MPRA

Munich Personal RePEc Archive

Sequential order as an extraneous factor in editorial decision

Sultan Orazbayev

20 March 2017

Sequential order as an extraneous factor in editorial decision[*]

Sultan Orazbayev[†]

Abstract

Academic journal editors reject a significant portion of first submissions without sending them out for peer review. This decision, desk rejection, is made to reduce the workload on associate editors and referees, to give the submitting author a head start on revision or pursuit of an alternative venue, as well as to achieve quicker turnaround time for the journal. Desk rejection is a judgement based on the manuscript's perceived quality, impact and fit with the journal's scope. Could extraneous factors, which are unrelated to the content of a manuscript, affect the editorial decision? This paper examines whether the sequential order in which manuscripts are submitted to a large academic journal affects the editorial decision. Becoming the first submission on the editor's list of manuscripts to review increases the probability of a desk rejection by up to 7% without any effect on the likelihood of a rejection after peer review.

Keywords: desk rejection; sequential order effect; decision fatigue; peer review.
JEL codes: D03; A19; D89; J01.

1 Introduction

When choosing a potential publication venue, an author considers several factors, including the manuscript's fit with a journal's scope, prestige of a journal, the perceived likelihood of eventual publication, and the estimated time to receive a decision on the submission. Editors, especially at journals with large submission volumes, will seek to reduce the reviewing burden on the associate editors and referees by rejecting some of the submissions without sending them for a peer review. This decision, desk rejection, is typically based on a judgement which is made under some time pressure: the editors usually have additional work commitments (e.g. teaching or conducting their own research) which limit the time budget for reviewing the submissions. The editors also have an incentive to make this decision quickly to improve the journal's turnaround time, as well as to give the authors more time for revision and finding an alternative outlet. Given these considerations, it's possible that the editor's decision to desk reject a manuscript could be biased by extraneous factors, factors unrelated to the manuscript's content.

This paper uses information on all submissions to a large academic journal during 2009–2013 to explore the potential impact of extraneous factors on editorial decisions. To focus attention on the hypothesis, rather than the journal, any explicit information identifying the journal has been removed and the journal will henceforth be referred to as 'Journal X' or, if appropriate, 'the journal'. Journal X is widely recognised as a leading journal in a broad range of subjects, occupying top positions in various journal rankings. During the sample period, the journal received 12 thousand submissions, of which more than 9'500 were first submissions. Given its wide scope, the journal's editorial board members are assigned to particular subject areas, between 1 and 5 editors per subject area. These subject-specific editors will be referred to as editors, but where relevant they will be distinguished from the editor-in-chief. Each editor can make an independent decision on the manuscript, specifically they can desk reject a manuscript without additional consultation with the editor-in-chief. The editor can also pass on the manuscript to an associate editor, who can recommend a rejection or proceeding with peer review. Depending on the combined judgement of referees, associate editors, subject editors and the editor-in-chief, the manuscript is either accepted,

rejected or sent back for revision.

As the manuscript passes each stage of the review process, there is a greater scrutiny of its content by multiple individuals who make an independent assessment of the manuscript's perceived quality, impact and fit with the journal's scope. It is reasonable to expect any extraneous factors to have a smaller effect (if any) at the later stages of the review. However, the first decision in the review process, to desk reject or not, involves just one individual who, despite being an expert, is only human. The editor could suffer from the same behavioural biases as other human experts, with relevant examples provided in Section 2. The evidence provided in this paper suggests that the editorial decision to desk reject a manuscript is influenced by the manuscript's sequential order, with the manuscripts appearing earlier in the (imputed) list being up to 7% more likely to be desk rejected. The analysis includes exploration of several competing explanations for this behavioural pattern.

The remainder of the paper is organised as follows. Section 2 contains a review of relevant literature. Section 3 describes the available data and provides key descriptive statistics. The empirical strategy for identifying the impact of sequential order is explained in Section 4. Section 5 presents the main results and Section 6 concludes the paper. Robustness of the results is explored in the Appendix.

2 Related literature

The current paper examines whether the editor's decision to desk reject is influenced by the order in which manuscripts are presented. There are two related sets of prior research: the literature on behavioural biases, specifically order effect and decision fatigue, and the literature on the editorial and peer review processes.

There is empirical evidence on behavioural biases of human experts in multiple contexts, including high-stake decisions (Danziger et al. 2011; Feenberg et al. 2017; Berger 2016). One particular bias, order effect, arises from a sequential or a list-based presentation of available options.

In a recent study, Feenberg et al. (2017) show that the position of a link to an academic document on an email newsletter has a positive causal effect on the number of times the article is viewed, downloaded and even cited. Being the top

3

item on the list is associated with approximately 30% increase in document views, downloads and citations (Feenberg et al. 2017). While there are multiple explanations that could generate such outcomes, Feenberg et al. (2017) argue that the most consistent explanation is offered by a model of skimming, in which the (time-constrained) readers allocate disproportionate attention to items in prominent positions. Berger (2016) shows that articles which appear earlier in an academic journal are cited more frequently and this effect persists over time. Berger (2016) exploits the alphabetic ordering of the journal's section to argue that editor's ordering of articles within a section has limited influence on the overall ranking of the article within the journal. Berger (2016) reports a significant improvement in citations from being the first article in the journal. Findings by Feenberg et al. (2017) and Berger (2016) are consistent with earlier research on the importance of leading positions, see Hamermesh (2017) for further references.

Prior research examined the potential impact of extraneous factors on the peer review and editorial decisions. A recent study by Kwan et al. (2016) examines whether presentation order of submitted manuscripts affects editorial decisions at an academic journal. By analysing almost 650 submissions and their detailed characteristics, Kwan et al. (2016) reassuringly find that the presentation order did not have an effect on the final editorial decisions about accepting a manuscript. However, Kwan et al. (2016) considered only manuscripts that passed peer review, and which were then discussed in a group setting. This contrasts with the focus of the present paper on decisions by a single individual, the editor, about desk rejecting a manuscript.

The editorial decision to desk reject requires a judgement, with incentive to make it a quick one. McAfee (2010) describes his experience as an editor: "I reject 10–15% of papers without refereeing … If, on reading a paper, I find that there is no chance I am going to publish a paper, why should I waste the referees' time and make the author wait? … The amount of time necessary to establish beyond a reasonable doubt that a paper is not suitable for a journal ranges from a few minutes … to many hours".

Improvement in the speed of first decision can make a journal more attractive to authors, with a resulting increase in the submission volume. Some editors wonder if the increased submissions to their journals are going to affect the quality or

accuracy of editorial decisions (Stewart et al. 2012). One of the questions asked by Stewart et al. (2012) is whether the editor's decisions are influenced by the number of manuscripts submitted on the same day? They report an increased desk rejection rate on days with large submission volume, specifically the desk rejection rate is 64% when the editor reviews 1 or 2 manuscripts a day, and it increases to 70% for 3+ reviewed manuscripts per day. Stewart et al. (2012) provide other evidence consistent with decision fatigue of the editors, but stop short of a formal investigation of this hypothesis. In a relevant study, Johnston et al. (2007) conducted a randomized trial to examine if desk rejection would significantly disadvantage the authors due to potential editorial bias. They find that desk rejections achieve a significant reduction in the time to decision, without adverse effects on reported reviewer ratings of scientific impact for accepted manuscripts.

The present manuscript contributes to the literature by showing the impact of sequential order on the editorial decision to desk reject a manuscript. It also contrasts predictions of order effect and decision fatigue hypotheses, with evidence in favour of editorial decision fatigue.

3   Data and descriptive statistics

Data provided by Journal X includes information on all submissions between late 2008 and early 2014. Each of 12 thousand entries in the dataset contains the computer-generated manuscript ID, date of manuscript submission, the editor's decision and its date, as well as the dates of receiving referee and associate editor's reports (if any).

The manuscript ID is generated by the editorial software when the author clicks the submit button and the submission is accepted by the editorial software.[1] The ID follows a fixed format of "JX-SEQ-Y.RN", where JX is abbreviation of Journal X, SEQ is a numeric sequential code, Y is the year of submission and RN is the revision number (only applicable to revised submissions).[2] The exact

---

[1]This and other software-related points were clarified through e-mail exchange with the software provider's technical support team and analysis of their technical documentation.

[2]The format was changed in 2009 by reordering parts of the manuscript ID and using two-digit year representation.

timing of submission is not known, but combining the day of submission with the sequential component of the manuscript ID allows imputing the order in which the manuscripts were submitted (see Appendix for robustness of this approach).

The main dependent variable used in this paper is a dummy for a first submission manuscript that is desk rejected by the (subject) editor. The dummy is set to 1 for rejected first submission manuscripts if between the submission date and decision date there were no reports submitted by the associate editor or any number of reviewers. If a report by an associate editor or at least one reviewer was submitted then the dummy is set to 0. For robustness, a weaker definition of desk rejection also includes rejections by associate editor (i.e. no reviewer reports were submitted). If an associate editor and any number of reviewers submitted a report, then the manuscript is considered to have gone through peer review process, regardless of the number of reports (and there is no information on their possible content). Table 1 provides an overview of the editorial decisions, showing the high rejection rate for the journal — about 80% of all first submissions are rejected.

The editor's goal is to select the best manuscripts for the limited journal space, so the manuscripts are judged not only against some journal-specific benchmark, but also against other submitted manuscripts. When the number of submissions increases, the editor would like to pick a limited number of 'the best' manuscripts, which implies increased desk rejection rate. However, the editor also has a limited time budget and might not be able to allocate the same amount of scrutiny when facing a large inflow of manuscripts. In this case, to avoid mistakingly desk rejecting the best manuscripts, the editor can respond by reducing the desk rejection rate, delegating closer scrutiny to the associate editor and referees. Ultimately, the limited availability of referees and fixed journal space imply that once the submission volume reaches a critical level, any editor will respond by increasing desk rejection rate. Below such critical level, the editor could demonstrate different behaviour depending on their workload, as well as the journal's 'spare capacity' in terms of associate editor and referee availability and remaining journal space.

Table 2 shows that actual decisions by editors appear to be consistent with the limited time budget interpretation: as the number of first submissions (per subject per day or week) increases, the editors pass on the detailed scrutiny of submissions to associate editors and reviewers. This broad pattern however masks a significant

6

Table 1: Editorial decisions over time and revision cycle.

(a) Decisions on first submissions.

	Submission year							
	2008	2009	2010	2011	2012	2013	2014	All years
Accept	0	0	0	0	0	0	0	0
Minor revision	1	1	1	1	1	1	0	1
Major revision	21	20	17	18	16	15	2	16
Reject	69	54	60	57	60	59	22	56
Desk reject	9	25	21	24	22	23	20	22
No decision	0	0	0	0	1	3	56	4
Total	100	100	100	100	100	100	100	100

(b) Decisions on all revisions.

	First submission	Revision round				
		1	2	3	4	5
Accept	0	11	50	76	78	100
Minor revision	1	28	32	19	17	0
Major revision	16	32	8	2	0	0
Reject	56	23	6	1	3	0
Desk reject	22	0	0	0	0	0
No decision	4	6	4	1	2	0
Total	100	100	100	100	100	100

Notes: these tables show editorial decision on first submissions over time or on all submission over revision version; desk reject is defined as a rejection without requesting referee or associate editor reports; rejection and 'other' editorial decisions are conditional on passing the desk rejection stage; 'other' editorial decisions are primarily minor and major revisions (the sample contains fewer than 5 acceptances out of all first submissions).

Table 2: Decisions at different volumes of first submissions.

(a) By daily volume of first submissions.

| # of submissions per subject per day | Decision type | | | Relative frequency of occurence |
	Desk rejection	Reject	Other	
1	24.0	58.2	17.8	54.9
2	24.6	58.0	17.3	25.5
3	19.8	61.3	18.9	10.8
4	17.9	65.1	17.1	3.4
5+	7.4	75.3	17.3	5.3

(b) By weekly volume of first submissions.

| # of submissions per subject per week | Decision type | | | Relative frequency of occurence |
	Desk rejection	Reject	Other	
1	25.0	57.0	18.0	8.6
2	24.3	57.7	18.0	13.1
3	29.0	55.7	15.3	14.5
4	25.3	59.3	15.4	14.9
5	21.9	57.1	21.0	11.4
6	24.1	58.6	17.4	7.9
7	20.4	59.3	20.3	9.3
8	23.4	57.7	18.9	5.1
9	16.8	63.3	19.9	4.0
10+	10.0	72.7	17.3	11.2

Notes: the table shows relative frequency of editorial decisions depending on the number of first submissions per unit of time per subject between late 2008 and early 2014.

heterogeneity in editor response to increased first submission volume, see Figure 1. Given that the journal's subjects are distinct and have non-overlapping pools of referees, this heterogeneity is not surprising. Also note that the relative frequency of 'Other' decisions (primarily minor and major revisions) remains quite stable even when the rate of desk rejections declines. This could mean that during peer review the manuscript is evaluated against some relatively fixed benchmark, especially since referees typically review only one submission at a time. However, the data doesn't contain information on the reviewer recommendations (only dates of submitted reports), so this pattern can also be explained by the editor accepting only a given number of the best manuscripts for the limited journal space.
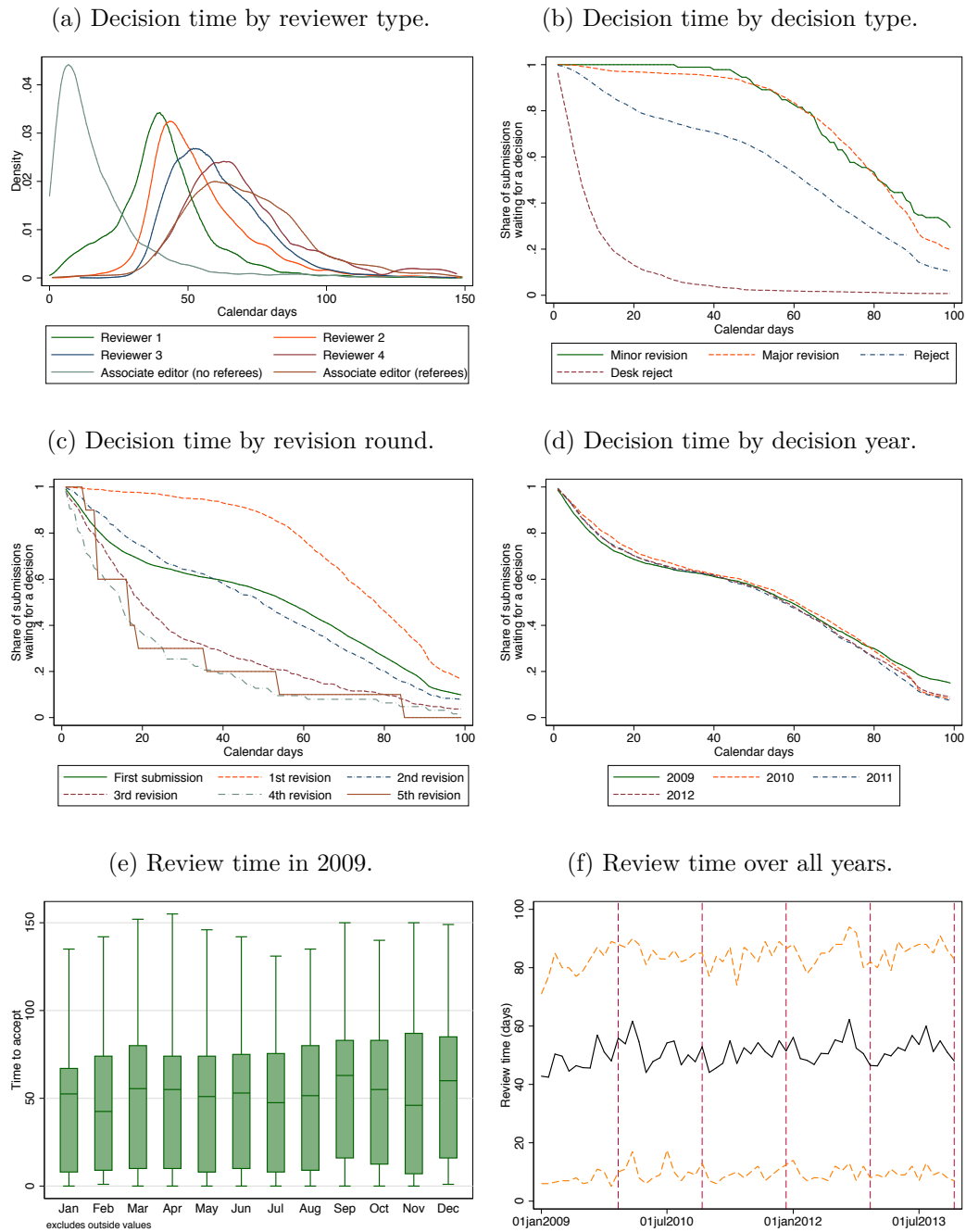
Figure 1: Differences in editor response to increased submission volumes.
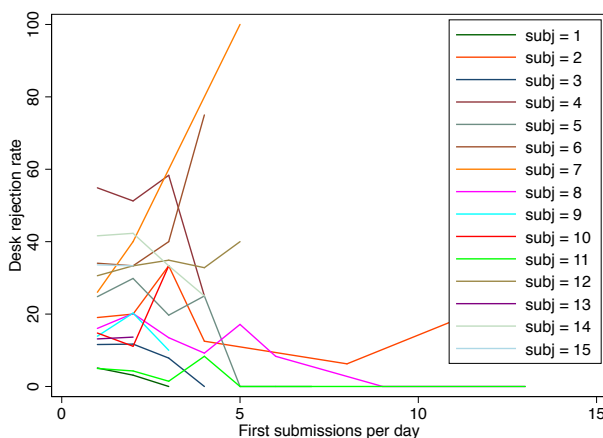
Note: includes all first submissions between 2009–2013.

Figure 2 shows the distribution of decision time by various characteristics. The distributions of decision time are different by decision type, suggesting quite different underlying processes (esp. presence of peer review). Panels (a) and (b) show that desk rejections are communicated quickly. Note that the decision on desk rejection could be taken soon after the submission, while the decision date could indicate the earliest convenient time that the editor (or her assistant) could enter the decision into the editorial system. Also note that the differences in distributions of referee timings in panel (a) are an artefact, since the dates of referee decisions are entered in the order in which they were received. Panel (c)

Figure 2: Distribution of review time for first submissions by selected characteristics.

(a) Decision time by reviewer type.

(b) Decision time by decision type.

(c) Decision time by revision round.

(d) Decision time by decision year.

(e) Review time in 2009.

(f) Review time over all years.

Note: includes all submissions, including desk rejects, between 2009–2013.

10

shows that the analysis must take into account the stage of submission, since revised submissions follow quite different time distributions reflecting involvement of multiple referees, along with the editor and associate editor, at advanced stages of peer review process. Panel (d) shows that the overall patterns are broadly the same over the sample period, 2009–2012. Panels (e) and (f) show that there is no obvious aggregate-level seasonal pattern in the review times.

4   Empirical strategy

Manuscript- and journal-level patterns described in the previous section suggest that an empirical analysis of editorial decision-making must take into account the different underlying processes of different stages of peer review. Given the limited data on submitted manuscripts and individuals involved in the review process, the analysis will focus on the very first editorial decision — to desk reject a manuscript or not.

The hypothesis explored in this paper is whether factors other than the submitted manuscript's content can affect the corresponding editorial decision. Specifically, does the sequential order in which the manuscripts were viewed by the editor influence the editor's decision to desk reject a manuscript?

The order in which manuscripts were viewed by the editor is not known. Given the importance of priority claim in academic research, as well as the default setting in the editorial software of presenting the manuscripts in order of their receipt, it's reasonable to expect that processing of first submissions will be done in the first-in-first-out sequence (See Appendix for an examination of heterogeneity in subject-level preferences). The order of submission is imputed from the manuscript ID and the day of submission. It is not feasible for authors to anticipate or influence the sequential rank of their manuscript at the time of submission, hence the manuscript's actual rank in the list of submissions is an exogenous variable.[3]

Another important consideration is the frequency with which the editor reviews the submissions. The journal's editorial board made a commitment to achieve

---

[3]Dietrich (2008) discusses an intriguing possibility of geography-determined bias due to differences in time zones, but finds no empirical support for this hypothesis. See Appendix for additional discussion of the rank exogeneity.

quick turnaround times, with a specific time limit for (subject) editors to assign the paper to an associate editor within five days of submission. This means that the editor is likely to check frequently for new submissions. The default assumption used in this paper is that the submissions are checked on a daily basis, but alternative frequencies are also explored in Appendix.

Consider first submission of manuscript i to the journal's subject j which is reviewed by the editor at time t, and let $x_{ij,t}$ be a dummy variable equal to 1 if the submission is desk rejected and 0 otherwise (regardless of the eventual decision on the submission). Suppose that the manuscript's sequential rank (either in terms of submission time or the editor's viewing order) is given by $R_{ij,t}$. Then the causal impact of sequential order in editorial decision will be tested by estimating the following logit specification:

$$\text{Prob}(x_{ij,t} = 1) = \Lambda\left(\alpha + \beta \times R_{ij,t} + \gamma_{j,t}\right), \tag{1}$$

where $\Lambda(y) = (1 + exp(-y))^{-1}$, $\gamma_{j,t}$ captures unobserved heterogeneity[4] specific to subject j at time t, and α is a constant term. A non-zero estimate of β will indicate that sequential order matters in desk rejection or other editorial decisions (if the dependent variable is substituted with another dummy). This specification is the most flexible approach allowing identification of the impact of the sequential order among all submissions viewed by the editor in subject j at time t.

A linear probability formulation of Equation 1 yields similar magnitude and significance of rank variables. Additional specifications explore the importance of the first and last positions in the list. The robustness of results is examined in the Appendix.

Multiple hypotheses are consistent with a non-zero β, but two specific explanations, decision fatigue and order effect, will give different predictions about the sign of β. These predictions will also depend on what is the 'default' choice of an editor when in doubt. The editor could be 'hard' and desk reject a manuscript unless clearly convinced to do otherwise, this could be especially true in the context

---

[4]Hausman specification test suggests that random effects estimation is consistent, but given the more restrictive identifying assumptions of random effects, estimation uses fixed effects. However, random effects estimates are very similar in sign and significance, but have a slightly smaller magnitude, see a sample table included in the Appendix.

of a journal that is operating at full capacity (there is only a limited journal space left and there are no spare referees). At the other extreme, the editor could be 'soft' and always seek a second opinion (from associate editor or referees) to avoid rejecting a "Market for lemons" (McAfee 2010; Shugan 2007; Gans and Shepherd 1994).

Decision fatigue predicts that as the editor works through the list of submissions, she exhausts the limited amount of energy and effort. The fatigued editor is then more likely to choose a default option, a 'safe' choice. The 'hard' editor's β would be then be positive, while the 'soft' editor would have a negative one. Conditional on the editor type, there would be no change in sign as the editor works through the list.

Order effect predicts that the editor facing a list would allocate disproportionate attention to prominent positions on the list, such as the top and bottom positions (Feenberg et al. 2017). The 'hard editor' would then be more likely to reject manuscripts which are not in prominent positions, since she allocates less time to those manuscripts and hence is more likely to miss an important footnote or misread a particular figure. A 'soft editor'[5] is also more likely to be in doubt about manuscripts in non-prominent positions, but by definition of her type she will avoid desk rejecting such manuscripts and will seek a second opinion from the associate editor or referees.[6] Conditional on the editor type, we would expect the manuscripts in prominent positions to have the same β compared to other manuscripts. The sign of β would be negative for a 'hard' editor and positive for a 'soft' editor.

These two hypotheses give different predictions about the importance of sequential rank. By comparing the estimated β for linear rank, as well as for the first and last positions, it's possible to discard one of these hypotheses.

---

[5]This term is taken from an interview with QJE editor: `https://blogs.worldbank.org/impactevaluations/qa-with-larry-katz-editor-of-qje`.

[6]Shugan (2007) provides an example of the editorial calculus on p. 594. Assuming, for example, that an article in a journal gets two citations on average, "… missing one highly cited article causes substantial opportunity cost. In this example, publishing one article that gets 300 citations compensates for publishing 149 articles that get zero citations."

5 Main results

Equation 1 was estimated for dummies indicating a desk rejection, a weak desk rejection (i.e. rejection by the editor or associate editor, but without referees), rejection by associate editor, and rejection by referees.[7] Table 3 shows that as a manuscript's sequential rank increases, its probability of a (weak) desk rejection decreases, without any effect on the decision after peer review. Regardless of the underlying causal mechanism, decision fatigue or order effect, the negative sign of β indicates that the editors at Journal X are of 'soft' type, seeking a second opinion when in doubt.

Table 3: The role of manuscript's rank at different stages of the manuscript review.

| | | Probability of rejection by | | |
	editor	editor or associate editor	associate editor	referee(s)
Rank	-0.043***	-0.014***	-0.008**	-0.008
	(0.010)	(0.005)	(0.004)	(0.005)
Observations	1435	2212	1257	916
Groups	550	801	422	330
Min group	2	2	2	2
Avg group	2.61	2.76	2.98	2.78
Max group	35	35	34	25

Notes: *** ($p<0.01$), ** ($p<0.05$), * ($p<0.1$); this table shows marginal effects of the variables estimated using 'xtlogit, fe' with bootstrapped standard errors provided in the parentheses (2000 replications); all first submissions to a particular subject on the same day are grouped together and manuscript rank within the group is imputed using manuscript ID.

Also note that if the editors were more likely to make a mistake as they work through the list of submissions, then the average quality of manuscripts they pass on to the associate editor and referees would decrease with increasing rank. This would be reflected in a changing sign of the rank variable as the manuscript receives additional scrutiny from the associate editor and referees. Hence, there is no evidence of persistent mistakes.

To distinguish decision fatigue and order effect, Equation 1 was estimated with dummies for the first and last rank. The sign of the dummies in panels (a)

---

[7]There is no information on the content of referee recommendations, but it is assumed that the editor follows their advice. Hence, the referee recommendations can be proxied by the final decision on the manuscript.

Table 4: Additional rank measures.

(a) First rank.

| | | Probability of rejection by | | |
	editor	editor or associate editor	associate editor	referee(s)
First rank	0.070***	0.074***	0.076***	-0.005
	(0.022)	(0.018)	(0.025)	(0.032)
Observations	1435	2212	1257	916
Groups	550	801	422	330

(b) Last rank.

| | | Probability of rejection by | | |
	editor	editor or associate editor	associate editor	referee(s)
Last rank	-0.064***	-0.072***	-0.072***	-0.014
	(0.022)	(0.018)	(0.026)	(0.029)
Observations	1435	2212	1257	916
Groups	550	801	422	330

(c) Comparing first and last ranks.

| | | Probability of rejection by | | |
	editor	editor or associate editor	associate editor	referee(s)
First rank	0.057	0.047	0.049	-0.036
	(0.039)	(0.031)	(0.039)	(0.048)
Last rank	-0.016	-0.034	-0.035	-0.039
	(0.039)	(0.031)	(0.041)	(0.044)
Observations	1435	2212	1257	916
Groups	550	801	422	330
Difference in coefficients	0.073***	0.081***	0.084***	0.004
Equality test (p-value)	0.001	0.000	0.002	0.907

Notes: *** (p<0.01), ** (p<0.05), * (p<0.1); this table shows marginal effects of the variables estimated using 'xtlogit, fe' with bootstrapped standard errors provided in the parentheses (2000 replications); all first submissions to a particular subject on the same day are grouped together and manuscript rank within the group is imputed using manuscript ID.

and (b) of Table 4 suggests that decision fatigue is the more likely explanation, since for an editor who is influenced by both the first and last positions on a list, the last ranked manuscript would have the same sign as the first ranked one.[8] The two rank dummies are strongly correlated, Spearman correlation is -0.62, so including both of them increases standard errors and leads to estimates that not significantly different from zero. However, a formal coefficient equality test in panel (c) suggests that the first rank is significantly different from the last rank, which is also consistent with the decision fatigue interpretation.

Robustness of these patterns is explored in the Appendix.

6 Conclusion

The analysis presented in this paper shows that first submissions on a given day are more likely to be desk rejected. There is no doubt that manuscript's content is the primary determinant of editorial decisions, but the sequential order appears to also matter. Given the evidence on exogeneity of sequential order, at least two hypotheses are consistent with this behaviour: decision fatigue and order effect. Decision fatigue hypothesis states that editors are more likely to go with a 'safe' decision, such as to seek a second opinion, as they work through the submissions. Order effect hypothesis suggests that when presented with a list of options, the journal editors might allocate disproportionate attention to 'prominent' positions, such as the top or bottom items on the list. The evidence presented in this paper suggests that the editors are prone to decision fatigue, although of course other explanations could be possible.

The manuscript's sequential order does not have a persistent effect on the submitted manuscript, since conditional on passing a desk rejection there is no statistically significant relationship between order and the probability of rejection after peer review. Desk-rejected manuscripts could, in general, benefit from additional comments by the associate editors and referees, however there is no guarantee that the comments would be sufficiently different from the brief message that usually accompanies a desk rejection. Moreover, some authors might instead prefer receiv-

---

[8]Assuming that the first and last positions are the prominent ones.

ing a quick answer, so that they can seek alternative outlets or have more time for improving their work.

Editorial decision fatigue suggests that the manuscript review process could be improved by allocating only a specific number of manuscripts per editor or to permit a degree of randomisation in presentation order of submitted manuscripts to the editor. The priority claim is important in academia, so the exact submission date must continue to be recorded, but modern editorial software can allow for a randomised presentation order. This wouldn't have particular advantages (or disadvantages) to the editor, but would reduce any potential incentive for strategic timing of submissions. It would also reduce other potential extraneous influences, for example the geographical bias hypothesis, which could not be tested explicitly in this paper, suggests that there could be exogenous factors affecting sequential order of submissions. The impact of extraneous factors which might influence an author's exposure to the sequential order effect would be diminished through randomisation.

For authors, being first might not necessarily be 'good'. There is no impact on the ultimate decision on the manuscript, which is the main outcome of interest for most authors. The authors that prefer feedback to a quick decision, might wonder if they can time their submission strategically to place it lower on the editor's list. However, estimating the likely sequential order of the manuscript upon submission is difficult, if not impossible, moreover there is no guarantee that a submission early in the day won't be appended to the previous day's submissions, or that the editor won't decide to take a few days off and let the manuscripts accumulate. Overall, the evidence suggests that the time and effort are likely to be more productive if spent on the manuscript, rather than timing its submission.

More broadly, this research contributes to the literature by providing another case in which human experts are susceptible to behavioural biases. Involvement of multiple experts in the review process appears to mitigate the effect of sequential order, which has implications for organisation of work in list-based or sequential processing environments. For example, the accuracy and reliability of quick decisions made on job or university admission applications might be affected by the order in which the applications are reviewed. While 'peer review' analogues might not be feasible in some contexts, involvement of additional independent evaluators

17

could be one way of reducing cognitive fatigue. A more thorough examination of the peer review process and how it reduces the impact of extraneous factors will require additional data. Such data is recorded and collected by modern editorial software and, hopefully, will become available to researchers in the future.

References

Berger, Jonah (2016). "Does Presentation Order Impact Choice After Delay?" *Topics in cognitive science* 8, pp. 670–684. DOI: `10.1111/tops.12205`.

Danziger, Shai, Jonathan Levav, and Liora Avnaim-Pesso (2011). "Extraneous factors in judicial decisions". *Proceedings of the National Academy of Sciences* 108 (17), pp. 6889–6892. DOI: `10.1073/pnas.1018033108`.

Dietrich, JP (2008). "Disentangling visibility and self-promotion bias in the arxiv: astro-ph positional citation effect". *Publications of the Astronomical Society of the Pacific* 120 (869), p. 801.

Feenberg, Daniel R, Ina Ganguli, Patrick Gaule, and Jonathan Gruber (2017). "It's Good to be First: Order Bias in Reading and Citing NBER Working Papers". *Review of Economics and Statistics* 99 (1), pp. 32–39.

Gans, Joshua S and George B Shepherd (1994). "How are the mighty fallen: Rejected classic articles by leading economists". *Journal of Economic Perspectives* 8 (1), pp. 165–179.

Hamermesh, Daniel S. (2017). "Citations in Economics: Measurement, Uses and Impacts". *Journal of Economic Literature.*

Johnston, S Claiborne, Daniel H Lowenstein, Donna M Ferriero, Robert O Messing, Jorge R Oksenberg, Stephen L Hauser, and Adam F Stewart (2007). "Early editorial manuscript screening versus obligate peer review: a randomized trial". *Annals of neurology* 61 (4).

Kwan, James, Libby Stein, Sean Delshad, Sunny Johl, Hannah Park, Bibiana Martinez, Lindsey Topp, and Brennan MR Spiegel (2016). "Does "Decision Fatigue" Impact Manuscript Acceptance? An Analysis of Editorial Decisions by the American Journal of Gastroenterology". *The American Journal of Gastroenterology.*

McAfee, R Preston (2010). "Edifying editing". *The American Economist* 55 (1), pp. 1–8.

Shugan, Steven M. (2007). "The Editor's Secrets". *Marketing Science* 26 (5), pp. 589–595.

Stewart, Adam F, Donna M Ferriero, S Andrew Josephson, Daniel H Lowenstein, Robert O Messing, Jorge R Oksenberg, S Claiborne Johnston, and Stephen L Hauser (2012). "Fighting decision fatigue". *Annals of neurology* 71 (1), A5–A15.

Appendix: robustness of the results

Exogeneity of rank and test for specification errors

A possible explanation for the main results is that there is another variable that simultaneously results in a higher rank and a lower quality of submitted manuscript (in turn leading to higher likelihood of desk rejection). In a different context, Dietrich (2008) suggested that geographic location of the submitting authors could lead to a correlation between their position on a list and other factors affecting the (perceived) quality of their work. For example, if the journal editor is based in Kazakhstan, then submissions from Europe or United States, which on average produce manuscripts of 'higher quality' (at least in terms of desk rejection probability), will be received late in the day. In this case, there would be a correlation between higher rank and lower likelihood of desk rejection.

The data from Journal X does not have manuscript-specific information, such as the number of authors, their identities (which could be linked to prior work) or location. As a result, it's not possible to control for a geographical bias in ranking. However, two steps were taken to address this concern.

First, for a small number of manuscripts that were eventually accepted (after multiple stages of review), it is possible to use the submission and acceptance dates to match the computer-generated first submission IDs to the Digital Object Identifiers of published manuscripts. After matching DOI to (accepted) manuscript ID, it's possible to examine if the sequential rank on the day of first submission has any correlation with some proxies for quality — the publication's citation count or the number of downloads. Table 5 shows that sequential rank on the first

19

submission day is not a predictor of manuscript quality.

Table 5: Bibliometric measures of (published) manuscript quality and the sequential order on the day of first submission.

| | Citation and download counts | | | |
	Citations		Downloads	
Rank (daily)	-0.745		-9.693	
	(0.486)		(7.200)	
No. of submissions (daily)	0.060	-0.467	4.851	-1.390
	(0.436)	(0.255)*	(5.684)	(2.137)
Rank (daily) = 1		-2.186		1.098
		(1.471)		(13.319)
N	690	690	690	690
Time FE	Yes	Yes	Yes	Yes
Subject FE	Yes	Yes	Yes	Yes

Notes: *** (p<0.01), ** (p<0.05), * (p<0.1); this table shows marginal effects of the variables based on Poisson estimation similar to Equation 1 but with the dependent variable equal to the citation or download count of the published manuscript (no longer a logit framework), heteroskedasticity-robust standard errors are provided in the parentheses; time fixed effects include dummies for year, week and day of the week derived from the manuscript submission date and also dummies for the volume and issue in which the manuscript was published (to capture potential differences in length of time since the manuscript was published); subject fixed effect is a set of dummies for the subject area of the assigned editor.

Second step is to test for a potential specification error using a link test. If the model is properly specified, then the square of the predicted value should have no explanatory power. This is confirmed using 'linktest' command in Stata after the estimations, all squared predicted values are insignificant. Prediction tables could not be generated given the fixed effects specification.

Measurement of rank: accuracy of manuscript ID-based approach

The rank variable is imputed from the manuscript ID-based ordering. This approach is a good proxy for the time of submission at long time scales, for example at the annual frequency the correlation between the manuscript ID-based rank and time of submission-based rank is 0.99 for both Pearson and Spearman correlations. However, as the time scale gets shorter the accuracy of this approach decreases. This is caused by the following feature of the editorial software: whenever an author modifies their submission the original manuscript ID is retained but the date

of 'first submission' is updated to the modification date. For example, if an author submitted a manuscript in August of 2010 and was assigned ID 'JX-001-10', but decided to modify the submission in January 2011, then their manuscript ID is retained with the 'new' submission date in January 2011.

Since the sequential nature of manuscript IDs is exploited as proxy for the submission order, such submission modifications introduce noise into the measurement of rank at high resolution time scales. For example, on a weekly scale, correlation between time of submission-based and manuscript ID-based ranks is 0.92 for Pearson and 0.90 for Spearman. Without knowing the exact submission time it's not possible to check how accurate this approach is on a daily scale. However, since the correlation decreased by about 10% when the time scale increased by a factor of 50, assuming that the decline in correlation is proportional, the accuracy on a daily scale is estimated to be well above 0.8 for Spearman correlation. This allows using the manuscript ID-based approach as a reasonable proxy for the actual time of submission.

Subject heterogeneity

Submission to different subject fields vary in their volume, some subjects are more popular and as a result have a larger number of editors to handle the submissions. To check whether the order effect varies depending on the number of subject editors, Equation 1 was modified by interacting the subject field dummies I_j with manuscript rank:

$$\text{Prob}(x_{ij,t} = 1) = \Lambda\left(\alpha + \beta \times I_j \times R_{ij,t} + \gamma_{j,t}\right). \tag{2}$$

Figure 3 shows the marginal effects for rank measures over different subjects. There is some heterogeneity in the rank effects, with strong patterns for subjects with 3–5 editors (subjects 6, 8 and 12). Subjects with 1–2 editors have low submission volumes, which can partly explain the large confidence intervals. Also, that the default setting in the editorial software is to sort incoming manuscripts in the order of submission, but an editor can reverse that order with a single click. Individual editors that do not need to coordinate with colleagues could have specific preference in processing the submissions, which would reduce the accuracy of

Figure 3: Marginal effect of a change in rank.
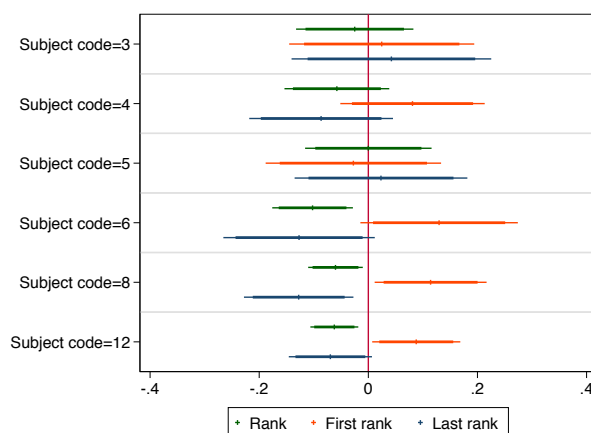
Notes: this figure shows the marginal effects of rank measures; each marginal effect was calculating by running 'xtlogit, fe' with one rank measure at a time; the sample includes only subjects with a large number of submissions; thick and thin lines represents 90% and 95% confidence intervals, respectively.

rank measure (e.g. if the editor checks only once or twice a week).

Measurement of rank: backlog accumulation

Each editor is likely to have their own schedule and time preferences, and hence daily rank might be an inaccurate measure of the order in which the manuscripts are viewed due to accumulation of manuscript backlogs. The editor's typical schedule could vary over the sample period, since the sample spans several academic terms. To visualise the editorial activity, the decisions on all submissions (including revisions) were aggregated to subject field-day level and relative intensity of activity on any particular day was calculated, see Figure 4. Panel (a) shows that a large number of editors (to process a larger number of submissions) had activity spread more evenly over the working week, panel (b) for the subject with just 1 editor shows more concentrated activity. This approach relies crucially on the number and types of decisions. Panel (c) shows the activity of a specific editor based on data for published manuscripts. This intensity is based only on decisions to accept, and hence unlikely to fully capture all of the journal-related activities.

Using this decision activity as a proxy for editor's activity on processing new

Figure 4: Intensity of daily activity for selected subject fields and individual editors.

(a) Subject 8 (5 editors)   (b) Subject 15 (1 editor)   (c) Editor A
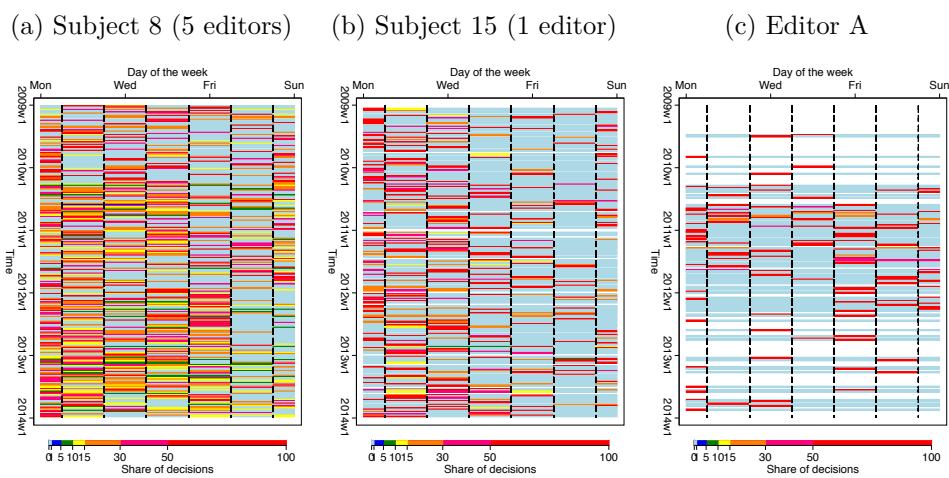
Notes: the figures show imputed work patterns of editors based on dates of decisions; light blue indicates days with no activity (during weeks with at least one active day), coloured lines correspond to the intensity specified on the legend bar, and white spaces indicate weeks with no activity; in panels (a) and (b) the activity is based on information on all submissions (including revisions); in panel (c) activity is derived from accepted manuscript only (due to data limitation).

23

submissions, manuscript backlogs were estimated by accumulating first submissions during the editor's inactive days. The resulting manuscript ranking is highly correlated with the daily ranking (R_i): Pearson 0.77, Spearman 0.61. It was expected that this adjusted manuscript ranking would improve estimation for single-editor subjects, however the rank variable is not significant if the sample is restricted to single-editor subjects. This could indicate that the decision data is not an accurate reflection of the timing of first submission review by the editors at the single-editor subject areas, so the rank variable is measured with a lot of noise. It could, of course, also indicate that due to the low submission volume the single editor could have more flexibility in integrating submission review into their existing work schedule reducing any decision fatigue or order effect.[9]

Large number of submissions

During the sample period the journal had three special issues and a call for papers for another one (deadline was beyond the sample period). A potential concern is that the rank effects could be driven by a number of high-quality submissions which were prepared in advance and submitted before the deadline. In contrast, manuscripts of lower quality could be submitted just before or on the deadline. This would introduce correlation between manuscript quality and imputed rank. The data doesn't allow distinguishing a special issue submission from a regular submission, moreover special issues could include submissions from multiple subjects. To reduce potential bias caused by these episodes of large submission inflows, two sample restrictions were checked. The first sample restriction excluded all submissions in the months containing a special issue deadline (excluding submissions on the deadline day only could bias the sample by excluding potentially lower-quality submissions). The second sample restriction excluded all submissions on days with more than 5 submissions. The results from these estimations are consistent with the main results.

---

[9]Variations to the backlog accumulation procedure were also considered, including a model of editor's working days that assumed a fixed work pattern within any year-month combination. For example, if the editor worked on a Tuesday in January 2010, then every Tuesday in that month was considered to be a working day, even if no activity was observed in the decision data on that day. These modifications did not lead to any meaningful changes in the estimates.

Table 6: Restricting the sample to 2–5 manuscripts per day.

(a) Sequential rank.

	editor	Probability of rejection by editor or associate editor	associate editor	referee(s)
Rank	-0.056***	-0.051***	-0.043***	-0.000
	(0.014)	(0.011)	(0.016)	(0.020)
Observations	1305	1919	982	736
Groups	542	782	404	312

(b) Comparing first and last ranks.

	editor	Probability of rejection by editor or associate editor	associate editor	referee(s)
First rank	0.063	0.048	0.045	-0.047
	(0.040)	(0.031)	(0.042)	(0.050)
Last rank	-0.009	-0.033	-0.039	-0.044
	(0.041)	(0.032)	(0.044)	(0.048)
Observations	1305	1919	982	736
Groups	542	782	404	312
Min group	2	2	2	2
Avg group	2.41	2.45	2.43	2.36
Max group	5	5	5	5
Difference in coefficients	0.072***	0.080***	0.084***	-0.004
Equality test (p-value)	0.001	0.000	0.002	0.914

Notes: *** (p<0.01), ** (p<0.05), * (p<0.1); this table shows marginal effects of the variables estimated using 'xtlogit, fe' with bootstrapped standard errors provided in the parentheses (2000 replications); all first submissions to a particular subject on the same day are grouped together and manuscript rank within the group is imputed using manuscript ID; the sample is restricted to observations with at most 5 submissions per subject per day.

Table 7: Removing submissions during months with a deadline for a special issue.

(a) Sequential rank.

| | | Probability of rejection by | | |
	editor	editor or associate editor	associate editor	referee(s)
Rank	-0.061***	-0.050***	-0.035**	-0.012
	(0.013)	(0.010)	(0.014)	(0.019)
Observations	1239	1798	921	719
Groups	512	729	373	297

(b) Comparing first and last ranks.

| | | Probability of rejection by | | |
	editor	editor or associate editor	associate editor	referee(s)
First rank	0.060	0.035	0.027	-0.023
	(0.042)	(0.033)	(0.045)	(0.053)
Last rank	-0.014	-0.046	-0.051	-0.014
	(0.041)	(0.033)	(0.046)	(0.051)
Observations	1239	1798	921	719
Groups	512	729	373	297
Min group	2	2	2	2
Avg group	2.42	2.47	2.47	2.42
Max group	16	16	14	11
Difference in coefficients	0.074***	0.081***	0.078***	-0.009
Equality test (p-value)	0.002	0.000	0.006	0.777

Notes: *** (p<0.01), ** (p<0.05), * (p<0.1); this table shows marginal effects of the variables estimated using 'xtlogit, fe' with bootstrapped standard errors provided in the parentheses (2000 replications); all first submissions to a particular subject on the same day are grouped together and manuscript rank within the group is imputed using manuscript ID; the sample is restricted to submissions during months when there was no deadline for a special issue (the data does not allow identifying specific manuscripts submitted for a special issue).

Random effects specification

Using random effects specification moderates the magnitude of the coefficients, but sign and significance are similar to the fixed effects specification.

Table 8: Estimation with random effects specification.

(a) Sequential rank.

	editor	Probability of rejection by editor or associate editor	associate editor	referee(s)
Rank	-0.030***	-0.010**	-0.001	-0.001
	(0.006)	(0.004)	(0.003)	(0.002)
Observations	9278	9278	7016	5294
Groups	6878	6878	5405	4277
Min group	1	1	1	1
Avg group	1.35	1.35	1.30	1.24
Max group	35	35	34	25

(b) Comparing first and last ranks.

	editor	Probability of rejection by editor or associate editor	associate editor	referee(s)
First rank	0.040***	0.045***	0.017	0.016
	(0.010)	(0.012)	(0.012)	(0.015)
Last rank	-0.004	-0.022*	-0.029**	-0.016
	(0.010)	(0.012)	(0.012)	(0.015)
Observations	9278	9278	7016	5294
Groups	6878	6878	5405	4277
Min group	1	1	1	1
Avg group	1.35	1.35	1.30	1.24
Max group	35	35	34	25
Difference in coefficients	0.044***	0.067***	0.045***	0.031
Equality test (p-value)	0.001	0.000	0.008	0.148

Notes: *** ($p<0.01$), ** ($p<0.05$), * ($p<0.1$); this table shows marginal effects of the variables estimated using 'xtlogit, re' with bootstrapped standard errors provided in the parentheses (2000 replications); all first submissions to a particular subject on the same day are grouped together and manuscript rank within the group is imputed using manuscript ID; the sample is restricted to observations with at most 5 submissions per subject per day.