On journal rankings and researchers' abilities*

Wojciech Charemza† Michał Lewandowski‡ Łukasz Woźny§

March 2024

Abstract

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 answer these questions, we propose a simple model of optimal rewards for publication in academic journals. Based on a principal-agent model with researchers' hidden abilities, we characterize the second-best 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 from the observed journal ranking schemes.

Keywords: journal rankings, publication reward mechanisms, optimal categorization, journal quality

JEL classification: I23, D61, O31

1 Introduction

There is growing interest in introducing financial and administrative systems that could encourage publication and improve the performance of research-oriented institutions.

*We would like to thank Rabah Amir, Wojciech Olszewski, Klaus Ritzberger, Marzena Rostek, 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.

†Faculty of Business and International Relations, Vistula University, 02-787 Warsaw, Poland.

‡Department of Quantitative Economics, SGH Warsaw School of Economics, 02-554 Warsaw, Poland.

§Corresponding author. Department of Quantitative Economics, SGH Warsaw School of Economics. Address: al. Niepodleglosci 162, 02-554 Warszawa, Poland. E-mail: lukasz.wozny@sgh.waw.pl.

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 in-depth assessments, analyzes and comparisons between countries or disciplines.

These assessments, which can be more or less varied and detailed, are also used by researchers in many countries 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 as they were originally, 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.

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

[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); Thomas et al. (2020); Vogel et al. (2017);

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 that admits a closed-form solution, producing the constrained-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 adverse selection.[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

Zacharewicz et al. (2019) or Smit and Hessels (2021) at a more general level, and a recent contribution by Mogstad et al. (2022) analyzing journal ranks that aims to minimize the statistical uncertainty associated with the indexes of journal citations.

[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.

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 prestige, where prestige is measured here by a journal quality measure[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 prestigious 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, working from 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 from the observable 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.

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

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, higher-ability researchers should optimally publish in higher-quality journals. Second, 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 prestige. We assume that the rating schemes never assign higher-quality journals to a lower class, a property we call monotonicity, while researchers from both populations only care about the rating of the journal in which their work will be published. Moreover, for a given researcher, the higher the quality of the journal, the more difficult it is to get the article accepted.

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

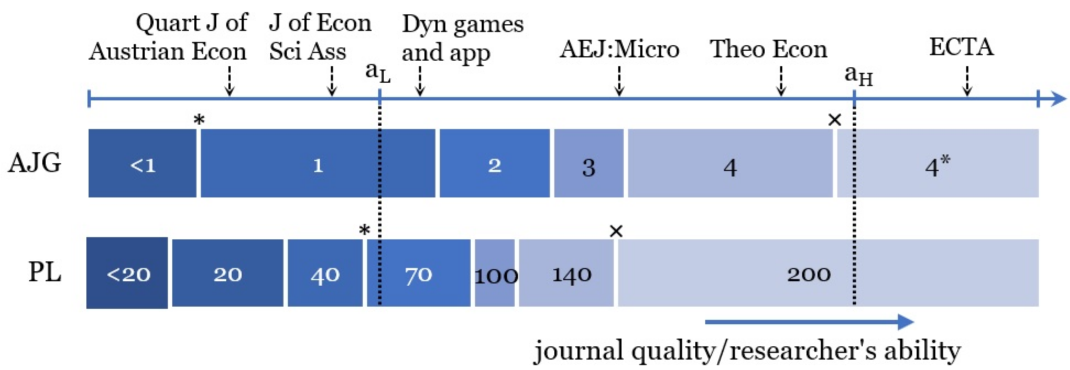
Journal	PL	AJG
Econometrica	200	4*
Theoretical Economics	200	4
AEJ: Microeconomics	200	3
Dynamic Games and Applications	70	1
Journal of the Economic Science Association	40	1
Quarterly Journal of Austrian Economics	20	1

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

[4]This paper combines rankings for two disciplines in the Polish rating: Economics & Finance, and Management.

would choose in the social optimum. Figure 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 size and position of the categories reflect 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.

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



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 Figure 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 prestige 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 Figure 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 model and its key assumptions are presented in Section 2. We characterize the optimal RSB solution for an arbitrary distribution of abilities and arbitrary probability of acceptance functions. We discuss both the hypothetical first-best scheme, in which each journal can be assigned a distinct reward or rank, and the second-best, where all the journals are grouped into a finite number of classes. The objective of the RSB for the general case does not allow a closed-form solution, so in Section 3 we use a simplified model where the acceptance probability functions are derived from a family of step-wise linear functions that satisfy some key stylized facts. This gives us a closed-form, optimal reward scheme. We do this for any distribution of ability. Section 4 shows how to reverse engineer the optimal solution obtained in Section 3 to calibrate the distribution of ability within the population.[5] We then compare several well-known journal ratings using this method. In Section 5 we discuss the limitations and possible extensions of the model. Appendix A contains proofs of the propositions from Section 2 and 3, and Appendix B contains 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 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:[6]

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 that researchers are risk-neutral and maximize expected reward, with the payoff for not publishing anything normalized to 0. That these assumptions can be relaxed without changing our qualitative results is shown in Section 5. The researcher's a problem is thus:

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

[6]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.

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. In the section 5.2 we show that these assumptions do not qualitatively affect our results. 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)*.*

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 fully monotonic in ϕ would create an unwarranted sense of precision (see Konig et al., 2022, p.2).[7] 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, and in our model with a continuum of journals, the second-best solution occurs when there is a finite number of journal categories. We consequently restrict the reward schemes in (2) to those that allow only $n \geq 1$ distinct non-zero rewards, where n is given exogenously.[8]

[7] 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 4.3.

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

The question is then how to partition the journals into categories and what reward levels should be set for each category. We start by proving that no quality is lost by restricting the policy R to be a family of non-decreasing, right-continuous step functions.

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:

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

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 optimal reward scheme in (2) in the family (3), meaning we find the set of weights $\alpha_1 < \ldots < \alpha_n$ and boundary journals $\phi_1, \ldots, \phi_n$.

Proposition 4. *If the reward schemes are restricted to the family given by* (3)*, the reward scheme that satisfies* $\alpha_i = \alpha\phi_i$, $i \in \{1, \ldots, n\}$ *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 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.[9] 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 may summarize our findings for the second-best problem in the following Corollary. As before, without loss of generality, we set $\alpha = 1$.

[9]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 $a_{i/i+1}$.

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

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

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

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

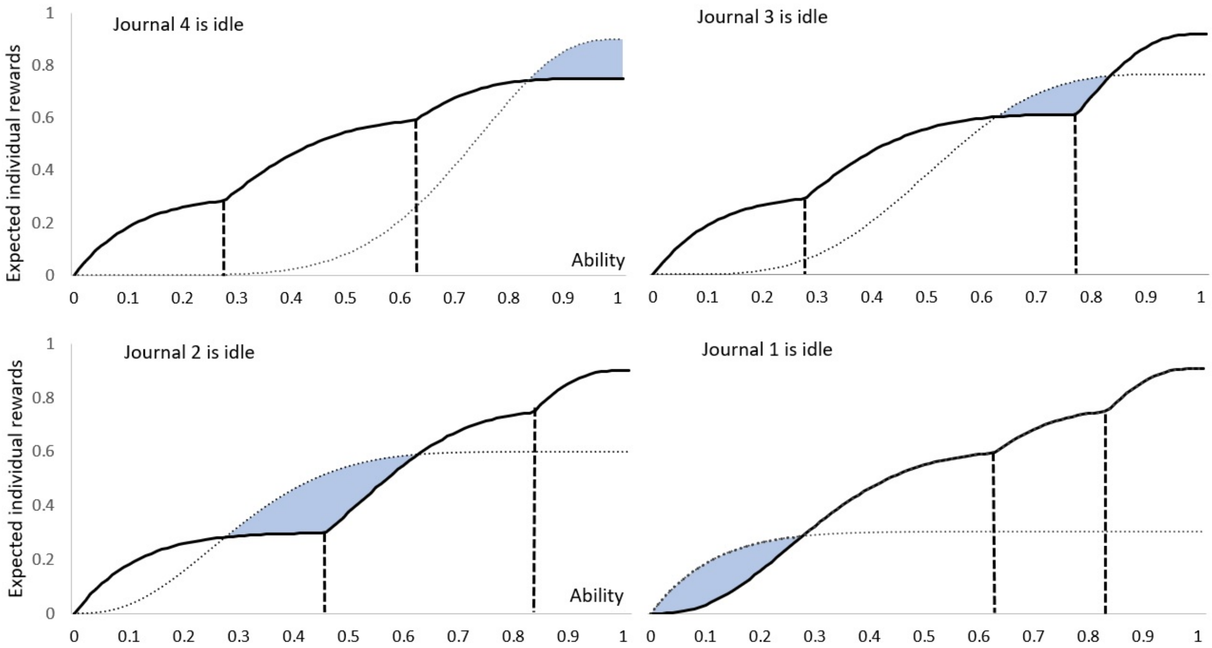
Supposing the probability of acceptance p satisfies Assumptions 1 and 2, we consider four journals with the quality levels $\phi_1, \phi_2, \phi_3, \phi_4 \in (0, 1)$, ordered from lowest to highest. If their rewards are given by $R_i = \alpha\phi_i$ for some $\alpha > 0$, then each researcher maximizes 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 prestige. Given our assumptions, the selection 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 Figure 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. 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 Figure 2, we notice that the removal of higher quality journals results in (point-wise) lower researcher's boundary types (dashed lines in Figure 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.

Figure 2: Second-best with three categories and four journals. Are shaded in blue corresponds to a loss of expected prestige as compared to the first-best case.



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:

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

In this, p takes the form of a CDF of a uniform distribution on $\left[\frac{\xi-1}{\xi}\phi, \phi\right)$.[10] 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[11] 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

[10] 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.

[11] See Appendix A.1 for a formal proof.

$\xi > 1$. 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 ETQ_I:

$$\text{ETQ}_\text{I} = \int_0^1 a dF(a) = \left[\frac{1}{2}a^2\right]_0^1 = \frac{1}{2}. \tag{8}$$

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):[12]

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

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:

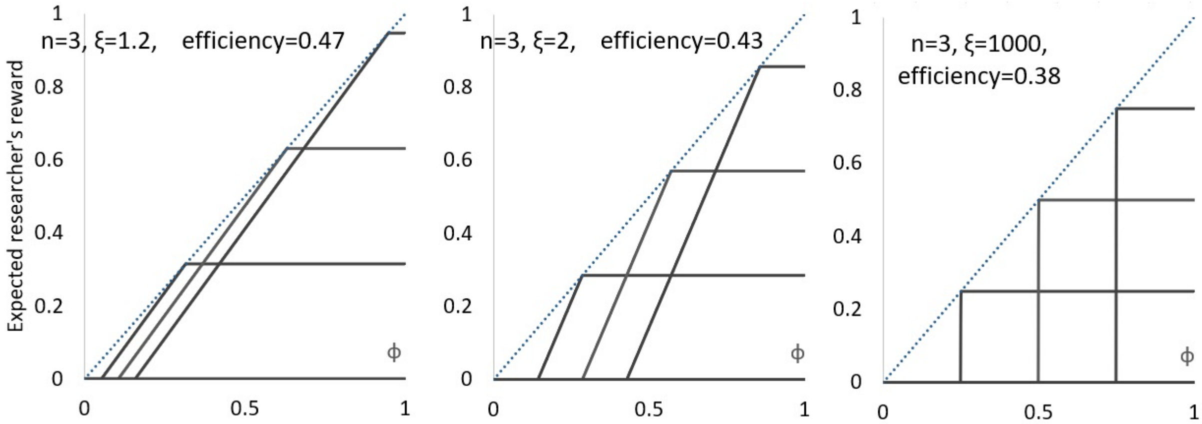
$$\text{ETQ}_\text{II}(\xi) = \max_{\phi_1,...,\phi_n} \sum_{i=1}^{n} \int_{a_{i-1/i}}^{\phi_i} (\xi a + (1-\xi)\phi_i) da + \sum_{i=1}^{n-1} \int_{\phi_i}^{a_{i/i+1}} \phi_i da + \int_{\phi_n}^{1} \phi_n da.$$

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

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

[12] Note that $a_{0/1}$ is technically not a crossing point, but if it is given the form of (7) it proves convenient in the notation. We therefore also use $a_{i-1/i}$ instead of $a_{i/i+1}$.

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



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. Figure 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 $\mathrm{ETQ}_{\mathrm{II}}(\xi)$, 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 $\mathrm{ETQ}_{\mathrm{II}}(\xi)$ 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.

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 and 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:

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

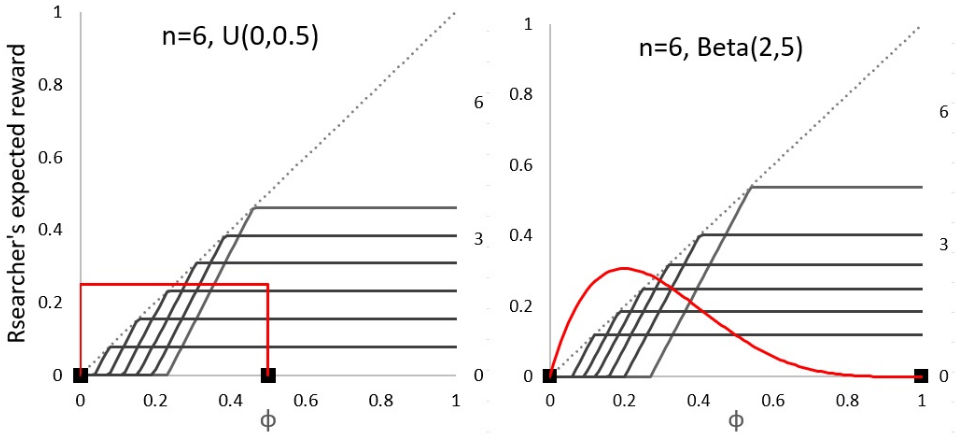
The optimal solution for these distributions satisfies: $F_{[a_L,a_U]}(\phi_i) = \frac{\xi i}{\xi(n+1)-1}$, or again following Proposition 4:

$$\phi_i = a_L + (a_U - a_L)F^{-1}\left(\frac{\xi i}{\xi(n+1)-1}\right),$$

while the objective function value remains identical for the original distribution F and the modified one $F_{[a_L,a_U]}$.

For illustration, We now compare two populations of researchers, for which we assume $\xi = 2$ and analyze the optimal solution for two selected distributions of the abilities. Figure 4 presents the optimal solution, with boundary journals and the researcher's envelope,

Figure 4: Optimal solution for different distributions of ability levels.



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.

We will now use the closed-form solution (10) to reverse engineer the distribution F from the observable reward schemes and journal quality measures used in practice.

4 Distributions of ability induced from journal ratings

4.1 Outline of empirical evaluation

Our data consist 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[13] 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

[13]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 and 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.

with the second-best policy (5).

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) (see also Section 5.2 for details), 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

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

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.[14]

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.[15]

Our procedure for finding a distribution of the ability of researchers $F : [0, 1] \to [0, 1]$ can be summarized as:

1. Inputs: the set of journals partitioned into classes $J = J_1 \cup ... \cup J_n$, , and a normalized

[14] Moreover, applied reward schemes are usually not fully monotonic in the measure of journal quality, as there are pairs of journals where the lower-quality journal is in a more prestigious class, see further in Section 4.3.

[15] 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.

measure of journal impact $\phi : J \to [0, 1]$.

2. Parameters: the slope ξ; the cut percentile k.
3. Set the cheapest journals $\phi_i := G_i^{-1}(k)$.
4. Set the corresponding quantile values $F(\phi_i) = \frac{\xi i}{\xi(n+1)-1}$.
5. Set the boundary values $F(0) = 0$ and $F(1) = 1$ and extrapolate the points of the CDF linearly.

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 (UK),

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), used by some economic departments in the US to support promotion and hiring decisions.

The CNRS and PL ratings are developed within the European PBRF program and explained briefly in Section 1 of this paper. 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 is widely used as an indicative measure of the quality of journals by business schools around the world.

Journals on the PL list are divided into six classes that are labelled 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[16] 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 B.

4.3 Measures of journal quality

Our model and how it is applied depends crucially on the index of journal quality ϕ. 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. Haddawy et al., 2016; Leydesdorff et al., 2019; Olszewski, 2020; Petersen et al., 2019; Wang et al., 2017 intensive discussion in Scientometrics in 2009-2012; and many others). The main problems in using JIMs as measures of the quality of journals are: (i) manipulability through excessive self-citation (Martin, 2016; Seeber et al., 2019), (ii) inclusion of *grey* journals or others of dubious reputation that inflate citations systematically in various ways on a large scale (Oviedo-García, 2021), (iii) delayed response to novelty papers and newly established quality journals that publish innovative frontier results (Wang et al., 2017), (iv) Tendency to ignore or under-represent papers published in limited-circulation journals, focused on country-specific issues, or published in languages other than English (Wuestman et al., 2019). To address some of these issues in our baseline example, we focus on comparing economics and management journals for which some alternative measures of journal quality using different criteria

[16]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.

have been proposed and calculated. One such method is 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 and Volij, 2014).This method weights citations by their importance within the field, thus circumventing problems (i) and (ii) listed above. We use the most recent updated version, with 319 journals with citations from 2014–19 (Konig et al., 2022), which includes most of the newly established high-quality journals in economics.[17] We call this measure the recursive impact factor, or RIF. Finally, problem (iv) is less of an issue in economics and management than in it is humanities or linguistic studies, where country-specific issues and language play a greater role. 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. We use these intervals for the robustness check of our results.

The quality indexes based on the invariant method are only available for a subset of all the journals that are listed in the many popular journal rating schemes (see Table B1 in Appendix B). We thus also report the results with the Source Normalized Impact per Paper (SNIP) as the index of journal quality, as data on a wide set of journals are freely available for this.[18] SNIP is described as a metric that "intrinsically accounts for field-specific differences in citation practices." We thus perform the robustness check for our results against a different measure of journal quality by comparing SNIP and RIF.

Unlike the RIF, for which data are available only for those journals that are listed in the economics category in the *Journal Citation Reports* and that have citable items in all of the years 2014–2019 (Konig et al., 2022, p.4), the SNIP data are in principle available for most journals that operate globally. For the baseline example where we use the RIF

[17]The method was originally applied for a sample of 37 economics journals with citations from 1993–99 (Palacios-Huerta and 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.

[18]We use SNIP 2020 available at `https://www.scopus.com/sources`. For more information on the metric see `https://www.elsevier.com/authors/tools-and-resources/measuring-a-journals-impact`.

data, we can consequently define the set of journals J for each reward scheme as the set of journals for which the RIF measure is calculated. 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. In the robustness analysis where we use SNIP data by contrast, the set of journals for a given reward scheme is defined as the set of all the journals listed in that scheme. Some of these journals do not have SNIP values, though this is usually a small fraction. We assign those journals a SNIP value of zero. We are aware that some of these journals are good quality new journals, but they represent only a small fraction of the journals with a zero SNIP, and therefore they cannot significantly bias the overall quality assessment.

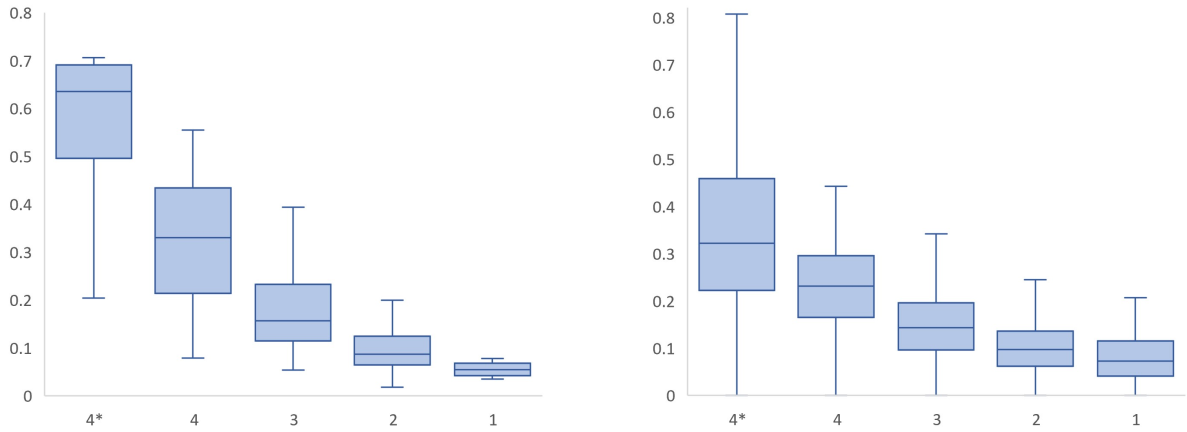
Figure 5 plots the distribution of the RIF (left panel) and SNIP (right panel) values into classes for the example of AJG. The values of the indexes are rescaled and normalized for better visibility. This does not affect our results, as in the model we are only interested in the ordinal properties of 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 Table B1 for more data). This reflects the two problems discussed above and indicates that most probably neither RIF nor SNIP is a perfect proxy of what an RSB, here the AJG, maximizes.

4.4 Empirical results

We now turn to our baseline example. 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 $j \in J$, 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 $k = 20$ (see section B.2 for robustness analysis using different values for ξ and k). Figure 6 presents the induced CDFs for each reward scheme. For

Figure 5: 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.



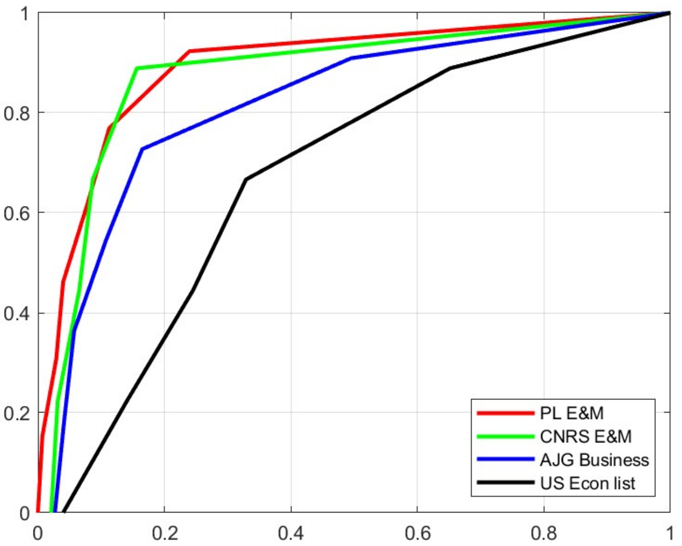
(a) normalized RIF (b) normalized SNIP

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 first-order stochastically dominates the remaining distributions, and the AJG dominates CNRS and PL. CNRS dominates PL 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 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 publishing in the 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 in the top rank in the UK 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 in 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

Figure 6: Induced distributions computed for four economics and management journal ratings. The values on the horizontal axis are given by $\sqrt{a}$



the RIF list are effectively out of their range. For Poland, the boundary researcher in 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.

To check the robustness of our results, we examined the impact of different values for the parameters ξ and k in Appendix B. Instead of a mean RIF value, we consider the minimum and maximum RIF values reported by Konig et al. (2022). We also consider SNIP as an alternative measure of journal quality instead of RIF. The robustness analysis performed in Appendix B indicates that our results are stable and do not change qualitatively due to model misspecification. This confirms our findings and validates the method of reverse engineering the ability distribution from the observable reward schemes.

5 Discussion and extensions

5.1 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. However, this requires finding 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.

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 21 thousand papers published by Polish authors. Official statistics show that there were around 45 thousand 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 14.5 thousand papers in MDPI journals that were officially ranked in the third class, which is 4.5% of the total number of articles, and over 3 thousand articles or 2.1% in the fourth-ranked MDPI journals.

5.2 Extensions of the basic model

In this section, we discuss a few natural extensions of the benchmark model.

Risk aversion and general utility function In the benchmark model, researchers are risk neutral. This implies that they linearly weigh the probabilities of acceptance with the reward for publishing in a journal of a given rank. Researchers who are risk averse may, however, be willing to choose *safer* journals that have a higher probability of acceptance rather than risking long shots with their submissions. We now discuss how the optimal journal reward scheme and journal ranking are affected by risk aversion in the first and second-best schemes. For this we suppose that the preferences of researchers are now: $u(R(\phi))p(\phi, a)$ for some strictly increasing utility $u : \mathbb{R} \mapsto \mathbb{R}$. For risk aversion this utility is assumed to be strictly concave, and $h := u^{-1}$ is set as the inverse of this utility. Without loss of generality, we can now consider an RSB that chooses not the reward scheme $R : \Phi \mapsto \mathbb{R}$, but the utility values for publishing in these journals, or $\phi \mapsto u_\phi \in \mathbb{R}$. Then the objective of the RSB is still $\int_A \phi^u(a)p(\phi^u(a), a)F(da)$, where $\phi^u(a) \in \arg\max_{\phi\in[0,1]} u_\phi p(\phi, a)$. We observe that the optimal journal rank in the second-best scheme is unchanged from our baseline model. What does change under this generalization is the reward scheme for publishing in journal ϕ in both the first and second-best schemes, as: $u_\phi = \alpha\phi$ and so $R(\phi) = h(\alpha\phi)$. This means that although the optimal reward scheme is affected by the shape of utility u and particularly by its risk aversion, the calibrated distributions of abilities importantly are not affected by this generalization. We also assumed in the benchmark model that the reward for not getting a paper published is normalized to 0. This assumption also does not entail a loss of generality. Indeed, suppose that the reward system gives each researcher a flat wage R_0 and a bonus $R(\phi)$ for a successful publication. Now the researcher's objective is: $u(R(\phi))p(\phi, a) + u(R_0)$ but the choice of the optimal journal is unaffected. Similarly, the ranking of optimal journals remains unchanged and so do our calibrated distributions of abilities.

Participation constraint We did not have a participation constraint in the benchmark model, this being a condition that guarantees that some researchers do not prefer an outside option. Our optimal reward model can be generalized to include the addition of such a constraint as well. To do this, $w(a)$ denotes an outside option for a researcher with ability a. Here we still use the general utility function as considered in the previous paragraph. The participation constraint of researcher a is then: $u(R(\phi))p(\phi, a) + u(R_0) \geq w(a)$. Here $\underline{a}$ denotes the lowest level of ability of a researcher that decides to publish in the first, non-zero category in the second-best scheme by choosing journal ϕ_1. The RSB needs to ensure that the participation constraint is satisfied for all $a \geq \underline{a}$. Recall that $u(R(\phi)) = \alpha\phi$ and denote $\beta := u(R_0)$. Then the participation constraint becomes: $\alpha\phi p(\phi, a) + \beta \geq w(a)$, and for each $a > 0$ the RSB can choose β such that:

$$\beta := \max_{a \in [\underline{a}, 1]} \{w(a) - \alpha v(a)\}, \tag{12}$$

where $v(a) := \max_{\phi_1, \phi_2, \ldots, \phi_n} \phi p(\phi, a)$ is the maximal utility of the researcher a in the second-best scheme. As defined by the maximization problem (12), each researcher with $a \geq \underline{a}$ accepts the reward scheme proposed and stays in the academia market. This analysis clearly abstracts from the cost of reward or the RSB budget constraint, but as our main goal is to propose and analyze the optimal system for ranking journals, we leave those considerations for further studies.

Noisy signal of journal quality Journal quality rankings are typically only noisy signals of the true quality levels of journals. This generalization supposes that the RSB and the researchers do not observe journal quality ϕ directly, but observe it only through a noisy signal s. Suppose further that the distribution of the journal quality levels ϕ conditional on receiving signal s admits a density $h(\phi, s)$. Now $p(s, a)$ is the observed probability of acceptance in a journal with signal s, and since the true ϕ is unobserved, the conditional probability of acceptance can no longer depend on ϕ. Similarly, the RSB can now set rewards that are based only on the observed signal s, meaning R is a function of s now. Then the researcher's objective is given by: $R(s)p(s, a)$ with the argmax $S_R(a)$. The RSB objective is

$$\int_A \left(\int_\Phi \phi p(s_R(a), a) h(\phi, s_R(a)) d\phi \right) F(da)$$

for some measurable selection s_R from S_R. Denoting the expected quality of the journal with signal s by $EQ(s) := \int_\Phi \phi h(\phi, s) d\phi$, the RSB sets optimally $R(s) = \alpha EQ(s)$, in the first-best scheme for any $\alpha > 0$. Then the researcher's objective becomes: $\alpha \int_\Phi \phi p(s, a) h(\phi, s) d\phi$, while the RSB objective is: $\int_A EQ(s_R(a)) p(s_R(a), a) F(da)$. It is clear that this is the same problem that we analyzed in the benchmark model with the change of variables from ϕ to $EQ(s)$. However, in order to recover our main results, both our Assumptions 1 and 2 must be now imposed on $p(s, a)$, instead of $p(\phi, a)$. Under these assumptions the second-best journal rank that is obtained and the calibrated moments of the distribution of abilities remain unchanged.

Quality vs quantity We assumed in the benchmark model that each researcher, say researcher a, has a single paper and $p(\phi, a)$ is the probability of acceptance for this single paper in journal ϕ. Our model can accommodate both a generalization to more than a single paper per researcher and the choices about quality versus quantity. As the optimal choices of the researchers and of the RSB depend on the *ratio* of p, we can easily allow $p(\phi, a)$ to be greater than one for all researchers or only for some of them, and for all the journals or only some of them. Under this interpretation $p(\phi, a)$ is the expected number of publications in journal ϕ in the period considered, or the evaluation window. To illustrate this point, if the ratio of p for researcher a and two journals ϕ' and ϕ is 0.25, it implies that within the period considered, the researcher and the RSB can get four times as many publications in journal ϕ' as they can in journal ϕ. This is irrelevant for the optimal journal rank and the optimal reward scheme if we assume that $p(\phi', a) = 1/8$ and $p(\phi, a) = 1/2$ and interpret p as probabilities or if we let $p(\phi', a) = 1$ and $p(\phi, a) = 4$ and interpret p as the expected number of papers published.

Relaxing the "cheapest" journal assumption In the benchmark model, we imposed Assumption 2, which implies that each researcher deciding to publish in a given category or rank chooses the same, that is, the "cheapest", journal. This is clearly a simplifying assumption. We can relax the assumption that p is decreasing in ϕ, and allow the researcher a to choose $\phi(a) = \arg\max_\phi p(\phi, a)$ within a category or rank that they decide to publish in. The second-best solution that is obtained, and hence the calibrated distributions of abilities, can then be interpreted as the *worst-case* scenario in such case. That is, as the specific journal choices are unknown to the RSB, the regulator can maximize the objective under the assumption that all researchers will select the cheapest journal in each category or rank. Here the *cheapest journal*, say ϕ_{cheapest}, within a rank is the journal that comes lowest on the quality index within a given rank. This can be shown formally by following the inequalities: $\phi_{\text{cheapest}} p(\phi_{\text{cheapest}}, a) \leq \phi(a) p(\phi_{\text{cheapest}}, a) \leq \phi(a) p(\phi(a), a)$, where the first inequality holds because of the above definition of the "cheapest" journal and the second inequality holds because of researcher maximization. Consequently, when we relax Assumption 2, we can interpret the RSB objective that we consider in the paper as the lower bound of the expected, quality-weighted number of publications within a class, and so our calibration results can be interpreted as the worst-case or most pessimistic scenario.

Probability of acceptance in the algebraic example The algebraic example presented in section 3.2 implies that researchers in the optimal solution choose journal ϕ, for which the probability of acceptance equals 1. This feature of the model is a by-product of the specific form of the probability of acceptance function given by (7). Figure 3 makes it clear that there is a kink at ϕ for each $\phi p(\phi, a)$ so that researcher $a = \phi$ is the lowest-ability researcher for whom $p(\phi, a) = 1$. We can relax this assumption by allowing for the possibility that a paper can be rejected no matter how high the researcher's level of ability is. This can be done without departing from the tractable piece-wise linear form of $p(\phi, \cdot)$. We can replace $p(\phi, \cdot)$ given by (7) by $p^*(\phi, \cdot) = \zeta p(\phi, \cdot)$ for some $\zeta < 1$. The optimal solution does not change, and the only change is that the objective function value at the optimal solution is reduced by ζ.

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 reward for the authors. At the same time, this choice should contribute to maximizing the expected prestige 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. Our theoretical model is parsimonious and, hence, based on simplifying assumptions. Out of few 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. 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) 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 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.

Our results can be of use 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 prestige of the discipline. This could be achieved if *forward engineering* is applied rather than reverse engineering, meaning if the distribution of the ability to publish is better known and catered to. The way to do this would be to conduct a detailed analysis of publications by authors from a particular population by tracing their citations, implementations and, particularly for newly published papers that have not yet been able to accumulate citations, journal impact measures. This we are leaving for further research.

References

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*, 102197.

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

Baccini, A. and G. De Nicolao (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.

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

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

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

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

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

Haddawy, P., S.-U. Hassan, A. Asghar, and S. Amin (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., J. Wright, and Z. Ye (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., T. P. Mamuneas, and T. Stengos (2003). Rankings of academic journals and institutions in economics. *Journal of the european economic association 1*(6), 1346–1366.

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

Laffont, J.-J. and D. Martimort (2001). *The Theory of Incentives: The Principal-Agent Model.* Princeton University Press.

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

MacLeod, W. B. and M. Urquiola (2021). Why does the United States have the best research universities? Incentives, resources, and virtuous circles. *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., J. Romano, A. Shaikh, and D. Wilhelm (2022). Statistical uncertainty in the ranking of journals and universities. *AEA Papers and Proceedings 112*, 630–34.

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

Ochsner, M., E. Kulczycki, A. Gedutis, and G. Peruginelli (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. and O. Volij (2004). The measurement of intellectual influence. *Econometrica 72*(3), 963–977.

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

Petersen, A. M., R. K. Pan, F. Pammolli, and S. Fortunato (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., M. Cattaneo, M. Meoli, and P. Malighetti (2019). Self-citations as strategic response to the use of metrics for career decisions. *Research Policy 48*(2), 478–491.

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

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

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

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

Wuestman, M. L., J. Hoekman, and K. Frenken (2019). The geography of scientific citations. *Research Policy 48*(7), 1771–1780.

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

Appendices

A Proofs

Proof of Proposition 1

(i. $\Rightarrow$ ii.) Consider any reward scheme R and any researcher $a_1 < 1$ such that a_1 weakly prefers ϕ_2 over ϕ_1, i.e. $\frac{p(\phi_2,a_1)}{p(\phi_1,a_1)} \geq \frac{R(\phi_1)}{R(\phi_2)}$. Take any $a_2 > a_1$. By journal monotonicity it follows that $\frac{p(\phi_2,a_2)}{p(\phi_1,a_2)} > \frac{R(\phi_1)}{R(\phi_2)}$, so that researcher a_2 strictly prefers ϕ_2 over ϕ_1.
(ii. $\Rightarrow$ i.) Consider any ability level $a_1 < 1$ and any pair of journals $\phi_1 < \phi_2$. Let the reward scheme R be such that $\frac{p(\phi_2,a_1)}{p(\phi_1,a_1)} = \frac{R(\phi_1)}{R(\phi_2)}$. Since the researcher's preferences are given by (1) this means that a_1 weakly prefers ϕ_2 over ϕ_1. Consider another researcher a_2 such that $a_2 > a_1$. By ii). $\frac{p(\phi_2,a_2)}{p(\phi_1,a_2)} > \frac{R(\phi_1)}{R(\phi_2)}$, hence $\frac{p(\phi_2,a_2)}{p(\phi_1,a_2)} > \frac{p(\phi_2,a_1)}{p(\phi_1,a_1)}$. Since a_1, a_2 were chosen arbitrarily, journal monotonicity holds.

Proof of Proposition 2

If the reward scheme is the one given in the Proposition, then each researcher will maximize the part of the infinite sum of the RSB objective. Since this is so at any point of the strictly increasing distribution F, the total expected quality will be maximized as well. Any other reward scheme might change the researcher's decision about allocation, and whenever it does so, it will entail lower total expected quality unless the rewards are changed only for those journals that are never chosen (before or after the change) by the population of researchers given by F.

Proof of Proposition 3

Consider a non-monotonic reward scheme. Let $\phi_1 < \phi_2$ and $R(\phi_1) > R(\phi_2)$. Suppose there exists a researcher a' for whom ϕ_2 is optimal, i.e. $\phi_2 \in \Phi(a')$. This means that $\phi_2 p(\phi_2, a') \geq \phi p(\phi, a')$ for all ϕ and particularly for ϕ_1. To avoid trivialities we assume that this holds with a strict inequality: $\phi_2 p(\phi_2, a') > \phi_1 p(\phi_1, a')$. Note that $\frac{\phi_1}{\phi_2} < \frac{R(\phi_1)}{R(\phi_2)}$, so that we have two cases to consider. Either $\frac{p(\phi_2,a')}{p(\phi_1,a')} \geq \frac{R(\phi_1)}{R(\phi_2)} > \frac{\phi_1}{\phi_2}$ or $\frac{R(\phi_1)}{R(\phi_2)} > \frac{p(\phi_2,a')}{p(\phi_1,a')} > \frac{\phi_1}{\phi_2}$. In the former case, ϕ_2 will still be chosen by a' and hence, other things being equal, the total expected quality will not change. In the latter case, however, ϕ_1 will be chosen by researcher a' and, as compared to the optimal case, the total expected quality will decrease by $p(\phi_2, a')\phi_2 - p(\phi_1, a')\phi_1$. Since the distribution F is strictly increasing, this decrease is non-zero. If ϕ_2 is never chosen by any researcher in the first-best solution, setting $R(\phi_2) < R(\phi_1)$ does not matter for the total expected quality provided that the rewards for the remaining journals are set optimally. So a non-monotonic R is weakly dominated by a monotonic one.

Proof of Proposition 4

We argue by contradiction. Suppose the reward scheme is such that there is i, j with $\frac{\alpha_i}{\alpha_j} \neq \frac{\phi_i}{\phi_j}$. We have three groups of researchers to consider:

i. researchers with a such that $\frac{p(\phi_j,a)}{p(\phi_i,a)} > \max\left(\frac{\alpha_i}{\alpha_j}, \frac{\phi_i}{\phi_j}\right)$. These researchers choose ϕ_j over ϕ_i under the first-best scheme and under the reward scheme considered, so there is no change of allocation decisions here;

ii. researchers with a such that $\frac{p(\phi_j,a)}{p(\phi_i,a)} < \min\left(\frac{\alpha_i}{\alpha_j}, \frac{\phi_i}{\phi_j}\right)$. These researchers choose ϕ_i over ϕ_j under the first-best scheme and under the reward scheme considered, so there is no change of allocation decisions here;

iii. researchers a such that $\frac{p(\phi_j,a)}{p(\phi_i,a)} \in \left(\min\left(\frac{\alpha_i}{\alpha_j}, \frac{\phi_i}{\phi_j}\right), \max\left(\frac{\alpha_i}{\alpha_j}, \frac{\phi_i}{\phi_j}\right)\right)$. These researchers choose differently under the first-best scheme and the reward scheme considered, which entails a loss of efficiency. Note that if $\frac{\alpha_i}{\alpha_j} = \frac{\phi_i}{\phi_j}$ was true, this third case would be impossible.

This proves that setting $\frac{\alpha_i}{\alpha_j} = \frac{\phi_i}{\phi_j}$ is never worse than any other reward scheme.

A.1 Journal monotonicity condition

Suppose $p(\phi, a)$ is given by formula (7). Suppose $\phi' > \phi$ and $\xi > 1$. We show that the ratio $\frac{p(\phi',a)}{p(\phi,a)}$ is increasing in a on $[\frac{\xi-1}{\xi}\phi', \phi]$. To see that observe that on this range:

$$\frac{p(\phi', a)}{p(\phi, a)} = \frac{1 + \xi\frac{a-\phi'}{\phi'}}{1 + \xi\frac{a-\phi}{\phi}} = 1 + \frac{(\frac{\xi}{\phi'} - \frac{\xi}{\phi})a}{1 - \xi + \frac{\xi}{\phi}a}.$$

Denoting $\beta' := \frac{\xi}{\phi'}$ and similarly $\beta := \frac{\xi}{\phi}$ and differentiating with respect to a we obtain that $\frac{p(\phi',a)}{p(\phi,a)}$ is increasing in a if and only if

$$(1 - \xi + \beta a) - \beta(\beta' - \beta)a' = (\beta' - \beta)(1 - \xi) > 0.$$

Since $\beta' < \beta$ the above is satisfied whenever $\xi > 1$.

A.2 Derivation of FOCs and SOCs for the analytical example

The objective function is given by:

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

We now calculate the FOCs using the Leibniz integral rule and substituting for the crossing points (9):

$$\begin{aligned}
\frac{\partial \mathrm{ETQ}_{\mathrm{II}}(\xi)}{\partial \phi_i} &= \phi_i - \phi_{i-1}\tfrac{\xi-1}{\xi} + \int_{a_{i-1/i}}^{\phi_i} (1-\xi)da - \phi_i\tfrac{1}{\xi} + \phi_{i-1}\tfrac{\xi-1}{\xi} + \phi_i\tfrac{1}{\xi} - \phi_i + \int_{\phi_i}^{a_{i/i+1}} da \\
&= (1-\xi)\left[\phi_i - \tfrac{\phi_{i-1}+(\xi-1)\phi_i}{\xi}\right] + \left[\tfrac{\phi_i+(\xi-1)\phi_{i+1}}{\xi} - \phi_i\right]. \qquad (13) \\
\frac{\partial \mathrm{ETQ}_{\mathrm{II}}(\xi)}{\partial \phi_n} &= \phi_n - \phi_{n-1}\tfrac{\xi-1}{\xi} + \int_{a_{n-1/n}}^{\phi_n} (1-\xi)da + \phi_{n-1}\tfrac{\xi-1}{\xi} - \phi_n + \int_{\phi_n}^{1} da \\
&= (1-\xi)\left[\phi_n - \tfrac{\phi_{n-1}+(\xi-1)\phi_n}{\xi}\right] + (1-\phi_n).
\end{aligned}$$

Setting both to zero and rearranging yields:

$$\begin{aligned}
\phi_i &= \frac{\phi_{i-1}+\phi_{i+1}}{2}, \quad i \in \{1, ..., n-1\} \\
1-\phi_n &= (\phi_n - \phi_{n-1})\frac{\xi-1}{\xi}.
\end{aligned}$$

B Data and robustness analysis

B.1 Data

In the empirical investigation, we applied data from four rating lists:

1. Rating for economics and management journals published by Centre National de la Recherche Scientifique, France, abbreviated as CNRS; https://www.gate.cnrs.fr/IMG/pdf/categorisation37_liste_juin_2020-2.pdf.

2. Rating for economic and business journals Academic Journal Guide, published by the Chartered Association of Business Schools in the UK, abbreviated as AJG; see https://charteredabs.org/academic-journal-guide-2021/.

3. Official ratings published by the Ministry of Education and Science in Poland for two disciplines: economics & finance, and management, combined together, abbreviated as PL; https://czasopisma.webclass.co/.

4. For the A and B economic journals list from the US we used a list of 71 journals grouped into four classes of A+, A, A- and B+:

   - Class A+ contains top 5 general-interest journals,
   - Class A contains 17 top major-field journals: J Econ Theory, J Econometrics, J Labor Econ, J Monetary Econ, Game Econ Behav, AEJ Macro, Rand J Econ, Econ Theor, J Public Econ, Theor Econ., Rev Financ Stud, AEJ Micro, J Int Econ, J Financ Econ, Quant Econ, AEJ Applied Econ, AEJ Policy and 4 general-interest journals: Int Econ Rev, The Econ J., J. Euro Econ Assoc, Rev Econ Stat,
   - Class A- contains 4 general-interest/survey journals: Euro Econ Rev, J. Econ Lit., J. Econ Persp., Brookings Pap Econ Activity and 7 major-field journals: J Bus Econ Stat, J Applied Econometrics, J Hum Resource., Rev Econ Dyn,, J. Econ Growth, J. Money Credit Banking, J. Econ Dyn Control,

- Class B+ contains 5 general-interest journals: J. Econ Behav Org, Scand J. Econ., Canadian J Econ., Econ Inquiry., Oxford Econ Papers and 29 field journals: J Env Econ Mgmt, Experim Econ, J Risk Uncertainty, J. Urban Econ, Int J Game Theory, J Econ Geogr, J. Econ Mgmt Strat., J. Ind Econ., Int J. Ind Org, J. Math Econ., J. Dev Econ, Ecolog Econ, Soc. Choice Welfare, J. Public Econ Theory, J. Health Econ., Amer J. Ag Econ, J. Econ History, J. Law Econ Organ, J Law Econ., Public Choice, J Bank Fin, Land Econ, Oxford Bull Econ Stat., Econometric Theory, Econometrics J., Expl. Econ Hist, Quant. Marketing Econ., Reg. Sci Urban Econ., Health Econ.

All these ratings give titles and bibliographical characteristics (ISBNs and e-ISBNs) and ranks or classes for each journal. We matched the journals with the corresponding JIMs obtained from external sources. For SNIP, we used data from `https://www.elsevier.com/authors/tools-and-resources/measuring-a-journals-impact/`. For RIF, we used data from Konig et al. (2022). Table B1 gives the main characteristics of the rating lists and their matchings with JIMs. The total number of journals might be slightly smaller than that reported in the original sources, as there were a few cases where we could not clearly identify journals due to duplications, improper issns or e-issns, or similar.

B.2 Robustness analysis

We analyze the robustness of results presented in Figure 6. We only consider two schemes, AJG Business and US econ, for better visibility. Figure B1 presents how the induced distribution changes with the slope parameter ξ (panel a), the cutting percentile k (panel b), and the RIF confidence intervals reported in Konig et al. (2022) (panel c). To obtain intervals for the induced distribution in the case of the RIF confidence intervals, we repeated the same analysis that we did for ϕ, defined as the mean RIF value, with ϕ defined first as the minimum and then as the maximum RIF value. The different parameter values shift the distribution, but they preserve the order of the AJG and US distribution quantiles. This means that our comparative results for different reward schemes and populations are remarkably robust to the choice of parameter values.

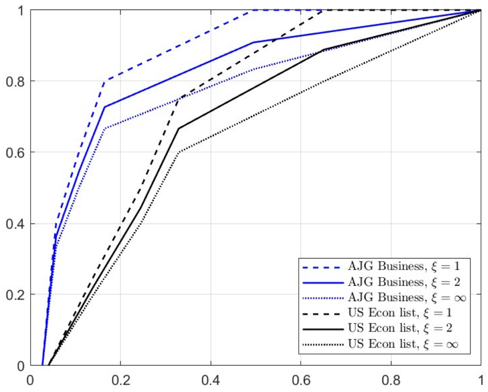
We also repeated the analysis of the four journal rating schemes for economics and management for ϕ defined as the SNIP value instead of the RIF. More precisely, the inputs are that for each of the four selected reward schemes $\beta \in \{\text{CNRS, AJG, ABDC, PL}\}$, we define the set of journals J^{β}, as the journals that are assigned a class by the scheme β. For each $j \in J^{\beta}$, we set $\phi(j)$ equal to the SNIP value of j, whenever it exists, normalized by the highest SNIP value in the dataset, which is that of Quarterly Journal of Economics, and zero otherwise. We set the parameter values of the slope as $\xi = 2$ and cut percentile as $k = 25$. Figure B2 presents the induced CDFs for the four reward schemes.

The results are similar to those in Figure 6, though the number of journals taken into account is significantly different for some reward schemes. The PL scheme for example covers 1967 journals, all of which were members of J^{PL} in the SNIP analysis, and 1308 of which had a non-zero SNIP assigned. In contrast, there are only 319 journals in the J set in the RIF analysis (see Table B1 for complete statistics). This implicitly means that many journals in the SNIP analysis are average-quality journals, whereas those in the RIF dataset are of better quality. Consequently, the RIF analysis focuses on the population of researchers publishing in journals that are included in the RIF database, whereas the SNIP analysis focuses on the whole population of researchers, including those publishing in some lesser-known local journals. Differences are

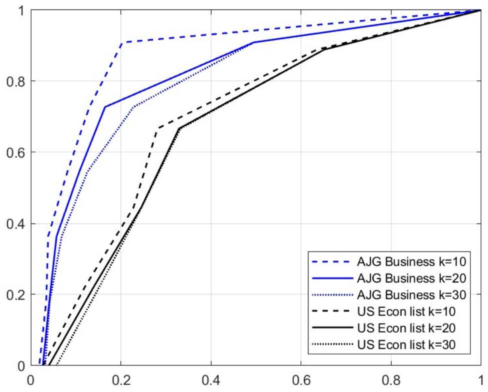
Table B1: Journals ratings and JIMs

Scheme	Class	No. Jrn	No of jrn with JIM>0	
			RIF	SNIP
CNRS	1	107	69	104
	2	193	85	158
	3	303	68	228
	4	231	23	149
	Total	**834**	**245**	**639**
AJG	4*	43	9	40
	4	93	27	83
	3	262	87	215
	2	301	74	252
	1	106	9	85
	Total	**805**	**206**	**675**
PL	200	146	36	131
	140	359	63	306
	100	587	73	465
	70	924	84	731
	40	904	50	647
	20	1585	13	850
	Total	**4505**	**319**	**3130**
US	A+	5	5	5
	A	21	21	21
	A-	11	11	11
	B+	34	34	34
	Total	**71**	**71**	**71**

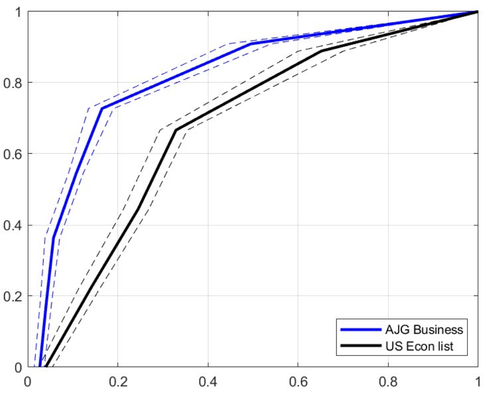
Figure B1: Robustness analysis. The values on the horizontal axis are given by $\sqrt{a}$.



(a) Different ξ



(b) Different k



(c) RIF intervals

of course also caused by different definitions of the journal quality index. Particularly so in the subset of journals assigned non-zero values by both RIF and SNIP, as the Spearman rank correlation between them is 0.72, and in the subset of journals assigned a RIF measure it is 0.56.

Figure B2: Induced distributions computed for four economics and management journal ratings using SNIP. The values on the horizontal axis are values of a