, the level of polishing, etc.

[3]In our data, however, the correlation between citations and time to the first decision is not significantly different from zero 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 polishing to be acceptable.

[4]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 suspicious.

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

bias. During the experiment papers were randomly allocated 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, i.e., 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 of 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, using data on 208 faculty members of the leading 17 economic departments who got their degrees between 1959 and 1983. They find that researchers are most productive in the first decade after graduation and slow down over time. However, early high productivity seems to be a characteristic of those who remain productive many years later.[6] They also obtain a random sample of submissions to a top general interest journal. This data suggests that the probability of acceptance does not 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 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 of our unique data set.

[6] However, it is unclear 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 better research environment) it is easier for them to stay at the forefront of research.

II The Model and the Data

We assume that each article has a quality, q_i, which cannot be observed directly. The purpose of the editorial process is to identify q and accept article i if q_i exceeds a threshold level, Q. We distinguish between factors that can be observed at the time of submission and those that cannot: 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.

A 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, with untenured faculty putting in more effort and so being more likely to submit high quality papers, other things constant. They might also be closer to the frontier, especially if they come from good programs, than faculty who are not as 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. We assume that the article i is published in the journal if its latent quality q_i exceeds a threshold level:

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

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

B The Data

We have several sources of data.

B.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. 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.

Submission Data. We have data on submissions from 1995 to 2004. For each submission, we observe the authors' names[7], 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.[8] Overall, the JIE received about 3032 submissions of which almost 600 articles (20%) were accepted for publication. At the same time acceptance rates have almost halved from 27% to 14% in 1995-2004, despite

[7] If a paper is co-authored, we collect information on all co-authors which allows us to check whether results are sensetive to the choice of the author characteristics used in estimation (average, best, and worst).

[8] 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.

doubling in the journal 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.[9]

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 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.[10] The backlog could affect the decisions of the co-editors if information on it is conveyed to the co-editors, who, 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, possibly raising the probability of acceptance. In our regressions we use the backlog in the previous month as an explanatory variable.

---

[9] 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 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.

[10] The backlog variable we calculate (recursively and based on the fact that as of September 1, 2002, the backlog was 73 papers) is accurate for later years of the sample (1999-2004), but is less so for 1995-1998 due to not observing all acceptances in the earlier years. 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.

Time to First Decision. We have the date of the paper submission and the date of the first decision. The length of the sample 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.

There is a slight increase in the time to the first decision from 134 days in 1995-1999 to 142 days in 2000-2004[11]. 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? Comparison of 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. For rejected papers, the time to the first decision remains about the same for the whole period (130 vs. 134 days)[12]. For articles not rejected after the first round reviewing time increased quite noticeably from 152 to 172 days[13]. This confirms the hypothesis of Ellison (2002a) of an increase in the polishing component of quality, r, rather than q component.

B.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 the authors' background from their curriculum vitaes (CVs). In the main body of the paper we use a simple probit approach. As a robustness check, we also present estimation results, where we deal with the potential bias of omitting individuals for whom CVs were not available. It turns out that the difference in the results between these two specifications are minor.

Ph.D. Vintage. This variable is not the year the Ph.D. was awarded, but the number of years since getting the Ph.D. at the time of submission. It helps capture how human capital

[11] The Anderson (1996) test allows us to reject the FOSD hypothesis, while the Kolmogorov-Smirnov test rejects the null hypothesis that the distributions are the same. See Appendix, Table 7, Column 2.

[12] The Anderson (1996) test allows us to reject FOSD hypothesis. See Appendix, Table 7, Column 4.

[13] The Anderson (1996) test does not let us to reject FOSD hypothesis. See Appendix, Table 7, Column 3.

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 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 the probability of acceptance. Which of these effects dominates 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 0 to 2 years ago, from 2 to 4, from 4 to 6, from 6 to 10, from 10 to 20, or that the Ph.D. is not completed at the time of submission. Scholars who got their Ph.D.s more than 20 years ago are used as the reference group.[14]

University Rank. This gives the ranking of the university that awarded the author's Ph.D. We use world-wide rankings of the best 200 economic schools from Kalaitzidakis et. al. (2003). Schools, which are not on this list, are labeled "non-ranked".

Employer Type. We specify whether the author was employed at the US, Canadian, United Kingdom, or European university[15], or at a university anywhere else on the date of submission. For authors employed outside 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 the US.

An obvious trend is the decrease in the share of submissions from the authors affiliated with the US universities from 50% in 1995 to 37% in 2004 and a corresponding increase in such number for researchers from the European universities from 12% to 28%, suggesting

[14] We tried but failed to collect good data on a tenure status at the time of submission.
[15] We treated Norwegian and Swiss universities as the EU ones though neither country is the EU member.

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% average. Very few submissions come from business employees but this could be partly due to the absence of their CVs.

Number of Previous Publications. These are broken down into those in the leading general interest journals (*Group 1*: The American Economic Review, Econometrica, The Journal of Political Economy, The Quarterly Journal of Economics, and The Review of Economic Studies), the number of publications in the second tier general interest journals (*Group 2*: The Journal of Monetary Economics, The Review of Economics and Statistics, The Journal of Economic Theory; *Group 5*: The International Economic Review, The European Economic Review, and The Economic Journal), and the number of publications in top field journals and general interest journals with a record of publishing papers in International Economics (*Group 3*: The Journal of Pubic Economics, The Rand Journal of Economics, The Scandinavian Journal of Economics, Economic Letters, Journal of Applied Econometrics, The Journal of Development Economics, The International Journal of Industrial Organizations; *Group 4*: Economic Theory, Econometric Theory, Games and Economic Behavior, The Journal of Econometrics, The Journal of Human Resources, The Journal of Labor Economics, Journal of Economic Dynamic and Control, The Journal of Environmental Economics and Management). We also track the number of publications in the JIE prior to submission and the number of papers in "*Network*" journals, which demonstrate that a person is well linked in the profession as the papers tend to be solicited, even if they are refereed. This group includes: AER Papers and Proceedings, The Journal of Economic Perspectives, The Journal of Economic Literature, Rochester Series, Brookings Papers on Economic Activity. Finally, we have the total number of papers in economic journals.

Native Language. Quite a few papers submitted to the JIE are written by non-native speakers, for whom it might be harder to get 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. In such cases we define language proficiency by treating a person as proficient in English if he obtained his bachelors and subsequent degrees from a university located in an English speaking country.

B.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.

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

Citation Data. The number of citations can be an indicator of the paper quality. Of course, there are problems here as well. Citations can be negative rather than positive due to the paper susceptibilities! Also, published papers are more likely to be cited just because they are published. However, since many published papers have almost 0 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.[16]

[16] One might be concerned that earlier papers would tend to be more heavily cited due to the time taken for word to spread. However, the distributions of citations per year for papers during earlier and later periods do not differ significantly.

III A Two Stage Procedure?

Here we first look at the determinants of acceptance. Then we see how well the model can predict acceptance. Table 8 in the Appendix summarizes the main findings. We run a probit model, that is, we estimate equation (1), where the error term ε is assumed to be i.i.d. normal with variance 1.

In Column 1 we report the estimated marginal effects of the probit model, where X includes Ph.D. vintage, a number of publications, which we separate into those published in the JIE and those in "Network" journals. Other variables in X are: language, university ranking, co-editors fixed effects, institutional affiliation, and the year dummies. In Column 2 we add citations per year to X. In Column 3 we add the number of other articles per year to the number of publications variables. In Column 4 we present the coefficient estimates of the specification in Column 2. Before we begin, note that as we are not estimating a structural model, we cannot interpret the estimated parameters as clearly. For example, a positive sign on prior publications in Group 1 journals can indicate that such people send better papers to the JIE or that the editor favors them. However, by looking at other correlations, one can sometimes argue for one interpretation over another.

A 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.

Ph.D. Vintage. The first block of coefficients in Table 8 corresponds to the results for Ph.D. vintage. Submission to the JIE is the authors' choice. 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. What we see in the probit equation results is that the acceptance probability increases with vintage up to Ph.D. vintage 2 to 4. Thereafter, the probit marginal effects decrease. Our findings are somewhat different to those by Hamermesh and Oster (1998), who use a random sample of submissions to one of major economic journals in 1991 and argue that "on average there is no decline with age in acceptance rate of papers submitted", after controlling for the author's quality and experience.

Co-editor Fixed Effects. Co-editors vary substantially in terms of the raw acceptance rates from 10 to 35% (17-51% for the subsample with CVs). 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 co-editors. For this reason, we allow for co-editor fixed effects in our regression, and when testing for differences in standards in Section IV below, look at these rather than the raw probabilities. As shown in Table 1 below, these dummies are significant for a number of co-editors, i.e., for co-editors 3, 5, 8, and 20. For the results presented in Table 8, we only include dummies of those co-editors. As we can see from the results in Table 1 and in Columns 1, 2, and 3 of Table 8, including only these 4 co-editors reduces the magnitude of the marginal effects somewhat. All of the coefficients, except for the one for co-editor 20, are strongly significant.

Experience. Good prior publications may reveal the ability to write good papers for the JIE, so as we expected, having publications in journals as good as or better than the JIE tends to raise the probability of acceptance at the JIE. Notice that only the number of publications in Group 2, 5, and "Network" journals is statistically significant. Statistical insignificance of other experience is likely to be a result of a multicollinearity problem: the numbers of publications in different groups are highly correlated! The first principal component[17] explains 60% of the variance as evident from Table 9.2 in the Appendix. If we

[17] This is just a linear combination of all papers an author has in different groups, for which all coefficients are positive.

drop the number of publications in Group 3 to 5, then the number of publications in Group 1 journals becomes positive and statistically significant at 5% or 1% level.[18]

Table 1: Co-editor fixed effects

Co-editor	% Accepted (Sample)	Probit marginal effect[19]	Quality difference[20]	Citations per year, Accepted	Citations per year, Rejected	Time to first decision
1	0.31	0.064	-0.13	4.6	2.0	187
2	0.28	-0.011	-0.08	5.8	2.0	115
3	0.49	0.183***	-0.07	6.8	1.7	124
4	0.26	-0.070	-0.04	4.7	2.0	127
5	0.38	0.229***	-0.22	5.3	1.4	156
6	0.37	—	—	11.4	3.2	80
7	0.32	0.053	-0.10	2.4	1.9	166
8	0.51	0.302***	-0.16	6.5	0.2	218
9	0.22	-0.008	-0.15	8.7	2.6	103
10	0.26	0.045	-0.16	3.4	1.0	191
11	0.32	-0.002	-0.05	6.6	0.8	101
12	0.30	0.002	-0.08	12.4	1.4	192
13	0.23	0.060	-0.20	2.7	2.4	128
14	0.20	0.066	-0.24	2.6	1.5	123
15	0.35	0.090	-0.11	9.4	1.8	117
16	0.23	0.071	-0.14	5.7	3.2	136
17	0.17	0.055	-0.20	3.7	0.7	107
18	0.17	-0.037	-0.16	9.0	3.6	187
19	0.18	-0.073	-0.13	14.9	2.6	180
20	0.40	0.194*	-0.17	5.7	0.7	122
21	0.41	0.032	0.01	7.8	2.1	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.

The number of publications per year in journals other than the above (other journals) has a negative impact on the probability of being published. This could be because writing

[18] The results are available upon request.

[19] The discrete change in the probability for a change of dummy variable from 0 to 1 is reported (evaluated at the means).

[20] See Section IV.C for a detailed description and discussion.

bad papers one more likely to keep doing so. One would expect this effect to be stronger for people being out of grad school for a while. To check for this, we also looked for the evidence of a differential impact on the probability of acceptance depending on vintage. While the interaction of Ph.D. vintage and the total number of publications not in top journals turns out to be negative, it is statistically insignificant.[21]

Language. The language dummy is positive in all specifications and significant except for the Probit specifications in Columns 1 and 3. This highlights the importance of good writing for publication.

University Rank. The distribution of submissions by the world-university ranking according to the author's Ph.D. is given in Table 2. 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 2: Submissions and acceptances by graduate school quality cohorts

	Submissions #	Submissions %	Accepted	Accepted / submissions
Top 20	1145	56%	394	34.4%
Top 50	1494	73%	465	31.1%
Top 200	1932	94%	531	27.5%
Sample Available	2051	100%	537	26.2%
"Population"	3032	—	600	19.8%

Our estimates (in all columns of Table 8) show that people who graduated from more highly ranked places are more likely to have their papers accepted[23]. 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 much 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 noting that submission decisions are endogenous. Authors

[21] The results are available upon request.

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 maybe such submissions actually tend to be of lower quality.

[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.

[23] The base group for the dummy variables is the universities ranked from 101 to 200.

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 a feedback on a paper, even if it has a 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 than Ph.D. vintage, we include interaction terms between university rankings and Ph.D. vintage. Table 3 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 for graduates of the top 10 universities. For other graduates the coefficients on the interaction terms are insignificant, which implies that initial differences seem to persist.

In Table 4, for each Ph.D. vintage we show the share of manuscript submissions for different university quality groups. Though the number of submissions falls with Ph.D. vintage, it clearly shows that the structure of submissions remains roughly the same as we vary Ph.D. vintage. 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.

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 higher quality of papers submitted by them, as a JIE publication could be more valuable to them in their careers.

Table 3: Divergence versus convergence

Rank of the university	Constant term	Slope of intersection with Ph.D. vintage
Top 10 universities	0.752	−0.030
	(0.148)***	(0.014)***
Top 20, excluding Top 10	0.490	0.012
	(0.158)***	(0.014)
Top 30, excluding Top 20	0.347	0.015
	(0.232)	(0.019)
Top 40, excluding Top 30	0.411	−0.013
	(0.260)	(0.023)
Top 50, excluding Top 40	0.469	−0.019
	(0.247)*	(0.021)
Top 100, excluding Top 50	0.321	−0.009
	(0.173)*	(0.016)
Not one of the 200 best universities	−0.250	0.006
	(0.135)**	(0.011)

Table 4: 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	31%	34%	33%	32%	31%
Top 10 - 20	24%	20%	22%	20%	17%
Top 20-30	7%	8%	7%	6%	10%
Top 30-40	6%	7%	9%	11%	10%
Top 40-50	6%	5%	7%	6%	7%
Top 50-100	15%	14%	12%	15%	19%
Top 100-200	12%	11%	10%	10%	7%
# of Papers	624	545	333	194	201

Backlog. The backlog variable turns out to be significant at the 10% level and has a negative coefficient[24].

[24]The detailed estimation results are available upon request.

B Time to First Decision and Survival Probabilities

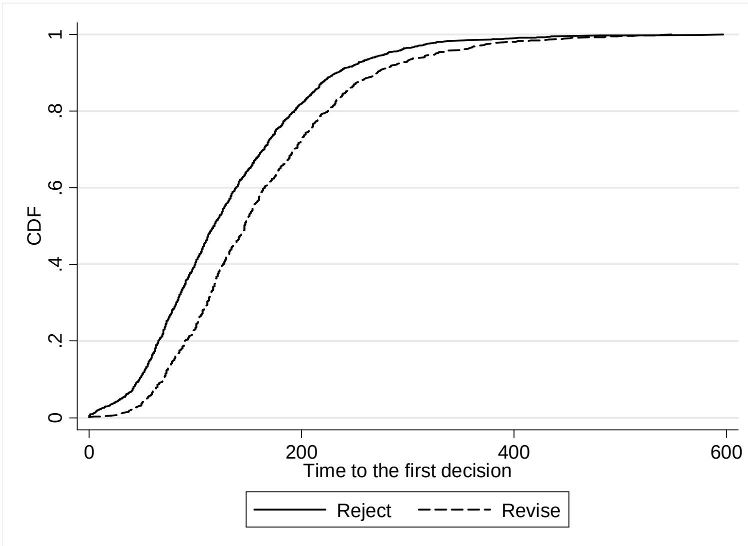
What does the time to the first decision say about the probability of rejection? Most accepted papers go through at least one revision: out of about 3,032 submissions, only 17 (0.6%) were accepted with no revision, about 770 (23%) were sent for revision, and about 600 (78%) of them were finally accepted. 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 distributions of waiting times for these 2 groups. The cumulative distribution for rejected papers clearly lies above that for accepted ones, so that the latter FOSD the former. Figure 1 also suggests[25] that "no news is good news"! This makes sense as it may take more time to review an acceptable paper 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 accepted without revision on average spent 130 days under reviewing, but with a very high standard deviation of 115 days. The probability that an article will not be rejected given that it has survived for X days from submission, i.e., the probability of acceptance conditional on survival is also increasing!

C Streamlining and its Costs

Finally, we ask, how well the model does in predicting final acceptance. Suppose the JIE 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 content, 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 2004 and sort the articles in descending order.

[25] Both the Anderson (1996) FOSD and Kolmogorov-Smirnov tests support our findings, see Appendix, Table 7, column 1.

Figure 1: Time to the first decision



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 2.

When $k = 60\%$, this number is 8%. Hence, without reading the papers, eliminating 40% of the submitted manuscripts will result in at most an 8% wrongful rejection rate relative to the current procedure. Note that this is without any information about the paper itself!

When we look at single authored papers only, we do even better: at $k = 60\%$, the error is zero! In other words, none of the worst 40% of papers as classified by our model were actually accepted and half of the best papers account for approximately 73% of articles accepted for publication. The quality of prediction on single-authored papers is far better than that for the whole sample, suggesting that co-authorship is befuddling the model.[26]

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

Figure 2: Costs of streamlining

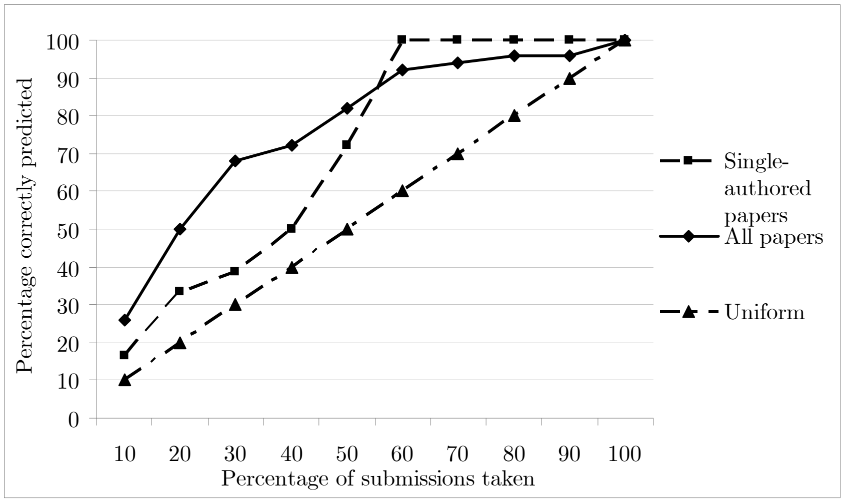


Table 5: Final publications for papers rejected by the JIE

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

IV Evaluating Performance

A 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 with 14% of them being published in journals ranked by Kalaitzidakis et. al.(2003) above the JIE (see Table 5). If we take publication in a journal ranked above the JIE as an evidence of a mistake, then type 1 error is 14%. This would be an overestimate, if most such papers were rejected for not being a good fit and not because of low quality, or an under-estimate, if rejection by the JIE discourages authors from submitting elsewhere. Table 6 shows that the average number of citations per year varies significantly across groups of articles. For papers published in better journals, this number is half that of those published in the JIE. It can be argued that articles rejected by the JIE take more time to get accepted in another journal, which leads to lower citations. However, most papers are available on-line and accumulate citations before publication. Moreover, even for the pre 2001 period, papers published in better journals are cited less often, suggesting that type 1 error is not so large. Figure 3 plots the cumulative distribution functions of the number of citations per year for 3 groups of papers. The citations for the papers published in the JIE FOSD all the others, which cannot be distinguished from each other.[27]

Table 6: Average citations per year for different groups of papers

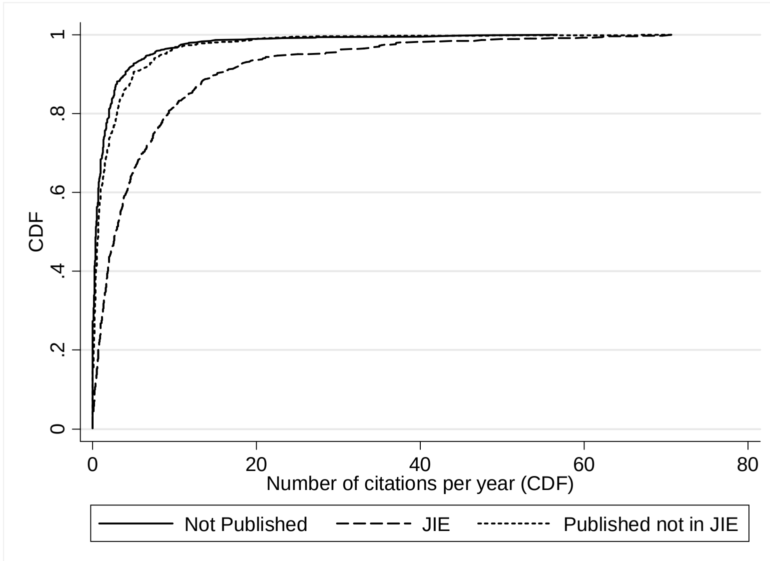
	Citations per year (average)	Maximum citations per year	Citations per year (average)	Maximum citations per year
	1995 - 2004	1995 - 2004	1995 - 06/2001	1995 - 06/2001
Published in JIE	6.33	70.75	6.30	69.54
Rejected by JIE, but published in a higher ranked journal	3.71	71.00	3.20	38.83
Rejected by JIE,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

B Acceptance, Rejection and Quality

Next, we look at the relation between quality and outcomes by considering the distributions of citations for accepted and rejected papers. The probability of acceptance equals that of rejection when citations per year are about 4.5-5. Thus, articles with 4.5-5 cites are equally likely to be accepted or rejected. Similarly, articles with 0.5-1 cites per year are 4 times more

[27] Results of the Anderson (1996) Kolmogorov-Smirnov tests are in Appendix, Table 7, Columns 5-7.

Figure 3: Number of citations per year (CDF)



likely to be rejected than accepted. In other words, they have an odds ratio (A/R) of .25. Hence, an accepted paper with this number of cites has $\frac{.25}{.75}$ or 1/3 chance of being wrongfully accepted. If 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 probabilities of acceptance and rejection (which add to unity), given that a paper has x citations per year or less. Figure 4 depicts the probability of accepting a paper conditional on that it has x citations per year or less. If we say that articles with a citation below x were wrongfully accepted, then even with $x = 0$, we have roughly 7% of such articles, suggesting that type 2 error is quite high. It is particularly high for earlier years of our sample, indicating that both overall quality of submissions and standards went up, and that the efficiency of reviewing system improved.

C Editorial Heterogeneity

The raw acceptance rates differ quite substantially across co-editors. How should we interpret this? Clearly there are at least two reasons for this difference. First, that co-editors get different quality papers, and second, that they have different standards. Suppose, for example, that the more interesting articles are retained by the managing editor who assigns papers but all co-editors have the same standards. This would lead to differences in raw acceptance rates that have nothing to do with differences in standards! How can we decompose acceptance rates into their components parts? By adding dummies for co-editors in conjunction with controlling for author characteristics, we can capture fixed effects associated with co-editors.[28] These fixed effects would capture differences in acceptance rates that are not due to quality differences. Moreover, since the difference in raw acceptance rates between two co-editors captures both the difference in standards and heterogeneity in quality, these two effects can be separated!

For co-editor i, denote the raw acceptance rate as A_i, the standards as S_i, and the average quality of papers he handles as Q_i. Consider the difference between the acceptance rates of co-editor j and the base editor 0. This can be decomposed as follows:

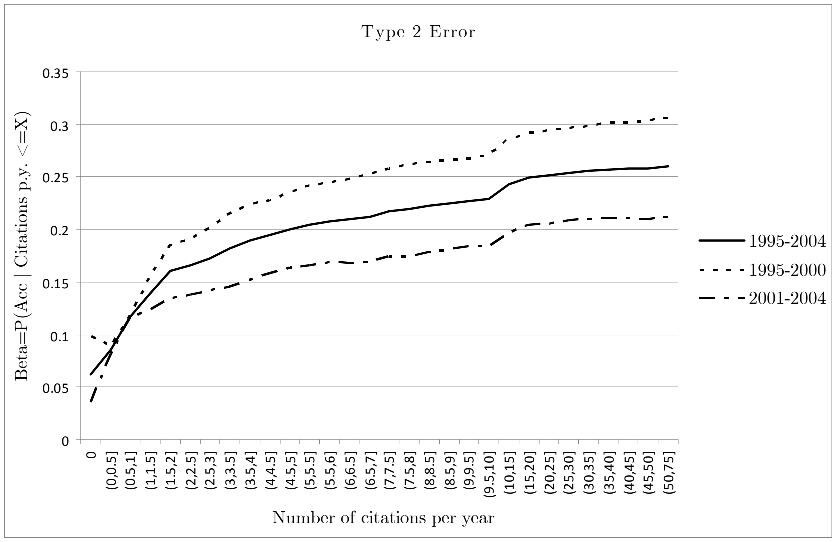
$$A(Q_j, S_j) - A(Q_0, S_0) = [A(Q_j, S_j) - A(Q_j, S_0)] + [A(Q_j, S_0) - A(Q_0, S_0)] .$$

The first term on the RHS corresponds to the effect of keeping the papers the same (i.e., as those handled by co-editor j) but asking how the acceptance rate would change if the base editor handled the papers rather than editor j. If this is positive then co-editor j would be more lenient than the base co-editor. This is, thus, just the marginal effect.

The second term on the RHS corresponds to the effect of keeping the co-editor the same (to be the base co-editor), but changing the quality of papers given to him. This can be obtained by subtracting the marginal effect from the difference in the raw acceptance rates (LHS). For example, we see that the raw acceptance rate of co-editor 1 is 31%, while that of co-editor 6 is 37% (see Table 1). The raw difference is -6%, suggesting that co-

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

Figure 4: Probability of accepting paper with X or less citations



editor 1 accepts less papers than co-editor 6. Does it mean that co-editor 1 is harsher than co-editor 6? In fact, he is more lenient as the marginal effect for co-editor 1 is 6%. This reveals that the quality of papers co-editor 1 gets is worse than that of co-editor 6 $(A(Q_j, S_j) - A(Q_0, S_0) - [A(Q_j, S_j) - A(Q_j, S_0)] = -12\%)$. Thus, co-editor 1 is both more lenient and gets lower quality papers than co-editor 6. Estimates of the second term of the decomposition are reported in Table 1, column 4. Heterogeneity in submissions is responsible for as much as 24, while differences in standards are responsible for up to 30 percentage points.[29]

Is it possible to interpret patterns in the decomposition in any way? What might explain these large differences in editorial standards? Could it be that what is happening is that co-editors who receive lower quality submissions are being overly generous, perhaps, because they compare each paper to the average they receive and aim for a target raw acceptance

[29] Unfortunately, we do not directly observe the fields the papers belong to. Differences in fields may affect probability of acceptance. Since different editors specialize in particular fields, the latter might be partially responsible for difference in marginal effects. However, we do not find any evidence of this! Co-editors who have significant and positive estimates of marginal effect specialize in different fields and each of them has at least one colleague co-editor, who works in the same field and does not have a significant marginal effect.

rate. If this was true, then more generous editors should be handling lower quality papers. This is exactly what we see: the rank correlation between the marginal effect estimate and the quality difference is -0.44 (significant at 10%). There is also a statistically significant (at the 10% level) rank correlation of 0.42 between the quality difference in assigned papers and the number of citations for accepted papers. Providing co-editors with information about the relative quality of the papers they are assigned may help correct this bias.

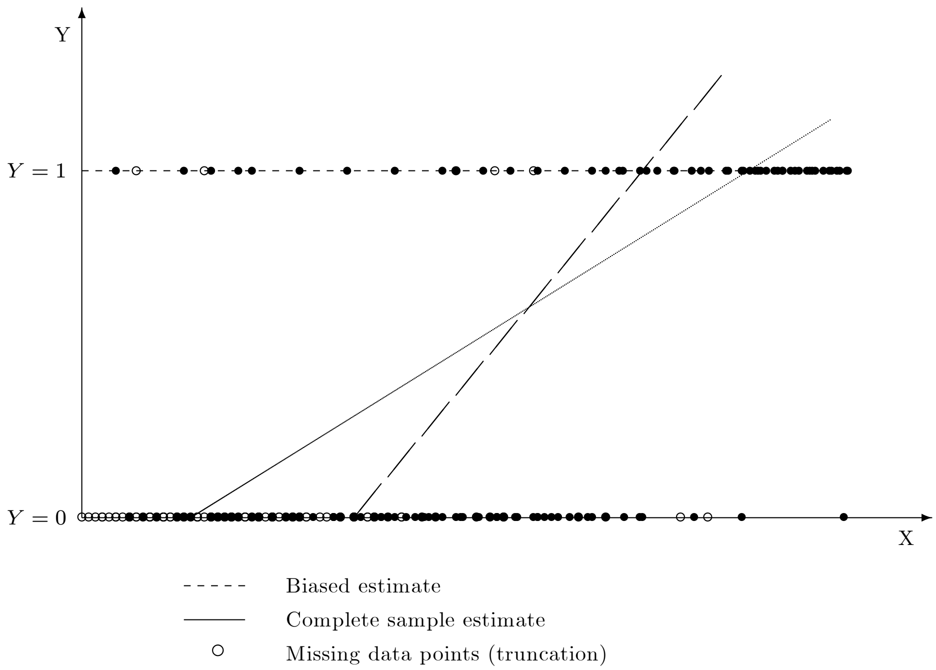
A lower cutoff for acceptance will reduce the average quality in both accepted and rejected papers. Thus, lower standards should reduce citations for both groups. As expected, there is a negative rank correlation of -0.3 ($-.55$) between the marginal effect and the number of citations for accepted (rejected) papers, consistent with our interpretation of the marginal effect. Moreover, for both accepted and rejected papers, the distribution of citations per year for four co-editors with positive and significant marginal effects is first order stochastically dominated[30] by that for all other co-editors! This once again supports the idea that there are significant differences in standards across co-editors.

V Robustness Checks

Simple binary choice models of acceptance could have problems when applied directly to our data as there is a selection problem. This comes from the fact that while we have data on all submissions, we only have CVs for authors with a web presence, and as a result, the set of submissions we can use is restricted. As the acceptance rate for authors with a CV on the web exceeds that of the entire population, there may be selection bias. It turns out that CVs are more commonly available over time, 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 4.4 times as likely to be accepted in 1995 relative to

[30] The FOSD test used is from Anderson (1996). PAT statistics are significant at 5% level for both groups of papers. All decile differences are of the same sign and at least 6 of them are significant at the 5% level. Details are available upon request.

Figure 5: Selection bias



those without a vita, and 4.75 times as likely to be accepted in 2004, suggesting that only the most marginal authors did not have CVs by 2004. While we have CV data for 85% of the authors of accepted articles, we only have 57% of CVs for the authors with unsuccessful submissions. In other words, the authors of accepted papers are over-represented in our sample.

The intuition for the expected bias is evident from Figure 5. In Figure 5, think of X as the explanatory variable in the model estimated. Y could be either 0 or 1. Solid points represent observations with CVs while hollow ones represent that without CVs. 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.

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 does not result in bias.[31] 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 biases 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 biased upward. Of course, if the selection was random, all points were equally likely to be removed, and then there would be no bias.

To account for this selection bias in our estimation, we incorporate the selection equation into the likelihood function as done below. Let X be the set of authors' characteristics that affect the likelihood of acceptance and Z be the set of the authors' characteristics that define a web presence. These sets can overlap to some extent. We assume that the article i is published in the journal if its latent quality q_i exceeds a threshold level:

$$Y_i = \begin{cases} 1, & \text{if } q_i = X_i\beta + \varepsilon_{1i} \geqslant 0, \\ 0, & \text{if } q_i = X_i\beta + \varepsilon_{1i} < 0, \end{cases} \tag{2}$$

where Y_i is an indicator for the paper being published ($Y_i = 1$) or not ($Y_i = 0$). Note that under such 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)

[31] 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.

only if $Z_i\gamma + \varepsilon_{2i} \geqslant 0$:[32] This is a critical component since the standard Heckman approach would require that both X and Z are observed for the whole sample.

$$(Y_i, X_i, Z_i) = \begin{cases} (Y_i, X_i, Z_i)\,, & \text{if } Z_i\gamma + \varepsilon_{2i} \geqslant 0, \\ (Y_i, Not\ observed)\,, & \text{if } Z_i\gamma + \varepsilon_{2i} < 0. \end{cases} \tag{3}$$

It is natural to expect that ε_{1i} and ε_{2i} are positively correlated: authors with a better chance of being published are also more likely to have an established name in profession and have CVs easily accessible on the web.[33] 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. (2006)). 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 the 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, and we discuss this further in the Appendix.

In Column 5 and 6 of Table 8, we report the coefficient estimates of the full model, where we estimate the truncation equation (2) jointly with the probit equation. We assume that the errors $(\varepsilon_1, \varepsilon_2)$ are jointly normally distributed with variance 1 and covariance ρ. In Column 6 we add citations to the independent variables in Column 5.

We find that most of the patterns in coefficient estimates of the 2-equation model are not very different from those of the probit model. Thus, our conclusions above remain valid. The coefficients on Ph.D. vintage do differ from those of the probit specification. This difference illustrates the potential problems that arise from only estimating the probit model so it is useful to see what drives it.

[32] This needs to be qualified as citation data is observed even for some submissions that lack CVs and are not observed for some that have CVs.

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

In the 2-equation model, the acceptance probability increases uniformly with vintage, except for the coefficient estimate of the Ph.D. vintage 4 to 6 year dummy. If we look at the coefficient estimates of the first stage truncation equation, which are presented at the last block of Table 8, we see that Ph.D. vintage increases web presence: its coefficient is 0.132 and significant, when citations are not added as the second stage independent variables, and 0.176 and significant when they are. Furthermore, the correlation between the error term of the truncation equation and the acceptance equation is significantly positive. Together, this implies that scholars with a higher Ph.D. vintage have on average lower value of ε_1, which is why the acceptance probability decreases with vintage for vintages over 2-4 years.

VI Conclusion

A better understanding of the way journals operate can help all parties involved. It can help authors understand how well or badly a journal performs, identify strengths and weaknesses to those operating the journal, and provide editors feedback on their performance. Such evaluations could also shed light on the biases, if any, inherent in the existing system.

Our main conclusions are that overall 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 a difference in standards and performance across co-editors. This might be reduced by providing feedback to co-editors on their relative performance and information on the quality composition of the papers they receive relative to the average. In the future, with electronic data bases being kept by journals, such feedback can also be provided to referees. Editors can also get more information on the referee performance. 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).

References

[1] Anderson, Gordon. 1996. "Nonparametric Tests of Stochastic Dominance in Income Distributions." *Econometrica,* 64(5): 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): 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): 778-784.

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

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

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

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

[8] Jenkins, Stephen P., Lorenzo Cappellari, Peter Lynn, Annette Jackle, and Emanuela Sala. 2006. "Patterns of Consent: Evidence From a General Household Survey." *Journal of the Royal Statistical Society, Series A*, 169(4): 701-722.

[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): 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): 193-204.

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

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

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

VII Appendix

A The Likelihood Function

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

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right), \tag{4}$$

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.

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:

$$Prob[X_i\beta + \varepsilon_{1i} > 0 \;\mid\; Z_i\gamma + \varepsilon_{2i} > 0] = \frac{Prob[\varepsilon_{1i} > -X_i\beta \;\cap\; \varepsilon_{2i} > -Z_i\gamma]}{Prob[\varepsilon_{2i} > -Z_i\gamma]} = \tag{5}$$

$$= \frac{\int_{-Z_i\gamma}^{\infty}\int_{-X_i\beta}^{\infty} g(E_1, E_2)dE_1 dE_2}{\int_{-Z_i\gamma}^{\infty} g_2(E_2)dE_2} = \frac{1 - \Phi(-X_i\beta) - \Phi(-Z_i\gamma) + G(-X_i\beta, -Z_i\gamma)}{1 - \Phi(-Z_i\gamma)} \tag{6}$$

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

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

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

B Kolmogorov-Smirnov and Anderson Tests: Detailed Results

Table 7: FOSD (Anderson) and Kolmogorov-Smirnov tests

FOSD (Anderson) Test

Column #	1	2	3	4	5	6	7
	Time to the first decision for:				Number of citations per year for:		
Distrib. A	Papers to be revised	All papers 1995-1999	Papers to be revised 1995-1999	Rejected papers 1995-1999	Papers published in JIE	Papers published in JIE	Papers published anywhere else
Distrib. B	Rejected papers	All papers 2000-2004	Papers to be revised in 2000-2004	Rejected papers 2000-2004	Papers never published	Papers published anywhere else	Papers never published
Decile 1	-8.0***	-1.7*	6.7***	-4.29***	-23.0***	-11.3***	-12.6***
Decile 2	-12.8***	0.7	12.1***	-2.12	-25.2***	-20.6***	-11.2***
Decile 3	-15.0***	5.2***	14.1***	1.61	-37.3***	-30.9***	-10.9***
Decile 4	-16.6***	6.3***	15.0***	3.29*	-39.6***	-34.2***	-10.3***
Decile 5	-14.5***	7.4***	13.5***	4.48**	-41.9***	-33.0***	-7.7***
Decile 6	-12.9***	8.1***	12.1***	7.34***	-40.0***	-28.9***	-7.4***
Decile 7	-11.4***	6.1***	7.7**	5.07***	-37.0***	-27.8***	-4.9**
Decile 8	-8.9***	4.1***	6.7**	4.75***	-28.9***	-21.7***	0.0
Decile 9	-5.8***	1.8*	2.6	1.58	-15.8***	-13.3***	—
Decile 10	0.0	0.0	0.0	0.0	0.0	0.0	—
PAT($\chi^2_{(9)}$)	85.7***	34.5***	23.1***	34.42***	306.2**	162.3***	32.2*** $\chi^2_{(7)}$
Test result:	FOSD	Mixed Result	FOSD	Mixed Result	FOSD	FOSD	FOSD

*, **,*** denote significance at 10, 5, and 1 percent level, respectively.

Kolmogorov - Smirnov Test

Column	1	2	3	4	5	6	7
	Time to the first decision for:				Number of citations per year for:		
Distrib. A	Papers to be revised	All papers 1995-1999	Papers to be revised 1995-1999	Rejected papers 1995-1999	Papers published in JIE	Papers published in JIE	Papers published anywhere else
Distrib. B	Rejected papers	All papers 2000-2004	Papers to be revised in 2000-2004	Rejected papers 2000-2004	Papers never published	Papers published anywhere else	Papers never published
H_0 : The two samples come from a common distribution							
P-value	0.000	0.000	0.000	0.005	0.000	0.000	0.621
H_0 : $F_A(X) > F_B(X)$, where F stands for CDF							
P-value	0.000	0.480	0.964	0.072	0.000	0.000	0.519
H_0 : $F_A(X) < F_B(X)$, where F stands for CDF							
P-value	0.998	0.000	0.000	0.000	0.999	1.000	0.346

C Estimation Results

Table 8: 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*[34]	*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	0.226	0.214	—	0.577	-1.134	-1.145
	[3.08]***	[2.54]**	—	[2.54]**	[2.42]**	[2.04]**
Ph.D. vintage: (0, 2]	0.272	0.229	0.232	0.619	-0.840	-0.878
	[3.81]***	[2.83]***	[2.81]***	[2.83]***	[2.00]**	[1.72]*
Ph.D. vintage: (2, 4]	0.361	0.346	0.346	0.922	-0.501	-0.424
	[5.12]***	[4.34]***	[4.27]***	[4.34]***	[1.30]	[0.90]
Ph.D. vintage: (4, 6]	0.233	0.213	0.218	0.578	-0.630	-0.547
	[3.45]***	[2.77]***	[2.78]***	[2.77]***	[1.78]*	[1.25]
Ph.D. vintage: (6, 10]	0.231	0.193	0.200	0.53	-0.417	-0.328
	[3.71]***	[2.72]***	[2.78]***	[2.72]***	[1.35]	[0.83]
Ph.D. vintage: (10, 20]	0.071	0.026	0.031	0.077	-0.376	-0.298
	[1.36]	[0.44]	[0.51]	[0.44]	[1.67]*	[0.98]
	Co-editors fixed effects					
Co-editor 3	0.132	0.162	0.139	0.442	0.272	0.344
	[3.21]***	[3.26]***	[2.60]***	[3.26]***	[2.83]***	[2.85]***
Co-editor 5	0.165	0.218	0.253	0.588	0.385	0.496
	[4.20]***	[4.82]***	[5.12]***	[4.82]***	[4.18]***	[4.44]***
Co-editor 8	0.286	0.28	0.272	0.735	0.627	0.642
	[3.59]***	[3.24]***	[2.94]***	[3.24]***	[3.81]***	[3.24]***
Co-editor 20	0.170	0.175	0.196	0.469	0.365	0.385
	[2.00]**	[1.81]*	[1.91]*	[1.81]*	[1.85]*	[1.70]*
	Experience: number of publications in various journals					
# of articles: Group 1	0.005	0.004	0.004	0.013	0.032	0.023
	[0.65]	[0.51]	[0.50]	[0.51]	[1.63]	[0.98]
# of articles: Group 2	0.017	0.016	0.015	0.048	0.050	0.051
	[1.95]*	[1.56]	[1.44]	[1.56]	[2.11]**	[1.79]*
# of articles: Group 3	-0.01	-0.008	-0.004	-0.022	-0.022	-0.016
	[1.42]	[0.96]	[0.46]	[0.96]	[1.18]	[0.74]
# of articles: Group 4	0.011	0.023	0.030	0.068	0.020	0.050
	[0.77]	[1.47]	[1.73]*	[1.47]	[0.57]	[1.23]
# of articles: Group 5	0.027	0.017	0.019	0.05	0.072	0.050
	[2.17]**	[1.14]	[1.21]	[1.14]	[2.17]**	[1.28]
# prior JIE publications	0.024	0.02	0.022	0.06	0.068	0.065
	[2.62]***	[1.88]*	[2.02]**	[1.88]*	[2.74]***	[2.21]**
# in network journals	0.098	0.061	0.066	0.182	0.227	0.136
	[4.13]***	[2.20]**	[2.30]**	[2.20]**	[3.84]***	[1.85]*
# other articles per year	—	—	-0.052	—	—	—
	—	—	[3.88]***	—	—	—
	Language effect					
Language dummy	0.05	0.074	0.057	0.22	0.150	0.207
	[1.60]	[1.98]**	[1.40]	[1.98]**	[1.85]*	[2.14]**
	Proxies for article quality					
Citations per year	—	0.022	0.022	0.066	—	0.053
	—	[9.23]***	[8.87]***	[9.23]***	—	[8.63]***

[34] See page 31, Table 10 for the defenition of marginal effect

Table 8: 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					
Grad. from top 10	0.176	0.163	0.168	0.47	0.407	0.373
	[5.77]***	[4.50]***	[4.39]***	[4.50]***	[5.18]***	[3.89]***
Grad. from top 10 - 20	0.189	0.21	0.216	0.583	0.437	0.482
	[5.90]***	[5.61]***	[5.49]***	[5.61]***	[5.48]***	[4.82]***
Grad. from top 20 - 30	0.117	0.146	0.147	0.4	0.259	0.330
	[2.61]***	[2.73]***	[2.62]***	[2.73]***	[2.29]**	[2.38]**
Grad. from top 30 - 40	0.047	0.077	0.074	0.218	0.104	0.192
	[1.05]	[1.40]	[1.28]	[1.40]	[0.89]	[1.36]
Grad. from top 40 - 50	0.085	0.105	0.061	0.294	0.196	0.249
	[1.75]*	[1.88]*	[1.07]	[1.88]*	[1.64]*	[1.74]*
Grad. from top 50 -100	0.068	0.08	0.078	0.228	0.153	0.178
	[1.94]*	[1.91]*	[1.73]*	[1.91]*	[1.73]*	[1.75]*
Not ranked university	-0.056	-0.061	-0.043	-0.184	-0.125	-0.130
	[2.28]**	[2.07]**	[1.36]	[2.07]**	[1.85]*	[1.63]
	Institution affiliation variables					
Affiliated with US univ.	0.079	0.004	-0.004	0.011	0.198	0.009
	[2.55]**	[0.10]	[0.11]	[0.10]	[2.43]**	[0.09]
Affiliated with CA univ.	0.120	0.087	0.091	0.244	0.286	0.196
	[2.38]**	[1.50]	[1.51]	[1.50]	[2.50]**	[1.39]
Affiliated with UK univ.	0.058	-0.013	0.037	-0.039	0.138	-0.042
	[1.21]	[0.24]	[0.61]	[0.24]	[1.09]	[0.28]
Affiliated with EU univ	0.021	-0.043	-0.041	-0.129	0.049	-0.122
	[0.62]	[1.11]	[0.99]	[1.11]	[0.56]	[1.22]
Affiliated with organiz.	0.124	0.063	0.068	0.181	—	—
	[3.18]***	[1.44]	[1.50]	[1.44]	—	—
	Fixed year effects					
Year 1996	-0.047	-0.082	-0.057	-0.262	-0.120	-0.222
	[0.97]	[1.36]	[0.86]	[1.36]	[0.92]	[1.36]
Year 1997	-0.114	-0.136	-0.150	-0.462	-0.279	-0.346
	[2.48]**	[2.37]**	[2.44]**	[2.37]**	[2.14]**	[2.13]**
Year 1998	-0.128	-0.147	-0.146	-0.507	-0.360	-0.422
	[2.81]***	[2.58]***	[2.32]**	[2.58]***	[2.58]***	[2.45]**
Year 1999	-0.101	-0.135	-0.123	-0.45	-0.260	-0.348
	[2.31]**	[2.50]**	[2.10]**	[2.50]**	[2.07]**	[2.24]**
Year 2000	-0.082	-0.127	-0.116	-0.424	-0.186	-0.320
	[1.75]*	[2.28]**	[1.90]*	[2.28]**	[1.46]	[2.05]**
Year 2001	-0.081	-0.12	-0.092	-0.394	-0.200	-0.331
	[1.75]*	[2.13]**	[1.46]	[2.13]**	[1.52]	[2.06]**
Year 2002	-0.092	-0.135	-0.138	-0.451	-0.262	-0.385
	[2.02]**	[2.44]**	[2.30]**	[2.44]**	[2.06]**	[2.47]**
Year 2003	-0.122	-0.183	-0.187	-0.636	-0.328	-0.517
	[2.88]***	[3.62]***	[3.43]***	[3.62]***	[2.74]***	[3.54]***
Year 2004	-0.125	-0.166	-0.194	-0.569	-0.343	-0.474
	[2.90]***	[3.19]***	[3.47]***	[3.19]***	[2.77]***	[3.17]***
	Constant term					
Constant term	—	—	—	-1.542	-2.088	-1.671
	—	—	—	[5.21]***	[5.84]***	[6.51]***
	First stage (truncation) equation					
Constant term	—	—	—	—	-3.003	-2.682
	—	—	—	—	[3.26]***	[2.28]**

Table 8: 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 (continued)					
Organiz. dummy	—	—	—	—	-0.509	-0.242
	—	—	—	—	[3.14]**	[1.12]
Ph.D. vintage	—	—	—	—	0.132	0.176
	—	—	—	—	[3.57]***	[2.73]***
Correlation (ρ)	—	—	—	—	0.647	0.620
	—	—	—	—	[99.2]***	[4.89]***
Pseudo R^2	0.16	0.21	0.22	0.21	N/A	N/A
Number of obs.	1792	1476	1082	1476	1792	1476

Estimates of month specific effects in probit equations are not reported to save space.
Robust Z statistics in parentheses for simple probit estimates, t-ratios for the full model with truncation equation; ***, **, * denote significance at 1%, 5% and 10%, respectively.

Marginal effect is the change in the probability for an infinitesimal change in each independent and continuous variable. For dummy variables marginal effect is a discrete change in the probability for change of dummy variable from 0 to 1. All marginal effects are evaluated at the means.

D Principal Components Analysis

Table 9.1: Correlation between number of publications in different groups of journals

	Group 1	Group 2	Group 3	Group 4	Group 5	JIE	Network
Group 1	1						
Group 2	0.61	1					
Group 3	0.41	0.53	1				
Group 4	0.23	0.30	0.24	1			
Group 5	0.50	0.39	0.31	0.26	1		
JIE	0.52	0.54	0.37	0.12	0.34	1	
Network	0.49	0.24	0.15	0.08	0.28	0.28	1

Table 9.2: Principal components: eigenvalues of variance / covariance matrix

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	11.5	8.4	0.60	0.60
Comp2	3.1	1.3	0.16	0.76
Comp3	1.7	0.4	0.09	0.85
Comp4	1.3	0.6	0.07	0.92
Comp5	0.8	0.2	0.04	0.96
Comp6	0.6	0.4	0.03	0.99
Comp7	0.2	.	0.01	1.00

Table 9.3: Principal components: eigenvectors of variance / covariance matrix

Variable	Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6	Comp 7	Unexplained
Group 1	0.66	-0.57	-0.44	-0.02	-0.17	0.07	-0.11	0.00
Group 2	0.49	0.11	0.63	-0.57	-0.10	-0.10	0.03	0.00
Group 3	0.44	0.80	-0.37	0.13	-0.06	-0.01	0.01	0.00
Group 4	0.08	0.04	0.00	-0.16	0.52	0.83	0.02	0.00
Group 5	0.16	-0.07	-0.06	0.00	0.83	-0.53	-0.04	0.00
JIE	0.31	-0.06	0.51	0.79	0.04	0.10	-0.02	0.00
Network	0.06	-0.08	-0.05	0.04	0.01	-0.03	0.99	0.00