

The Economics of Research Reproducibility*

Jean-Edouard Colliard[†] Christophe Hurlin[‡] Christophe Pérignon[§]

July 15, 2021

Abstract: Empirical evidence documents a relatively low level of research reproducibility in economics. In this paper, we investigate why this is the case and what can be done to move out of this low-reproducibility equilibrium. We study the supply and demand for research reproducibility, provide empirical evidence on authors' preferences for reproducibility, and estimate the cost of verifying reproducibility. We theoretically show that competition between journals to attract authors leads to a suboptimally low level of reproducibility. Leading journals with sufficient market power can set higher reproducibility standards, which is consistent with recent changes in data availability policies.

Keywords: reproducibility, peer-review process, data availability policy, confidential data

JEL Codes: C80, C81, C88.

*This version supersedes our previous work circulated under the title "Reproducibility Certification in Economics Research". We thank seminar participants at HEC Paris and participants at the 2021 HEC Liege Research Day for their comments and suggestions. We are grateful to Sebastien Saurin for valuable research assistance. Disclosure: Hurlin and Pérignon are the founders of *casca*d (Certification Agency for Scientific Code And Data, www.cascad.tech), a non-profit academic organisation promoting and verifying research reproducibility.

[†]HEC Paris, 1 rue de la Libération, 78351 Jouy-en-Josas, France. Email: colliard@hec.fr.

[‡]University of Orléans, Rue de Blois, 45067 Orléans, France. Email: christophe.hurlin@univ-orleans.fr.

[§]HEC Paris, 1 rue de la Libération, 78351 Jouy-en-Josas, France. Email: perignon@hec.fr.

Introduction

Like many scientific disciplines, economics has experienced a data and computational revolution. Today, most academic papers are empirical and rely on complex scripts analysing rich and sizable datasets. As computational results now account for a large part of the output and contribution of an academic paper, it is more important than ever that empirical results are *reproducible*, that is, one can check whether “same data + same code = same results” (Buckheit and Donoho (1995); Barba (2018)).¹ However, in contrast to the swift evolution of research practices and the growing importance of reproducibility in science, until recently economics journals have evolved slowly on this front.

In this paper, we aim to rationalize, and potentially improve, this situation by conducting a comprehensive study of the economics of research reproducibility. Our study first encapsulates the existing literature in a simple theoretical framework that clarifies the economic determinants of the current level of reproducibility in economic journals. We then fill important gaps in the literature. We provide evidence on authors’ preferences for reproducibility and estimates of the costs of reproducing empirical results based on our experience at *casca*d, a reproducibility verification agency.² We also show that in our theoretical framework competition between journals leads to a suboptimally low level of reproducibility, thus calling for remedial actions.

To understand the economics of research reproducibility, it is useful to consider that the path towards publishing reproducible follows three key stages. Stage 1 is the default situation of no reproducibility policy. Stage 2, or unverified reproducibility, is reached when the journal introduces its first Data Availability Policy (DAP). Stage 3, or verified reproducibility, requires conducting a

¹In contrast to reproducibility, replicability refers to the ability of a researcher to duplicate the results by implementing the same methodology in another context or time period (“same code + new data = same results”) or a different methodology to the same data (“new code + same data = same results”) (Peng *et al.* (2006)).

²*casca*d stands for Certification Agency for Scientific Code And Data and its website is www.casca.d.tech. It is a non-profit scientific organization promoting and verifying research reproducibility in economics and business. It is funded by the French National Center for Scientific Research (CNRS) along with several research institutions and universities.

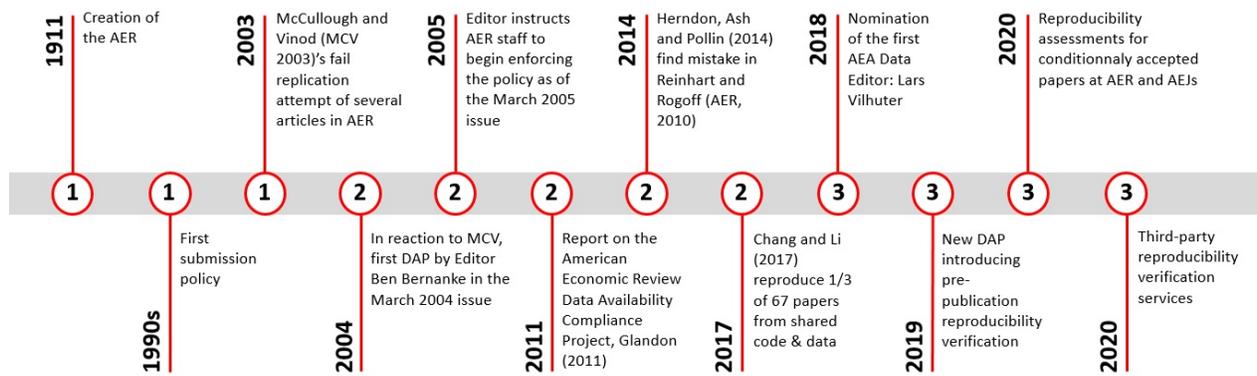


Figure 1: Making Research Reproducible: The *American Economic Review* Timeline

systematic verification of the reproducibility of all accepted papers prior to publication. Figure 1 illustrates how the *American Economic Review* gradually moved from Stage 1 to Stage 3 in the past 20 years.

Interestingly, in economics, the reproducibility seed was planted very early on. In 1933, Ragnar Frisch, founding editor of *Econometrica* and first Nobel laureate in economics, wrote that “*in Econometrica the original raw data will, as a rule, be published, unless their volume is excessive. This is important in order to stimulate criticism, control, and further studies*” (Frisch (1933)).³ This rule was gradually abandoned until pioneer journals such as the *Journal of Money, Credit, and Banking* introduced their DAP and hence entered Stage 2 as early as in the late 1990’s (see Figure 1 in McCullough (2009a)). The *American Economic Review* was the first “top 5” economics journal to introduce a DAP back in 2004 (Bernanke, 2004), whereas the *Quarterly Journal of Economics* was a late adopter in 2016. Höfler (2017) reports that 54% of a sample of 343 economics journals included in the Thomson Reuters Social Science Citation Index have a DAP.⁴

Despite the generalization of DAP, the capacity for other researchers to reproduce published

³We are grateful to Duvendack *et al.* (2015) for resurrecting this inspiring quote.

⁴These policies vary extensively across journals: from vague comments in some journals to strict requirements in others (e.g., *Journal of Applied Econometrics*). Höfler (2017) shows that only 7.5% of economics journals have a DAP that explicitly requests data from authors, 32% simply recommend that data be disclosed on a repository, and 14% suggest their authors share their data if requested by a reader. Stodden *et al.* (2018) find that 44% of researchers publishing in *Science* provide their code when politely asked for.

results has remained surprisingly low in economics (McCullough *et al.* (2008); Chang and Li (2017); Gertler *et al.* (2018)). Economic journals' DAP have been quite vague and only partly enforced in practice (Duvendack *et al.* (2017)). The result is that for many papers numerical resources are unavailable, improperly documented, or of insufficient quality, so that well-trained economists are often unable to reproduce the results of papers published in leading economics journals, even when code and data are publicly available.

The solution to this problem is to move to Stage 3, verified reproducibility. The verification process takes place before the final acceptance of a manuscript and it follows two steps. The first step consists in checking whether the authors have complied with the guidelines aiming to ease the duplication of the findings. These guidelines cover the presentation and structure of code and data and aim to make these resources findable, interpretable, and reusable.⁵ The second step aims at ensuring that the numerical results (tables and figures) included in the scientific article correspond to the numerical results generated by the computer code and data of the authors. The two-step verification is conducted by a reproducibility referee under the supervision of a data editor or reproducibility editor. The tasks of the reproducibility referee are to check that both the code and data comply with the guidelines, to execute the code, to compare the output with the results presented in the tables and figures of the article, and to list any potential discrepancy.

To date, only a handful of economics journals have moved to Stage 3 reproducibility, a significant exception being the AEA Journals. We argue in this paper that this situation is best understood as an equilibrium phenomenon. We propose a conceptual framework for the demand and supply of research reproducibility. Following the existing literature (e.g., Jeon and Rochet (2010)), we model journals as platforms intermediating between authors (the suppliers of reproducible research) and readers (the consumers). An explanation often given for the low level of reproducibility is the high

⁵Such guidelines and best practices can come from scientific associations (AEA data standards in Vilhuber (2019)), journals' data editors (Social Science Data Editors (2020)), or researchers (Koenker and Zeileis (2009), Orozco *et al.* (2020)).

cost reproducibility imposes on authors. However, to our knowledge the only evidence on these costs is based on surveys. We analyze the propensity of authors of papers accepted in the *Journal of Financial Economics* to shed light on these costs. In particular, we do not find any evidence that authors are more reluctant to share their code when they are more senior, more cited, or affiliated with more prestigious universities. This evidence does not support the received idea that journals should not increase their reproducibility standards as this would discourage the best authors from submitting their work (see, e.g., [Harvey \(2014\)](#)).

We then use our theoretical framework to study the competition between two academic journals that choose submission fees, subscription fees, and a reproducibility level. We show that in equilibrium this competition leads to a suboptimally low level of reproducibility. The reason is that journals are competing fiercely to attract authors who form a “competitive bottleneck” ([Armstrong, 2006](#)), and lowering reproducibility is a way to lure authors from competing journals. Hence, we argue that the low level of reproducibility observed in most economic journals cannot be assumed to optimally balance the supply of and demand for reproducibility.

Finally, we discuss different paths to be explored to move out of a low reproducibility equilibrium: imposition of new “reproducibility standards” by journals with sufficient market power, appointment of data editors, recourse to third party verification services, and lowering the reproducibility costs for journals. On the last point, we use our experience at *cascaD* to provide concrete estimates of the costs of verifying the reproducibility of empirical research in economics. We show that these costs exhibit very significant economies of scale, due in particular to the large costs for the verification team to accessing new commercial or administrative datasets. In particular, we estimate that significantly exploiting these economies of scale would lead to an average cost around USD 350 per paper. We conclude from this exercise that reaching Stage 3 reproducibility is achievable in economics but will require optimizing the verification process.

1 Benefits and Costs of Reproducibility

In this section, we present a simple conceptual framework to think about the demand and the supply of reproducibility in academic research. We then discuss the costs and also the potential benefits of reproducibility for authors and for readers.

1.1 Setup

In order to apprehend the economics of reproducibility, we rely on the literature conceptualizing academic journals as platforms intermediating between authors and readers, in particular [Jeon and Rochet \(2010\)](#). For the moment, we consider the case of a single journal. A journal publishing n_A articles chooses a level of reproducibility $q \in \mathbb{R}^+$. The three stages of reproducibility defined above can be thought of as three discrete values of q , but we allow for a continuum between those stages. The journal charges a fee p^A to each author upon submission and a fee p^R to each reader for subscribing to the journal. As in [Jeon and Rochet \(2010\)](#), we can have $p^A \leq 0$ (authors being paid to publish) but not $p^R < 0$.⁶

There is a continuum $[0, 1]$ of articles indexed by i , each article is characterized by a vector of characteristics X_i (e.g., topic, quality, authors, etc.). We abstract away from the issue of selecting papers through the work of referees and editors, and assume that all articles in $[0, 1]$ are good enough to pass the journal's screening process.⁷

There is a continuum $[0, +\infty)$ of potential readers, indexed by j . Each reader is characterized by some characteristics Y_j (e.g., different taste for different types of articles). Each reader chooses whether to subscribe to the journal.

⁶A journal may want to charge negative subscription fees to attract more readers and hence citations. However, the journal cannot control that a reader who gets the subsidy indeed reads the journal, so that negative subscription fees would lead to having many “fake” readers.

⁷The growing theoretical literature on academic journals has focused on screening, see [McCabe and Snyder \(2005\)](#), [Jeon and Rochet \(2010\)](#), [Wang \(2018\)](#), and [Gehrig and Stenbacka \(2021\)](#). While screening and reproducibility can both be seen as a quality variable, a crucial difference is that a high level of reproducibility imposes a cost on authors and not only on the journal.

The author of a published article receives a payoff that depends on the readers of the journal and their characteristics, the reproducibility level q , the article's characteristics X_i , and the submission fee paid p^A . Namely, if $\mathcal{S} \subset \mathbb{R}^+$ is the set of readers who subscribe to the journal, we assume the payoff of article i 's authors to be:

$$\int_{j \in \mathcal{S}} u^A(q, Y_j, X_i) dj - C^A(q, X_i) - p^A. \quad (1)$$

For the moment, we simply assume that u^A is always positive: all else equal, authors prefer to be published in journals with a larger number of readers. We discuss in Section 1.3 below how u^A could also depend on q and X_i , and how this affects the economics of reproducibility. C^A is the cost of reaching reproducibility level q for the authors of article i , independently of who reads the journal.

A reader who subscribes to a journal receives a payoff that depends on the number and characteristics of the published papers, the reproducibility level q , and the subscription fees p^R . Namely, denoting $\mathcal{P} \subset [0, 1]$ the set of published articles, we assume this payoff to be:

$$\int_{i \in \mathcal{P}} u^R(q, X_i, Y_j) di - p^R. \quad (2)$$

We assume that u^R is always positive and increasing in q : all else equal, readers prefer journals with more articles and a higher level of reproducibility.

The journal serves as a platform for authors and readers, and attracts an endogenous number n^A of articles and n^R of readers. Reaching reproducibility level q has a cost $C^J(q, n_A)$ for the journal, with $C_1^{J'} \geq 0$. We will assume throughout the paper that the journal aims at maximizing citations,⁸ and that citations are proportional to the number of readers n^R . The journal faces a break-even

⁸See Card and DellaVigna (2020) for evidence on editors' objectives.

constraint:

$$n^A p^A + n^R p^R - C^J(q, n_A) \geq 0. \quad (3)$$

In this framework, the readers express a demand for reproducibility, which is costly to journals and to authors. The prices p_A and p_R can be used to compensate the authors and the journal for these costs. The equilibrium level of reproducibility will then depend on the balance between supply and demand, and hence on the utility functions u^A and u^R .

1.2 Reproducibility for Readers

To better understand the demand for reproducibility, we discuss why the readers of academic research care about the reproducibility of results. Below, we show that the three main answers were already articulated in Ragnar Frisch's 1933 editorial of *Econometrica*: to “*stimulate criticism, control, and further studies*”.

First, as reproducible research calls for the disclosure of code, it stimulates the criticism of existing results by simplifying the conduct of replication studies running the original methodology on another dataset. Considering a sample of articles published in top-50 economics journals, [Mueller-Langer et al. \(2019\)](#) show that a mandatory data-disclosure policy has a positive effect on the replication probability by six percentage points. They conclude that replication efforts could be incentivized by promoting data disclosure and hence reducing the cost of replication.

Second, reproducibility serves as a control that the results reported in an article come from the methodology described. In their survey of transparency and reproducibility in economics research, [Christensen and Miguel \(2018\)](#) claim that this basic standard should be expected of all published economics research, as it is the first step toward a more thorough assessment of the validity of a scientific claim. In particular, reproducibility allows to conduct an in-depth analysis of the code

and data, which allows one to spot coding errors (see Herndon *et al.* (2013) on Reinhart and Rogoff (2010)), cases of specification searching and p-hacking (Harvey (Forthcoming); Christensen *et al.* (2019); Brodeur *et al.* (2020)), or possible cases of data fraud (Simonsohn (2013)). Moreover, given the high reputation cost for a researcher whose publication is found to be erroneous, requiring data and code to be publicly available should encourage researchers to exert more effort in detecting such errors in the first place.

Third, reproducibility can stimulate further studies by communicating more information to the academic community on how exactly to conduct a given analysis. This generates economies of scale as different authors working on the same data do not have to repeat time-consuming procedures, for instance those necessary to clean up the data.⁹

1.3 Reproducibility for Authors

1.3.1 Theory

While unambiguously desirable from the perspective of the readers, reproducibility imposes costs on authors. In a survey, Stodden (2010) reports that almost half of the respondents state that the lack of incentives and direct benefits is an important reason for researchers not to make their computer code publicly available. We briefly survey different costs, but also benefits, for authors, and how they relate to our framework.

The first and perhaps main cost is the opportunity cost of the time spent cleaning data, documenting code, and providing technical support to other researchers using the shared material (Harvey, Forthcoming). The pressure to publish is indeed ranked first among all the reasons put forward by scientists when surveyed by *Nature* about the impediments to reproducible research (Baker (2016)). These costs are independent of the journal's readership and are captured by $C^A(q, X_i)$.

⁹Examples of popular datasets and code include the Fama and French portfolios used in empirical finance (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html) and the GARCH toolbox provided by Kevin Sheppard (<https://www.kevinshppard.com/code/matlab/ucsd-garch/>).

A second cost is reputational risk. Making one's code and data available makes it easier for others to spot errors, which is socially beneficial but privately costly. This cost could depend on the journal's readership and enter $u^A(q, Y_j, X_i)$.

A third cost is that sharing one's code and data will make the authors face more competition when writing follow-up papers. These costs depend again on an article's characteristics X_i . Some types of research may be particularly discouraged by a tough reproducibility policy, in particular research using proprietary datasets (Harvey, 2014). The costs could also be larger for more productive authors and/or more innovative papers.¹⁰

Reproducibility may also have benefits for researchers. A first benefit is that journals with a stricter reproducibility policy may attract more readers, and hence publishing in such a journal may attract more citations: in our framework, journals with a higher q may endogenously have a higher n^R and this benefits authors.¹¹

A second benefit is that a high level of reproducibility may signal the high quality of a paper and its authors to the relevant audience - in the present case, peers, universities, and research funding agencies. Lerner and Tirole (2002) argue that this type of signaling motive is an important driver of the open-source software development community. Theoretically, $C^A(q, X_i)$ could thus decrease in q , at least for certain paper characteristics X_i .

1.3.2 Empirics

We are not aware of empirical evidence other than surveys on the benefits for authors of making their research reproducible. To fill this gap, we analyze authors' decisions to voluntarily make their paper reproducible, and how this decision correlates with various authors' characteristics. Our

¹⁰The *Journal of Financial Economics* recently implemented a new policy to mitigate this concern: authors have to disclose code and data but can optionally choose to keep both hidden for up to two years after publication (Whited, 2021). See also Hill and Stein (2020) for empirical evidence that the race to publish first can lead to lower quality research.

¹¹The empirical evidence for a positive causal impact of open data on citations is at best weak, see McCabe and Mueller-Langer (2019).

empirical analysis focuses on all papers published in the *Journal of Financial Economics* (JFE) between January 2010 and September 2020, which corresponds to all issues between the first one of volume 95 and the third one of volume 137.¹² We end up with a total number of 1,347 papers written by 2,231 authors.

We focus on the JFE for two reasons. First, over the past decade, this journal has encouraged authors to share code and data associated with their papers but never made it mandatory. Hence, by studying the decision of whether to share numerical resources, we can identify researchers' characteristics that act as an impetus or impediment to research reproducibility. Second, by focusing on a single journal, we neutralize the effect of the reputation of a journal on an author's decision to make a given paper reproducible. That is, different decisions can only come from the articles' characteristics X_i .

We collected information about available code and data from the JFE data and program webpage.¹³ Over our sample period, 67 published papers or 4.97% of all published papers are open, i.e., having code or data, or both, available for download from the JFE website.¹⁴ This very low percentage is surprising given the broad definition we use to identify open papers but it is in line with previous evidence showing that sharing code and data is not widespread in economics (see Section 2.1). In order to mitigate any imbalanced dataset problem that may arise when studying rare events, we also consider a subsample only including papers published in monthly issues including at least one paper with code and/or data available. In this subsample, there are 544 papers and the fraction of open papers is 12.13%.

¹²With a 5.731 impact factor, JFE is ranked in the top-3 in finance and in the top-10 in economics (2019 SCImago Journal Rank). [Schwert \(2021\)](#) reports that 88% of the papers published during our sample period were empirical. Moreover, even theoretical papers commonly use computer code to solve numerical problems or produce theoretical results.

¹³<http://jfe.rochester.edu/data.htm>

¹⁴Information about code and data were retrieved on October 7, 2020. We excluded ten papers with sharable material but published prior to 2010 and did not include ten forthcoming papers with available sharable material but scheduled to be published after the end of our sample period. As we did not check the personal website of all 2,231 authors or other data repositories for downloadable material associated with JFE papers, the reported 4.97% frequency should be seen as a lower bound.

We estimate the following logistic regression model:

$$\ln\left(\frac{\Pr(Open_{it})}{1 - \Pr(Open_{it})}\right) = \alpha + \beta_1 \cdot Top10_{it} + \beta_2 \cdot International_{it} + \beta_3 \cdot Seniority_{it} + \beta_4 \cdot Citations_{it} + \varepsilon_{it} \quad (4)$$

We define the different variables in Table 1. Panel A gives summary statistics for the four explanatory variables, and Panel B reports the results of different regression specifications.

Table 1: The Drivers of Code and Data Sharing: Evidence from the *Journal of Financial Economics*.

Panel A displays the summary statistics of the four explanatory variables. Panel B displays the estimated coefficients and the associated p-values of the logistic regression (4). The explained variable *Open* is a binary variable taking the value of one for any paper with code and/or data available, and zero otherwise. *Top10* takes the value of one if at least one coauthor is affiliated with a top-10 university (see Heckman and Moktan (2020)) and *International* equals one if at least one coauthor is affiliated with a university outside North-America. *Seniority* is equal to the time between the year of a given JFE publication and the year when the author got his or her first citation in the *Web of Science* database. *Citations* is the annual number of citations of a given author in *Web of Science*, measured in the year of a given JFE publication. For papers with more than one author, we use the average *Seniority* and *Citations* across all coauthors. The two continuous variables have been manually constructed for all the authors of the 544 papers included in the subsample.

Panel A: Summary Statistics

	Mean	Std-dev	Min	Q1	Median	Q3	Max	Obs.
International	0.318	0.466	0	0	0	1	1	1,347
Top 10	0.232	0.422	0	0	0	0	1	1,347
Citations	98.70	226.67	0	17.00	41.42	112.33	3,756.50	544
Seniority	9.14	6.62	-4	4.67	8.50	12.67	45.50	544

Panel B: Logistic Regression

	(1)	(2)	(3)	(4)	(5)	(6)
International	0.5089 (0.0658)	0.5306 (0.0579)	0.4470 (0.0857)	0.4609 (0.0817)		
Top 10	0.0197 (0.9527)	0.0015 (0.9965)	-0.0386 (0.9023)	-0.0237 (0.9406)		
Log Citations	0.1306 (0.3334)	0.1210 (0.3769)			0.0793 (0.3958)	0.0797 (0.3951)
Seniority	-0.0098 (0.7354)	-0.0061 (0.8348)				
Year FE	No	Yes	No	Yes	No	Yes
Obs.	544	544	1,347	1,347	544	544

We draw several insights from our analysis. First, as 4.97% of papers are open, it seems that for some authors at least the benefits of sharing code or data outweigh the costs. Second, the regression results show that the cost-benefit analysis does not seem to depend on authors' seniority, number of citations, or affiliation with a top 10 institution.¹⁵ This suggests that opportunity costs and reputational costs may not be first order, as we would expect them to be larger for “better” researchers.¹⁶ Third, in specifications (1)-(4), we also find that authors affiliated with universities located outside of North-America are significantly more likely to share code or data. This finding is consistent with the signaling benefit: new entrants (i.e., international researchers) send an additional costly signal to a predominantly North-American research community.¹⁷

According to Harvey (2014), finance journals were reluctant to adopt mandatory data sharing policies by fear of losing submissions by the most senior and/or productive authors. Our evidence does not support this view.¹⁸ In columns (5) and (6), we check directly whether the best-cited authors are more reluctant to share their data or code, and again find no effect. Finance journals may thus have less to lose from mandating a higher level of reproducibility than previously thought.

2 The Unverified Reproducibility Equilibrium

Having shed some light on the demand for reproducibility by authors and the supply of reproducibility by authors, we turn to the role of journals in matching the two. We summarize the existing evidence about the level of reproducibility in economic journals. We show that, until the recent

¹⁵The results on seniority and citations are in line with the survey evidence reported in Swanson *et al.* (2020).

¹⁶Instead of the mean across coauthors, we also used the median, minimum, and maximum values. We also contrasted researchers with tenure (i.e., *Seniority* > 6 years) and without tenure. The *Citation* variable was used alternatively with and without log transformation. We estimated the regression for North-American researchers only. In all cases, results remained qualitatively unchanged.

¹⁷Schwert (2021) indicates that over the past decade the percentage of US authors (respectively referees) at JFE was greater than 65% (85%). He also shows in the context of a logit model that US authors have, all else equal, a higher acceptance rate than their international peers.

¹⁸A limitation of our analysis is that we only observe published papers. It could be for instance that a paper is accepted because the data and code are available and despite its characteristics being associated with fewer citations. While we think this effect is unlikely to be first-order, we cannot rule it out based on our data.

introduction of systematic pre-publication reproduction of the results by some journals (Vilhuber, 2019), the level of reproducibility has generally been low in economics. We then show in the context of our conceptual framework that competition between journals can be expected to lead to a suboptimally low level of reproducibility.

2.1 Empirical Evidence

When they exist, data policies have been only partially enforced. The study of McCullough *et al.* (2008) on several economics journals with compulsory data-sharing policies reveals that the fraction of the papers that (1) are concerned by the policy (i.e., papers not using confidential data) and (2) actually have a companion data file available in the journal archive goes from 12% for the *Economic Journal* to close to 100% for the *Journal of Applied Econometrics*. These rates also vary over time. For instance, the rate for the *Federal Reserve Bank of St. Louis Review* drops from 100% when the policy was first introduced in 1993 to 50% ten years later.¹⁹

Is the unverified-reproducibility policy sufficient to guarantee reproducibility, or do journals need to move to stage 3, i.e., verified reproducibility? This can be tested by checking which fraction of articles in stage-2 journals can actually be reproduced from the numerical resources available in the journal's archives. McCullough *et al.* (2006) analyzed the data archive of the *Journal of Money, Credit and Banking* between 1996 and 2003.²⁰ Out of 266 published papers, 193 were empirical by nature and, according to the data policy of the journal, all should have their code and data publicly available. In practice, only 69 of them, or 37%, had their resources available for download. Only 14 papers, or 7.25%, shared material of sufficient quality for McCullough *et al.* (2006) to successfully reproduce the results.²¹

¹⁹See also Vlaeminck and Herrmann (2015) and Duvendack *et al.* (2017) for more recent evidence.

²⁰See Dewald *et al.* (1986) for an earlier attempt to reproduce empirical articles published in the *Journal of Money, Credit, and Banking* in which the authors were directly contacted by the replication team to obtain code and data.

²¹See Gertler *et al.* (2018) and Mueller-Langer *et al.* (2019) for more recent estimates of the compliance rate with DAPs, across a broader set of journals.

Is the situation any better for the top-5 economics journals? In 2008, the *American Economic Review* launched an audit to assess the quality of the data and code contained in its online data archive. A replication team selected a sample of 39 empirical articles, out of the 135 published articles subject to the data policy between 2006 and 2008. The results, published by Gandon (2011), indicate an 80% compliance rate with the policy: out of the selected 39 articles, 11 were based on proprietary data and 20 had the appropriate code and data posted. McCullough (2018) draws much more negative conclusions from this replication experiment, noting that Gandon's study only investigates nine papers, of which only five have been fully reproduced. Chang and Li (2017) study a broader set of 67 articles published in top economics journals. They were able to reproduce the results for one-third of these papers from the code and data available on the journals' repositories. For another 10% of these articles, the results were reproduced with some help from the original authors.

All these results point toward an unverified reproducibility policy being insufficient, so that there is a case for economics journals to engage in reproducibility verification or stage 3. More generally, an important question, which we are going to address in the next section, is whether all journals should increase their level of reproducibility, and have an incentive to do so.

2.2 Is Low Reproducibility Socially Optimal?

The observation of a low level of reproducibility in economics research does not necessarily imply that it is inefficient. Indeed, competition between academic journals may lead to the level of reproducibility that equates readers' demand with authors' supply.

We show in this section that the two-sided nature of competition between academic journals does not guarantee that the equilibrium level of reproducibility corresponds to a social optimum. To emphasize this point, we will show a theoretical example in which journals are competitive and

choose a suboptimally low level of reproducibility. Hence, the low level of reproducibility observed in economics may be due to a market failure and call for corrective actions.

We enrich the setup of Section 1.1 with a model of competition between two academic journals, indexed by $k \in \{1, 2\}$. Each journal aims at maximizing its impact, represented by the number of readers n_k^R , subject to the break-even constraint (3).²² The game then plays as follows:

Step 1: Each journal k simultaneously chooses the authors' fee p_k^A , readers' fee $p_k^R \geq 0$, and reproducibility level $q_k \geq 0$.

Step 2: The authors and readers observe (p_1^A, p_1^R, q_1) and (p_2^A, p_2^R, q_2) . Each author chooses whether to submit to journal 1, journal 2, or to no journal (a paper cannot be submitted to two journals). Each reader chooses whether to subscribe to journal 1, journal 2, both journals, or no journal. All players make their decisions simultaneously.

We provide a formal analysis of this game and the proofs of our main results in the appendix to this paper. To keep the model tractable, and in line with the literature on two-sided markets,²³ we reduce the articles' characteristics to a single dimension $x_i \leftrightarrow \mathcal{U}([0, 1])$. More specifically, we assume that the authors of an article i published in journal k obtain:

$$u_1^A(n_1^R, q_1, x_i) - p_1^A = B + \alpha n_1^R - a q_1 - \frac{t}{2} x_i \text{ if } k = 1 \quad (5)$$

$$u_2^A(n_2^R, q_2, x_i) - p_2^A = B + \alpha n_2^R - a q_2 - \frac{t}{2} (1 - x_i) \text{ if } k = 2 \quad (6)$$

This assumption corresponds to a linear Hotelling specification, with some authors having a preference for journal 1 and others for journal 2. A higher t implies that the two journals are more

²²In our setup maximizing the number of readers is equivalent to maximizing the welfare of the journal's readers. While other specifications are possible, we are adopting the one that seems the least likely to bias the outcome towards a low level of reproducibility.

²³Most notably [Armstrong \(2006\)](#). This section can be seen as an extension of Section 5 in his paper, with reproducibility as an additional "quality" variable. See also [Armstrong \(2015\)](#) for an application to academic journals. Also note that quality is a choice variable for the journal. If it were a characteristic of authors, journals would face the problem of excluding some types (as in [Hagiu \(2009\)](#)) or sorting them with prices (as in [Damiano and Li \(2008\)](#)).

differentiated and have more market power over the authors closer to them.

Symmetrically, readers' characteristics are reduced to a single dimension $y_j \in \mathbb{R}^+$, distributed such that y readers have $y_j \leq y$. Subscribing to journal k gives reader j the payoff:

$$u^R(n_k^A, q_k, y_j) - p_k^R = \beta n_k^A + bq_k - y_j - p_k^R. \quad (7)$$

Note that authors care only about the number of readers, but not about their characteristics, while symmetrically readers care only about the number of authors.²⁴ The parameters α and β measure the sensitivity of each side to the number of agents on the other side. The parameter $a > 0$ measures the cost of reproducibility for authors, and the parameter $b > 0$ the gain for readers.

Finally, we use a simple quadratic specification for the reproducibility cost faced by journals:

$$C^J(q, n_A) = n_A \frac{\kappa}{2} q^2. \quad (8)$$

We make the following assumptions on the parameters:

Assumption 1. $t > \alpha\beta$

Assumption 2. $\beta < \frac{a}{b} < \alpha$

Assumption 3. $\kappa > 2b^2$

Assumption 1 is a stability condition, standard in the literature, that ensures that both journals are sufficiently differentiated to coexist in equilibrium. The right-hand side of Assumption 2 means that reproducibility is desirable at least if the journal's costs are null. The left-hand side means that reproducibility is not "too desirable", which helps reduce the number of cases to consider. Assump-

²⁴The critical assumption here is that the submission decisions of authors are not driven by characteristics that readers also care about. Note that our empirical analysis in Section 1.3.2 does not reject this assumption. Assuming instead for instance that authors with potentially more cited papers are also more sensitive to the level of reproducibility should intuitively reinforce the market failure exhibited in this section.

tion 3 means that the journal’s costs are sufficiently high relative to the benefits of reproducibility for readers, which also helps ruling out some possible equilibrium configurations.

We look for a subgame perfect Nash equilibrium. Authors and readers choose a journal so as to maximize their utility, rationally anticipating the behavior of other players. The two journals choose their fees and reproducibility levels so as to maximize their readership, rationally anticipating the future behavior of authors and readers. In addition, journals have to break even. Finally, we restrict our attention to symmetric equilibria with full coverage: in such an equilibrium $(p_1^A, p_1^R, q_1) = (p_2^A, p_2^R, q_2)$ and all authors submit to a journal.

A critical feature of this market is that a given article can only be published in one journal at most (“single-homing”). Hence, an article i is published in journal 1 if and only if $u_1^A(n_1^R, q_1, x_i) - p_1^A \geq \max(0, u_2^A(n_2^R, q_2, x_i) - p_2^A)$, and symmetrically for journal 2. In contrast, readers can subscribe to different journals (“multi-homing”). Reader j subscribes to journal k if and only if $u^R(n_k^A, q_k, y_j) - p_k^R \geq 0$. Since we focus on equilibria with full coverage, the number of articles submitted to each journal for given prices and reproducibility levels is determined by solving for the cutoff type x_i indifferent between both journals. In the Appendix, we solve in closed-form the equilibrium cutoff type, the numbers of readers and authors at both journals, and then the equilibrium choice of prices and reproducibility levels by journals:²⁵

Proposition 1. *A symmetric equilibrium with full coverage exists if and only if $t \leq \bar{t}$, where the value of \bar{t} is given in the Appendix. In such an equilibrium both journals choose (p_A^*, p_R^*, q^*) , with $p_R^* = 0$, $p_A^* = C^J(q^*, 1)$, and*

$$q^* = \frac{2b(t - \alpha\beta) + \beta(\alpha b - a)}{\beta\kappa}$$

Is the equilibrium level of reproducibility socially optimal? To answer this question, we con-

²⁵In this equilibrium, readers can subscribe to each journal for free (“open access”), and the costs of reproducibility are fully borne by the authors. This result is due to the assumption that $\beta < a/b$ and is not generic. McCullough (2009b) observes that open access journals are less advanced than others in promoting reproducibility. Our model highlights that this may be due to the greater necessity for these journals to attract authors.

sider the program of a social planner who would choose p_k^A, p_k^R, q_k in order to maximize the total number of readers across both journals (or, equivalently, who would maximize the total welfare of readers), under the constraints that both journals break even and all authors submit their paper (full coverage). We obtain the following solution:

Proposition 2. *For any $t \leq \bar{t}$, the social planner implements $(p_A^{**}, p_R^{**}, q^{**})$ in both journals, with $p_R^{**} = 0, p_A^{**} = C^J(q^{**}, 1)$, and*

$$q^{**} = \frac{\alpha b - a + \sqrt{(\alpha b - a)^2 + \kappa(2B - t + \alpha\beta)}}{\kappa}. \quad (9)$$

We can now compare the level of reproducibility achieved under competition and with a social planner:

Proposition 3. *For any $t \leq \bar{t}$, q^* increases in t and q^{**} decreases in t . Moreover, $q^* \leq q^{**}$ with an equality in $t = \bar{t}$.*

Hence, we obtain that the social planner always chooses a higher level of reproducibility than the one we obtain under competition, as illustrated by Figure 2. The intuition is the following. Because the authors are “single-homing”, they form what [Armstrong \(2006\)](#) calls a “competitive bottleneck”: the journals are competing over attracting the marginal article, whereas for a given number of readers the demand of readers for a journal does not depend on the strategy of the other journal. Since reproducibility is costly to authors, the journals reduce their reproducibility level to attract authors. The social planner instead does not face the “competitive bottleneck” problem and does not have to leave any surplus to the author with $x_i = 1/2$.²⁶ As t increases the social planner needs to choose a lower level of reproducibility in order to keep all authors submitting, hence q decreases. On the contrary, under competition as t increases the journals are more and

²⁶The result can also be compared to the case of competition between platforms owned by associations in [Rochet and Tirole \(2003\)](#), which does not lead to the first-best due to a business-stealing effect.

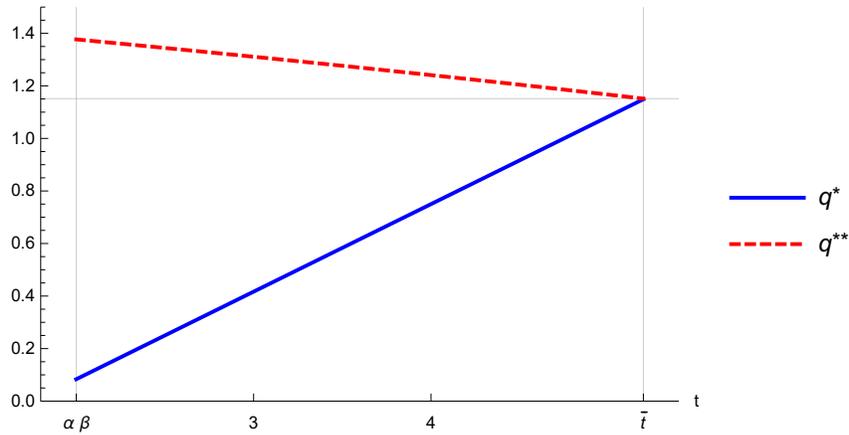


Figure 2: **Reproducibility level under competition (q^*) and under a social planner (q^{**})**

more differentiated and can choose a higher q . As t approaches \bar{t} , competition between journals gives close to zero surplus to the author with $x_i = 1/2$. The “competitive bottleneck” effect disappears and the outcome of competition converges to the planner’s solution.

3 Removing Barriers to Verified Reproducibility

As of today (July 2021), only a handful of economic journals moved to verified reproducibility. The previous section showed that this may correspond to a suboptimal equilibrium of the competition between journals. This section discusses several avenues to move out of this equilibrium.

3.1 Changing Journals’ Incentives

The effect of competition between journals is similar to situations in which competition between producers leads to a suboptimally low quality (e.g., [Kranton \(2003\)](#)). Classical solutions to this problem are the establishment and enforcement of industry standards, in this case a common reproducibility policy across journals, or an initiative to increase reproducibility taken by a journal with sufficient market power (when t is large in our model). As we will now discuss, the top 3 finance journals illustrate the first possibility, and the *American Economic Review* the second.

As explained in [Harvey \(2014\)](#), an initiative to increase reproducibility at the top 3 finance journals emerged in 2010. The idea was to adopt a common reproducibility policy at all journals simultaneously, akin to fixing a common level q in both journals in our model. Ultimately this initiative was not adopted. Among the different reasons mentioned in [Harvey \(2014\)](#), an important one is competition. In particular, the top 3 finance journals do not have such an important market power over authors, as they need to compete with top economics journals to attract the best finance papers.

In economics journals instead, the marginal benefit of publishing a paper in a top-5 journal is so large ([Heckman and Moktan \(2020\)](#)) that researchers have strong incentives to comply with any standard or disclosing requirement imposed by these journals.²⁷ There is of course competition among top-5 journals. Despite this competitive pressure, the market power of the *American Economic Review* seems to have been high enough to keep its ambitious 2004 data policy, and even taking the next step of verified reproducibility in 2018 for conditionally accepted papers (as shown in Figure 2). Moreover, the 2004 data policy became a standard that was adopted by the other journals. Indeed, all top-5 journals now have a similar data policy, and two of them (*Journal of Political Economy* and *Quarterly Journal of Economics*) explicitly mention that they adopted the AER's 2004 policy.

3.2 Lowering Costs for Journals

Verified reproducibility is easier to achieve if it is conducted by people or organizations with the right expertise and incentives. In theory, three different models could be envisioned: A first possibility could be to add verification to the tasks of the regular editors and referees of the journal.

However, editors and referees may not have the time, expertise, and data access to check the repro-

²⁷See [Ductor et al. \(2020\)](#) for a theory of why generalist journals became dominant in economics. Note that in their paper, there are externalities only among authors publishing papers to the same journal, whereas in our model there are also externalities between readers and authors.

ducibility of all accepted papers.²⁸ Importantly, they also do not have the right incentives. In the suboptimal equilibrium discussed in Section 2.2, even if both journals announce a minimum level of reproducibility \underline{q} , each journal has an incentive to renege on this level ex post. Concretely, one can imagine the situation of an editor and referees who have accepted a promising paper for publication after multiple rounds of revision. If at this stage, the editorial team discovers that the data policy is not strictly adhered to by the authors, there seems to be a large cost and little benefit to stop the publication process. On the contrary, the editorial team may rightly consider that the benefit for the journal of publishing an impactful paper will be larger than the cost of not fully enforcing the data policy.

A second possibility is for the journal to appoint a special editor in charge of implementing the verification policy, thus avoiding the conflict of objectives that arises when the same editor is in charge of selecting impactful papers and verifying reproducibility.²⁹ For instance, in 2017 the AEA appointed Lars Vilhuber from Cornell University as the data editor for all the journals operated by the Association. On its website, the AEA defines the role of the data editor as follows: “[*The data editor*] will work with the AEA journal editors and Executive Committee to develop and implement methods to maximize replicability and reproducibility of research findings published in AEA journals. Such methods may involve some pre-publication verification of materials provided by authors but will also encourage incorporating basic principles of replicability into researchers’ workflows and address the increasing reliance on restricted-access data”. Since then, similar positions have been created at the *Review of Economics Studies*, the *Economic Journal*, and *Management Science*.

A third possibility for a journal is to use the services of a trusted third party dedicated to verifying research reproducibility. The latter can either complement a journal internal replication

²⁸Leek and Peng (2015) observe that editors and reviewers at medical and scientific journals often lack the training and time to rerun a data analysis. This problem is compounded by the fact that datasets and data analyses are becoming increasingly complex and the number of submissions to journals continues to increase.

²⁹To the best of our knowledge, the first journal to follow such a policy was *Biostatistics* (Peng (2009); Peng (2011)). For more examples of journal policies to verify computational research, see Willis and Stodden (2020).

team for some types of verification or replace it. In any case, internal and external validators conduct similar tasks: they check the submitted material, rerun the code, contrast the results obtained with those in the paper, write a reproducibility report presenting all the steps and any discrepancies. An example is the *cascaD* certification agency, which conducted 21 verifications for journals managed by the *American Economic Association* (Vilhuber, 2021). Another interesting example is the partnership between the *American Journal of Political Science* and the University of North Carolina's Odum Institute. This institute provides the journal with a specialized data review service to guarantee the quality of replication datasets (Christian *et al.*, 2018).³⁰

3.3 Allowing Verification of Confidential Data

The use of confidential data is often mentioned as a major impediment to the implementation of reproducible research (Christensen and Miguel, 2018). Without a solution to handle papers using confidential data, only two outcomes are possible: (1) Exclusion: the journal publishes only papers based on non-confidential data, and may have to pass on many interesting papers; (2) Exemption: papers using confidential data are exempted from the DAP, which may leads to many articles being non-reproducible, and may even encourage authors to work on confidential data so as to bypass the policy.

As an example of exclusion, the DAP of the review PLOS states that whenever the data cannot be accessed by other researchers, the manuscript must include an additional analysis based on public data that validates the conclusions so that others can reproduce the results.³¹ Exemption is the most common policy among economics journals. Among the 49 DAP considered by Vlaeminck and Herrmann (2015), 34 offer exemptions to the policy for confidential datasets. Christensen and Miguel (2018) show that the share of empirical papers published in the *American Economic Review*

³⁰See Willis and Stodden (2020) for more examples.

³¹See the unacceptable data access restrictions considered in PLOS data policy, <https://journals.plos.org/plosone/s/data-availability>.

that fall under these exemptions rose sharply from 10% to around 40% between 2005 and 2015.

These numbers show that exclusion is not a realistic possibility in economics. Confidential data on consumers or firms allow researchers to address new research questions or provide innovative answers to traditional ones. It provides to those who can access such data a comparative advantage and greatly increases chances to publish in top journals. Moreover, the extension of the legal frameworks protecting confidentiality implies that a growing fraction of the data used in economics now has to be treated as confidential.³²

If the exclusion of papers using confidential data imposes too high a cost on the progress of economics, making 40% of published articles non-reproducible does not seem satisfactory either. Hence there is a need to find a reproducibility solution for papers using confidential data. In principle, editors or referees of a journal could access such data for purpose of reproducibility by following a specific accreditation process for each confidential-data provider. The reason why journals typically do not use this option is that each journal would have to follow an often long and tedious accreditation process for each confidential data provider.

Here the use of a third party verification team can prove particularly efficient: once accredited by a confidential-data provider, the third party can use this accreditation for all papers using the data, regardless of the journal they are published in. As an example, in 2018 *casca*d partnered with the Centre d'Accès Sécurisé aux Données (CASD), a public research infrastructure enabling researchers to access individual data from the French Institute of Statistics and Economic Studies (INSEE), and from various French public administrations and ministries (Pérignon *et al.*, 2019). In total, CASD hosts data from 378 sources and offers a data provider service to 742 user institutions. In October 2018, the French Statistical Secrecy Committee granted *casca*d a permanent accreditation

³²A recent example is the EU General Data Protection Regulation enforced in 2018. Unlike the European Union, the US does not have a single law on data protection but instead a system of federal and state laws and regulations, including among others the Federal Trade Commission Act, the Financial Services Modernization Act, the Health Insurance Portability and Accountability Act, the Electronic Communications Privacy Act, etc.

for all their verification reviewers.

3.4 Achieving Economies of Scale

An obvious way of favoring the move towards verified reproducibility is to decrease its cost, represented by $C^J(q, n_A)$ in our theoretical framework. In this section, we summarize some quantitative information regarding this cost and discuss the implications for how to best organize the verification of reproducibility.

We consider a given level of reproducibility \bar{q} , corresponding to Stage 3 on the reproducibility scale displayed in Figure 1. We take as given the number n_A of articles to verify, and the number n_D of distinct non-public databases to access. The total cost can be represented as:

$$\bar{C}(n_A, n_D) = c_F + n_A(c_L + c_C) + n_D c_D \quad (10)$$

We briefly discuss and give a tentative estimate of each cost. Our estimates are based on the actual experience of the *cascad* verification agency. However, these estimates are necessarily quite rough and are only provided to give an order of magnitude of the costs of verifying reproducibility.

- Fixed costs (c_F) reflect the cost of setting up a Swiss-army-knife IT infrastructure allowing the replicating team to run any code provided by an author. Based on the actual expenses faced by *cascad*, we set c_F to 50,000 euros. This includes expenses related to dedicated hardware, storage capacity, cloud resources, software, etc. In addition, the costs include building an online platform allowing data editors to manage manuscripts and to communicate with reviewers and researchers, and covering legal and administrative costs, as well as a fraction of the salary of the two reproducibility editors.

- Labor costs (c_L) reflect the compensation of the technical staff in charge of checking the compliance of the submitted material to the guidelines, running the code, comparing the results

with the ones in the paper, and writing an execution/reproducibility report to be provided to the data editor. By looking at the actual time spent by the reproducibility reviewers in 2020, we set the average number of hours per verification to 10 hours.³³ Given the salaries actually paid by *cascad* in 2020, we use an hourly rate of 15 euros. Hence c_L is approximately equal to 150 EUR on average.

- Computing costs (c_C) need to be paid when the code is run on a commercial cloud. Our estimate of C_C is 5 EUR per article.

- The cost of accessing data (c_D) varies a lot across databases. Many commercial databases are already available “for free” via a campus license. Finding the fee for other commercial databases is easy, but providing the fair monetary cost of establishing a partnership with a restricted-data access center like CASD is almost impossible. Yet, not including it would lead to massively underestimate the cost of setting up a verification service. Averaging across commercial and administrative databases, we use $c_D = 5,000$ EUR as a rough (and conservative) estimate.³⁴

We now have estimates for all the parameters of the function $\bar{C}(n_A, n_D)$. Given the large share of data costs in the total reproducibility costs, a critical issue is how many new databases become necessary as the number of papers to be verified grows. As in Maurice Allais’ famous “Calais traveller metaphor”, the marginal cost of a new paper using a dataset that is not currently available to reviewers is much higher than the cost of a paper using already available datasets. Furthermore, bringing a new data source enriches the data portfolio of the reviewing team, which makes less likely the need to access an additional data source for the next papers to be verified (as the pool of available data sources is now larger).

³³Our estimate is higher than the 5 hours reported by [Vilhuber \(2019\)](#) at Cornell University. We believe the reason the reviewing time is on the high side at *cascad* is that (1) the proportion of papers using confidential data is larger at *cascad* and (2) the level of compliance of the submitted material with the guideline remains moderate.

³⁴In the case of CASD, access to the data requires formal approval from the French Statistical Secrecy Committee, which is a 3-6 month process. We found 134 articles published over 2016-2020 and acknowledging using CASD data, in 91 different academic journals. In order to verify the reproducibility of all these articles without *cascad*, a total of 91 different journals would have had to go through the same lengthy accreditation process. In some cases, access to the data by academic journals would have been simply impossible, as access is restricted to users based in France (e.g., data from the French Tax authority).

To obtain a rough estimate of how the number of necessary datasets grows with the number of papers, we use the following model. Assume there is a total \bar{n}_D of non-public datasets that can be used in economic research. A fraction $1 - \theta$ of papers use public data or no data, and a fraction θ pick one dataset at random among the \bar{n}_D of non-public datasets. We show in the Appendix that the expected number of distinct non-public datasets that will be used by n_A papers is equal to:

$$n_D^*(n_A, \bar{n}_D) = \bar{n}_D \times \left[1 - \left(\frac{\bar{n}_D - 1}{\bar{n}_D} \right)^{\theta n_A} \right]. \quad (11)$$

In our quantification exercise below, we use $\theta = 0.4$ (estimate of Christensen and Miguel (2018) for the *American Economic Review*) and tentatively set $\bar{n}_D = 50$. We finally arrive at the following cost function:

$$C^J(\bar{q}, n_A) = \bar{C}(n_A, n_D^*(n_A, \bar{n}_D)) = c_F + n_A(c_L + c_C) + c_D n_D^*(n_A, \bar{n}_D). \quad (12)$$

Finally, we take into account that there might be multiple verification teams. If there are n_C such teams and they share the n_A articles to be verified equally, we can compute the average cost per article $AC(n_A, n_C)$ as:

$$AC(n_A, n_C) = \frac{C^J(n_A/n_C)}{n_A/n_C}. \quad (13)$$

Figure 3 shows the average cost per article for different values of n_C , using our numerical estimates for our parameters. The graph makes it clear that economies of scale are very large, driven both by the fixed cost c_F and the data costs c_D . The economic feasibility of reproducibility thus critically depends on the number of articles n_A .

To give an example, assume we consider verifying all the articles published by the AEA Journals (except JEP, JEL, AEA P&P) and amortize the fixed costs cover 3 years of activity. According to

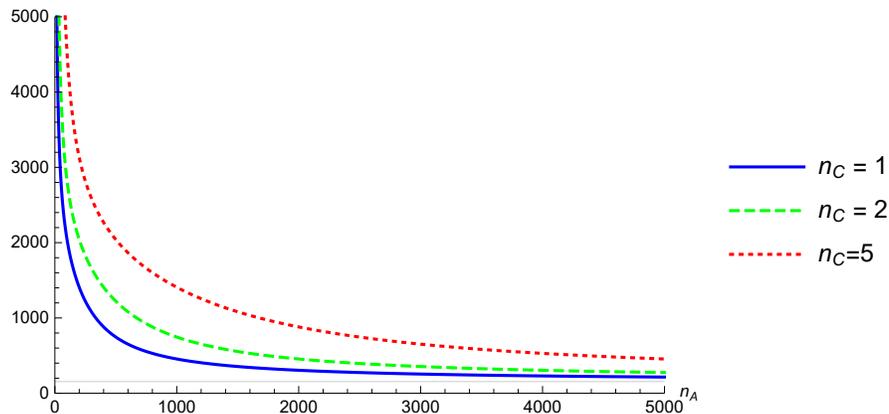


Figure 3: **Average costs per article, as a function of the number of articles n_A and number of verification teams n_C**

the 2020 editorial reports, these journals together published 344 papers in 2019. This would give $n_A = 1,032$. At this level of activity, the average cost per article is 446 EUR for $n_C = 1$, and 729 EUR for $n_C = 2$, hence economies of scale are still very significant. Assume we go further and the verification team also verifies the papers from the four other top-5 journals. Together, these four journals published 261 papers in 2019. Multiplying by three, we reach a total of $n_A = 1,815$ for three years. At this level, the average cost per paper falls to 320 EUR for $n_C = 1$, and 485 EUR for $n_C = 2$. Obviously the costs fall further if one adds other journals or the fixed costs can be amortized over more years, and the average cost reaches 155 EUR per article in the limit.

Assuming there is only one verification team, how large is the cost of 320 EUR \simeq 358 USD (2019) per article? One way to answer this question is to estimate how much different sources of income for the AEA would have to be increased to compensate the costs of verifying the reproducibility of all articles published by AEA journals.³⁵ One possibility is to ask the authors of the papers reproduced to pay the cost. This would increase the submission fees for authors of accepted papers from 200 USD to 558 USD, a 179% increase. A second possibility would be to increase the submission fees for all papers submitted. Given a 7% acceptance rate for AEA journals on average, one would need

³⁵We use numbers from the AEA Financial Statements 2019, available here: <https://benny.aeaweb.org/content/file?id=12235>.

to raise submission fees by $0.07 \times 358 = 25.06$ USD, a 12.5% increase from the current submission fees. Finally, a third possibility would be to raise the costs on readers. According to the financial statements of the AEA, the AEA earned 5.931 million USD in licensing fees and subscription fees for its journals in 2019. In order to absorb an extra cost of $344 \times 358 = 123,152$ USD, these fees would have to increase by 2.07%.

The lesson we draw from this quantitative exercise is that the costs of verified reproducibility are far from negligible, but still manageable if one strives to reduce implementation costs and finds the right economic model. Very practically, our estimates mean that the total cost of verifying the reproducibility of all articles published in the top-5 economics journals and other AEA journals would be less than USD 220,000 per year, close to the average annual salary of one full professor in economics (Scott and Siegfried, 2021). However, this relatively low number assumes operating at scale. This is certainly a factor in the success of the policy adopted by the AEA journals. Achieving the same outcome will be more difficult for standalone journals, unless they are able to pool resources or resort to third-party verification.

4 Discussion

4.1 Externalities and Government Intervention

In our framework, the journals internalize the demand for reproducibility coming from “readers”, which could be the subscribers of the journals (who pay subscription fees) and/or academics who cite the articles published by the journal in their research. However, an article i published in a journal may have a social utility, denoted $u^S(q, X_i)$, in addition to the readers’ utility. Even a monopoly journal maximizing the utility of readers will fail to take this additional social utility into account. If u^S is increasing in q , then this is another channel through which the equilibrium level of reproducibility might be suboptimally low.

A traditional externality of research is its application to create new products or processes and generate economic growth. Reproducibility of research may strengthen such spillovers: the availability of the code and data used by academics may greatly facilitate the transformation of scientific discoveries into economic forces. This echoes the “stimulation” function of reproducibility of Ragnar Frisch described in Section 1.2.

In recent years, the demand for “evidence-based” policymaking has also greatly increased (Foundations for Evidence-Based Policymaking Act of 2018; Cole *et al.* (2020)). As governments and policy institutions increasingly want to base their policies on recent academic research, there is logically a demand for this research to be as trustworthy as possible. This echoes the “control” function of reproducibility of Ragnar Frisch discussed in Section 1.2. This function was very prominent during the Covid-19 crisis, with the emphasis put by various public decision-makers on the use of *peer-reviewed* medical research.

Such externalities may rationalize the use of government intervention to increase the level of reproducibility. While the government has no regulatory power over academic journals, it can considerably influence research policy via its control of public universities and various funding sources. For instance, public research funding agencies now impose open research standards and data management plans for the projects they fund.³⁶

How well-intended these policies might be, centralized authorities are likely to be in a worse position than academics themselves to evaluate the costs of reproducibility on researchers and strike the right trade-off.³⁷ Scientific associations and journals may be seen as “self-regulatory organizations” that set the rules of research independently of the government, but with government

³⁶In 2014, the NSF proposed a framework to improve the reproducibility and replicability in funded research, including data sharing policy and data management plans (National Science Foundation, 2014). In the European Union’s Horizon 2020 funding program, research data underlying a publication has to be made available, in addition to the requirement to create a data management plan (European Commission, 2020).

³⁷See for instance the critical discussion in Armstrong (2021) of the EU’s Plan S initiative for open-access journals, relying on the “competitive bottleneck” effect.

intervention remaining a last resort possibility. This implies that academics have a collective interest in setting an appropriate level of reproducibility in research, in order to make additional government intervention unnecessary.

4.2 Refereeing

Our finding that competition between journals may lead to a suboptimally low level of reproducibility stands in sharp contrast with the view expressed by Ellison (2002a) and Ellison (2002b) that there exists a “race to the top” in the requirements imposed by referees on authors, leading to longer delays in publication, more robustness checks, longer appendices, etc. However, there is no contradiction between these two mechanisms. The “race to the top” in Ellison (2002a) stems from the behavioral biases of referees, who mistakenly infer from their own submission history that journals require a very high quality of execution. The requirements on reproducibility are for the moment not in the hands of referees but of editors, and it is hard to see how the mechanism on which Ellison (2002a) relies could be transposed at this level.

Conversely, our “competitive bottleneck” mechanism could imply that editors may care less about quality of execution (which may be another interpretation of q in the model) and more about the idea of a paper. This view is summarized by Matthew Spiegel, reflecting on his editorship of the *Review of Financial Studies*, as “reviewing less and progressing more” (Spiegel (2012)).

Conclusion

Writing empirical papers in economics takes a great deal of time - literally years of data cleaning, merging, coding, debugging, analyzing and re-analyzing - yet, for decades, data and code played no role in the peer review process. When receiving a new manuscript to review, editors and referees had to assume that the results outlined in the paper were actually resulting from running the researchers’

computer code on their data. Over the past 15 years, economics journals have introduced and then gradually enriched their reproducibility policy: first inviting and then forcing authors to share their code and data. There is now growing empirical evidence to show that this period, which we call the unverified reproducibility policy stage, led to low compliance rate with DAP, low quality of the shared resources, and in turn low reproducibility rates.

In this paper, we show that this unfortunate situation can be the suboptimal outcome of two-sided competition between economic journals. If so, it is possible to improve the situation by increasing the level of reproducibility. Finally, we show that one way to mitigate this market failure is to conduct a systematic verification of the results prior to publication (verified reproducibility stage). Such pre-publication verification could be conducted either internally by journals or outsourced to a trusted third party. Interestingly, this is the strategy followed by the American Economic Association, as well as by a handful of other leading actors of the scientific publishing industry.

While the focus of this paper is on economics research, the situation described and some of the solution put forward are by no means specific to our field. In particular, our analysis resonates well with current discussions in the medical studies in the context of the Covid crisis. Indeed, both economics and medical studies are extremely data and computationally intensive, focus on causality, make extensive use of confidential data, and play a vital role in informing the policy and societal debates. The publication in May 2020 in *The Lancet* of an article on the effects of Hydroxychloroquine in the treatment of Covid-19, followed by its quick retraction, illustrates both the high societal demand for reproducibility and the difficulties journals face in supplying it. The recent policy changes and initiatives discussed in this paper could be sources of inspiration for the medical community.

References

- ARMSTRONG, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, **37** (3), 668–691. 4, 15, 18
- (2015). Opening access to research. *The Economic Journal*, **125** (586), F1–F30. 15
- (2021). *Plan S: An Economist’s Perspective*. Tech. rep. 29
- BAKER, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, **533** (7604), 452–455. 8
- BARBA, L. A. (2018). Terminologies for reproducible research. *arXiv preprint arXiv:1802.03311*. 1
- BERNANKE, B. S. (2004). Editorial statement. *The American Economic Review*, **94** (1), 404–404. 2
- BRODEUR, A., COOK, N. and HEYES, A. (2020). Methods matter: p-hacking and publication bias in causal analysis in economics. *The American Economic Review*, **110** (11), 3634–60. 8
- BUCKHEIT, J. B. and DONOHO, D. L. (1995). Wavelab and reproducible research. In *Wavelets and statistics*, Springer, pp. 55–81. 1
- CARD, D. and DELLAVIGNA, S. (2020). What Do Editors Maximize? Evidence from Four Economics Journals. *The Review of Economics and Statistics*, **102** (1), 195–217. 6
- CHANG, A. C. and LI, P. (2017). A preanalysis plan to replicate sixty economics research papers that worked half of the time. *The American Economic Review*, **107** (5), 60–64. 3, 14
- CHRISTENSEN, G., FREESE, J. and MIGUEL, E. (2019). *Transparent and Reproducible Social Science Research: How to Do Open Science*. University of California Press. 8
- and MIGUEL, E. (2018). Transparency, reproducibility, and the credibility of economics research. *Journal of Economic Literature*, **56** (3), 920–80. 7, 22, 26
- CHRISTIAN, T.-M. L., LAFFERTY-HESS, S., JACOBY, W. G. and CARSEY, T. (2018). Operationalizing the replication standard. *International Journal of Digital Curation*, **13** (1), 114–124. 22
- COLE, S., DHALIWAL, I., SAUTMANN, A. and VILHUBER, L. (2020). *Handbook on Using Administrative Data for Research and Evidence-based Policy*. 29
- DAMIANO, E. and LI, H. (2008). Competing matchmaking. *Journal of the European Economic Association*, **6** (4), 789–818. 15
- DEWALD, W. G., THURSBY, J. G. and ANDERSON, R. G. (1986). Replication in empirical economics: The journal of money, credit and banking project. *The American Economic Review*, **76** (4), 587–603. 13
- DUCTOR, L., GOYAL, S., VAN DER LEIJ, M. and PAEZ, G. N. (2020). *On the Influence of Top Journals*. Cambridge Working Papers in Economics 2029, Faculty of Economics, University of Cambridge. 20
- DUVENDACK, M., PALMER-JONES, R. W. and REED, W. R. (2015). Replications in Economics: A Progress Report. *Econ Journal Watch*, **12** (2). 2
- , — and — (2017). What is meant by ”replication” and why does it encounter resistance in economics? *The American Economic Review*, **107** (5), 46–51. 3, 13
- ELLISON, G. (2002a). Evolving standards for academic publishing: A q-r theory. *Journal of Political Economy*, **110** (5), 994–1034. 30
- (2002b). The slowdown of the economics publishing process. *Journal of Political Economy*, **110** (5), 947–993. 30

- EUROPEAN COMMISSION (2020). *Reproducibility of scientific results in the EU: scoping report*. Tech. rep. 29
- FRISCH, R. (1933). Editor's note. *Econometrica*, **1** (1), 1–4. 2
- GEHRIG, T. and STENBACKA, R. (2021). Journal competition and the quality of published research: Simultaneous versus sequential screening. *International Journal of Industrial Organization*, **76**, 102718. 5
- GERTLER, P., GALIANI, S. and ROMERO, M. (2018). How to make replication the norm. *Nature*, **554** (7693), 3, 13
- GLANDON, P. J. (2011). Appendix to the report of the editor: Report on the american economic review data availability compliance project. *The American Economic Review*, **101** (3), 696–699. 14
- HAGIU, A. (2009). *Quantity vs. Quality and Exclusion by Two-Sided Platforms*. Tech. rep. 15
- HARVEY, C. R. (2014). Reflections on Editing the Journal of Finance, 2006 to 2012. *Secrets of Economics Editors*, p. 67. 4, 9, 12, 20
- (Forthcoming). Replication in financial economics. *Critical Finance Review*, - (-). 8
- HECKMAN, J. J. and MOKTAN, S. (2020). Publishing and promotion in economics: The tyranny of the top five. *Journal of Economic Literature*, **58** (2), 419–70. 11, 20
- HERNDON, T., ASH, M. and POLLIN, R. (2013). Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff. *Cambridge Journal of Economics*, **38** (2), 257–279. 8
- HILL, R. and STEIN, C. (2020). *Race to the bottom: Competition and quality in science*. Tech. rep., MIT. 9
- HÖFFLER, J. H. (2017). Replication and economics journal policies. *The American Economic Review*, **107** (5), 52–55. 2
- JEON, D.-S. and ROCHET, J.-C. (2010). The pricing of academic journals: A two-sided market perspective. *American Economic Journal: Microeconomics*, **2** (2), 222–55. 3, 5
- KOENKER, R. and ZEILEIS, A. (2009). On reproducible econometric research. *Journal of Applied Econometrics*, **24** (5), 833–847. 3
- KRANTON, R. E. (2003). Competition and the incentive to produce high quality. *Economica*, **70** (279), 385–404. 19
- LEEK, J. T. and PENG, R. D. (2015). Opinion: Reproducible research can still be wrong: Adopting a prevention approach. *Proceedings of the National Academy of Sciences*, **112** (6), 1645–1646. 21
- LERNER, J. and TIROLE, J. (2002). Some simple economics of open source. *The Journal of Industrial Economics*, **50** (2), 197–234. 9
- MCCABE, M. J. and MUELLER-LANGER, F. (2019). *Does Data Disclosure Increase Citations? Empirical Evidence from a Natural Experiment in Leading Economics Journals*. JRC Working Papers on Digital Economy 2019-02, Joint Research Centre (Seville site). 9
- and SNYDER, C. M. (2005). Open access and academic journal quality. *The American Economic Review*, **95** (2), 453–458. 5
- MCCULLOUGH, B. (2009a). Open access economics journals and the market for reproducible economic research. *Economic Analysis and Policy*, **39** (1), 117 – 126. 2
- (2009b). Open access economics journals and the market for reproducible economic research. *Economic Analysis and Policy*, **39** (1), 117–126. 17

- (2018). Quis custodiet ipsos custodes? despite evidence to the contrary, the american economic review concluded that all was well with its archive. *Economics : the Open-Access, Open-Assessment e-Journal*, **12**, 1–13. [14](#)
- MCCULLOUGH, B. D., MCGEARY, K. A. and HARRISON, T. D. (2006). Lessons from the jmcba archive. *Journal of Money, Credit and Banking*, **38** (4), 1093–1107. [13](#)
- , — and — (2008). Do economics journal archives promote replicable research? *The Canadian Journal of Economics / Revue canadienne d'Economique*, **41** (4), 1406–1420. [3](#), [13](#)
- MUELLER-LANGER, F., FECHER, B., HARHOFF, D. and WAGNER, G. G. (2019). Replication studies in economics - How many and which papers are chosen for replication, and why? *Research Policy*, **48** (1), 62–83. [7](#), [13](#)
- NATIONAL SCIENCE FOUNDATION (2014). *A Framework for Ongoing and Future National Science Foundation Activities to Improve Reproducibility, Replicability, and Robustness in Funded Research*. Tech. rep. [29](#)
- OROZCO, V., BONTEMPS, C., MAIGNÉ, E., PIGUET, V., HOFSTETTER, A., LACROIX, A., LEVERT, F. and ROUSSELLE, J.-M. (2020). How to make a pie: Reproducible research for empirical economics and econometrics. *Journal of Economic Surveys*, **34** (5), 1134–1169. [3](#)
- PENG, R. D. (2009). Reproducible research and biostatistics. *Biostatistics*, **10** (3), 405–408. [21](#)
- (2011). Reproducible research in computational science. *Science*, **334** (6060), 1226–1227. [21](#)
- , DOMINICI, F. and ZEGER, S. L. (2006). Reproducible epidemiologic research. *American journal of epidemiology*, **163** (9), 783–789. [1](#)
- PÉRIGNON, C., GADOUCHE, K., HURLIN, C., SILBERMAN, R. and DEBONNEL, E. (2019). Certify reproducibility with confidential data. *Science*, **365** (6449), 127–128. [23](#)
- REINHART, C. M. and ROGOFF, K. S. (2010). Growth in a time of debt. *The American Economic Review*, **100** (2), 573–578. [8](#)
- ROCHET, J.-C. and TIROLE, J. (2003). Platform Competition in Two-Sided Markets. *Journal of the European Economic Association*, **1** (4), 990–1029. [18](#)
- SCHWERT, G. W. (2021). The remarkable growth in financial economics, 1974–2020. *Journal of Financial Economics*, **140** (3), 1008–1046. [10](#), [12](#)
- SCOTT, C. E. and SIEGFRIED, J. J. (2021). American economic association 2020–2021 universal academic questionnaire summary statistics. *AEA Papers and Proceedings*, **111**, 647–49. [28](#)
- SIMONSOHN, U. (2013). Just post it: The lesson from two cases of fabricated data detected by statistics alone. *Psychological science*, **24** (10), 1875–1888. [8](#)
- SOCIAL SCIENCE DATA EDITORS (2020). Unofficial guidance on various topics by social science data editors. [3](#)
- SPIEGEL, M. (2012). Reviewing less—progressing more. *The Review of Financial Studies*, **25** (5), 1331–1338. [30](#)
- STODDEN, V. (2010). Data sharing in social science repositories: Facilitating reproducible computational research. Communication at the NIPS workshop: Computational Science and the Wisdom of Crowds. [8](#)
- , SEILER, J. and MA, Z. (2018). An empirical analysis of journal policy effectiveness for computational reproducibility. *Proceedings of the National Academy of Sciences*, **115** (11), 2584–2589. [2](#)
- SWANSON, N., CHRISTENSEN, G., LITTMAN, R., BIRKE, D., MIGUEL, E., PALUCK, E. L. and WANG, Z. (2020). Research transparency is on the rise in economics. *AEA Papers and Proceedings*, **110**, 61–65. [12](#)

- VILHUBER, L. (2019). Report by the AEA Data Editor. *AEA Papers and Proceedings*, **109**, 718–29. [3](#), [13](#), [25](#)
- (2021). Report by the AEA Data Editor. *AEA Papers and Proceedings*, **111**, 808–17. [22](#)
- VLAEMINCK, S. and HERRMANN, L.-K. (2015). Data Policies and Data Archives: A New Paradigm for Academic Publishing in Economic Sciences? *EconStor Open Access Articles*, pp. 145–155. [13](#), [22](#)
- WANG, J. (2018). Quality screening and information disclosure in two-sided markets. *Economics Letters*, **171**, 183–188. [5](#)
- WHITED, T. M. (2021). Editorial. *Journal of Financial Economics*. [9](#)
- WILLIS, C. and STODDEN, V. (2020). Trust but verify: How to leverage policies, workflows, and infrastructure to ensure computational reproducibility in publication. *Harvard Data Science Review*, **2** (4). [21](#), [22](#)

A Proofs

A.1 Proof of Proposition 1

We first solve for the equilibrium numbers of authors and readers on each journal for given fees and quality levels. Given our assumptions, these numbers have to satisfy the following system:

$$n_1^A = \frac{1}{2} + \frac{\alpha(n_1^R - n_2^R) - a(q_1 - q_2) - (p_1^A - p_2^A)}{2t} \quad (\text{A.1})$$

$$n_2^A = 1 - n_1^A \quad (\text{A.2})$$

$$n_1^R = \beta n_1^A + bq_1 - p_1^R \quad (\text{A.3})$$

$$n_2^R = \beta n_2^A + bq_2 - p_2^R \quad (\text{A.4})$$

Solving for this system yields the following equilibrium allocation of authors and readers, for given prices and reproducibility levels:

$$n_1^A = \frac{1}{2} - \frac{\Delta}{2(t - \alpha\beta)} \quad (\text{A.5})$$

$$n_2^A = \frac{1}{2} + \frac{\Delta}{2(t - \alpha\beta)} \quad (\text{A.6})$$

$$n_1^R = \left(\frac{\beta}{2} + bq_1 - p_1^R \right) - \frac{\beta\Delta}{2(t - \alpha\beta)} \quad (\text{A.7})$$

$$n_2^R = \left(\frac{\beta}{2} + bq_2 - p_2^R \right) + \frac{\beta\Delta}{2(t - \alpha\beta)} \quad (\text{A.8})$$

$$\text{with } \Delta = p_1^A - p_2^A + \alpha(p_1^R - p_2^R) - (\alpha\beta - a)(q_1 - q_2) \quad (\text{A.9})$$

Note that Assumption 1 ensures that n_k^A and n_k^R are both decreasing in p_k^A and p_k^R , while Assumption 2 ensures that all else equal n_k^A and n_k^R are both increasing in q_k .

We can now solve for the equilibrium in Step 1. Taking (p_2^A, p_2^R, q_2) as given, we write the following Lagrangian for journal 1:

$$\mathcal{L} = n_1^R + \lambda[n_1^A(p_1^A - C^J(q_1)) + n_1^R p_1^R] + \mu p_1^R + \nu q_1 \quad (\text{A.10})$$

We then differentiate with respect to p_1^A , p_1^R , and q_1 :

$$\frac{d\mathcal{L}}{dp_1^A} = \frac{dn_1^R}{dp_1^A} + \lambda \left[\frac{dn_1^A}{dp_1^A} (p_1^A - C^J(q_1)) + n_1^A + \frac{dn_1^R}{dp_1^A} p_1^R \right] = 0 \quad (\text{A.11})$$

$$\frac{d\mathcal{L}}{dp_1^R} = \frac{dn_1^R}{dp_1^R} + \lambda \left[\frac{dn_1^A}{dp_1^R} (p_1^A - C^J(q_1)) + n_1^R + \frac{dn_1^R}{dp_1^R} p_1^R \right] + \mu = 0 \quad (\text{A.12})$$

$$\frac{d\mathcal{L}}{dq_1} = \frac{dn_1^R}{dq_1} + \lambda \left[\frac{dn_1^A}{dq_1} (p_1^A - C^J(q_1)) - n_1^A C'(q_1) + \frac{dn_1^R}{dq_1} p_1^R \right] + \nu = 0 \quad (\text{A.13})$$

In a symmetric equilibrium, these derivatives have to be zero at the equilibrium prices and reproducibility levels $p_1^A = p_2^A = p^A$, $p_1^R = p_2^R = p^R$, and $q_1 = q_2 = q$. We obtain:

$$0 = -\frac{\beta}{2(t - \alpha\beta)} + \lambda \left[\frac{1}{2} - \frac{1}{2(t - \alpha\beta)} (p^A - C^J(q)) - \frac{\beta p^R}{2(t - \alpha\beta)} \right] \quad (\text{A.14})$$

$$0 = -1 - \frac{\alpha\beta}{2(t - \alpha\beta)} + \lambda \left[-\frac{\alpha(p^A - C^J(q))}{2(t - \alpha\beta)} + p^R \left(-1 - \frac{\alpha\beta}{2(t - \alpha\beta)} \right) + \frac{\beta}{2} + bq - p^R \right] + \mu \quad (\text{A.15})$$

$$0 = b + \frac{\beta(\alpha b - a)}{2(t - \alpha\beta)} + \lambda \left[-\frac{\gamma q}{2} + \frac{\alpha b - a}{2(t - \alpha\beta)} (p^A - C^J(q)) + \left(b + \frac{\beta(\alpha b - a)}{2(t - \alpha\beta)} \right) p^R \right] + \nu \quad (\text{A.16})$$

In addition, we have the constraints $\lambda \geq 0$, $\mu \geq 0$, $\nu \geq 0$, $n^A(p^A - C^J(q)) + n^R p^R \geq 0$, $p^R \geq 0$, $q^R \geq 0$, $\lambda[n^A(p^A - C^J(q)) + n^R p^R] = 0$, $\mu p^R = 0$ and $\nu q = 0$. Finally, we need to check that the solution satisfies the assumption of full coverage, meaning that an article with $x_i = \frac{1}{2}$ gives positive surplus to its authors, which gives:

$$B + \alpha n^R - aq - p^A - \frac{t}{2} \geq 0. \quad (\text{A.17})$$

We are going to show that under our assumptions any solution to this problem has $\lambda > 0$, $\mu > 0$, and $\nu = 0$. We then solve analytically for this equilibrium, derive the expression of \bar{t} and show that our equilibrium holds if and only if $t \leq \bar{t}$.

Note first that immediately follows from (A.14) that the budget constraint is binding and $\lambda > 0$.

Step 1 - $\nu = 0$: Assume $\nu > 0$ and hence $q = 0$. Then the budget constraint becomes $n^A p^A + n^R p^R = 0$, which gives:

$$p^A = p^R \left[p^R - \frac{\beta}{2} \right]. \quad (\text{A.18})$$

Replacing p^A in (A.16) gives:

$$\nu = -b(1 + \lambda p^R) - \frac{\alpha\beta - a}{2(t - \alpha\beta)} (\beta + \lambda p^R(p^R + (\beta/2))). \quad (\text{A.19})$$

This quantity is necessarily negative, which is a contradiction. Hence, we cannot have $q = 0$.

Step 2 - $\mu > 0$: Assume a solution with $\mu = 0$ and $\nu = 0$. We then solve for the system of equations formed by (A.14), (A.15), (A.16), and the binding budget constraint, to be solved in p^A, p^R, q, λ . In particular, we obtain:³⁸

$$q = \frac{-a(\kappa - b^2)^2 + b\sqrt{X}}{\kappa(\kappa - b^2)(\kappa - 2b^2)} \quad (\text{A.20})$$

$$p^R = \frac{(\beta\kappa - ab)(\kappa - b^2) - \sqrt{X}}{2\kappa(\kappa - b^2)} \quad (\text{A.21})$$

$$\text{with } X = (\kappa - b^2)^2[a^2b^2 + \kappa(\kappa - 2b^2)(2t - \alpha\beta)]. \quad (\text{A.22})$$

We need both q and p^R to be positive. Given Assumption 3 this is equivalent to having:

$$\frac{a}{b}(\kappa - b^2)^2 \leq \sqrt{X} \leq (\kappa - b^2)(\beta\kappa - ab). \quad (\text{A.23})$$

However, it is easily shown that the left-hand side term is lower than the right-hand side term if and only if $\beta b \geq a$, which violates Assumption 2. Hence, under our parametric assumptions we cannot have a solution with $\mu = 0$.

Step 3 - Candidate solution: we now consider the only remaining candidate solution, which is to have $\lambda > 0$, $\mu > 0$, and $\nu = 0$. We set $p^R = 0$ and $\nu = 0$. The budget constraint then gives us $p^A = C^J(q)$. Replacing p^A with $C^J(q)$, we then solve for the system of equations formed by (A.14), (A.15), (A.16), to be solved in q, λ, μ .

(A.14) immediately gives:

$$\lambda = \frac{\beta}{t - \alpha\beta} > 0. \quad (\text{A.24})$$

We then plug this value of λ into (A.16) and obtain:

³⁸There is another solution to the system, in which q has the same expression with a negative coefficient in front of \sqrt{X} . q is then obviously negative, so that this solution can be discarded.

$$q = \frac{2b(t - \alpha\beta) + \beta(\alpha b - a)}{\beta\kappa} > 0. \quad (\text{A.25})$$

Finally, we replace λ and q in (A.15) to obtain:

$$\mu = \frac{2t(\kappa - 2b^2) + 2\beta b(a + \alpha\beta) - \kappa\beta(\alpha + \beta)}{2\kappa(t - \alpha\beta)}. \quad (\text{A.26})$$

We need $\mu > 0$, which is equivalent to $t > \underline{t}$, with:

$$\underline{t} = \alpha\beta - \frac{\beta}{2}[(\alpha - \beta)(\kappa - 2b^2) + 2b(a - \beta b)]. \quad (\text{A.27})$$

The term in brackets is positive due to Assumptions 2 and 3. Hence, $\underline{t} < \alpha\beta$. Assumption 1 thus guarantees that $t > \underline{t}$.

Step 4 - Full coverage: The last point to check is the assumption of full coverage. We need to have $t \leq \bar{t}$, where \bar{t} is such that the author of an article with $x_i = 1/2$ makes zero surplus by submitting to a journal. This gives:

$$\frac{\bar{t}}{2} = B + \alpha \left(\frac{\beta}{2} + bq \right) - aq - p^A \quad (\text{A.28})$$

$$= B + \frac{\alpha\beta}{2} + (\alpha b - a) \frac{2b(\bar{t} - \alpha\beta) + \beta(\alpha b - a)}{\beta\kappa} - \frac{\kappa}{2} \left(\frac{2b(\bar{t} - \alpha\beta) + \beta(\alpha b - a)}{\beta\kappa} \right)^2. \quad (\text{A.29})$$

The unique positive root to this equation gives us:

$$\bar{t} = \alpha\beta + \beta \frac{\sqrt{(4b(a - b\alpha))^2 + 32b^2\kappa B + \beta^2\kappa^2} - \beta\kappa}{8b^2} > \alpha\beta. \quad (\text{A.30})$$

In particular, we have $\bar{t} > \alpha\beta$ so that the range of values t such that a symmetric equilibrium with full coverage exists is always non empty.

A.2 Proof of Proposition 2

The social planner chooses p_A, p_R, q symmetrically for both journals under the constraint that an article with $x_i = 1/2$ gets submitted. Since the two journals are symmetric, we can write the

planner's Lagrangian as:

$$\mathcal{L} = n^R + \lambda[n^A(p^A - C^J(q)) + n^R p^R] + \mu p^R + \nu q \quad (\text{A.31})$$

$$+ \rho[B + \alpha n^R - aq - p^A - (t/2)] \quad (\text{A.32})$$

$$\text{with } n^A = \frac{1}{2} \quad (\text{A.33})$$

$$n^R = \frac{\beta}{2} + bq - p^R \quad (\text{A.34})$$

We then differentiate with respect to p^A , p^R , and q to get:

$$\frac{\lambda}{2} - \rho = 0 \quad (\text{A.35})$$

$$-1 + \lambda \left[-p^R + \frac{\beta}{2} + bq - p^R \right] + \mu - \alpha\rho = 0 \quad (\text{A.36})$$

$$b + \lambda \left[-\frac{\kappa q}{2} + bp^R \right] + \nu + \rho(\alpha b - a) = 0 \quad (\text{A.37})$$

Assumption 2 and condition (A.37) immediately give us $q > 0$ and hence $\nu = 0$. Using (A.37) again gives us $\lambda > 0$ and hence $\rho > 0$ using (A.35).

We first show that we cannot have a solution with $p^R > 0$. Assume this were the case. Then we have $\mu = 0$. We then solve the system formed by (A.35), (A.36), (A.37), the budget constraint, and the constraint that all articles are submitted, where we need to solve for $p^A, p^R, q, \lambda, \rho$. In particular, we get the following solution for p^R :

$$p^R = \frac{-1}{4\sqrt{\kappa}(\kappa - b^2)} \left(\sqrt{\kappa}[(\alpha - \beta)(\kappa - 2b^2) + 2b(a - \beta b)] + (\kappa - 2b^2)\sqrt{(\alpha + \beta)^2 + 4(2B - t)(\kappa - b^2) + 4a(a - b(\alpha + \beta))} \right). \quad (\text{A.38})$$

Under this form, we immediately see that Assumptions 2 and 3 imply that $p^R < 0$. A contradiction.

The only possible solution to the planner's program thus involves $\mu > 0$ and $p^R = 0$. We then need to solve (A.35), (A.36), (A.37), the budget constraint, and the constraint that all articles are submitted, in $p^A, q, \lambda, \mu, \rho$. We obtain:

$$q = \frac{\alpha b - a + \sqrt{X}}{\kappa} \quad (\text{A.39})$$

$$p_A = \frac{\kappa}{2} \left(\frac{\alpha b - a + \sqrt{X}}{\kappa} \right)^2 \quad (\text{A.40})$$

$$\lambda = \frac{2b}{\sqrt{X}} \quad (\text{A.41})$$

$$\rho = \frac{b}{\sqrt{X}} \quad (\text{A.42})$$

$$\mu = \frac{(\kappa - 2b^2)\sqrt{X} + b[(\alpha - \beta)(\kappa - 2b^2) + 2b(a - \beta b)]}{\kappa\sqrt{X}} \quad (\text{A.43})$$

$$\text{with } X = (\alpha b - a)^2 + \kappa(2B - t + \alpha\beta) \quad (\text{A.44})$$

Let us show that this is always a solution when $t \leq \bar{t}$. We first prove that $X > 0$. To see this, denote \bar{q} the value of q^* obtained in Proposition 1 when $t = \bar{t}$. By definition, at this point the author with $x_i = 1/2$ is indifferent between submitting and not submitting, which gives:

$$2B + \alpha(\beta + 2b\bar{q}) - 2a\bar{q} - \kappa\bar{q}^2 = \bar{t}. \quad (\text{A.45})$$

Since $t \leq \bar{t}$, we can write:

$$X \geq (\alpha b - a)^2 + \kappa(2B - \bar{t} + \alpha\beta) \quad (\text{A.46})$$

$$\Leftrightarrow X \geq (\alpha b - a)^2 + \kappa(\kappa\bar{q}^2 - 2\bar{q}(\alpha b - a)) \quad (\text{A.47})$$

$$\Leftrightarrow X \geq (\alpha b - a - \kappa\bar{q})^2 > 0. \quad (\text{A.48})$$

Using this result and Assumptions 2 and 3 implies that $\lambda, \mu, \rho, p^A, q$ are all well-defined and positive.

A.3 Proof of Proposition 3

It immediately follows from the analytical expressions of q^* and q^{**} that the former increases in t and the latter decreases in t . For any t , q^{**} is uniquely determined by the condition that an author with $x_i = 1/2$ has zero surplus, with $p^A = C^J(q)$ and $p^R = 0$. Under competition, \bar{t} is defined as the level of t such that $q^*, p^A = C^J(q^*), p^R = 0$ give zero surplus to an author with $x_i = 1/2$. Hence, in that point q^* and q^{**} need to satisfy the same condition, which can admit only one solution. This

shows that $q^* = q^{**}$.

A.4 Proof of Equation (11)

For a given n_A , denote X the number of different databases that the n_A papers will use. Denote L_i the event “database i is used by at least one paper”. We have $\mathbb{E}[X] = \bar{n}_D \mathbb{E}[L_1]$ and:

$$\mathbb{E}[L_1] = 1 - \Pr(\text{“no paper uses database 1”}) = 1 - \left(\frac{\bar{n}_D - 1}{\bar{n}_D}\right)^{\theta n_A}$$

We then obtain equation (11).