Journal of Informetrics 18 (2024) 101559

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Research Paper

On journal rankings and researchers' abilities

Wojciech Charemza [a], Michał Lewandowski [b], Łukasz Woźny [b,*]

[a] *Faculty of Business and International Relations, Vistula University, 02-787 Warsaw, Poland*
[b] *Department of Quantitative Economics, SGH Warsaw School of Economics, 02-554 Warsaw, Poland*

A R T I C L E I N F O

JEL classification:
I23
D61
O31

Keywords:
Journal rankings
Publication reward mechanisms
Optimal categorization
Journal quality

A B S T R A C T

Over the last few years, ranking lists of academic journals have become one of the key indicators for evaluating individual researchers, departments and universities. How to optimally design such rankings? What can we learn from commonly used journal ranking lists? To address these questions, we propose a simple, theoretical model of optimal rewards for publication in academic journals. Based on a principal-agent model with researchers' hidden abilities, we characterize the optimal journal reward system, where all available journals are assigned to one of several categories or ranks. We provide a tractable example that has a closed-form solution and allows numerical applications. Finally, we show how to calibrate the distribution of researchers' ability levels implied by the observed journal ranking schemes.

1. Introduction

There is growing interest in introducing rating systems that could encourage publication and improve the performance of research-oriented institutions. Such systems are commonly used in many countries and universities in the hiring of new faculty members and promotion decisions, although this is usually done informally or indirectly. Rating grades are often labeled as 4-star, 3-star, 2-star and 1-star, or in other countries, may be A+, A, B, C, and occasionally D. These grades are typically awarded for quality and number of publications at the individual, departmental or university level. Such systems have long been used in many countries, usually at the level of individual universities, and are often subject to analyzes and comparisons between countries or disciplines.

These ratings, which can be more or less varied and detailed, are also used by researchers as an unofficial support tool when looking for the most appropriate place to publish their academic output, and by universities when assessing the performance of their current employees before promotions or potential employees before hiring them. In many countries, particularly the US, UK and Australia, there is a tendency to officially avoid such rankings as part of regular reviews of universities and faculties. Unofficially, however, they are still used to quickly assess the quality of researchers' output.

In countries that use the Performance-Based Research Funding Program (PBRF) metric, these ratings are no longer indicative, but have become more directive as university funding is allocated based on rankings of journals in which articles by affiliated researchers are published. The rating lists for journals consequently achieve official status in the metric scheme, and some countries have then upgraded the role of the rating lists for journals even further. Many universities in Poland, for instance, have used a publication bonus or a reward scheme that entitles the authors to receive a financial reward that is proportional to the rank of the journal they have published in.

* Corresponding author at: al. Niepodleglosci 162, 02-554 Warszawa, Poland.
E-mail address: lukasz.wozny@sgh.waw.pl (Ł. Woźny).

https://doi.org/10.1016/j.joi.2024.101559
Received 25 March 2024; Received in revised form 30 June 2024; Accepted 9 July 2024
1751-1577/

Although the rating lists for journals are popular, relatively little attention has been paid in the literature to a formal characterization of the optimal journal rating, to the associated reward schemes or to the institutional context that explains some key differences observed between the systems applied in various institutions or countries.[1] Such characterization might help answer some reasonably obvious questions. Do rating schemes and the associated reward schemes encourage researchers to publish in journals that best match authors' potential? Should the optimal system incentivize researchers to publish a smaller number of articles in top journals only, or should it instead incentivize researchers to produce a high number of lower-quality publications? Why do some universities, countries or even academic fields seem to use rankings that are *steeper* at the top, while others have schemes that are more lenient at the top and steeper at the bottom? Given how competitive the academic market is nowadays, the answers to these questions may be important for institutions aiming to stimulate academic research performance and for researchers looking to maximize the rewards for their work output.

Main goals With these questions in mind, the paper has three aims. The first is to propose a parsimonious theoretical model that allows us to address some of the key trade-offs that arise in designing the optimal reward scheme for journals. The second is to propose a tractable algebraic example of this model that has a closed-form solution (i.e. the optimal journal ranking), and thus allows comparative statics with respect to model parameters. And the third is to apply this solution to compare a few well-known journal rating schemes by matching the implied distribution moments of the researcher population for which the ratings were designed.

Any evaluation of academic reward systems should be preceded by the construction of a theoretical reward model and a characterization of an optimal publication reward mechanism. With these in mind, we propose a simple principal-agent model of hidden information.[2] Agents in such a mechanism, that is, the researchers, identified by their ability level (which is understood as a summary expression of their skills, education, experience, networking, willingness to work, and anything else that is needed to publish in high-quality journals), aim to maximize their reward from publications by choosing which journal they submit their research to. The rewards may be direct or indirect but are always related to the rank of the journal. The principal, which is here called the Research Supervisory Body or RSB (a ministry in some countries or research councils or panels of experts in others), knows the distribution of the levels of ability in the population of researchers and aligns the reward scheme with this distribution in the best way possible. The system is constructed to encourage researchers to allocate their output to journals with the highest possible quality.[3]

We formalize the objective for the RSB and characterize the optimal reward scheme. In doing so, we consider a number of specific issues. Firstly, the RSB would like to set up a system that leads to a large number of quality publications. Secondly, the RSB must take into account the probability of acceptance by the journal. An ambitious system that only rewards publications in top journals where the probability of acceptance is low may be inefficient, as the expected number of publications will be small. Thirdly, the number of distinct journal categories is typically limited, so the RSB must decide how to group journals into different categories, and how to reward the journals in these categories. Fourthly, the RSB must adapt the reward system to the distribution of abilities in the researchers' population. In a population of very good researchers, the reward system is likely to be very steep at the top, meaning it will distinguish between very good and exceptionally good journals and so encourage researchers to submit their papers to journals that are closer to their potential. If such a system is adopted in a population where the general level of ability is low, however, many researchers will become discouraged and will choose journals that do not live up to their potential.

Finally, basing on the insights gained from studying the optimal solution to the RSB problem, we propose a method of retrieving information on the distribution of researchers' abilities implied by the observed journal rating schemes. Before presenting the details, we consider a simple example that illustrates the key insights of our method and some key intuitions that underline our results.

A motivating example We consider two journal rating schemes that are used to incentivize researchers working in the broad field of business and economics. One is the Academic Journal Guide (AJG) rating, which is published by the Chartered Association of Business Schools in the UK, and the other is the rating of the Polish Ministry of Education and Science (PL).[4] Both rating schemes assign economics and business journals to one of several classes, with AJG using 4*, 4, 3, 2, and 1, and PL using 200, 140, 100, 70, 40, and 20, both in descending order of prestige. Table 1 lists six selected journals and their rating scores in the two schemes. It may be noted that the PL scheme seems to be flatter at the top than the AJG, and steeper at the bottom. This is reflected in PL being more sensitive to differences in the quality of journals at the lower end of the quality scale, and AJG to differences at the higher end. To understand why these two ratings differ, we will make some simplifying but intuitive assumptions. First, the purpose of AJG and PL ratings is to encourage researchers from a given population, which is also designated AJG or PL, to submit their work to journals with the highest expected quality. Second, researchers from both populations only care about their country rating of the journal in which their work will be published. Third, higher-ability researchers should optimally publish in higher-quality journals. Moreover, for a given researcher, the higher the quality of the journal, the more difficult it is to get the article accepted.

[1] The few exceptions include papers describing and analyzing PBRF funding schemes, e.g. Adam (2020); Baccini and De Nicolao (2022); De Boer et al. (2015); Vogel et al. (2017); Zacharewicz et al. (2019); Thomas et al. (2020); Vîiu and Păunescu (2021); Abramo et al. (2024a) or Smit and Hessels (2021) at a more general level. See also recent contributions by Mogstad et al. (2022) analyzing journal ranks that aims to minimize the statistical uncertainty associated with the indexes of journal citations and Kosyakov and Pislyakov (2024) studying journal quartile distributions across subject categories and topics. We also refer the reader to Abramo et al. (2020) for a recent study of a relative performance of Italian vs. Norwegian professors or Kulczycki et al. (2018); Korytkowski and Kulczycki (2019) analyzing, how the country-level science policy shapes publication patterns as well as Abramo et al. (2024b) for the quality vs. quantity choice analysis.

[2] See Laffont and Martimort (2001) for a textbook exposition and MacLeod and Urquiola (2021) for a recent application of principal-agent models in related problems.

[3] See Card and DellaVigna (2020) for a discussion on modeling and estimating the quality of papers.

[4] We combine rankings for two disciplines in the Polish rating: Economics & Finance, and Management.

Table 1
Ratings for selected journals according to the two rating schemes.

Journal	PL	AJG
Econometrica	200	4*
Theoretical Economics	200	4
AEJ: Microeconomics	200	3
Dynamic Games and Applications	70	1
Journal of the Economic Science Association	40	1
Quarterly Journal of Austrian Economics	20	1

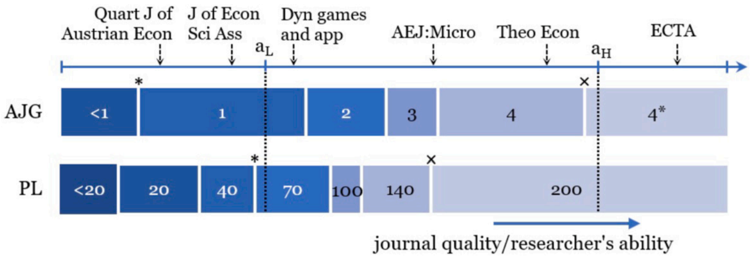

Fig. 1. Journals ordered by quality and assigned to classes of increasing prestige. Researchers with higher ability should publish in higher-quality journals.

Based on these assumptions, formally introduced later in the article, we can apply the same scale to a researcher's ability and the level of journal quality that that researcher would choose in the social optimum. Fig. 1 shows the division of the journal quality level (and corresponding level of researcher ability) into categories according to the two ratings (the scale has been changed for visibility, but the relative position of the categories reflects the actual ratings). We have added one extra class for each rating scheme in addition to the official journal categories,"<1" for AJG and "<20" for PL, and these contain journals that are not assigned to any class by the respective scheme and are deemed to be of lower quality than any of the journals that have a class assigned.

In our simple model, researchers only care about the journal's rank, so they rationally choose the "cheapest" journal in the class of journals containing their socially optimal choice. According to our assumptions, it is the lowest quality journal in this class, because it gives the highest probability of acceptance and the same prestige as other journals in this class. Researchers with a high level of ability, such as researcher a_H shown by the dashed line in Fig. 1, would aim for the cheapest '4*' journal under the AJG scheme, which is much closer to their socially optimal choice than the cheapest '200' journal (marked $\times$ on the appropriate scales) that they would aim for under the PL program. Similarly, lower ability researchers, e.g. a_L, in the AJG scheme will aim for the cheapest '1' journal, while in the PL scheme they will aim for the cheapest '70' journal (both marked with $*$ on their respective scales). The latter choice is much closer to their socially optimal choice and thus leads to smaller losses.

The expected loss of quality is consequently greater in the PL scheme than in the AJG scheme for high-ability researchers like a_H, and lower for lower-ability researchers like a_L. This is true on the *individual* level. However, the best rating scheme with a certain number of classes should determine journal classes in such a way that the *total* loss of quality is as small as possible. Since AJG accepts losses at the lower end of the ability/quality scale, while PL accepts losses at the higher end, the AJG population must have a larger mass concentrated in the higher ability levels than the PL population has.

Our strategy in the empirical part of the paper for deducing the unobservable distribution from the observable rating scheme is to reverse engineer the optimal solution for the RSB objective. We take rating schemes like those shown in Fig. 1 as input and ask what distribution of ability levels the scheme is optimized for.

Structure of the paper The rest of the article is organized as follows. The general model and its key assumptions are presented in Section 2. In Section 3 we use a simplified model allowing a closed-form solution. Section 4 shows how to reverse engineer the optimal solution, particularly the one obtained in Section 3, to calibrate[5] the distribution of ability within the population. In Section 4, we compare several well-known journal ratings using this method. Section 5 contains further discussion of the limitations and possible expansion of the model.

The supplementary material includes Appendix A with the proofs of the propositions from Section 2 and 3, Appendix B where we discuss possible extensions of the basic model, and Appendix C containing the results of the robustness analysis of the case presented in Section 4.

[5] We use the term calibration in an economic rather than a statistical sense, meaning we select the parameters of the theoretical model so that the model fits best with the empirical data and various simulation scenarios (see, e.g. Foster, 2011).

2. The model

The model consists of an RSB and a continuum of researchers. Each researcher, interpreted as a single author or (more loosely) as a group of co-authors, is identified by a private type $a \in A = [0,1]$, which is referred to as the ability level. Abilities are distributed in the population according to a strictly increasing CDF denoted by F. Each researcher has a single[6] paper and must decide which journal it should be submitted to. Journals are uniquely identified by the quality index $\phi \in \Phi = [0,1]$ and have a conditional probability of acceptance $p : \Phi \times A \to [0,1]$, where $p(a,\phi)$ is the probability that an article by researcher a will be accepted by journal ϕ if it is submitted there. It is assumed that p is common knowledge. We also assume that p is continuous and that the following assumption holds whenever probabilities are strictly positive[7]:

Assumption 1 *(Monotonicity of journals)*. The ratio $\frac{p(\phi',a)}{p(\phi,a)}$ is increasing in a for any $\phi < \phi'$.

The RSB does not know the individual abilities of researchers, but it knows their distribution in the population. It sets up a reward system $R : \Phi \to \mathbb{R}$, which is assumed to be upper semicontinuous. As usual, problems of this kind are solved backwards, starting with the researcher's problem.

2.1. The researcher's problem

For greater clarity, we assume[8] that researchers are risk-neutral and maximize expected reward, with the payoff for not publishing anything normalized to 0. The researcher's a problem is thus:

$$\max_{\phi \in \Phi} R(\phi)p(\phi,a). \tag{1}$$

Let $\Phi_R(a)$ denote the set of optimal solutions. It is nonempty by the standard arguments for any upper semicontinuous reward scheme. The next proposition expresses the journal monotonicity assumption in an equivalent observable form.

Proposition 1. *If the researcher's objective is given by (1), then the following are equivalent:*

i) *monotonicity of journals holds.*
ii) *for any reward scheme R, ability levels $a_1 < a_2$ and journals $\phi_1 < \phi_2$, if researcher a_1 weakly prefers ϕ_2 over ϕ_1 then a_2 strictly prefers ϕ_2 over ϕ_1.*

Since journal quality and researcher ability are not directly observable, the monotonicity of journals and other properties of p can be used to define one quantity relative to another. Assuming, for example, that ϕ is a good measure of journal quality, Proposition 1 implies that a can be understood as a researcher's ability to publish in a journal with a high ϕ. A direct corollary of this result is that researchers with greater ability choose higher-quality journals.

Corollary 1. *Given any reward scheme R, each selection ϕ_R from Φ_R is non-decreasing on A.*

2.2. The RSB problem

First-best policy

The RSB maximizes the total expected quality of the papers published in the population of researchers by setting a policy R for some measurable selection $\phi_R(a)$ from $\Phi_R(a)$. This implies incentive compatibility of the journal selection. The RSB problem is then:

$$\max_R \int \phi_R(a)p(\phi_R(a),a)dF(a). \tag{2}$$

For greater clarity, we assume there are no participation or budget restrictions.[9] The first best solution under incentive compatibility is therefore to establish a reward system that is proportional to the RSB's preferences and therefore linear in journal quality ϕ:

Proposition 2. *For any $\alpha > 0$, the reward scheme given by $R(\phi) = \alpha\phi$ for any ϕ, solves problem (2).*

[6] Our model easily encompasses a generalization to more papers. A researcher producing m papers in the evaluation period is represented, in our model, by m agents (with the same ability levels) each writing a single paper.

[7] This property is similar to the monotone likelihood ratio property. The difference is that in the present context, the monotone likelihood is a property of two density functions on the binary outcome space, i.e. accept or reject, while journal monotonicity is a condition of $p(\phi,a)$, that is the probability of acceptance with respect to two parameters.

[8] That these assumptions can be relaxed without changing our qualitative results is shown in Supplementary Material, Online Appendix B.

[9] In the Supplementary Material, Online Appendix B we show that these assumptions do not qualitatively affect our results.

The reward scheme given by Proposition 2 is actually a unique maximizer (up to normalization by α) if for each ϕ there is a such that $\phi p(\phi, a) \geq \phi' p(\phi', a)$ for each ϕ'. If there are some dominated journals where this is not the case, there is no loss of generality in setting their reward to 0 in the optimal solution.

Observe that the first-best solution does not depend on the distribution of the researchers' abilities. Since only relative, not absolute, rewards matter for optimal decisions, from now on we will assume that $\alpha = 1$. Let the researcher's solution under the first-best reward scheme be denoted by $\Phi(\cdot)$ and a single selection from it by $\phi(\cdot)$.

Second-best policies

The first-best solution given by Proposition 2 implies a unique reward for each level of journal quality, but such solutions are not actually used in practice. The commonly used measures of journal quality are only stochastic indicators of the underlying quality, so a reward system that is linear in ϕ would create an unwarranted sense of precision (see Konig et al., 2022, p.2).[10] Instead, the existing reward systems divide journals into a small number of classes, so that journals in different classes receive different rewards, but journals within a single class are treated equally. Journals with similar measures of quality are therefore combined into one class. We call this the second-best solution. Consequently, in our model with a continuum of journals, we restrict the reward schemes in (2) to those that allow only $n \geq 1$ distinct non-zero rewards, where n is given exogenously.[11] The question is then how to partition the journals into categories and what reward levels should be set for each category. We start with the following result.

Proposition 3. *For any distribution of abilities F, the set of reward schemes R maximizing the second-best RSB objective contains a non-decreasing R.*

So from now on, we will consider non-decreasing R. Combined with the conditions that R takes only n distinct non-zero values and that it is upper semicontinuous, this results in the following family:

$$R_{\phi_1,\dots,\phi_n,\alpha_1,\dots,\alpha_n}(\phi) = \begin{cases} 0 & \text{for } \phi \in [0, \phi_1), \\ \alpha_1 & \text{for } \phi \in [\phi_1, \phi_2), \\ \dots & \dots \\ \alpha_n & \text{for } \phi \in [\phi_n, 1], \end{cases} \tag{3}$$

where $\alpha_1 < \alpha_2 < \dots < \alpha_n$ and $0 = \phi_0 \leq \phi_1 < \phi_2 < \phi_3 < \dots < \phi_n \leq \phi_{n+1} = 1$, $n \geq 1$. The RSB problem then boils down to setting the *boundary journals* $(\phi_i)_i$ and the reward values $(\alpha_i)_i$ that will maximize (2). The following assumption, although not crucial to our main findings, will help in identifying the parameters of the model.

Assumption 2 *(Better journals are more expensive).* $p(\phi, a)$ is decreasing in ϕ.

This assumption implies that among the full set of journals that receive the same reward, the one with the highest probability of acceptance, or the "cheapest", will be the one with the lowest level of quality. So if the reward scheme is specified by (3) then only the boundary journals $\phi_1, \phi_2, \dots, \phi_n$ will be selected. All journals in between, meaning in the interval (ϕ_i, ϕ_{i+1}), will be dominated by the ϕ_i journal, and so will never be chosen. We will discuss the practical implications of this assumption in Sections 3 and 4. We may next consider a reward scheme $R_{\phi_1,\dots,\phi_n,\alpha_1,\dots,\alpha_n}$, denoted by R^* for simplicity. To determine $\Phi_{R^*}(a)$, we need to find the ability levels $a_{1/2}, \dots, a_{n-1/n}$ of the indifferent researchers, which are these for whom the cheapest journals in subsequent categories are equally good. These ability levels are obtained by solving the following system of equations:

$$\frac{p(\phi_i, a_{i/i+1})}{p(\phi_{i+1}, a_{i/i+1})} = \frac{R^*(\phi_{i+1})}{R^*(\phi_i)}, \quad i \in \{1, \dots, n-1\}. \tag{4}$$

A solution might generally not exist, but the assumption of journal monotonicity implies that $p(\phi_{i+1}, \cdot)R^*(\phi_{i+1})$ crosses $p(\phi_i, \cdot)R^*(\phi_i)$ only once, and it does so from below. The RSB can, in consequence, always set the reward scheme so that there is a unique solution and ϕ_i is optimal for researchers with a level of ability in the interval $[a_{i-1/i}, a_{i/i+1})$. A researcher with an ability level of $a_{i/i+1}$ is indifferent between ϕ_i and ϕ_{i+1}, while those below this level prefer ϕ_i and those above prefer ϕ_{i+1}. Having established $\Phi_{R^*}(a)$, we can now determine the set of optimal weights $\alpha_1 < \dots < \alpha_n$.

Proposition 4. *If the reward schemes are restricted to the family given by (3), the reward scheme that satisfies $\alpha_i = \alpha\phi_i$, $i \in \{1, \dots, n\}$ for some $\alpha > 0$ solves the second-best RSB problem.*

The same argument as in the first-best case also applies here. Any choice of a reward that is different from the positively-scaled quality of the cheapest journal in a given reward category would change the allocation decision of the researcher relative to the

[10] Differences in opinions and personals interests of the members of the RSB may result in problems when designing a continuous journal ranking. As a result, researchers affected by it, may not regard such a continuous ranking as fully legitimate. Using finitely many categories is hence a solution to soften these designing and legitimacy problems. For more discussion on measuring the quality of journals and how it impacts the optimal reward scheme see Section 5.1.

[11] The optimal number of classes in a journal rating scheme is a separate issue. See Mogstad et al. (2022) for the latest contributions.

objective pursued by the RSB. Proposition 4, together with equation (4) allows us to determine $a_{1/2}, a_{2/3}, \ldots, a_{n-1/n}$, which are the types of boundary researchers.[12] From Proposition 1, it follows that $a_{1/2} < a_{2/3} < \ldots < a_{n-1/n}$. What remains to be determined is the set of boundary journals or the cheapest journals for each class $\phi_1, \ldots, \phi_n$. We state the problem in the following Corollary, setting, as before and without loss of generality, $\alpha = 1$.

Corollary 2. *The second-best RSB problem can be written as:*

$$\max_{(\phi_i)_i} \sum_{i=1}^{n} \int_{a_{i-1/i}}^{a_{i/i+1}} p(\phi_i, a)\alpha\phi_i dF(a), \tag{5}$$

$$\text{s.t.} \quad p(\phi_i, a_{i/i+1})\phi_i = p(\phi_{i+1}, a_{i/i+1})\phi_{i+1}, \quad \text{for each } i \in \{1, ..., n-1\}, \tag{6}$$

where $a_{0/1} = 0$ *and* $a_{n/n+1} = 1$.

Unlike the first-best solution, the optimal solution here depends on the distribution of ability F. This is because there are only n categories available, and so we can fit the best solution for at most n boundary researchers. The other researchers necessarily incur a loss from what they had in the first-best solution because of the suboptimal allocation of papers to journals by the researchers (see Section 1), and it is the RSB job to decide how to minimize this loss, given the size of that loss for each type of researcher and the mass of researchers of that type. Problem (5) is generally analytically complex, so it is often impossible to give a solution in a closed form. However, important insights can be obtained by considering some specific cases. For this reason, the next section first illustrates the key trade-offs made when assigning four journals into three classes. It then considers a parameterized family of piecewise linear functions p, for which a solution in closed form is obtained.

3. The optimal categorization of journals

3.1. Efficiency trade-offs in the second-best solution

Supposing the probability of acceptance p satisfies Assumptions 1 and 2, we consider four journals with the quality levels $\phi_1, \phi_2, \phi_3, \phi_4 \in (0, 1)$, ordered from lowest to highest. If their rewards are given by $R_i = \alpha\phi_i$ for some $\alpha > 0$, then each researcher maximizes part of the RSB objective and so the total expected quality is also maximized. Any other choice of rewards would give different intersections between expected rewards and so there would be a different journal choice for some researchers. This would potentially lead to a loss of expected (publication) quality. Given our assumptions, the choice of the optimal journal is monotone in ability, meaning researchers with lower levels of ability will never find it optimal to publish in higher-quality journals.

Suppose that the RSB may set only *three* reward levels instead of four. With the continuum of abilities, it is never optimal to have fewer than three categories. Furthermore, Proposition 4 implies that the boundary journals in the second-best scheme receive rewards that are equal to their first-best rewards. Assumption 2 implies that reducing the reward of journal ϕ_i to the level R_{i-1} or below makes it *idle* because it is dominated by journal ϕ_{i-1}, so it is never selected. Our problem then comes down to finding the journal that contributes the least benefit, and downgrading its reward to that of one of the lower-quality journals.

The four panels of Fig. 2 show the impact of downgrading each of the four journals (causing them to become idle) on the researcher's typical envelope and the boundary ability levels compared to the first- best case: $\Pi(a) := p(\phi(a), a)\phi(a)$. The RSB compares the efficiency loss (areas shaded in blue) of sacrificing researchers who would optimally choose journal i in the first-best scheme but who now have to choose a different journal. If the researchers' abilities are distributed uniformly on A, the RSB should make journal no. 3 *idle* – this entails the least efficiency loss, as is evident by examining the Figure. The optimal rating is, hence, to group journals 2 and 3 in the middle category (and leave the worst category with the worst journal and the top category with the top journal only). For the general distribution, the loss in efficiency for a given level of ability should be weighted by its density.

Examining all the cases in Fig. 2, we notice that the removal of *higher* quality journals results in (point-wise) *lower* researcher's boundary types (dashed lines in Fig. 2). So, if the selected reward system is set optimally, it can inform us about the distribution of researchers' abilities. If the distribution is left-skewed, we expect higher-quality journals to be *idle*, while lower-quality ones would be with a right-skewed distribution. This means that a second-best reward scheme for a given set of journals that is flat for higher-quality journals and steep for lower-quality journals indicates a less able population of researchers, while one that is flat for lower-quality journals and steep for higher-quality ones indicates a more able population.

3.2. A parametric example and a closed-form solution for uniform distribution of ability

We now consider the general setup with a continuum of journals and a continuum of researchers and assume the following specification for probabilities of acceptance conditional on the level of ability a. Let $\xi \in [1, \infty)$ be a slope parameter:

[12] Let a_i^* be such that $\phi_i = \phi(a_i^*)$ for each i, so a_i^* denotes the type that chooses journal ϕ_i in the first-best scheme. Note that these types are not in the optimization problem, only the types $a_{i/i+1}$.

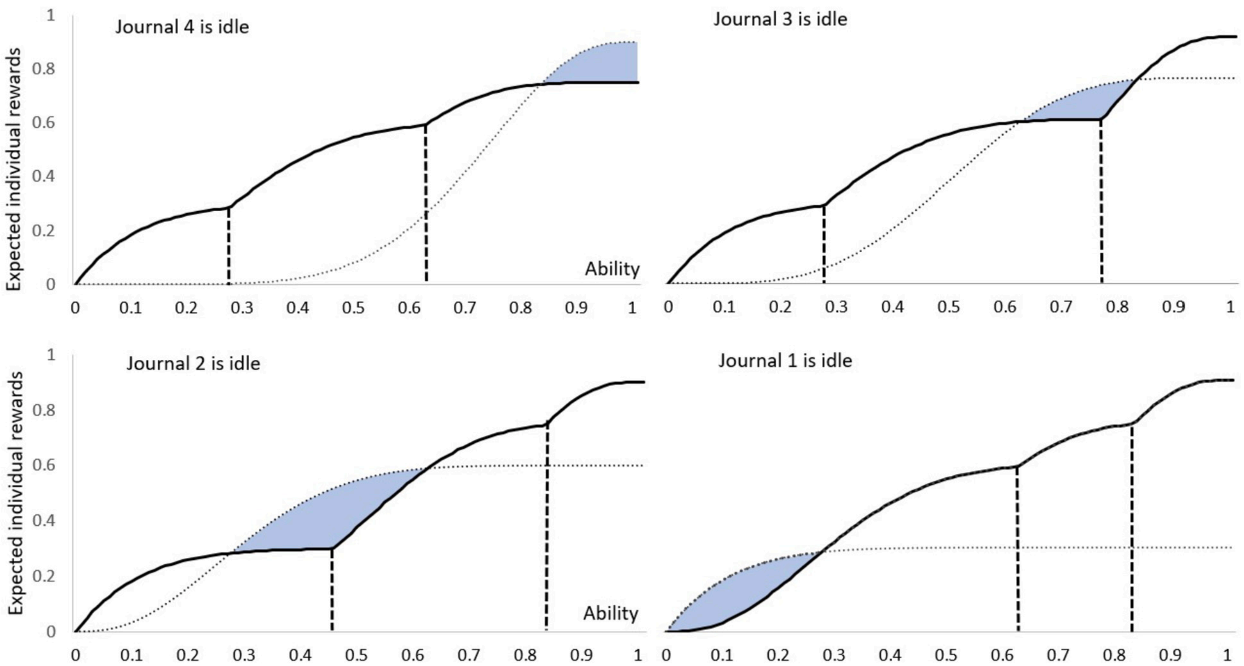

Fig. 2. Second-best with three categories and four journals. Areas shaded in blue correspond to a loss of expected publications' quality as compared to the first-best case.

$$p(\phi,a)=\begin{cases} 0, \text{ for } a\in\left[0,\frac{\xi-1}{\xi}\phi\right), \\ 1+\xi\frac{a-\phi}{\phi}, \text{ for } a\in\left[\frac{\xi-1}{\xi}\phi,\phi\right), \\ 1, \text{ for } a\in[\phi,1]. \end{cases} \tag{7}$$

In this, p takes the form of a CDF of a uniform distribution on $\left[\frac{\xi-1}{\xi}\phi,\phi\right)$.[13] Assuming the set of journals is rich enough and there exists a reward scheme such that optimal journals for different abilities do not coincide, this interval is also the set of abilities for which journal ϕ is the optimal choice for some increasing reward scheme. Parameter ξ controls the level of segregation so that when $\xi=1$, all researchers of non-zero ability have a positive chance of acceptance even in the top journals. As ξ tends to infinity at the other extreme, only the best researchers have a non-zero chance in the top journals.

The probability of acceptance given by (7) captures some common-sense intuition. As ϕ gets larger, the fraction of types who have no chance of success increases, the fraction of types for whom acceptance is certain decreases, and higher-quality journals require a greater increase in ability for the same increase in the probability of acceptance. Given that neither ability nor journal quality are directly observable, (7) is not as restrictive an assumption as it seems since it defines one measure relative to another. For example, under (7) the common percentage change in a and ϕ leaves the value of $p(\phi,a)$ unaffected, meaning $\frac{d\log(a)}{d\log(\phi)}=1$. This produces testable implications as soon as one of the two quantities is given observable meaning. When we calibrate our model to the actual data in the next section, we assume that ϕ is well approximated by the invariant method index proposed by Palacios-Huerta and Volij (2004). If this is so, (7) implies that for the chances of acceptance to remain the same, a given percentage change in the journal index requires the same percentage change in the ability level.

It is easy to verify that $\phi\to p(\phi,a)$ is a non-increasing function and is decreasing on its support (for a given ϕ, we define a support of $p(\phi,\cdot)$ as a set of all a for which $0<p(\phi,a)<1$). Moreover, the ratio $\frac{p(\phi',a)}{p(\phi,a)}$, whenever defined, is non-decreasing in a, whenever $\phi'>\phi$. Whenever $\xi>1$, this ratio is also increasing[14] in a on a joint support on $p(\phi',\cdot)$ and $p(\phi,\cdot)$. As a result p satisfies Assumptions 1 and 2 on its support whenever $\xi>1$. This is sufficient for our conclusions from section 2.

For now we assume that abilities are distributed according to the *uniform distribution* on $[0,1]$. Proposition 2 in the first-best solution implies that $R(\phi)=\alpha\phi$, where $\alpha>0$, for any ϕ. Given (7), the researcher's problem has a unique solution $\phi_R(a)=a$, which we can plug into the RSB objective to get the maximum expected total quality of $\mathrm{ETQ_I}$:

[13] This form of probability of acceptance can be interpreted as follows. Suppose journal ϕ accepts only one article. If two articles are submitted, the one submitted by the researcher with the higher ability level will be accepted. Suppose one researcher with an ability level uniformly distributed in the interval $\left[\frac{\xi-1}{\xi}\phi,\phi\right)$ submits to journal ϕ. Then $p(\phi,a)$ is the probability that the article of another researcher with an ability level of a will be accepted by ϕ if submitted there.

[14] See Supplementary Material, Online Appendix A.1 for a proof.

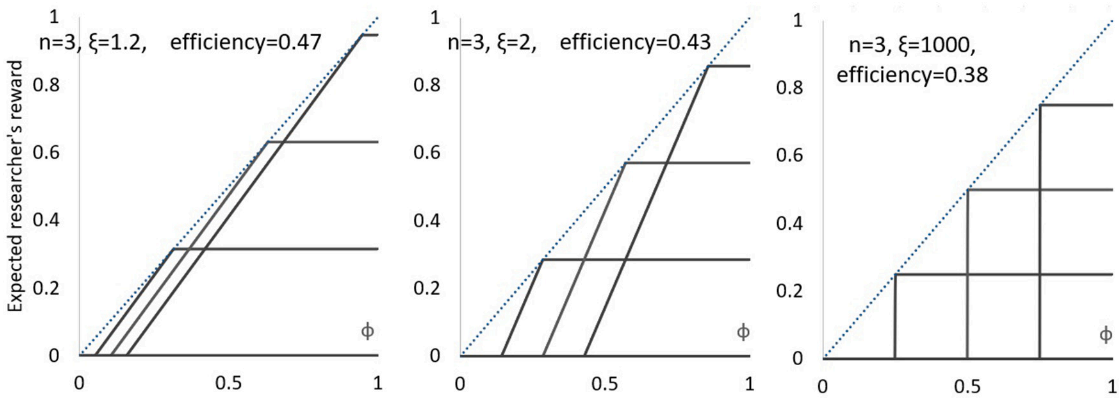

Fig. 3. Optimal solution and selected efficiency values for $\xi = 1.2, 2$, and 1000 and for $n = 3$. The ability is distributed uniformly on $[0, 1]$.

$$\mathrm{ETQ_I} = \int_0^1 a\,dF(a) = \left[\frac{1}{2}a^2\right]_0^1 = \frac{1}{2}. \tag{8}$$

Note that the researcher's problem's envelope, $a \to \Pi(a)$, is linear in a. We now consider the second-best, for which we first fix the number of categories $n \geq 1$. We know from Proposition 4 that the optimal reward scheme has the form of (3) with $\alpha_i = \alpha\phi_i$ for $\alpha > 0$ for each $i \in \{1, ..., n\}$. Since p is decreasing in ϕ and given the reward scheme in (3), the cheapest journal in each category is ϕ_i.

The crossing points are obtained by substituting (7) in (6)[15]:

$$a_{i-1/i} = \frac{\phi_{i-1} + (\xi - 1)\phi_i}{\xi}, \quad i \in \{1, ..., n\}. \tag{9}$$

So Φ_R is specified as follows: the researcher with a level of ability in the interval $[a_{i-1/i}, a_{i/i+1})$ will optimally choose journal ϕ_i. Plugging this into (5) we get:

$$\mathrm{ETQ_{II}}(\xi) = \max_{\phi_1,\dots,\phi_n} \sum_{i=1}^{n} \int_{a_{i-1/i}}^{\phi_i} (\xi a + (1-\xi)\phi_i)da + \sum_{i=1}^{n-1} \int_{\phi_i}^{a_{i/i+1}} \phi_i da + \int_{\phi_n}^{1} \phi_n da.$$

This function has an interior maximum as verified by SOCs, and its FOCs are (details of the derivation are given in Appendix A.2):

$$\frac{\partial \mathrm{ETQ_{II}}(\xi)}{\partial \phi_i} = 0 \iff \phi_i = \frac{\phi_{i-1} + \phi_{i+1}}{2}, \quad i \in \{1, ..., n-1\}$$

$$\frac{\partial \mathrm{ETQ_{II}}(\xi)}{\partial \phi_n} = 0 \iff (\phi_n - \phi_{n-1})\frac{\xi - 1}{\xi} = 1 - \phi_n,$$

with a convention that $\phi_0 = 0$. After rearranging we obtain the following solution together with the corresponding crossing points:

$$\phi_i = \frac{\xi i}{\xi(n+1) - 1}, \quad i \in \{1, ..., n\},$$

$$a_{i-1/i} = \frac{\xi i - 1}{\xi(n+1) - 1}, \quad i \in \{1, ..., n\}.$$

The optimal boundary journals vary from $\frac{i}{n+1}$ for $\xi \to \infty$ to $\frac{i}{n}$ for $\xi \to 1$. The larger the number of categories, the smaller the difference between the lower and upper bounds. Fig. 3 shows the optimal boundary journals and the resulting envelope for researchers, which is the maximum expected reward for researchers over the n cheapest journals for various levels of ξ and n. The area below the envelope equals $\mathrm{ETQ_{II}}(\xi)$, indicating the expected total quality or simply efficiency. It may be recalled that $1/2$, or the area below the identity function, is the efficiency of the first-best solution. We observe that as n gets large, the optimal $\mathrm{ETQ_{II}}(\xi)$ value approaches the first-best value. Moreover, efficiency decreases with ξ, so that if all the researchers publish in a single journal ($n = 1$), for example, the maximum efficiency is 0.25 in the worst case ($\xi \to \infty$) and 0.5 in the best case ($\xi \to 1$). The boundary journals are equally spaced because the distribution of abilities is uniform. The journals will generally adjust optimally to the distribution of abilities so that there are relatively more categories in ability regions with greater mass and relatively fewer where the ability mass is smaller.

[15] Note that $a_{0/1}$ is technically not a crossing point, but it proves convenient in our example. We therefore also use $a_{i-1/i}$ instead of $a_{i/i+1}$.

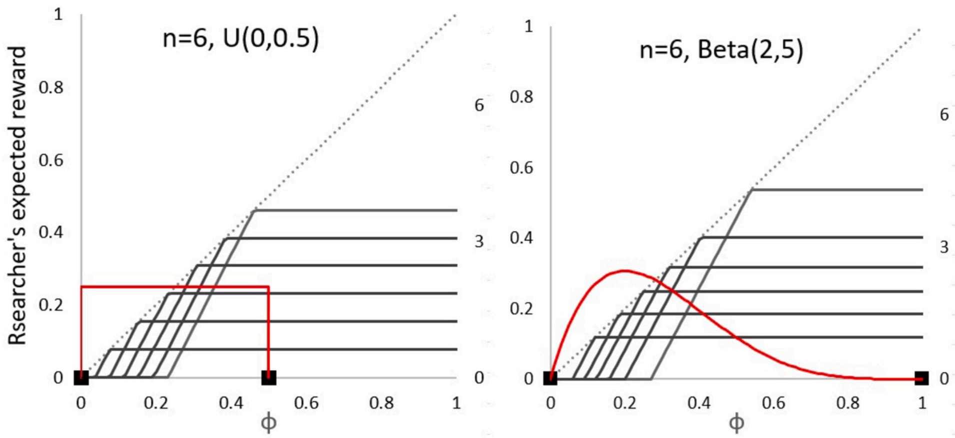

Fig. 4. Optimal solution for different distributions of ability levels.

3.3. A general distribution of abilities

We now consider a *general distribution* of abilities, given by the CDF F. We make two observations. First, the probability integral transform implies that even if abilities a are not uniformly distributed, the $F(a)$ values are (Casella & Berger, 2002, Theorem 2.1.10, p.54), and so our solution for the uniform case can be applied. Second, Proposition 4 implies that the optimal journal for a boundary researcher is $\phi(a_{i-1/i}) = a_{i-1/i}$. We thus apply a change of variables to get the optimal solution for the general case:

$$F(\phi_i) = \frac{\xi i}{\xi(n+1)-1} \quad \text{or } \phi_i = F^{-1}\left(\frac{\xi i}{\xi(n+1)-1}\right), \quad \text{for } i \in \{1, ..., n\}. \tag{10}$$

The same technique can be used to obtain the solution for distributions that have less than full support. Let F be any distribution on the support $[0,1]$, then for any pair $a_L < a_U$ in $[0,1]$, we define $F_{[a_L,a_U]}$ as another CDF where:

$$F_{[a_L,a_U]}(x) = \begin{cases} 0, \text{ if } a < a_L, \\ F\left(\frac{a-a_L}{a_U-a_L}\right), \text{ if } a \in [a_L, a_U), \\ 1, \text{ if } a \geq a_U. \end{cases} \tag{11}$$

The optimal solution for these distributions satisfies: $F_{[a_L,a_U]}(\phi_i) = \frac{\xi i}{\xi(n+1)-1}$, or again following Proposition 4: $\phi_i = a_L + (a_U - a_L)F^{-1}\left(\frac{\xi i}{\xi(n+1)-1}\right)$, while the objective function value remains identical for the original distribution F and the modified one $F_{[a_L,a_U]}$.

For illustration, we compare two populations of researchers, for which we assume $\xi = 2$ and analyze the optimal solution for two selected distributions of the abilities. Fig. 4 presents the optimal solution, with boundary journals and the researcher's envelope, for uniform distributions on a given support and a left-skewed beta distributions. The density functions of the distributions are superimposed in the pictures, and the support of the distribution is depicted as the interval between the two black squares.

4. Distributions of ability induced from journal ratings

We will now show how to use the closed-form solution such as (10) to reverse engineer the distribution F from the observable reward schemes used in practice. The following procedure can be used to compare several journal rating schemes with different numbers of journal classes. This is done by estimating the distribution of abilities for each journal rating separately.

4.1. Outline of empirical evaluation

A single journal rating consists of a set of journals J partitioned into n classes $J_1, ..., J_n$. Each journal j in J is assigned a journal quality measure $\phi(j) \in \mathbb{R}$. We assume that Assumptions 1 and 2 hold, and the probabilities of acceptance are given by (7). This assumption conveniently designates ϕ as both the measure of journal quality and the measure of the ability level of a researcher who optimally[16] chooses journal ϕ in the first-best solution. Finally, we assume that the RSB sets the reward scheme R optimally in the family of (3), in line with the second-best policy (5).

[16] We are aware that the RSB might have more complex objectives in practice. It may, for example, artificially upgrade some journals by putting them in a class that is higher than that given by the measure of quality. This could reflect a policy of promoting some journals that are of particular relevance in the hope that such inflated grading might attract better papers, meaning those that are frequently cited, to the journal in the future. This might create the so-called Matthew effect (Drivas & Kremmydas, 2020). It is particularly relevant for promoting national journals by ranking them higher, so as to avoid their downgrading and eventual extinction in the long run.

Since the reward schemes are typically ordinal but they enter the researcher's objective in a cardinal way, as each researcher optimizes the expected reward, we assume that the ordinal rewards correspond to the cardinal utility of rewards in a way that is consistent with (6), so $R(j) = \phi_i$ for $j \in J_i$, where ϕ_i is the boundary or cheapest measure of journal quality in journal class i. Given the above assumptions we can reverse engineer the implied distribution of the ability levels of researchers from the reward scheme observed. To do this we determine n values of the CDF of the distribution according to (10), so $F(\phi_i) = \frac{\xi i}{\xi(n+1)-1}, \quad i \in \{1, ..., n\}$.

This solution critically depends on ϕ_i, the cheapest journal in each class. Taking our assumptions literally, we would set ϕ_i as equal to the lowest value for journal quality in class i. Behaviorally, this reflects the assumption that each researcher knows all the journals in J and can potentially submit their paper there. In practice, specializations, incomplete information or simply the desire to avoid journals with a low academic reputation mean that a given researcher only considers a small subset of all journals.

Consequently, instead of setting ϕ_i as equal to the lowest value for journal quality in class i, we set $\phi_i := G_i^{-1}(k)$, where G_i is the empirical distribution of the values for journal quality in the i-th class, meaning $\{\phi(j) : j \in J_i\}$, and $k \in [0, 1]$ is the percentile value of the distribution. This assumption is a simplified version of the idea that each researcher considers only m journals from each class and that the *cheapest* journal in class i is the lowest-quality journal among those m journals.[17] Our procedure for finding a distribution of the ability of researchers $F : [0, 1] \to [0, 1]$ is summarized as:

Inputs: the set of journals partitioned into classes $J = J_1 \cup ... \cup J_n$, and a normalized measure of journal impact $\phi : J \to [0, 1]$.
Parameters: the slope ξ; the cut percentile k.
Procedure:
1. Set the cheapest journals $\phi_i := G_i^{-1}(k)$.
2. Set the corresponding quantile values $F(\phi_i) = \frac{\xi i}{\xi(n+1)-1}$.
3. Set the boundary values $F(0) = 0$ and $F(1) = 1$
4. Complete the graph of the CDF by connecting the points with lines.

Note that the above procedure can be applied separately for each reward

4.2. Ratings systems for journals

In our empirical example, we focus on four specific, country-oriented ratings of journals for the disciplines of economics and management. These are:

CNRS: Comité National de la Recherche Scientifique journal rating in economics and management (France),
AJG: Academic Journal Guide published by Chartered Association of Business Schools,
PL: Polish Ministry of Education and Science journal index for the combined disciplines economics & finance and management,
US: the US economic journals list (A and B journals).

In what follows we will use the above names both to refer to the journal's rating and to describe its target population. The CNRS and PL ratings are developed within the European PBRF program. However, their roles are slightly different, as the French assessment system within the PBRF is primarily based on peer reviews, and so the ratings of journals have an indicative role, while the Polish system is close to the ideal type of the *metric* PBRF (see, e.g. Ochsner et al., 2021), and its rating is official and directive in that it is part of the calculation for assigning funds to universities and grading their departments. The AJG rating is widely used as an indicative measure of the quality of journals by business schools around the world. The US list is used by some economic departments in the US to support promotion and hiring decisions.

Journals on the PL list are divided into six classes that are labeled by the number of *ministerial points* awarded, which can be 200, 140, 100, 70, 40 or 20. AJG partitions its journals into five *ratings* of 4*, 4, 3, 2 and 1. The other two lists divide the journals into four classes; the CNRS[18] has four *categories* from 1 as the highest to 4 as the lowest, and the US list has four *ratings* of A+, A, A- and B+. In the US Econ list A+ consists of the top 5 general-interest economic journals; A consists of 17 top major-field journals and 4 general-interest journals, A- is composed of 4 general-interest/survey journals and 7 major-field journals, and B+ are 5 general-interest journals and 29 field journals. Further details and data sources are given in Appendix C.

4.3. Recursive impact factor

In our baseline example we use the recursive impact factor (RIF) as the index of journal quality ϕ. This index is obtained using the invariant method proposed by Narin et al. (1976) and derived axiomatically from a few intuitive properties by Palacios-Huerta and Volij (2004) (see also Palacios-Huerta & Volij, 2014). We use the most recent updated version, with 319 journals that are listed in the

[17] So if the ϕ values in class i were distributed uniformly over the interval $[a, b]$ for example, then the mean of the minimum of samples of size m from this distribution is given by $a + (b - a)\frac{1}{n+1}$, which is the $\frac{1}{n+1}$ percentile of the original distribution.
[18] In 2021 a wider list incorporating the CNRS was published (the HCÉRES list, https://www.hceres.fr/en/publications/liste-des-revues-et-des-produits-de-la-recherche-hceres-pour-le-domaine-shs1-1). However, the HCÉRES list divides journals into only three classes, which makes it less informative for our purposes.

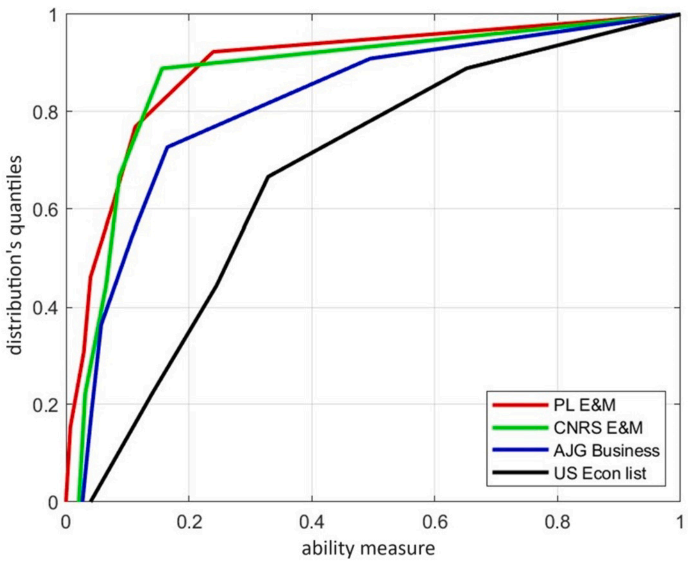

Fig. 5. Induced cumulative distribution functions of researchers' abilities computed for four economics and management journal ratings (PL,CNRS,AJG,US). The ability values on the horizontal axis are scaled using the order-preserving transformation $\sqrt{a}$ for better visibility. The values on the vertical axis are the quantiles of the corresponding distributions.

economics category in the *Journal Citation Reports* and that have citable items in all of the years 2014–2019 (Konig et al., 2022).[19] This list includes most of the newly established high-quality journals in economics.

A great advantage of the Konig et al. (2022) ranking is that it recognizes the uncertainty that is inherently present in the measurement of journal quality. Instead of giving only point estimates for the measures of journal quality and the quality ranks, it reports the confidence intervals. See also Lyhagen and Lyhagen (2020) for a related recent study. We use these intervals for the robustness check of our results.[20]

4.4. Empirical results

For the inputs for each of the four selected reward schemes for economics and management {CNRS, AJG, US, PL}, we define the set of journals, J, as the journals that are assigned a RIF measure. For each reward scheme, we create an additional class consisting of all the journals in J that are not assigned a rank by this reward scheme. For each $j \in J$, we set $\phi(j)$ as equal to journal j's RIF value. We set the values for the parameters as the slope $\xi = 2$ and the cut percentile $k = 20$.

Fig. 5 presents the induced CDFs for each reward scheme. For better visibility, we have transformed the ability values on the horizontal axis with the square root (in fact any strictly increasing transformation preserves the order). The US distribution stochastically dominates the remaining distributions (i.e. it is shifted to the right in relation to them), and the AJG distribution dominates the CNRS and PL distributions. The CNRS distribution dominates the PL distribution except for quantile values in the interval 0.75–0.9. Our model indicates that of the four schemes, the US population has the highest induced ability and the PL and CNRS the lowest, while the AJG is somewhere in between; the CNRS is actually better than the PL distribution for most quantile values except the quantile values between 0.75 and 0.9.

The algebraic example we consider allows us to interpret the distribution of abilities computed through the abilities of the boundary researchers and the best-quality journals that are within their range. The boundary researcher, that is, a researcher who is indifferent between choosing the top and the second top rank, according to the US list, for example, has a positive probability of publishing in all the top five journals apart from QJE. The boundary researcher publishing according to the AJG distribution has a chance of publishing in JPE (0.075), RES (0.15) and AER (0.21), but the probability of publishing in ECTA or QJE is zero. For France, the boundary researcher of the top rank has a positive probability of reaching Economic Theory (0.17), J Labour E (0.28) or J Risk & Uncer (0.18), for example, but the top 50 journals from the RIF list are effectively out of their range. For Poland, the boundary researcher of the top rank can publish in RAND J of Economics (0.019), or Review of Economic Dynamics (0.05), but the top 25 journals from RIF are out of range. A similar interpretation can be provided for other ability levels.

[19] The method was originally applied for a sample of 37 economics journals with citations from 1993–99 (Palacios-Huerta & Volij, 2004), then extended to 159 journals with citations from 1994–98 (Kalaitzidakis et al., 2003), 261 journals with citations from 2003–05 (Ritzberger, 2008), and 376 journals with citations from 2015–2019 (Ham et al., 2021). See also Amir and Knauff (2008) for an interesting application of this method for the ranking of economics departments and a recent axiomatic characterization of journal rankings by Csató (2019).

[20] See Supplementary Material, Online Appendix C.

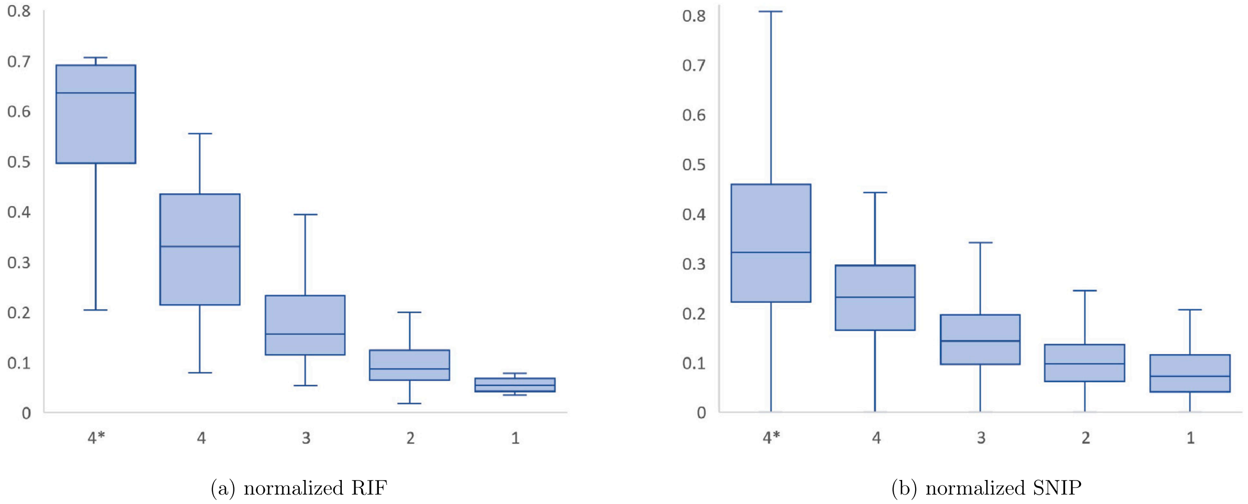

Fig. 6. Box-plots of the values for journal quality for the AJG Business classes; the lines of the box mark the quartiles and the whiskers mark the minimum and the maximum of the data points, excluding outliers.

To check the robustness of our results, we examined the impact of different values for the parameters ξ and k in the Supplementary Material, Online Appendix C. Instead of a mean RIF value, we consider the minimum and maximum RIF values reported by Konig et al. (2022). The robustness analysis indicates that our results are stable and do not change qualitatively due to model misspecification.

5. Discussion

5.1. Measures of journal quality

Our model crucially depends on the index of journal quality. There are clearly no universal standards for measuring journal quality. The commonly-used journal impact measures, or JIMs, that are based on the frequency of citations of papers published in a journal have been regularly criticized, and numerous alternatives have been proposed (see e.g. Bornmann et al., 2018; Haddawy et al., 2016; Wang et al., 2017; Petersen et al., 2019; Leydesdorff et al., 2019; Olszewski, 2020; Franceschet & Colavizza, 2017; Sjögårde & Didegah, 2022; Gorraiz et al., 2022 intensive discussion in Scientometrics in 2009-2012; and many others).

In our baseline example we focus on comparing economics and management journals for which some alternative measures, such as the RIF, have been proposed and calculated. The RIF circumvents many common problems with the standard impact factors. In particular, it weights citations by their importance within the field, and thus is immune to manipulability through excessive self-citation (Martin, 2016; Seeber et al., 2019), or the inclusion of *grey* journals that inflate citations (Oviedo-García, 2021).

However, RIFs are only available for a subset of all the journals that are listed in the many popular journal rating schemes (see Table B1 in Supplementary Material, Online Appendix C). We thus also report the results with the Source Normalized Impact per Paper (SNIP) as the index of journal quality, as SNIP data is freely available for most journals that operate globally.[21] Our results reported in Online Appendix C show that our qualitative results remain stable with respect to this modification. This further strengthens validation of our method of reverse engineering the ability distribution from the observable reward schemes.

However, we are aware that neither RIF nor SNIP perfectly captures the real RSB objective. Fig. 6 plots the distribution of the RIF (left panel) and SNIP (right panel) values into classes for the example of AJG. The values of the indexes are rescaled and normalized for better visibility. This does not affect our results, as in the model we are only interested in the ordinal properties of the measures of journal quality.

Although the means and the medians of the measures increase for the higher-quality classes, the reward schemes are not fully monotonic in the measures of journal quality. The problem seems to be more pronounced for the SNIP data, which also contain more journals; of the 805 journals listed in the AJG scheme, 675 had a SNIP value, and only 206 had RIF data assigned (see Supplementary Material, Online Appendix Table B1 for more data).

Fig. 6 illustrates the problem. For SNIPs, the lower quartile of the box for a higher class is below the upper quartile of the box to the right, indicating substantial non-monotonicity. It is much less evident for RIFs, where such overlap is minimal. If a lower-quality journal is misplaced to a higher category, it causes the calibrated distribution function to become steeper for low quantiles, as low-ability researchers can reach highly rewarded journals, which are not necessarily of high quality. If, for one ranking list, such non-monotonicities are plentiful and, for another one, rare, the distribution function of the induced ability will be initially steeper

[21] We use SNIP 2020 available at https://www.scopus.com/sources. It is described as a metric that "intrinsically accounts for field-specific differences in citation practices." For more information on the metric see https://www.elsevier.com/ authors/tools-and-resources/measuring-a-journals-impact.

than the latter one. Thus, the effect will be similar to enlarging the upper classes at the expense of lower ones, exemplified in Section 1 of the paper.

5.2. Do people follow the incentives provided by the RSB?

Our model crucially depends on the response of researchers to the incentives provided by the RSB. In this section, we seek empirical confirmation that this is indeed the case, and in particular that researchers aim for the "cheapest" journals. To what extent this is true can be checked by observing the change in the publication strategy of researchers in Poland in response to the introduction by the Ministry of Education and Science in 2019 of the official ranking list of journals.[22]

Among the highly-ranked journals, there are some open-access mass publication journals that publish a very large number of articles online in each issue and are quite lenient in their acceptance policy. Of particular note are the journals owned by the Multidisciplinary Digital Publishing Institute (MDPI). This makes it likely that such highly ranked MDPI journals would be regarded as the "cheapest" ones, in the sense that the probability of acceptance would be substantially higher for them than for the other journals in this class. Proposition 4 states that researchers should aim to publish in these journals, as they are likely to have a higher probability of acceptance than other journals in this class. The counterargument is that researchers might avoid publishing in mass-publication journals because of their poor academic reputation (see e.g. Oviedo-García (2021)). However, the evidence from Poland overwhelmingly supports the strategy described by our model. In the ranking of the Polish Ministry of Education and Science, 11 MDPI journals have been assigned the second-highest of the six ranks, while 35 have the third-highest rank, and 26 have the fourth rank. There are no MDPI journals in the first rank.

Before the first information about the contents of the new list became available in 2019, the percentage of papers in these 11 MDPI journals that were authored or co-authored by researchers with affiliation at Polish universities was 3.3%, making 1709 papers. Between 2019 and May 2023, this fraction rose to 9.7%, which corresponds to over 21K papers published by Polish authors. Official statistics show there were around 45K academics working at Polish universities between 2019 and 2022, and so it appears that on average, nearly half of all Polish academics published a paper in one of these journals. Our model is further supported by the evidence that Polish researchers were substantially less keen to publish in lower-ranked MDPI journals, and their keenness was further reduced as the rank assigned to these journals decreased. In 2019-2023, they published 14.5K papers in MDPI journals that were officially ranked in the third class, which is 4.5% of the total number of articles, and over 3K articles or 2.1% in the fourth-ranked MDPI journals.

6. Conclusions

Our paper looks into the role of rankings of academic journals in incentivizing the efficient dissemination of research output through publications. An optimally constructed ranking of journals and the related reward system should encourage authors to direct their output to journals that are appropriate to their abilities. This can be done by setting the thresholds for ranks so that they maximize the expected quality for the authors. At the same time, this choice should contribute to maximizing the expected publication quality of the entire population of researchers. Our theoretical model shows how to construct such a system of rewards, and the algebraic example proves that this is feasible and intuitively convincing. The model is parsimonious and, hence, based on simplifying assumptions. Out of the extensions, we plan to work on in the future, the most important is to consider how the academic journals rankings shape the long run distribution of abilities in the population by providing incentives to improve ones abilities (especially these of younger researchers) as well as by selection of agents with most suitable abilities to the academia (including researchers' mobility between the countries). Apart from that, allowing for endogenous effort and hence ability to improve the quality of submitted paper (for example in the revision or resubmission process) or simply increase the quantity of the produced paper's in the evaluation window seems to be another important generalization that can affect the derived optimal academic journal ranking. These extensions require, however, to model a publication strategy as an outcome of a dynamic game which is beyond the scope of the current paper.

We have applied reverse engineering to calibrate the model and constructed the implied distribution of the abilities of authors for different populations of researchers. For economics and management, we have found out that the creators of the Academic Journal Guide ranking list in the UK see their population as more able, than those who make the equivalent rankings for France and Poland. The list for Poland is the most lenient, meaning it is the flattest of all the rankings compared.

The results of our model lead to important direct and indirect policy conclusions. Firstly, research supervisory bodies should use a good measure of journal quality and construct the rating that is monotone with this measure. A nonmonotonic increase in the rating of even few low-quality journals may lead to a massive increase in publications in these journals and result in a large loss of overall quality. Secondly, due to wide differences in publication practices and standards, connections with business, trends, and the popularity of specific areas of scientific research, it is very difficult to identify a measure of journal quality that is appropriate for all scientific disciplines at the same time. This, in turn, suggests that separate reward schemes for journals from a given discipline are better than a large centralized journal evaluation system. Thirdly, comparing the rankings applied in various countries by the calibrated distribution of abilities gives incentives to reevaluate the research policies used. More specifically, estimating and comparing the abilities through calibration of the indirect distribution functions is also cheaper and less questionable than the complex evaluation

[22] See also Yuret (2017) for a related study and evidence from Turkey and Śpiewanowski and Talavera (2021) from UK.

of the academic output of the entire population of researchers. It also makes it feasible and easy to repeat every time the rankings are changed. Such repetition facilitates a straightforward evaluation of changes in the position of the country's research level in relation to other countries, which could have direct implications regarding the intensity of the research policy used.

In further perspective, our results can be of use as a starting point to research supervisory bodies, which can put their efforts into constructing ranking lists that will better motivate authors to direct their output to journals that maximize the overall publications' quality of the discipline. This could be achieved if *forward* rather than reverse engineering is applied. This can be done by conducting a detailed analysis of publications and citations in a given population. A database of publications collected for such an analysis would also allow to test empirically the publication' incentives applied in various countries. This is left for further research.

CRediT authorship contribution statement

Wojciech Charemza: Writing – original draft, Methodology, Formal analysis, Conceptualization. **Michał Lewandowski:** Writing – original draft, Methodology, Formal analysis, Conceptualization. **Łukasz Woźny:** Writing – original draft, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Acknowledgements

We would like to thank Rabah Amir, Wojciech Olszewski, Klaus Ritzberger, and Richard Tol for valuable discussions during the writing of this paper. We are also grateful to Tomasz Jurkiewicz and Jędrzej Siciński for their help in data collection. Wojciech Charemza gratefully acknowledges the support granted by the Ministry of Education and Science of Poland to Vistula University. Łukasz Woźny thanks the National Science Center, Poland [NCN grant number UMO-2019/35/B/HS4/00346] for the financial support of this project. Michał Lewandowski thanks the National Science Center, Poland [NCN grant number NCN grant number UMO-2021/41/B/HS4/03429] for the financial support during this project.

Appendix. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.joi.2024.101559.

References

Abramo, G., Aksnes, D. W., & D'Angelo, C. A. (2020). Comparison of research performance of Italian and Norwegian professors and universities. *Journal of Informetrics*, *14*(2), Article 101023.

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2024a). Do research assessment systems have the potential to hinder scientists from diversifying their research pursuits? *Scientometrics*, 1–21.

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2024b). The moderating role of personal characteristics of authors in the publications' quality for quantity trade-off. *Journal of Informetrics*, *18*(1), Article 101495.

Adam, E. (2020). 'Governments base performance-based funding on global rankings indicators': A global trend in higher education finance or a global rankings literature fiction? A comparative analysis of four performance-based funding programs. *International Journal of Educational Development*, *76*, Article 102197.

Amir, R., & Knauff, M. (2008). Ranking economics departments worldwide on the basis of PhD placement. *Review of Economics and Statistics*, *90*(1), 185–190.

Baccini, A., & De Nicolao, G. (2022). Just an artifact? The concordance between peer review and bibliometrics in economics and statistics in the Italian research assessment exercise. *Quantitative Science Studies*, *3*(1), 194–207.

Bornmann, L., Butz, A., & Wohlrabe, K. (2018). What are the top five journals in economics? A new meta-ranking. *Applied Economics*, *50*(6), 659–675.

Card, D., & DellaVigna, S. (2020). What do editors maximize? Evidence from four economics journals. *Review of Economics and Statistics*, *102*(1), 195–217.

Casella, G., & Berger, R. L. (2002). *Statistical inference*. Duxbury: Thomson Learning.

Csató, L. (2019). Journal ranking should depend on the level of aggregation. *Journal of Informetrics*, *13*(4), Article 100975.

De Boer, H., Jongbloed, B., Benneworth, P., Cremonini, L., Kolster, R., Kottmann, A., Lemmens-Krug, K., & Vossensteyn, H. (2015). *Performance-based funding and performance agreements in fourteen higher education systems*. Center for Higher Education Policy Studies.

Drivas, K., & Kremmydas, D. (2020). The Matthew effect of a journal's ranking. *Research Policy*, *49*(4), Article 103951.

Foster, J. (2011). *Evolutionary macroeconomics: A research agenda*. Springer.

Franceschet, M., & Colavizza, G. (2017). Timerank: A dynamic approach to rate scholars using citations. *Journal of Informetrics*, *11*(4), 1128–1141.

Gorraiz, J., Ulrych, U., Glänzel, W., Wenceslao, A.-M., & Torres-Salinas, D. (2022). Measuring the excellence contribution at the journal level: An alternative to Garfield's impact factor. *Scientometrics*, *127*(12), 7229–7251.

Haddawy, P., Hassan, S.-U., Asghar, A., & Amin, S. (2016). A comprehensive examination of the relation of three citation-based journal metrics to expert judgment of journal quality. *Journal of Informetrics*, *10*(1), 162–173.

Ham, J. C., Wright, J., & Ye, Z. (2021). New rankings of economics journals: Documenting and explaining the rise of the new society journals. Available at https://ssrn.com/abstract=3606030.

Kalaitzidakis, P., Mamuneas, T. P., & Stengos, T. (2003). Rankings of academic journals and institutions in economics. *Journal of the European Economic Association*, *1*(6), 1346–1366.

Konig, J., Stern, D. I., & Tol, R. S. (2022). *Confidence intervals for recursive journal impact factors*. CESifo Working Paper No. 9780.

Korytkowski, P., & Kulczycki, E. (2019). Examining how country-level science policy shapes publication patterns: The case of Poland. *Scientometrics*, *119*(3), 1519–1543.

Kosyakov, D., & Pislyakov, V. (2024). "I'd like to publish in Q1, but there's no Q1 to be found' ": Study of journal quartile distributions across subject categories and topics. *Journal of Informetrics*, *18*(1), Article 101494.

Kulczycki, E., Engels, T. C. E., Pölönen, J., Bruun, K., Dušková, M., Guns, R., Nowotniak, R., Petr, M., Sivertsen, G., Starčič, A. I., & Zuccala, A. (2018). Publication patterns in the social sciences and humanities: Evidence from eight European countries. *Scientometrics*, *116*, 463–486.

Laffont, J.-J., & Martimort, D. (2001). *The theory of incentives: The principal-agent model*. Princeton University Press.

Leydesdorff, L., Bornmann, L., & Adams, J. (2019). The integrated impact indicator revisited (I3*): A non-parametric alternative to the journal impact factor. *Scientometrics*, *119*(3), 1669–1694.

Lyhagen, J., & Lyhagen, P. (2020). Uncertainty and the ranking of economics journals. *Scientometrics*, *125*, 2545–2560.

MacLeod, W. B., & Urquiola, M. (2021). Why does the United States have the best research universities? Incentives, resources, and virtuous circles. *The Journal of Economic Perspectives*, *35*(1), 185–206.

Martin, B. R. (2016). Editors' JIF-boosting stratagems–which are appropriate and which not? *Research Policy*, *45*(1), 1–7.

Mogstad, M., Romano, J., Shaikh, A., & Wilhelm, D. (2022). Statistical uncertainty in the ranking of journals and universities. *AEA Papers and Proceedings*, *112*, 630–634.

Narin, F., Pinski, G., & Gee, H. H. (1976). Structure of the biomedical literature. *Journal of the American Society for Information Science*, *27*(1), 25–45.

Ochsner, M., Kulczycki, E., Gedutis, A., & Peruginelli, G. (2021). National research evaluation systems. In *Handbook bibliometrics* (pp. 99–106). De Gruyter Saur.

Olszewski, W. (2020). A theory of citations. *Research in Economics*, *74*(3), 193–212.

Oviedo-García, M. A. (2021). Journal citation reports and the definition of a predatory journal: The case of the Multidisciplinary Digital Publishing Institute (MDPI). *Research Evaluation*, *30*(3), 405–419.

Palacios-Huerta, I., & Volij, O. (2004). The measurement of intellectual influence. *Econometrica*, *72*(3), 963–977.

Palacios-Huerta, I., & Volij, O. (2014). Axiomatic measures of intellectual influence. *International Journal of Industrial Organization*, *34*, 85–90.

Petersen, A. M., Pan, R. K., Pammolli, F., & Fortunato, S. (2019). Methods to account for citation inflation in research evaluation. *Research Policy*, *48*(7), 1855–1865.

Ritzberger, K. (2008). A ranking of journals in economics and related fields. *German Economic Review*, *9*(4), 402–430.

Seeber, M., Cattaneo, M., Meoli, M., & Malighetti, P. (2019). Self-citations as strategic response to the use of metrics for career decisions. *Research Policy*, *48*(2), 478–491.

Sjögårde, P., & Didegah, F. (2022). The association between topic growth and citation impact of research publications. *Scientometrics*, *127*(4), 1903–1921.

Smit, J. P., & Hessels, L. K. (2021). The production of scientific and societal value in research evaluation: A review of societal impact assessment methods. *Research Evaluation*, *30*(3), 323–335.

Śpiewanowski, P., & Talavera, O. (2021). Journal rankings and publication strategy. *Scientometrics*, *126*, 3227–3242.

Thomas, D. A., Nedeva, M., Tirado, M. M., & Jacob, M. (2020). Changing research on research evaluation: A critical literature review to revisit the agenda. *Research Evaluation*, *29*(3), 275–288.

Vîiu, G.-A., & Păunescu, M. (2021). The citation impact of articles from which authors gained monetary rewards based on journal metrics. *Scientometrics*, *126*(6), 4941–4974.

Vogel, R., Hattke, F., & Petersen, J. (2017). Journal rankings in management and business studies: What rules do we play by? *Research Policy*, *46*(10), 1707–1722.

Wang, J., Veugelers, R., & Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, *46*(8), 1416–1436.

Yuret, T. (2017). Do researchers pay attention to publication subsidies? *Journal of Informetrics*, *11*(2), 423–434.

Zacharewicz, T., Lepori, B., Reale, E., & Jonkers, K. (2019). Performance-based research funding in EU member states - comparative assessment. *Science & Public Policy*, *46*(1), 105–115.