arXiv:2601.17035v1 [cs.DL] 19 Jan 2026

延迟接受算法改进同行评审流程

ChristophBartneck[1*†]、Richard Watt[2†]、Etienne Borde[1†]、Pattara Klinpibul[2†]

[1*]计算机科学与软件工程，坎特伯雷大学，新西兰基督城私人信箱4800，邮编8041。[2]经济与金融系，坎特伯雷大学，新西兰基督城私人信箱4800，邮编8041。

*通讯作者。电子邮箱：chirstoph.bartneck@canterbury.ac.nz；贡献作者：richard.watt@canterbury.ac.nz；etienne.borde@canterbury.ac.nz；pattara.klinpibul@pg.canterbury.ac.nz；

†这些作者对本研究的贡献相同。

摘要

同行评审流程对科学的成功至关重要，但也会延迟出版物发表，并耗费大量精力。期刊越来越难以招募合格审稿人。本研究展示了采用基于智能体的仿真对当前同行评审流程进行建模的结果。我们将其与一种新同行评审流程的仿真进行了比较；该流程使用延迟接受算法（DAA）将论文与期刊进行匹配。匹配效果同样良好，同时大幅减少了所需的审稿意见数量和延迟。结果表明，科学界有可能显著优化同行评审流程。

关键词： 延迟决策算法、同行评审、模拟、延迟、效用

1 引言

同行评审流程耗费大量资源，但对科学至关重要。尽管其缺点已广受讨论，它仍是

研究（Lee、Sugimoto、Zhang 和 Cronin，2013年；Smith，2006）。这一流程的主要问题之一是发表延迟，即从投稿到录用所需的时间，以及从录用到发表所需的时间。根据学科不同，论文从投稿到录用平均可能需要长达一年（Bj¨ork 和 Solomon，2013年）。论文可能需要经过多轮修改才能录用，编辑稿件确实需要时间，但大多数延迟很可能是由于审稿人完成审稿所需的时间造成的。

同行评审流程的第二个主要问题是，它本质上非常耗费资源。假设每篇论文平均由三位审稿人评审，那么所需的审稿意见总数就是投稿论文数量的三倍。鉴于科学研究的生产力不断提高（Larsen 和 von Ins，2010年），所需的审稿意见也越来越难以获得。编辑们努力寻找愿意再审一篇论文的审稿人。造成这一情况的一个因素是，编辑和审稿人所做的所有工作都没有任何经济回报。整个系统依靠研究人员的善意和投入来维系。

近年来，人们测试了一些改进同行评审流程的创新方法，例如在保持流程“盲审”特性的同时，对审稿人的贡献予以认可，以此激励他们（Ravindran，2016年）；此外，也尝试了更开放的做法，例如开放式同行评论。在 Peters 和 Ceci (1982) 关于同行评审流程可靠性的实证研究论文中，开放式同行评论所占篇幅是原目标文章的十倍。这项具有里程碑意义的研究还凸显了另一项挑战：对同行评审流程进行实证研究可能会被视为一种举报行为，因此作者可能因此承受后果。科罗拉多大学的 Jeffrey Beall 曾维护一份可能存在掠夺性行为的期刊名单，并多次收到诉讼威胁，尤其来自OMICS集团（Chappel，2013年）。2019年，OMICS集团因欺骗性做法被判支付5000万美元损害赔偿金。例如，OMICS运营的一场会议曾接受一篇由 iOS 自动完成功能生成的荒谬论文（Hunt，2016年）。Bohannon (2013年) 对开放获取出版商进行了钓鱼式测试，OMICS集团再次出现在这份出版商黑名单上；这些出版商不采用值得信赖的同行评审流程。

Jeffrey Beall 的名单已被下线，据报道，Beall“因受到威胁和政治因素的影响，被迫关闭了博客”（Chawla，2017）。这表明，对同行评审制度进行试验可能会带来负面后果。

然而，除了围绕出版过程的政治和法律斗争之外，要针对实际运行的审稿流程开展实验也需要投入大量资源。要完整检验同行评审流程的质量和可靠性，一个挑战是还必须考虑被拒论文（Bartneck，2017）。通常，只有已接收论文是公开可知的，因此需要编辑团队的配合。2015年，神经信息处理系统（NIPS）会议的编辑团队[1] 开展了一项实验，将投稿至该会议的论文中166篇（10%）进行两轮同行评审（Langford，2015；E. Price，2014年）。E.Price（2014年）得出结论：“如果重新进行会议评审流程，NIPS的大多数论文都会被拒绝。”NIPS

[1] https://nips.cc/

2015年，该会议收到1838篇论文投稿，1524位审稿人共提交了10,625份审稿意见。管理如此大型会议的审稿流程十分复杂，但这也使研究人员能够研究审稿流程本身，尽管代价相当高昂。

开展此类初步实证研究当然很有价值，但要改变科学出版的政策、流程和传统依然困难。系统性操纵影响这一过程的因素并不现实。基于智能体的社会模拟为我们提供了测试科学出版过程新方法的机会。Yu、Yu 和 Li (2005)开展了一项模拟，以更好地理解发表延迟，并得出结论：向某一期刊投稿的数量会显著影响发表延迟。

在本研究中，我们探讨了延迟接受算法（DAA）在多大程度上可用于将科学稿件投稿与科学期刊进行匹配。Lloyd Shapley 和 Alvin Roth 因此获得了2012年诺贝尔经济学奖。Gale 和 Shapley (1962)早已证明，当两组规模相同时，他们的算法能够产生稳定匹配。多年来，他们成功地将该算法应用于各种问题，例如学生与学校的匹配，或器官捐献者与受者的匹配（Roth，2015）。

科学出版过程也是一种匹配形式，传统上并不通过金钱来调节市场。至少在单个研究人员和期刊这一层面上是如此。作者将论文提交给期刊或会议，在此时作者和期刊都无需支付任何费用。不过，期刊出版商与大学图书馆之间存在一个由金钱调节的市场，但我们在此不探讨这个市场，因为对于必须选择向何处投稿的单个研究人员而言，它通常没有直接关系。然而，一些大学开始为研究人员在知名期刊上发表论文提供激励措施，例如 Scopus 收录的期刊。

近年来，收取费用就发表任何论文的掠夺性期刊和会议不断增加，这可能会扰乱这一过程。有人已经指出，凡是写出来的论文最终都会发表（Bartneck，2010年）。掠夺性期刊进一步强化了这一观点，因为连完全荒谬的论文也能发表（Hunt，2016；Lindsey& Boyle，2017；Safi，2015；Sokal，1996年）。欧盟和美国的新立法也可能使市场略有变化，因为这些法律要求由特定资助机构资助的作者在开放获取平台上发表论文（美国国立卫生研究院，2009年）。所有传统出版商都已提供开放获取选项，不过费用相当高昂。

本研究关注 DAA 能为稿件与期刊的更好匹配作出哪些贡献。目前，作者只会将论文投稿给一家负责审稿的刊物。如果审稿意见不理想，编辑就会拒绝论文，作者随后会再次将稿件投给同一期刊或一家全新的期刊。作者有望根据审稿人的意见改进论文，但这并不能保证。这个串行流程会持续进行，直到稿件最终与某一期刊匹配。由此造成的延迟是一个相当严重的问题（Bj¨ork & Solomon，2013年）。

DAA 可以通过一步将所有投稿论文与所有期刊进行匹配，从而改进这一流程。作者将稿件提交至中央仓库，并按照自己的偏好对期刊进行排序。提交的稿件被视为商品；中央仓库负责组织评估每篇投稿论文，所有期刊都可以查阅审稿意见，这类似于如今商品的分类方式（Roth，2015年）。随后，期刊将其偏好列表提交至中央仓库，DAA 则将稿件与期刊进行匹配。这种方案有望同时减少所需的审稿意见数量和审稿延迟。

为了将这两种 DAA 情景与现状进行比较，我们使用基于 Python 的框架 Mesa 创建并运行了基于智能体的仿真（Kazil、Masad 和 Crooks，2020年；Masad、Kazil 等，2015年；ter Hoeven 等，2025年）。我们将由此产生的审稿人工作负荷和审稿延迟作为效率指标进行测量，并将研究人员的匹配契合度和满意度作为质量指标进行测量。本文其余部分的结构如下：第 2 节介绍现状下的投稿流程微观经济学模型；第 3节概述总体仿真框架；第 4节将现状投稿模型应用于仿真；第 5 节将 DAA 匹配算法应用于仿真；第 6节讨论所选的模拟参数；第 7 节呈现模拟结果并讨论其影响；最后，第 8 节作出总结。

2 期刊投稿的一个简单微观经济学模型

我们首先介绍一种用于构建传统论文投稿流程模型的微观经济学方法。设想这样一种情形：一位作者撰写了一篇文章，正在寻找愿意发表该文的期刊。作者会对可选期刊进行排序，并首先向排名最高的期刊投稿。如果文章被接受，博弈便随之结束。如果文章被拒绝，作者便向其排名中下一家最高的期刊投稿。这个过程会一直持续，直到论文被接受，或排名列表中的所有期刊都拒绝该论文。

这种情形在形式上具有与封闭式第一价格拍卖相似的特征，也具有与序贯搜索问题相似的其他特征。基本上，作者是卖方，文章是待售商品。期刊是买方（竞标者）。每本期刊对作者都有一个已知的“价值”或收益（即文章被该期刊录用的价值），并且每本期刊都有一个（主观上）已知的录用文章的概率。因此，每本期刊都向作者提供一个“前景”，即一个二维向量 $\pi = (p,\ v)$，其中 p 是录用概率，v 是作者的论文被录用后对作者的价值。我们假设，对于录用概率为 p_k的期刊 k，如果文章被录用，收益为 $v_k = 1/p_k$。此外，假设期刊构成一个按 $p \in (0, 1]$排序的连续统。因此，排名最低的期刊必定录用，其收益为 1；而排名最高的期刊，其录用概率无限趋近于 0，估值则趋于无穷大。

首先，考虑一个单次博弈：作者只有一次向期刊投稿的机会。如果所选期刊拒稿，论文就会直接由排名最低的期刊发表（$p = 1$）。假设期刊审阅论文需要一个“时期”，因此，如果期刊拒稿，作者就进入新的时期，相应地，效用按因子 $\beta \in [0, 1]$.[2]折现。

将 $u(x)$ 定义为效用函数，以概率 p选择某一期刊的期望效用为

$$Eu(p) = pu\left(\frac{1}{p}\right) + (1 - p)\beta u(1), \tag{1}$$

使得，

$$p^* = \underset{p}{\operatorname{argmax}}\, Eu(p). \tag{2}$$

如果我们假设风险厌恶（$u'' < 0$），则满足二阶条件（SOC）。为此，我们假设效用函数满足恒定相对风险厌恶（CRRA），[3]

$$u(x) = \frac{x^{1-r} - 1}{1 - r}, \tag{3}$$

其中 r 是相对风险厌恶的恒定水平[4]，不等于 1——当 u（x）趋于 $r \to 1$ 时的极限为 $\ln(x)$。将（3）代入（2）可得

$$p^*(r) = r^{\frac{1}{1-r}}. \tag{4}$$

注意，当 $r \to 1$ 时，我们得到 $p^* = 1/e \cong 0.3679$。更重要的是，单次博弈中的最优概率与 β无关。另请注意，p^* 是 r的增函数；其直观解释是，随着相对风险厌恶程度提高，作者应选择向更有可能接受该论文的期刊投稿，因为这是风险较低的选择。

2.1 多个投稿时期

现在，我们将前面的例子扩展到多个投稿时期。我们构建一个递归式的市场：如果被最受青睐的期刊拒稿，作者就会转投较不受青睐的期刊，但仍不是默认期刊。因此，作者需要决定从哪本期刊开始，例如 p_1^*，以此确定可能的投稿序列。比如，假设只有时间尝试向两本拒稿概率均小于 1 的期刊投稿（之后再投默认期刊），那么作者面临的问题是

$$p_1^*, p_2^* = \underset{p_1, p_2}{\operatorname{argmax}}\; p_1 u\left(\frac{1}{p_1}\right) + (1 - p_1)\beta \underbrace{\left[p_2 u\left(\frac{1}{p_2}\right) + (1 - p_2)\beta u(1)\right]}_{\text{one-shot game}}. \tag{5}$$

[2] β 用于量化一个人对未来事物的估值会低多少；它代表个人的耐心程度，其中 $\beta = 1$ 的极端情况意味着无限的耐心。
[3]这是经济学模拟实验中一个非常常见的假设，因为它以非常简单直观的方式体现了风险厌恶程度可能发生变化所带来的影响。[4]这有时也称为消费边际效用弹性（EMU）。

请注意，方括号中的项（5）与单次博弈（1）和（2）的情况完全相同。因此，只需递归代入 p_2^* 单次博弈的解，再求解博弈的第一阶段，就能轻松找到双阶段博弈的解。

据此，我们可以轻松求出任意投稿周期数的多期博弈中的最优期刊序列。假设总共有 n 个投稿周期，包括作者可能投稿给 $p = 1$的最后一个周期。那么，第1期是首次投稿，第2期是第二次投稿（若第1期遭到拒稿），依此类推。考虑作者到达第 i期时会怎么做。第 i 期之前发生的拒稿对于决定 p_i^*已无关紧要。我们定义

$$U_i^*(p_i) \equiv \operatorname{argmax} p_i u\left(\frac{1}{p_i}\right) + (1-p_i)\beta U_{i+1}^*(p_{i+1}), \quad \text{for } i \in \{1, ..., n-1\}, \tag{6}$$

以及 $U_n^*(p_n) = u(1)$。

（6）的一阶条件为

$$u\left(\frac{1}{p_i^*}\right) - u'\left(\frac{1}{p_i^*}\right)\frac{1}{p_i^*} = \beta U_{i+1}^*, \quad \text{for } i \in \{1, ..., n-1\}. \tag{7}$$

这就确定了构成投稿计划的期刊序列，即 $\mathbf{p}^* = (p_1^*, ..., p_{n-1}^*, 1)$。该序列通过递归求解：先计算 p_{n-1}^*，再计算 p_{n-2}^*，依此类推，直到最后计算 p_1^*。在（7）中，由于 $u'' < 0$，我们可以注意到 βU_{i+1}^* 是 $1/p_i$的减函数。因此，βU_{i+1}^* 越大（越小），$1/p_i^*$就越小（越大）——也就是说，p_i^*越大（越小）。因此，给定 $\beta > 0$，如果对所有 i都有 $U_i^* > U_{i+1}^*$，那么最优投稿计划 $\mathbf{p}^*$ 中的概率序列就是递增的，即对所有 i都有 $p_i^* < p_{i+1}^*$。[5] 图1展示了在 β 和 r取不同值时，最多包含8个投稿周期的最优投稿计划。

一个值得深入探讨的问题是，作者希望用多少个时期来投稿才是最优的。应允许作者选择在接受普遍接受的期刊之前投稿多少次。直观来说，作者每花一个时期寻找期刊，就有一定概率获得更高的收益；但如果新的搜索遭到拒稿，作者就要承担收益延迟的成本。最优投稿次数会在边际上平衡这两种影响。简言之，如果存在一个时期 $i \in \{1, ..., n\}$，使得 $U_i^*(p_i) < \beta U_{i+1}^*(p_{i+1})$，作者就不会希望用完全部 n 个投稿周期；但如果对于所有 i，都有 $U_i^*(p_i) > \beta U_{i+1}^*(p_{i+1})$，那么增加投稿时期（即增大 n）将对作者有利。

[5] 一个有趣但极不可能出现的情况是 $\beta = 0$，在这种情况下，每个时期面临的问题都相同，因此作者会为所有 $i < n$ 选择相同的 $p_i^* = p^*$（当然，$p_n^* = 1$也是如此）。其他情况也有可能出现，例如单调递减序列，或作者在第一个时期就选择最优的 $p^* = 1$。

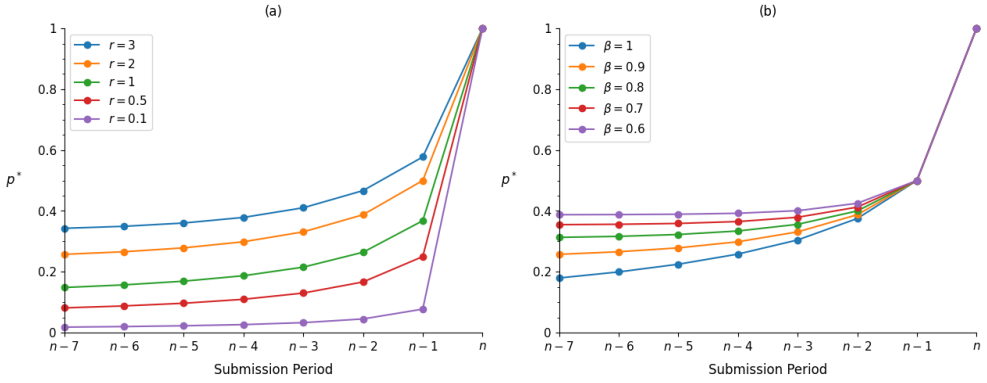


图1 最多包含 $n=8$ 个投稿周期的最优投稿计划：(a) 不同水平的 r，当$\beta=0.9$时；(b) 不同水平的 β，当 $r=2$时

3 出版模拟模型

我们的模拟模型聚焦于由研究人员和期刊代理组成的一个学术领域。研究人员撰写并评审论文，期刊则发表这些论文。主体在每个时间步依次行动。时间步长 Δt 定义为 1 天，我们使用随机调度器来确定每个时间步的行动顺序。该模型将用于比较传统（现状）发表过程与延迟接受算法（DAA）。

在我们的模型中，研究人员（R）撰写论文并审阅他人的论文。每位研究人员 $i \in \{1,...,N_R\}$ 都被赋予一个质量值 $q_i^R \in (0,1)$，其中 N_R 是研究人员总数。质量评级越高，代表研究人员越优秀，撰写的论文也越好。我们假设每位研究人员都能感知自己的真实质量；也就是说，研究人员的感知质量属于完全信息。此外，正如我们稍后将看到的，我们还假设研究人员能够正确感知自己论文的质量，因此不会受到邓宁-克鲁格效应的影响。研究人员撰写论文的速度各不相同，因此，我们为每位研究人员设定一个写作生产率 η_i^w，表示该研究人员撰写一篇论文所需的天数。我们不考虑实际开展研究所需的时间，因为这可能有很大差异。请注意，我们假设研究人员一次只撰写一篇论文。每位研究人员还有一个响应时间 η_i^{resp} 和一个审稿生产率 η_i^{rev}，分别表示研究人员响应（接受/拒绝）审稿邀请所需的天数，以及接受邀请后审阅一篇论文所需的天数。为简化起见，我们假设 $\eta_i^{\text{rev}} = \eta^{\text{rev}}$ 和 $\eta_i^{\text{resp}} = \eta^{\text{resp}}$。最后，我们假设每位审稿人一次最多只能审阅三篇论文。[6]

研究人员产出的论文（P）质量为 $q_{i,j}^P \in (0,1)$，相对于其自身质量存在一定偏差 $\varepsilon_{i,j}^P$，因此，对于论文 j 和研究人员 i，

$$q_{i,j}^P = q_i^R + \varepsilon_j^P. \tag{8}$$

[6]一个局限是，在现实中，研究人员的评审容量也可能受到其他工作负荷的影响。

与研究人员不同，论文的感知质量属于不完全信息。我们将用“帽子”表示感知信息，并用“噪声”产生不完全信息。对于每篇由研究人员审阅的论文，他们都会根据自己对该论文感知到的质量为其打分，$\hat{q}_{i,j}^{P} \in (0, 1)$。每位研究人员对每篇论文质量的评估都略有不同，存在变动 $\varepsilon_{i,j}^{\text{rev}}$，因此，对于论文 j 和审稿人 i，

$$\hat{q}_{i,j}^{P} = q_j^{P} + \varepsilon_{i,j}^{\text{rev}}. \tag{9}$$

我们模型中的第二类主体是期刊（J）；与研究人员类似，在我们的模型中，每本（标准的同行评审）期刊 $k \in \{1, ..., N_J\}$ 都有一个质量 $q_k^J \in (0, 1)$，其中 N_J 是同行评审期刊的总数，质量评级越高，期刊就越好。与研究人员不同，每当期刊接收一篇新论文时，其质量都会更新；每本期刊都有一个初始质量 $q_k^{J_0}$。期刊 k 的质量重新计算为已接收论文的平均质量，其中初始质量的权重相当于 1,000 篇论文。我们将$A_k(t)$定义为期刊 k当前已接收论文的集合，使得

$$q_k^J(t) = \frac{1,000 q_k^{J_0} + \sum_{j \in A_k} q_j^P}{1,000 + |A_k(t)|}, \quad k \in \{1, ...N^J\}. \tag{10}$$

此外，我们定义了一种特殊期刊 $k = 0$，称为全接收期刊，它会（立即）接收所有投稿论文，而不进行同行评审。与其他期刊不同，它的质量固定为 $q_0^{J_0} = q_0^J = 0$。全接收期刊早已以存储库的形式存在。举例来说，感到沮丧的研究人员可能会将论文发表在其所在大学的研究存储库中，或发布在自己的网站上。如果作者将论文发表在预出版平台上，例如 Arxiv，[7] 他们可能就会把论文留在那里，不再继续发表。我们也假设研究人员对期刊质量的感知是不完全信息。研究人员 i 在给定时间 t 对期刊 k 质量的感知为

$$\hat{q}_{i,k}^J(t) = q_k^J(t) + \varepsilon_{i,k,t}^J. \tag{11}$$

请注意，我们假设全接收期刊的质量是完全信息（$\hat{q}_0^J = q_0^J = 0$）。

4 现行出版流程

我们使用第 2 节中介绍的期刊投稿微观经济学模型，创建第一个模拟情景。现状模拟对当前采用的传统同行评审流程进行建模。研究人员撰写论文并将其投稿至期刊。期刊将论文送交研究人员评审。收到所有审稿意见后，期刊决定录用还是拒稿。当然，这是对流程的略微简化，因为论文在最终录用前可能会经过多轮修改；不过，我们将这一问题留待未来研究。

[7] http://arxiv.org/

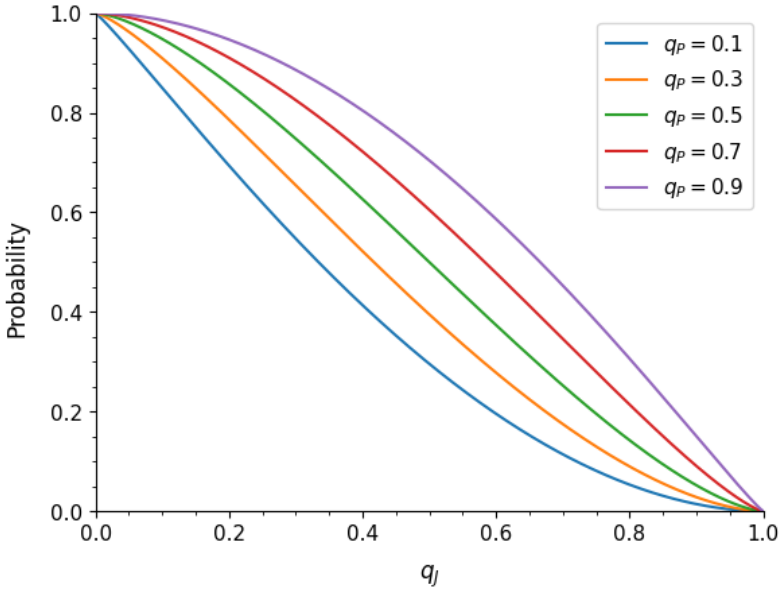


图 2 研究人员对期刊录用概率的信念。

4.1 研究人员的信念系统

我们首先定义一个支配研究人员行为的通用信念系统。研究人员认为，对于质量为 q_P 的论文，若期刊质量为 q_J，其被拒稿的可能性服从贝塔分布，其概率密度函数 (PDF) 为

$$f_{\text{beta}}(q^J; 1+q^P, 2-q^P) = \left[q^J\right]^{q^P}\left[1-q^J\right]^{1-q^P}. \tag{12}$$

因此，研究人员认为论文被录用的概率为

$$p(q^P, q^J) = 1 - F_{\text{beta}}(q^J; 1+q^P, 2-q^P), \tag{13}$$

其中 $F_{\text{beta}}(\cdot)$ 是贝塔分布对应的累积分布函数（CDF）——由于我们假设信息不完全，$\hat{}p(q^P, \hat{}q^J)$ 是相应的感知概率。请注意，对于全接收期刊，我们假设录用概率为 1 是完全信息（$\hat{}p_0 = p_0 = 1$）。图 2展示了不同论文质量下所认为的录用概率。

4.2 论文投稿流程

我们按照第2.1节概述的多期博弈，对论文投稿流程进行建模。根据多期投稿模型，我们知道，研究人员会有一个最优投稿计划$\mathbf{p}^*(r) = (p_0^*, ..., p_{n_{\text{att}}}^* = 1)$，该计划根据接受概率，为每次投稿尝试 $l \in \{1, ..., n_{\text{att}}\}$ 确定最优期刊

。在决定向何处投稿时，研究人员会估计各期刊的接受概率。我们假设研究人员只会向同一期刊投稿一次。也就是说，第一次投稿尝试时，有 $N_J + 1$ 个潜在期刊（N_J 家同行评审期刊加上全接收期刊）；经过 l 次投稿尝试后，还剩下 $n_J \equiv (N_J + 1) - (l - 1)$ 家期刊。对于某次投稿尝试，论文 j 的（估计）接受概率定义为 $\hat{\mathbf{p}}_j = (\hat{p}_{j,0}, ..., \hat{p}_{j,n_J})$。在第 l 次投稿尝试时，研究人员希望向满足 $\hat{p}_{j,k} = p_l^*$ 的期刊 k 投稿。然而，我们必须认识到，由于模拟中期刊数量有限，未必总能找到（估计）概率与最优投稿计划完全吻合的期刊。不失一般性，假设期刊按以下顺序排列：$\hat{p}_{j,n_J} \leq ... < \hat{p}_{j,0} = 1$——请注意，期刊 $k = 0$ 表示全接收期刊。对于给定的 p_l^*，研究人员会选择满足 $\hat{p}_{k+1} < p_l^* \leq \hat{p}_k$（对 $k \in \{0, ..., n_J - 1\}$ 而言）的期刊 k；若遇到 $k = n_J$ 这一边界情况，则选择满足 $\hat{p}_k$ 的期刊 $0 < p_l^* \leq$。请注意，如果研究人员针对某篇论文用尽了所有可能的同行评审期刊，就只能转而在全接收期刊发表。最后，我们设定一个时限：如果一份稿件在 N_{wait} 年后仍未被期刊录用，那么下一次投稿将投向全接收期刊。

4.3 同行评审流程

期刊收到论文投稿后，会发出三份审稿邀请。[8]受邀审稿人将从全体研究人员库中随机选出（不包括作者）。[9] 此外，期刊绝不会多次邀请同一位研究人员评审同一份稿件。经过 η^{resp} 天后，受邀研究人员会回复接受或拒绝审稿邀请。如果该研究人员的审稿容量已满，就会拒绝邀请；否则，当且仅当满足以下条件时，该研究人员会接受审稿邀请：

$$\hat{q}^J \geq q^R - \theta^R, \tag{14}$$

其中，θ_R 是某个容忍水平，我们稍后将对此进行定义。

如果接受邀请，审稿人需要 η^{rev} 天时间，按照（9）中的定义返回审稿评分；如果拒绝邀请，期刊则必须从缩减后的研究人员库中（随机）另选一位审稿人。如果研究人员库中的人选已用尽，该审稿名额的审稿评分将默认为 0。

期刊收到三份审稿评分后，当且仅当满足以下条件时，才会接受该论文：

$$\frac{\sum_{i=1}^{3} \hat{q}_i^P}{3} \geq q^J - \theta^J, \tag{15}$$

其中 θ_J 是我们稍后将定义的某个容忍水平。

论文被期刊接收后，我们会评估：(i) 匹配质量，以及 (ii) 研究人员对结果的满意度——研究人员

[8]Huisman 和 Smits (2017) 指出，各科学领域每篇投稿平均收到的审稿报告约为 2.2 份；Fox (2017) 也发现了类似的数据。[9]我们将此视为模型的一项局限，并将开发更复杂的审稿人遴选流程留待未来研究。

如果能在更好的期刊上发表论文，会感到更加满意。为确定匹配质量，我们按如下方式计算质量拟合度 δ：

$$\delta_{j,k} = 1 - |q_j^P - q_k^J|. \tag{16}$$

对于研究人员的满意度，我们直接使用（3）中定义的效用函数，并计算 $u(1/p_{k^*})$，其中 k^* 是论文发表的期刊，而 $p_{k^*}(q^P,\ q_{k^*}^J)$ 是相应的接受概率（基于所假设的信念系统）。此外，为了考虑相关的发表延迟，我们还计算折扣效用

$$\bar{u}\left(\frac{1}{p_{k^*}}; n^*\right) = \beta^{(n^*-1)} u\left(\frac{1}{p_{k^*}}\right), \tag{17}$$

其中 n^* 是投稿尝试次数。[10]

5 延迟接受算法

在本节中，我们将 DAA 匹配应用于模拟，以取代现有的同行评审流程。我们设立一个负责将稿件分配给期刊的中央仓库。研究人员将完成的稿件连同期刊偏好列表一并提交至中央仓库。仓库将论文送交评审，期刊则可以查阅审稿意见。随后，期刊将其（针对稿件的）偏好列表提交给仓库。仓库根据这些偏好列表，使用 DAA 将论文与期刊进行匹配。我们采用的是多对一匹配博弈。这对应于医院—住院医师（HR）分配问题（Gale & Shapley，1962年；Roth，1984年）。对于 HR 匹配博弈，存在论文最优解和期刊最优解；我们将采用期刊最优解。我们使用 Python 的 Matching 库（Wilde, Knight,&Gillard，2020年）来实现这一方法。

5.1 投稿与评审流程

论文连同按优劣排序的期刊偏好列表一起提交至中央仓库——这类似于研究人员在投稿时对期刊的排序。此外，假设 $0 = \hat{}q_0^J < ... \leq \hat{}q_{N_J}^J$，使我们能够将 $\hat{}\mathbf{q^J} = (\hat{}q_{N_J}^J, ...\hat{}q_0^J)$ 定义为期刊偏好列表——再次提醒，期刊 $k = 0$ 是全接收期刊。[11]

收到一份稿件后，仓库会发出三份审稿邀请。与现状一样，审稿人从全体研究人员库中随机选出（不包括作者）。与现状不同的是，现在研究人员只有在审稿容量已满时才会拒绝邀请。我们定义

$$\bar{q}_P = \frac{\sum_{i=1}^{3} \hat{q}_P^i}{3} \tag{18}$$

[10]请注意，(17) 是一种粗略计算，因为它没有考虑到每次投稿尝试的结果可能以不同速度确定这一事实。[11]值得注意的是，全接收期刊始终是最不受偏好的期刊。

为给定论文的汇总评审分数。[12] 我们定义 n_P 为仓库库存中已完成评审、待分配论文的数量，并将 $\bar{\mathbf{q}}_P = (\bar{q}_P^1, ..., \bar{q}_P^{n_P})$ 定义为论文分数的排序列表（从高到低）。由于仓库集中组织评审并对稿件进行排序，所有期刊都会有相同的偏好列表。[13] 当有足够的论文分配给每家期刊（包括全接收期刊 $n_P \geq N_J + 1$）时，仓库将求解相应的匹配博弈，以确定最优分配。在每一轮分配中，每家标准期刊都会得到同等考虑，并被指定一个容量

$$C_k = \text{Floor}\left(\frac{n_P}{N_J + 1}\right), \quad k \in \{1, ..., N_J\}, \tag{19}$$

此时，所有剩余论文都将分配给全接收期刊，因此

$$C_0 = n_P - \sum_{k=1}^{N_J} C_K. \tag{20}$$

分配确定后，每本期刊都会发表分配给它的论文。一个关键特点是，借助 DAA，每篇稿件都能在一次投稿尝试中发表。我们将再次通过质量拟合度（δ）和效用（u 和 $\bar{u}$）指标来衡量匹配质量以及研究人员的满意度。

6 模拟参数

现在我们已经对模拟进行了正式描述，接下来需要定义其参数。我们尽可能依据现有文献和常识得出结论。

6.1 研究人员和期刊的数量

Mabe 和 Amin（2001年）和Mabe（2003年）指出，作者群体增长约 100–150 名研究人员，便会促成一本新期刊的创办。这表明一本期刊可容纳约 100–150 名研究人员，因此，在模拟中采用 125:1 的研究人员与期刊比例是合理的近似值。我们认为，100 本（同行评审）期刊（$N_J = 100$）足以恰当地体现不同的期刊质量，同时又不会使模拟规模过大。与之配套的研究人员数量为 12,500 名（$N_R = 12,\ 500$）。

6.2 质量与生产力

遵循Lotka(1926年）、D.J.D.S. Price (1963）和Seglen (1992年），显然，科学在不同维度上（包括但不限于

[12]与之前一样，如果审稿名额空缺且潜在评审人已全部用尽，该审稿名额将默认得分为 0。[13]实际上，如果审稿意见是主观评论而非客观分数，期刊的偏好列表可能会有所不同。

生产力、论文和期刊质量）都存在显著的偏斜。[14,15] Shockley (1957年）指出，科学可以视为极端人类努力的一个例子，需要多种特质的累积。极值分布通常用于模拟极端人类能力（最优秀者中的最优秀者）。Gumbel分布（双指数分布）就是其中一种。广义右偏Gumbel分布（最大值情形）的概率密度函数（PDF）为

$$f_{gumbel_R}(x;x_0,\alpha)=\frac{1}{\alpha}\exp\left[-(z_R+e^{-z_R})\right],\quad z_R=\frac{x-x_0}{\alpha},\quad x\in(-\infty,\infty),\tag{21}$$

其中，x_0 是位置参数，$\alpha>0$ 是尺度参数。广义左偏 Gumbel 分布（最小值情形）的 PDF 为

$$f_{gumbel_L}(x;x_0,\alpha)=\frac{1}{\alpha}\exp\left[-(z_L+e^{-z_L})\right],\quad z_L=\frac{x_0-x}{\alpha},\quad x\in(-\infty,\infty).\tag{22}$$

相应的累积分布函数为

$$F_{gumbel_R}=\exp\left[-e^{-z_R}\right],\tag{23}$$

$$F_{gumbel_L}=1-\exp\left[-e^{-z_L}\right].\tag{24}$$

我们定义一个（99%）有界 Gumbel 分布

$$f_{bounded}(x)=\frac{f(x)}{F(b)-F(a)},\tag{25}$$

$$F_{bounded}(x)=\frac{F(x)-F(a)}{F(b)-F(a)},\tag{26}$$

其支撑集为（a，b），其中 $a=F^{-1}(0.005)$，$b=F^{-1}(0.995)$。因此，如果我们希望分布的支撑集为（x_1，x_2），可以通过计算

$$\alpha=\frac{x_2-x_1}{b-a},\quad x_0=x_1-\alpha a.\tag{27}$$

图3 展示了有界 Gumbel 分布的形状。[16]

对于期刊和研究人员的质量，我们使用取值范围为 (0,1) 的右偏分布——这对应于 0.3199 的平均值。至于研究人员的写作生产率，Kyvik(2003) 观察到，在 1998—2000年这一三年期间，终身教职学术人员平均每人每年发表三篇出版物。[17] 然而，Newman(2001) 和 Kyvik (2003) 认为，每篇论文大约有 3 位作者。因此，我们

[14]关于较新的实证研究，例如参见 Albarr´an、Crespo、Ortu˜no 和 Ruiz-Castillo（2011年），以及 Ruiz-Castillo 和 Costas（2014年）。[15]关于科学引用度偏斜成因的理论分析，还可参见 Dorogovtsev 和 Mendes（2001年）、Jackson 和 Rogers（2007年），以及 Peterson、Press´e 和 Dill（2010年）。[16]我们注意到，这种偏斜程度可能不足以体现科学计量指标中的偏斜；有研究指出，这种偏斜类似于幂律关系，在双对数图或半对数图中趋于线性（Lotka，1926年；Seglen，1992年）。这是我们模型的另一项局限，我们将留待未来研究探讨。[17]不过请注意，该图还包括书籍和报告。

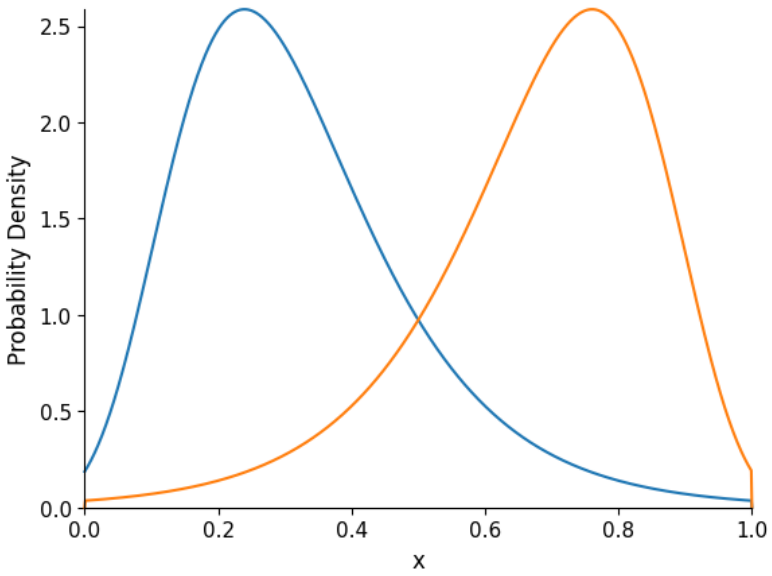


图 3 支撑集为 $[0,1]$ 的有界 Gumbel 分布；蓝色曲线表示右偏最大值情形，橙色曲线表示左偏最小值情形。

希望研究人员的平均生产率约为每年一篇论文。由于我们的写作生产率 η^w 表示撰写一篇论文所需的天数，因此我们使用左偏分布。我们选择了 (60,500) 的支撑集，这对应于 360 天的平均写作时间。对于这一分布，我们希望写作时间以整天计，因此使用向上取整函数来获得准离散分布。

6.3 审稿时长

为确定适当的审稿时间和审稿邀请响应时间，我们将参考有关首次回复时间（即首轮审稿时长）的实证数据。首次回复时间（FRT）包括期刊初步评估投稿并寻找审稿人所用的时间、审稿人审阅论文所用的时间，以及编辑根据审稿报告作出最终决定并通知作者所用的时间（Huisman & Smits，2017）。Huisman 和 Smits（2017）观察到，FRT 平均为 13 周（91 天）。

在我们的模拟中，我们假设 FRT 等于审稿时间与审稿邀请响应时间之和：

$$\mathrm{FRT} = \eta^{\mathrm{resp}} + \eta^{\mathrm{rev}}. \tag{28}$$

我们将 FRT 定为 90 天，其中研究人员需要 70 天审阅论文，20 天回复审稿邀请（$\eta^{\mathrm{resp}} = 20$ 和 $\eta^{\mathrm{rev}} = 70$）。

6.4 风险厌恶、耐心与投稿行为

我们将采用 $\beta = 0.9$ 作为跨期贴现因子，这是经济学中的一个常见假设。至于（恒定水平的）相对风险厌恶 r，我们将遵循

文献中的估计值。 Groom 和 Maddison Pr (2019) 估计，在英国， $r = 1.5$，而 Evans (2005) 估计 r 在 20 个经济合作与发展组织国家中约为 1.4。[18] 我们将采用 $r = 1.5$。

Toroser 等人 （2017）观察到，虽然大多数论文在 1–2 次投稿尝试内就被录用，但有些论文可能需要最多 6 次投稿尝试才能发表。同样，Azar （2004）估计，一篇论文平均需要 3–6 次投稿尝试才能发表。基于此，我们设定 $n_{\text{att}} = 6$；也就是说，在第 6 次投稿尝试时，研究人员会将论文投稿至全接收期刊。此外，我们设定 $N_{\text{wait}} = 35$ 年；这一数值是对研究人员职业生涯长度的粗略估计。[19]

6.5 噪声与偏差

回顾一下，在我们的模拟中，为了计算 q^P、$\hat{}q^P$ 和 $\hat{}q^J$，我们定义了噪声ε^s: $s \in \{P,\text{rev } J\}$。它们反映了研究人员产出的论文质量差异，以及研究人员在评估论文或期刊质量时产生的判断误差。我们将质量评级 q 定义在域（0，1）内。简单的高斯（正态）噪声可能会使论文质量或感知质量值超出指定域。因此，我们改用截断高斯噪声。截断正态分布的 PDF 如下所示

$$f_{\text{truncnorm}}(x;\mu,\sigma,a,b) = \frac{1}{\sigma}\frac{\phi\left(\frac{x-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}, \quad x \in (a,b), \tag{29}$$

其中， $\phi(\cdot)$ 和 $\Phi(\cdot)$ 分别是标准正态分布的 PDF 和累积分布函数。我们按如下方式对噪声进行采样

$$\varepsilon^s \sim \text{TruncNorm}(0,\sigma^s, 0-q, 1-q), \quad s \in \{P,\text{rev},J\}, \tag{30}$$

使得 q^P，$\hat{}q^P$，$\hat{}q^J \in$(0，1)。这等同于直接采样

$$q^P \sim \text{TruncNorm}(q^R,\sigma^P,0,1), \tag{31}$$
$$\hat{q}^P \sim \text{TruncNorm}(q^P,\sigma^{\text{rev}},0,1), \tag{32}$$
$$\hat{q}^J \sim \text{TruncNorm}(q^J,\sigma^{\text{J}},0,1) \tag{33}$$

对于书面论文质量的波动，我们假设其波动程度适中，并设定 $\sigma^P = 0.1$。另一方面，就审稿意见的波动而言，根据 NIPS实验，第一轮评审中获录用的论文在第二轮评审中被拒稿的概率约为 60%（Langford，2015年）。此外，Bonavia 和 Marin-Garcia (2023) 发现，在审查已接收论文时，由两名评审组成的委员会对发表决定意见不一致的概率约为

[18]请注意，相对风险厌恶会因子群体和/或所考虑的情境而有很大差异；不过，这些数值应能提供一个良好的起始范围。
[19]然而，应当认识到，在达到这一极限之前，每篇论文都应当发表——尽管是在全接收期刊上。

48%。这表明波动程度相当显著，因此我们假设波动较强，并设定 $\sigma^{\text{rev}} = 0.2$。最后，对于期刊质量，我们假设研究人员可以获取期刊质量的定量指标（例如 SJR、H 指数或影响因子），因此对期刊质量的认知只取决于不同定量指标之间的差异以及细微的主观偏好。基于此，我们假设波动较弱，并设定 $\sigma^{J} = 0.05$。

6.6 容忍度

对于审稿接受和发表容忍度（θ^{R}, θ^{J}），我们简单地假设容忍水平为 10%。因此，研究人员 i 的审稿接受容忍度为

$$\theta_i^R = 0.1 q_i^R, \tag{34}$$

期刊 k 的发表容忍度为

$$\theta_k^J = 0.1 q_k^J. \tag{35}$$

7 结果

我们运行了为期 100 年的模拟（其中爬坡阶段为 50 年）。表 1 和 2分别展示了现状模拟和 DAA 模拟的关键指标。[20] 在现状模拟中，100 年间共撰写了 1,369,718 篇论文。其中，99.05% 在模拟结束前发表——82.96% 的出版物发表于同行评审期刊。共完成 11,635,350 条审稿意见，平均每篇出版物对应 8.58 条审稿意见；共发出 26,345,917 份审稿邀请，平均每篇出版物对应 19.42 份邀请。这意味着审稿人接受率为 44.16%——Breuning、Backstrom、Brannon、Gross 和 Widmeier (2015年)，Fox(2017) 以及 Meyerson 、 Suzzi-Simmons 和 Simberloff (2025 年) 观察到的审稿人接受率约为 40–60%，这印证了我们的数据。平均发表延迟为 405.20 天（中位数：338 天）。再次说明，这是一个符合实际的数值： Bj¨ork 和 Solomon(2013年) 观察到的平均发表延迟为 12.18 个月。论文平均经过三次投稿尝试后发表，质量拟合度为 0.89。在满意度方面，作者每篇出版物获得的平均效用为 0.43。考虑到发表延迟，每篇出版物的平均折扣效用为 0.38。最后，全接收期刊所发表论文的平均质量为 0.14。

在 DAA 模拟中，共撰写了 1,355,358 篇论文，模拟结束时有 99.74% 已发表，其中 77.20% 的出版物发表于同行评审期刊。共完成 4,055,804 条审稿意见，平均每篇出版物对应 3.00 条审稿意见；共发出 4,146,250 份审稿邀请，平均每篇出版物对应 3.07 份邀请。这意味着审稿人接受率为 97.82%。平均发表延迟为 91.24 天（中位数：89 天）。所有论文都在一次投稿尝试后发表，符合预期，平均质量拟合度为 0.89。每篇出版物的平均效用为 0.35，折扣效用也为 0.35

[20]所有模拟和结果均可在我们的 GitHub 代码仓库中获取（https://github.com/Etienne13/ publication sim）。

因为每篇论文都在一次投稿尝试后发表。最后，全接收期刊所发表论文的平均质量为 0.17。

可以立即看出，采用 DAA 匹配后，发表过程的效率大幅提高。也就是说，每篇论文所需的审稿工作量减少了，同时发表过程也加快了。此外，DAA 的匹配质量几乎与现状相同；也就是说，加快发表过程并未降低稿件与期刊的匹配质量。此外，考虑到发表延迟后，DAA 匹配下的作者满意度（折扣效用）也与现状情形相当。最后，与现状一样，低质量论文仍会在全接收期刊上发表。

表 1 现状模拟的关键指标。

总体	**数量**		**/论文**		**/出版物**		
已撰写论文	1,369,718		100%		—		
发表:							
Total	1,356,639		99.05%		100%		
Peer-reviewed	1,125,470		82.17%		82.96%		
All-acceptance	231,169		16.88%		17.04%		
审稿意见	11,635,350		8.49		8.58		
审稿邀请	26,345,917		19.23		19.42		
已发表论文	**Mean**	**中位数**	**Min**	**Max**	**Q1**	**Q3**	**MAD**
发表延迟 [天]	405.20 (333.86)	338 (278)	89 (89)	1,965 (1,965)	218 (189)	616 (447)	189 (129)
提交尝试次数	3.03 (2.42)	2 (2)	1 (1)	6 (5)	2 (1)	4 (3)	1 (1)
质量拟合度	0.89 (0.90)	0.90 (0.91)	0.34 (0.34)	1.00 (1.00)	0.84 (0.85)	0.95 (0.96)	0.05 (0.05)
效用	0.43 (0.52)	0.46 (0.49)	0 (0.09)	1.44 (1.44)	0.33 (0.40)	0.58 (0.60)	0.12 (0.10)
折扣效用	0.38 (0.46)	0.37 (0.41)	0 (0.08)	1.42 (1.42)	0.27 (0.32)	0.51 (0.55)	0.12 (0.10)
全接收论文质量	0.14	0.13	0.00	0.63	0.07	0.19	0.06

NB: MAD 是中位数绝对偏差；括号内的数值不包括全接收期刊。

7.1 期刊质量

另一个值得关注的方面是两种匹配情景下的期刊质量。表3 概述了期刊质量分布的关键指标，图 4 展示了期刊质量分布的初始状态和最终状态。从表3 和图 4可以看出，经过爬坡阶段后，期刊质量趋于稳态，整体变化不大；更重要的是，期刊质量

表 2 DAA 模拟关键指标。

总体	**数量**	**/论文**	**/出版物**
已撰写论文	1,355,358	100%	—
出版物:			
Total	1,351,872	99.74%	—
Peer-reviewed	1,043,700	77.01%	77.20%
All-acceptance	308,172	22.74%	22.80%
审稿意见	4,055,804	2.99	3.00
审稿邀请	4,146,250	3.06	3.07

已发表论文	**Mean**	**中位数**	**Min**	**Max**	**Q1**	**Q3**	**MAD**
发表延迟 [天]	91.24 (91.28)	89 (89)	89 (89)	176 (155)	89 (89)	90 (91)	0 (0)
投稿尝试次数	1	1	1	1	1	1	0
质量拟合度	0.89 (0.90)	0.90 (0.92)	0.30 (0.38)	1.00 (1.00)	0.84 (0.86)	0.95 (0.96)	0.06 (0.05)
（折现）效用	0.35 (0.45)	0.36 (0.42)	0 (0.04)	1.35 (1.35)	0.20 (0.32)	0.50 (0.55)	0.15 (0.11)
全接收论文质量	0.17	0.16	0.00	0.70	0.09	0.24	0.07

NB: 中位数绝对偏差；括号中的数值不包括全接收期刊。

采用 DAA流程时明显更高（最终质量平均值：0.37），而现状下则较低（最终质量平均值：0.23）。此外，在传统匹配（现状）下，期刊质量往往略有下降，平均质量变化为 -3.60%；相较之下，采用 DAA 匹配时，期刊质量往往略有上升，平均质量变化为 1.19%。这些结果表明，DAA匹配不仅有助于优化同行评审流程，也可能有利于期刊，因为它能使期刊维持较高的质量标准。

表3 现状与DAA模拟的期刊质量分布摘要。

	Mean	**中位数**	**Min**	**Max**	**Q1**	**Q3**
Status quo:						
初始质量	0.30	0.26	0.05	0.87	0.18	0.39
起始质量*	0.24	0.24	0.05	0.88	0.18	0.27
最终质量	0.23	0.23	0.05	0.88	0.18	0.26
质量变化 [%]**	-3.60	-3.97	-8.72	1.01	-7.01	0.00
DAA:						
初始质量	0.35	0.33	0.07	0.74	0.22	0.45
起始质量*	0.37	0.34	0.18	0.78	0.25	0.45
最终质量	0.37	0.34	0.19	0.79	0.26	0.45
质量变化 [%]**	1.19	0.85	-2.39	8.66	-0.17	2.21

NB: 这些指标不包括全接收期刊。*这是模拟开始时的质量（爬坡阶段之后）。
**这是起始质量与最终质量之间的差值。

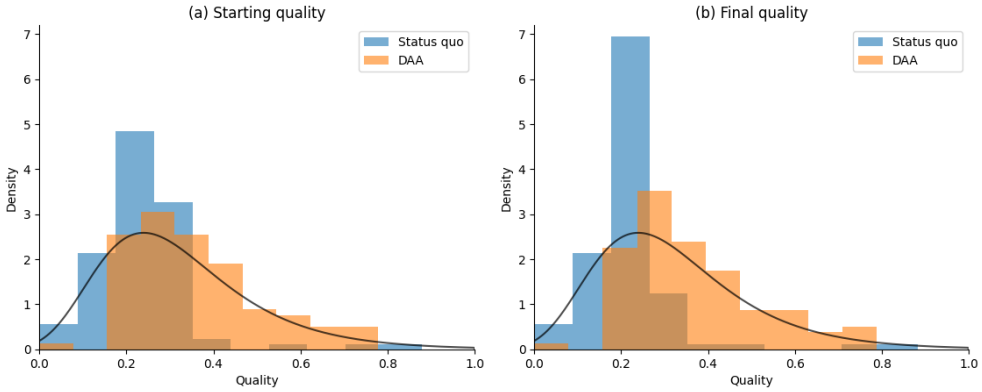


图 4 现状与 DAA 期刊质量分布（与底层 PDF 对比）。

8 结论

我们的研究表明，能够管理审稿意见并将论文与期刊匹配的科学仓库将极大地造福科学。鉴于科学研究事业的多样性和规模，人们或许会想知道如何建立这样一个科学仓库。如何鼓励所有研究人员和期刊参与？由谁来支付费用？乍看之下，这些挑战似乎难以克服。

不过，也有几个例子让我们抱有希望。我们已经成功为研究人员（ORCID）[21] 和出版物（DOI）建立了独特的标识符系统。[22] 这些系统由出版商和科学组织组成的联盟管理并维持。Open Knowledge 项目提供免费开源软件，用于运营科学期刊，而 OpenAlex[23] 则为 Scopus 等商业索引服务提供了替代方案。Arxiv[24] 等类似项目提供免费的（预）发表平台，并取得了惊人的发展。科学界拥有应对技术挑战的专业知识。

对研究人员和期刊来说，主要的推动力是效率的大幅提升。作者将能够更快地发表文章，期刊也不必再为寻找审稿人而苦恼。但最重要的是，DAA 鼓励诚实。参与者试图钻系统空子的行为不会带来任何好处。该算法会为所有人找到最佳匹配。

基于 DAA 的同行评审流程也存在一些局限性，至少在我们此处介绍的初始形式中是如此。作者只有一次提交论文的机会。目前，作者可以利用拒稿后收到的反馈来改进作品。因此，DAA 同行评审流程更类似于会议的审稿流程。在计算机科学领域，会议论文集与期刊同样受到认可，但会议有严格的时间安排。在主要审稿流程结束后，被接收的

[21] https://orcid.org [22]
https://www.doi.org [23]
https://openalex.org [24]
https://arxiv.org

作者可以利用反馈改进论文，但被拒稿的论文没有第二次机会。这就引出了第二个局限性。

低质量论文正在全接收期刊中发表（约占23%）。作者仍有机会利用收到的反馈改进论文，但将不再能够向其他期刊投稿。作者可能不会认可这一结果。不过，全接收论文的比例与未获得任何引用的论文比例相吻合（VanNoorden，2017）。我们可能不得不承认，同行评审流程无论如何都可能已经失去了其筛选功能，而仅作为一种排序系统运行（Bartneck，2010年）。

延迟接受算法已在许多不同的应用领域证明了其价值，因此，它也能改善论文与期刊之间的匹配或许并不令人意外。鉴于该算法带来的诸多益处，我们希望我们的模拟有助于推广这样一种理念：利用采用延迟接受算法的科学仓库，更好地将论文与期刊进行匹配。

参考文献

Albarr´an,P.,Crespo,J.A.,Ortu˜no, I., Ruiz-Castillo, J. (2011年)。科学的偏态：219个子领域及若干总合领域。《科学计量学》，*88*（2），385–397，https://doi.org/https://doi.org/10.1007/s11192-011-0407-9

Azar, O.H.（2004）。拒稿与首次回复时间的重要性。《国际社会经济学杂志》，*31*(3), 259–274，https://doi.org/https://doi.org/10.1108/03068290410518247

Bartneck, C.（2010）。全接收出版政策 [会议论文集]。第四届数字社会国际会议（*icds 2010*）（第 37-40 页）。IEEE。

Bartneck, C.（2017）。审稿人评分无法预测影响力——人机交互会议论文集的文献计量分析 [期刊文章]。《科学计量学》，*110*(1), 179-194，https://doi.org/https://doi.org/10.1007/s11192-016-2176-y

Bj¨ork, B.-C., & Solomon, D.（2013）。学术同行评审期刊的出版延迟 [期刊文章]。《信息计量学杂志》，*7*(4), 914-923，https://doi.org/https://doi.org/10.1016/j.joi.2013.09.001

Bohannon, J.（2013年）。谁害怕同行评审？ [期刊文章]。科学，*342*（6154），60-65，https://doi.org/https://doi.org/10.1126/science.342.6154.60

Bonavia,T.,&Marin-Garcia, J.A. (2023). 对一篇科学论文同行评审的噪声审计：WPOM期刊案例研究。*WPOM*-运营工作论文

管理学， *14*(2), 137–166, https://doi.org/https://doi.org/10.4995/wpom.19631

Breuning, M., Backstrom, J., Brannon, J., Gross, B.I., Widmeier,M. (2015年). 审稿人疲劳？为什么学者拒绝评审同行的作品。《政治学：政治科学与政治》， *48*(4), 595–600， https://doi.org/https://doi.org/10.1017/S1049096515000827

Chappel, B. (2013年). [网页]。取自 http://www.npr.org/sections/thetwo-way/2013/05/15/184233141/publisher-threatens-librarian-with-1-billion-lawsuit

Chawla, D.S. (2017). 争议性掠夺性出版商名单神秘消失[期刊文章]。科学，， https://doi.org/https://doi.org/10.1126/science.aal0625

Dorogovtsev, S.N., & Mendes, J.F. (2001年). 无标度演化网络的尺度特性：连续方法。《物理评论*E*》， *63*(5), 056125， https://doi.org/https://doi.org/10.1103/PhysRevE.63.056125

Evans, D.J. (2005). 消费边际效用弹性：对20个经济合作与发展组织国家的估算。《财政研究》， *26*(2), 197–224, https://doi.org/https://doi.org/10.1111/j.1475-5890.2005.00010.x

Fox, C.W. (2017). 招募审稿人的难度可预测生态学与进化学六种期刊的评审评分和编辑决策。《科学计量学》， *113*(1), 465–477,https://doi.org/https://doi.org/10.1007/s11192-017-2489-5

Gale, D., & Shapley, L.S. (1962). 大学招生与婚姻稳定性[期刊文章]。《美国数学月刊》， *69*(1), 9-15, https://doi.org/https://doi.org/10.2307/2312726

Groom, B., & Maddison Pr, D. (2019). 英国边际效用弹性的新估算。《环境与资源经济学》， *72*(4), 1155–1182,https://doi.org/https://doi.org/10.1007/s10640-018-0242-z

Huisman, J. 与 Smits, J.（2017年）。同行评审流程的持续时间与质量：作者视角。《科学计量学》， *113*（1）， 633–650， https://doi.org/https://doi.org/10.1007/s11192-017-2310-5

Hunt, E.（2016）。由 iOS 自动补全撰写的无意义论文被会议 [期刊文章]接收。《卫报》，，取自 https://www.theguardian.com/science/2016/oct/22/nonsense-paper-written-by-ios-autocomplete-accepted-for-conference

Jackson, M.O., & Rogers, B.W. (2007年). 陌生人和朋友的朋友相遇：社交网络有多随机？《美国经济评论》，*97*(3), 890–915，https://doi.org/https://doi.org/10.1257/aer.97.3.890

Kazil, J., Masad, D., Crooks, A. (2020年). 利用 Python 进行基于主体的建模：Mesa框架。国际社会计算、行为与文化建模、预测及行为表征建模与模拟会议（第308–317 页）。

Kyvik, S. (2003年). 大学教职人员在 1980—2000 年间发表行为的变化趋势。《科学计量学》，*58*(1), 35–48，https://doi.org/https://doi.org/10.1023/A:1025475423482

Langford, J. (2015年). *NIPS*实验。取自 http://cacm.acm.org/blogs/blog-cacm/181996-the-nips-experiment/fulltext

Larsen, P.O., & von Ins, M.（2010年）。科学出版物的增长速度及科学引文索引覆盖范围的下降 [期刊文章]。《科学计量学》，*84*（3），575-603，https://doi.org/https://doi.org/10.1007/s11192-010-0202-z

Lee, C.J., Sugimoto, C.R., Zhang, G., Cronin, B.（2013年）。同行评审中的偏见 [期刊文章]。《美国信息科学与技术学会会刊》，*64*（1），2-17，https://doi.org/https://doi.org/10.1002/asi.22784

Lindsey, J., & Boyle, P.（2017）。作为社会建构的概念性阴茎 [期刊文章]。《*Cogent*社会科学》，*3*（1），1330439，https://doi.org/https://doi.org/10.1080/23311886.2017.1330439

Lotka, A.J.（1926年）。科学生产力的频率分布。《华盛顿科学院学报》，*16*(12), 317–323，取自 https://www.jstor.org/stable/24529203

Mabe, M.（2003年）。期刊的增长与数量。《连续出版物》，*16*(2)，，

Mabe, M. 和 Amin, M.（2001年）。学术与科学期刊的增长动态。《科学计量学》，*51*(1)，147–162，https://doi.org/https://doi.org/10.1023/A:1010520913124

Masad, D., Kazil, J.L., 等（2015）。Mesa：基于智能体的建模框架。*SciPy*（第 51–58 页）。Meyerson, L., Suzzi-Simmons, A., Simberloff, D.（2025）。量化科学出版中的审稿人拒审：来自《生物入侵》2002–2024 年的二十一年数据。《生物入侵》，*27*（10），1–10，https://doi.org/https://doi.org/10.1007/s10530-025-03679-1美国国立卫生研究院（2009）。美国国立卫生研究院公共获取政策详情 （报告）。美国国立卫生研究院。取自 https://publicaccess.nih.gov/policy.htm Newman, M.E.J.（2001）。科学合作网络的结构 [期刊文章]。《美国国家科学院院刊》，*98*（2），404-409，https://doi.org/https://doi.org/10.1073/pnas.98.2.404Peters, D.P., & Ceci, S.J.（1982）。心理学期刊的同行评审实践：已发表文章再次投稿后的命运 [期刊文章]。《行为与脑科学》，*5*（2），187-195，

Peterson, G.J.，Press´e, S.，Dill，K.A.（2010年）。科学引文概率分布中的非普适幂律标度。《美国国家科学院院刊》，*107*（37），16023–16027，https://doi.org/https://doi.org/10.1073/pnas.1010757107

Price, D.J.D.S.（1963年）。《小科学，大科学》。哥伦比亚大学出版社。Price, E.（2014年）。*NIPS* 实验。取自 http://blog.mrtz.org/2014/12/15/the-nips-experiment.html Ravindran, S.（2016年）。因同行评审而获得认可 [期刊文章]。科学，，https://doi.org/https://doi.org/10.1126/science.caredit.a1600022Roth, A.E.（1984年）。医学实习生和住院医师劳动力市场的演变：博弈论案例研究。《政治经济学杂志》，*92*（6），991–1016，https://doi.org/https://doi.org/10.1086/261272

Roth, A.E.（2015年）。谁得到什么，以及为什么：匹配与市场设计的新经济学 [书籍]。Eamon Dolan/Houghton Mifflin Harcourt。Ruiz-Castillo, J., & Costas, R.（2014年）。科学生产力的偏态。《信息计量学杂志》，*8*（4），917–934, https://doi.org/https://doi.org/10.1016/j.joi.2014.09.006Safi, M.（2015年）。[报纸文章]。取自 http://www.theguardian.com/australia-news/2014/nov/25/journal-accepts-paper-requesting-removal-from-mailing-list Seglen, P.O.（1992年）。科学的偏态 [期刊文章]。《美国信息科学学会杂志》，*43*（9），628-638, https://doi.org/https://doi.org/10.1002/(SICI)1097-4571(199210)43:9⟨628::AID-ASI5⟩3.0.CO;2-0Shockley, W.（1957年）。研究实验室中生产力个体差异的统计研究。《*IRE*汇刊》，*45*（3），279–290, https://doi.org/https://doi.org/https://doi.org/10.1109/JRPROC.1957.278364

Smith, R.（2006年）。同行评审：科学与期刊核心环节中的一个有缺陷的过程[期刊文章]。《皇家医学会杂志》，*99*(4)，178-182, https://doi.org/https://doi.org/10.1177/014107680609900414

Sokal, A.（1996年）。一位物理学家对文化研究展开实验 [期刊文章]。*LinguaFranca*，五月/六月（1996年），，取自 http://linguafranca.mirror.theinfo.org/9605/sokal.html ter Hoeven, E.、Kwakkel, J.、Hess, V.、Pike, T.、Wang, B.、Kazil, J. 等（2025年）。Mesa 3：2025 年使用 Python 进行基于主体的建模。 *Journal of Open Source Software*，*10*（107），7668，https://doi.org/https://doi.org/10.21105/joss.07668Toroser, D.、Carlson, J.、Robinson, M.、Gegner, J.、Girard, V.、Smette, L.、……O'Kelly, J.（2017）。影响同行评审出版物录用与发表时间的因素。《当代医学研究与观点》，*33*（7），1183–1189, https://doi.org/https://doi.org/10.1080/03007995.2016.1271778Van Noorden, R.（2017年12月）。从未被引用的科学研究。《自然》，*552*（7684），162–164, https://doi.org/10.1038/d41586-017-08404-0 检索日期：2026年1月8日，来源：https://www.nature.com/articles/d41586-017-08404-0

Wilde, H., Knight, V., Gillard, J. (2020). Matching：用于解决匹配博弈的 Python 库。 *Journal of Open Source Software*, *5*(48), 2169, https://doi.org/https://doi.org/10.21105/joss.02169

Yu, G., Yu, D., Li, Y. (2005). 关于期刊出版延迟控制过程的模拟研究 [期刊文章]。《科学计量学》, *63*(1), 25-38, https://doi.org/https://doi.org/10.1007/s11192-005-0202-6