MACIE PAPER SERIES

Marburg Centre for Institutional Economics

MACIE

Nr. 2018/03

The Inspection Game in Science

Matthias Verbeck

MACIE, Philipps-Universität Marburg

Marburg Centre for Institutional Economics • Coordination: Prof. Dr. Elisabeth Schulte
c/o Research Group Institutional Economics • Barfuessertor 2 • D-35037 Marburg

Phone: +49 (0) 6421-28-23196 • Fax: +49 (0) 6421-28-24858 •
www.uni-marburg.de/fb02/MACIE • macie@wiwi.uni-marburg.de

Philipps Universität Marburg

The Inspection Game in Science *

Matthias Verbeck †
University of Marburg

April 26, 2018

Abstract

What are the conditions under which fraudulent or erroneous research will rise and survive in the scientific community? To answer this question, we build on the work of Lacetera and Zirulia (2011) and model the scientific approval process along the lines of an inspection game. A researcher publishes a possibly fraudulent or faulty result, which goes under the scrutiny of a (large) scientific readership. Scrutinizing scientific publications may constitute a public good for the scientific community, such that the volume of (unrevealed) faulty research can *increase* in the number of interested readers. In fact, an author might intentionally expand the level of fraud, so as to attract more readers, thereby aggravating the free rider problem and *reducing* the likelihood of getting caught.

JEL codes: A14, D82, K42, O31, Z1

Keywords: Scientific Misconduct, Reproducibility, False positives, Inspection Game, Public Goods, Economics of Science

*For very helpful comments and suggestions I thank Elisabeth Schulte, Rebecca Dietrich, the participants of the Economics department's internal research seminar of the University of Marburg, and the participants of the 2018 MAGKS research seminar in Rauischholzhausen. I also thank Guido Buenstorf and Stephan B. Bruns for having familiarized me with the work of Lacetera and Zirulia as part of a superb Ph.D course on the "Economics of Science"at the University of Kassel in 2016/17.

†Corresponding author: matthias.verbeck@wiwi.uni-marburg.de

1 Introduction

In the more recent past, several cases of flawed academic publications in highly respected journals have attracted the scientific community's and even the broader public's attention. Among the most notable ones was an article of Hwang et al. (2005), published in *Science*, in which the authors claimed to have succeeded in generating human embryonic stem cells through cloning. Several months later, after a couple of researchers had unsuccessfully tried to replicate the results, the article was retracted, and finally the findings were revealed to be fraudulent. Another much-noticed case from Economics did not involve scientific misconduct, but mere human error. "Growth in a time of debt"(Reinhart and Rogoff 2010), published in *American Economic Review: Papers & Proceedings* analyzed the connection between national debt levels and economic growth rates and concluded "for levels of external debt in excess of 90 percent"GDP growth was "roughly cut in half". Their article was widely-cited and provided a rationale for austerity measures for debt-ridden economies. In 2013 however, graduate student Thomas Herndon discovered that the reported effect size was highly exaggerated and the reason for this was as trivial as an improperly handled Excel-sheet (Herndon et al. 2014).

A further important reason why scientific publications might contain defective results is not due to fraud or error, but is rather the result of the scientific journal's prevailing selection process that favors the publicationm of statistically significant results. In his seminal paper on reporting bias, Ioannidis (2005) has argued, that the (vast) majority of claimed research findings is false, at least in the field of medical or medical related research. Recent studies (e.g. Baker 2016) confirm the existence of a "replication crisis", which indicates that trust in published results is unfortunately less justified than scientists would like to be the case. In summary it can be said, therefore, that a substantial number of papers manages to clear the hurdle of peer-review and get published despite they contain false, fraudulent or at least non-reproducible findings.

The crucial question then is whether academia succeeds to weed out such false results over time. In other words: Do the wrong results that made it into academic journals also finally find their way into academic textbooks? In fact, academia is a realm characterized by a high degree of autonomy and a rather low level of external intervention. The role of the individual researcher is therefore complex, as s/he is contributor, competitor and supervisor at the same time. Despite this complexity, the individual aspiration for reputation and the resulting competetive pressure among peers is often considered as sufficient to exterminate wrong or deficient findings over time (e.g. Merton, 1973). The model presented here aims to further clarify if and when this notion is justified.

Among the first theoretical analyses of scientific miscondiuct is the enlightening model of Lacetera and Zirulia (2011) (henceforth "L&Z"). Their findings include but are not limited to the following:

- Cases of *detetcted* fraudulent research are not representative of the *overall* amount of bogus research, since less innovative research papers are not likely to be scrutinized at all.
- A reduction of the individual costs for checking scientific results does not necessarily lead to an increase of detected fraudulent research.
- A high pressure to publish meaningful results may decrease (and not increase) scientific misconduct, since peers will then check results with increased probability.

While their analysis is extremely insightful, it leaves the (crucial) role of audience size and structure unmodeled. In their model, the scientific audience is assumed to consist of a single reader who may or may not check a published article for soundness. This is of course an extreme simplification, since a typical scientific publication will attract a wider readership, and will especially do so when the published result is of greater importance. Without a thorough analysis it is unclear however, if and how the existence of multiple readers affects the volume of (undetected) flawed findings. A cursory view inspired by the economic theory of crime (Becker 1968) suggests, that larger audiences will unambigiously help to keep science clean. If n readers check a given article, and disclose any flaw or fraud to the scientific community which they find - if present - with fixed individual probability $\tau \in (0, 1)$, then the overall probability of detection is $1 - (1 - \tau)^n$ and therefore strictly increases in n. This result would be a pleasant one, since it would mean that highly influential articles (those with many readers) survive in the scientific community, if and only if the published findings are valid. However, as the critique of Tsebelis (1989, 1990) has shown, a norm-enforcing authority should not be modeled as a fixed probability distribution, but rather as a rational player. If scrutinizing an article for flaws or fraud involves a cost for the reader, then - given the presence of a multitude of readers - this activity constitutes a public good for the scientific community and incenctives to free-ride on the efforts of peers must be taken into consideration. In fact also L&Z speculate that free riding could affect individual behavior when a multitude of readers is assumed (p. 594). Our analysis explicitly adresses this issue and can therefore be regarded as the conflation of a public good game and an inspection game.

Moreover, we use their framework to also analyze the prevalence of erronious or non-reproducible (but not fraudulent) findings in academia. The crucial action on the part of the researcher is his or her exerted effort level. A scientist who has invested a fair amount of time double-checking her results is less likely to unwittingly submit a flawed paper than her less diligent colleague. Or else, a researcher who makes a greater effort in data collection is less prone to spuriously produce a false positive result than her colleague who uses a smaller sample size.

The model presented here therefore allows to analyze both, fraudulent and flawed research in a common setting that sheds light upon the prevalence on problematic research findings in academia. It is mainly meant to capture aspects of empirical academic research, but in principle it is also possible to apply it to any kind of science where errors or deception can occur, e.g.

mathematics or theoretical physics.

The remainder of the paper is structured as follows: Section 2 provides an overview about the literature, related to our research. Section 3 contains the basic model. The publication process is modeled as an extensive-form game under incomplete information, where we first analyze the case of fraudulent research (the "deception game"), before we treat the problem of erroneous research (the "delusion game"). Section 4 extends the derived results in numerous ways. Most imporantly, we analyze how players' behavior is affected by competition among readers, the possibility of multiple audiences, heterogeneity among readers and the existence of an editor, who might also wish to check articles. Section 5 discusses our findings, and section 6 concludes. Detailed proofs can be found in the Appendix.

2 Related Literature

As stated above, our work is mainly an extension of the model of Lacetera and Zirulia (2011). We make extensive use of their framework and also adopt most of their notation. The crucial difference between their model and ours is that we do not limit the number of readers to 1 and explicitely address the positive externaility that any scrutinizing reader creates for the entire scientific community. Moreover, while their model mainly focuses on different types of research and their respective vulnerability for (undetected) fraud, we are interested in how size and structure of the academic readership influences the volume and persistence of fraudulent publications. Our work is also closely related to a follow-up paper by Kiri, Lacetera and Zirulia (2015). In this work, the authors explicitely model a researcher's effort decision when a colleague might wish to scrutinize a publication. In an extension of their model, they increase the number of peers to two, and show that this increase can reduce an author's incentive to strive for high quality research. Moreover, the overall probability of detecting deficient findings decreases. Even if their findings might seem similar to ours at first glance, the underlying mechanism is different and free riding is not considered in neither of the two papers. As we proceed, we will continually highlight the differences between our and their models, and discuss them again more closely in section 5.

Altogether, our paper is related to the theoretical literature that is concerned with questionable or fraudulent research practices, as well as the (problems of the) academic publication process. A still very readable overview of different forms and shades of academic misdemeanor is offered in LaFollette (1992). Wible (1998) provides a first formal analysis of the academic publication process that is situated in decision theory, rather than in game theory. Among others, Ioannidis (2005, 2012) and Bettis (2012) argue very forcefuly that a vast fraction of published research articles will contain false positive results which might let go unchallenged. Bobtcheff et al. (2017) present a formal analysis of academic publishing and shows how the researchers' striving for priority can undermine their incentives to care for quality. McElreath and Smaldino (2015) and Nissen et al. (2016) model the academic approval process and anaylze conditions under which incorrect claims will falsely be adopted from the scientific community. Gall and Maniandis (2015) provide a model in which competing authors can choose between different levels of transgression, such as omission of data as opposed to overt data fabrication. They find that policies that aim to prevent mild forms of misdemeanor are also suited to prevent more severe forms of scientific misconduct, but not vice versa. Furthermore, our work is related to the class of "persuasion games"that analyze the difficulties in honestly transfering scientific findings between asymmetricaly informed parties (e.g. Felgenhauer and Schulte 2014, Henry and Ottaviani 2017, Di Tillio et al. 2017).

Moreover, our model features characteristics of a typical "inspection game"(Tsebelis 1989, Andreozzi 2004). In its most simple version, a potential wrongdoer can either act in a way preferred by the inspector (e.g. working hard), or else act counter to the inspector's wishes (e.g. shirking hard work). The inspector, for his part, either does or does not engage in costly monitoring, and will discover misbehavior only if he decided to monitor. Typically this game has only one equilibrium in mixed strategies, in which the inspector only sometimes checks the agent, who in turn cheats with positive probability. Interestingly, a higher level of punishment for the perpetrator does not affect the likelihood of cheating, but instead reduces the inspector's incentives to engage in monitoring. In our model, each reader represents a potential "inspector"who can check the soundness of a colleague's work. Scrutinizing an author's work then constitutes a public good for the scientific community and provokes a "volunteer's dilemma"(Diekmann 1985) among all colleagues. Such a dilemma is characterized by the fact that the provision of a (public) good only takes place with certainty, as long as the number of potential contributors is restricted to 1. Once there exist a multiplicity of potential contributors, in the symmetric equilibrium, every player provides the public good with a probability scrictly smaller than 1. Hence a diffusion of responsibility takes place and the provision of the public good could completely fail. We are explicitely interested in how far this free rider problem affects a rational author's incentives to cheat in the first place.

Empirical works that deal with unreplicable, questionable or fraudulent research are abundant and contributions cited here are only exemplary. The problem of non-reproducable research is especially well documented in medicine (e.g. Begley and Lee (2012)) and psychology (e.g. Simmons et al. (2011), Wagenmakers et al. (2011)), but recent studies call into question also the replicability of other disciplines, such as (experimental) economics (Camerer et al., (2016), Brodeur et al. (2016)) and management science (Goldfarb and King (2016)). Bruns et al. (2017) find evidence for errors and biases in reported significance levels in innovation research. Necker (2014) provides a survey, conducted among members of the European Economic Association, a non-negligible fraction of which admitted to have already engaged in questionable research practices. Furman and co-authors (2012) show that after an academic publication gets retracted, it will be much less cited in the future, therefore supporting the idea that word spreads fast among the scientific community. Azoulay et al. (2015) show that fraudulent articles may contaminate whole fields of research and are therefore suited to shift future research activities.

3 The Model

3.1 Fraudulent Research: The Deception Game

In the model there are two types of players: a (male) author (A) who produces a scientific article and a (possibly large) readership[1], consisting of $n \geq 2$ (female) readers (R) who can scrutinize a published article.

The game consits of three stages. At stage 1, nature decides the researcher's output level. The researcher's output Y is modeled as binary, where a success is labeled as S, and a failure is denoted as F. The probability for success is denoted by β, whereas $1-\beta$ defines the probability for failure. Although not explicitely modeled here, we can understand this probability of succcess as the result of some positive effort level, that the researcher found optimal to invest at some earlier (unmodeled) stage. The probability of success is common knowledge to all players. The *realized* output level however is private information to A and cannot be directly observed by any of the other players. At stage 2, A decides whether and how to present the research output to the scientific community. The set of actions depends on the observed output level Y. For output level S, A can decide to publish (*pub* S) or not to publish (*no pub* S) the article. For output level F, however, a publication of the resulting article - in the current form - is not possible. Hence, A might decide to pretty up the results. This means that the true type of output Y and the announced type of output $\widehat{Y}$ can possibly differ. He therefore chooses between (*pub* $\widetilde{S}$) which he plays with probability p, and (*no pub* F) that is played with $(1-p)$. The former action refers to the practice of "overselling"a result, i.e. to make it appear like a success, although in fact it is none. Naturally, we assume that this behavior is not in accordance with the scientific community's code of conduct. We can think of actions like deliberately applying inappropriate statistical methods, unjust deletion of outliers, outright fabrication of data or any other wrongful behavior that yields a higher level of statistical significance or bestows the work an unduly amount of scientific recognition. If A decides to publish the fraudulent article, the game enters stage 3. At this stage, every reader i simultaneously chooses either to check (*check*) the article with individual probability q_i, or not to check (*no check*) the article with probability $1-q_i$. If a reader decides to check an article, she detects scientific misconduct with probability $\tau > 0$.[2] We assume, that as soon as at least one researcher detects fraud, word spreads within the scientific community and the fact becomes common knowledge. This assumption seems justified, when the checking reader informs the journal editor, who then would usually retract the article. Or the reader writes an own article that makes the author's misdemeanor public. Like in L&Z, the *check*-action describes different behaviors, e.g. spot-checking statistical figures for obvious

[1] We will use the terms audiences and readerships interchangeably.

[2] This assumption differs from L&Z, who assume that a check will uncover fraudulent behavior with certainty.

inconsistencies, but also replicating the results with the auhor's original data, or conducting an experiment similar to that of the author's etc.

Next we describe the payoff structure. If A publishes a non-fraudulent article, he receives a benefit $B > 0$. This benefit represents gains in reputation, advanced career opportunities and the like. If instead, the author publishes a faked article that remains unchallenged, she obtains a benefit B' and we assume that $B \geq B' > 0$. Should the author decide to publish no article at all, he reiceves a zero payoff. The assumption that $B' > 0$ implies that the author's motivation to publish articles is mainly driven by career concerns rather than by promoting the state of the art.

The publication of a research article also generates a payoff for every reader. In contrast to L&Z, we assume that this payment is state-dependent and let W denote any reader's payoff in case the author's contribution is valid, whereas W' denotes the payoff when the published results are fake. We assume that $W \geq W'$. Moreover, W can take any value in $\mathbb{R}$, whereas $W' < 0$. Whether W is positive or not generally hinges on the readers' perception of the published result. If a reader perceives the work as complementary to own prior work, or if the published result simply supports a view which the reader approves, W will be positive. If the contribution is regarded as a substitute to own prior or future research however, W will be negative. Hence a publication that yields negative values for W can be considered as a finding, that is at odds with the audience's preferences and contradicting esixting theories. High values for W rather represent results, that fit well into existing research, do not limit the readers' room for own contributions and teach the reader something that could be of interest for her own research. Our assumptions that $W' < 0$ and $W' < W$ can be motivated by the fact that most (honest) researchers - as people in general - do not like to be cheated and clearly prefer both, no publication and a sound publication to any false publication. In addition, the rejection of faked results can also be justified with the reader's position as a producer of new research. Fraudulent results can certainly mislead a reader to pursue wrong or unpromising paths in her future research and therefore have a negative effect on a reader's utility.

If any R decides to check a published result, she has to bear cost $k > 0$.[3] If a reader checks an article and finds scientific misconduct, she receives a benefit $E(G) > 0$. In general this benefit is likely to depend on the number of other researchers who also managed to successfuly discover the fraud. A single reader's gain (in reputation) might be smaller, when further colleagues successfully uncover a wrongdoer. This assumption can be justified with the priority principle in science. Only if a reader succeeds to be the first one to show the invalidity of a previously accepted result, she will gain in reputation, whereas otherwise she will usually come away empty-handed. We will explicitely deal with this issue in section 4.1, where we will model the competition among readers in the form of a contest. In the following baseline model how-

[3]This cost can be understood as the obvious cost of data collection or conducting an experiment, but also involve the reader's opportunity cost of not doing original research instead of reviewing already existing research.

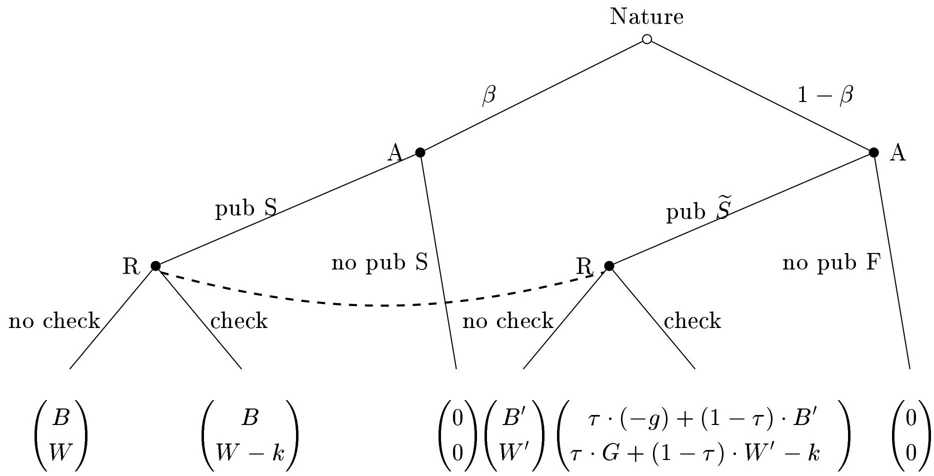

Figure 1: The deception game under the simplified assumption of $n = 1$ reader

ever, we make the simplifying assumption, that $E(G) = G$ and therefore is independent from the (expected) number of successful readers. The list of payoff parameters is completed with the cost that the author experiences when caught cheating (i.e. the loss of reputation, monetary fines, etc.). We denote this cost by $g > 0$. Should A prefer to cheat when his research project remaines fruitless, his expected payoff equals

$$\beta \cdot B + (1-\beta) \cdot \left(\left(\prod_{i=1}^{n} (1 - q_i \cdot \tau) \right) \cdot B' + \left(1 - \left(\prod_{i=1}^{n} (1 - q_i \cdot \tau) \right) \right) \cdot (-g) \right). \tag{1}$$

Likewise, the total expected utility of any reader, who is willing to check a publication is

$$\mu \cdot W + (1-\mu) \cdot \left(\left(\prod_{i=1}^{n} (1 - q_i \cdot \tau) \right) \cdot W' + \tau \cdot G \right) - k \tag{2}$$

and μ denotes the readers' updated probability on the author's likelihood for success if an article has been released.

Unlike in the contributions of L&Z and Kiri et al. (2015), a checking reader creates a positive externality for the whole scientific community. With each additional scrutinizing colleague, the expected gain of any reader who will check a publication herself, will decrease. We add one final assumption, that restricts our analysis to cases, where this externality is large enough, such that not *all* readers strictly prefer to check. This is guaranteed if

$$\begin{aligned} \beta \cdot W + (1-\beta) \cdot (1-\tau)^{n-1} \cdot W' &> \\ \beta \cdot W + (1-\beta) \cdot \left(\tau \cdot G + (1-\tau)^{n} \cdot W' \right) - k & \\ \Leftrightarrow G < (1-\tau)^{n-1} \cdot W' + \frac{k}{(1-\beta) \cdot \tau}&. \end{aligned} \tag{3}$$

Figure 1 shows the game tree under the simplifying assumption that there is only one reader.

The appropriate solution concept for the presented game is that of a perfect Bayesian equilibrium. We solve for all *symmetric* perfect Bayesian equilibria and distinguish between equilibria in pure and mixed strategies.

An equilibrium is fully characterized by (a) the author's publication decision in case of success, (b) the author's publication decision in case of failure, (c) the action chosen by any reader in case an article gets published, (d) the readers' posterior belief μ about the author's success in case of a publication.

Proposition 1. *For the deception game, evry parameter constellation yields exactly one symmetric equilibrium, such that*

1. *For* $G < W' + \frac{k}{(1-\beta)\cdot\tau}$*:* $p = 1$, $q_i = 0$, $\mu = \beta$. *The probability that a fraudulent article gets published is* $(1-\beta)$ *and the probability that a fraudulent article gets caught (if published) is 0. We call this equilibrium "Pooling I".*

2. *For* $G \geq W' + \frac{k}{(1-\beta)\cdot\tau}$ *and* $(1-\tau)^n > \frac{g}{B'+g}$*:* $p = 1$, $q_i = \left(1 - \sqrt[n-1]{\frac{1}{W'} \cdot \left(G - \frac{k}{(1-\beta)\cdot\tau}\right)}\right) \cdot \frac{1}{\tau}$, $\mu = \beta$. *The probability that a fraudulent article gets published is* $(1-\beta)$ *and the probability that a fraudulent article gets caught (if published) is* $1 - (1 - q_i \cdot \tau)^n = 1 - \left(\frac{1}{W'} \cdot \left(G - \frac{k}{(1-\beta)\cdot\tau}\right)\right)^{\frac{n}{n-1}}$. *We call this equilibrium "Pooling II".*

3. *For* $G \geq W' + \frac{k}{(1-\beta)\cdot\tau}$ *and* $(1-\tau)^n \leq \frac{g}{B'+g}$*:* $p = \frac{\beta}{1-\beta} \cdot \frac{k}{\tau \cdot \left(G - W' \cdot \left(\frac{g}{B'+g}\right)^{\frac{n-1}{n}}\right) - k}$, $q_i = \left(1 - \sqrt[n]{\frac{g}{B'+g}}\right) \cdot \frac{1}{\tau}$, $\mu = \frac{\beta}{\beta + p\cdot(1-\beta)}$. *The probability that a fraudulent article gets published is* $(1-\beta)\cdot p$ *and the probability that a fraudulent article gets caught (if published) is* $1 - (1 - q_i \cdot \tau)^n = 1 - \frac{g}{B'+g}$. *We call this equilibrium "Semi-Separation".*

In case of a success, the author will publish an article in any of the equilibria.

Proof: See Appendix.

Similar to L&Z, in equilibrium fraud will occur with positive probability. The existing equilibria can be characterized along different parameter thresholds. If $G < W' + \frac{k}{(1-\beta)\cdot\tau}$, not a single reader wants to check, as the expected profit from doing so does not cover the cost. Since all readers will abstain from checking a publication, a rational author will never stop short of scientific misconduct, as the probability to be debunked equals zero.

If $G \geq W' + \frac{k}{(1-\beta)\cdot\tau}$ the readers will check a published article with positive probability.[4] Then, the size of B' relative to g and τ determines the author's strategy. For $(1-\tau)^n > \frac{g}{B'+g}$, the author's expected punishment is not severe enough to deter him from releasing a fraudulent article and we observe pooling behavior once more ("Pooling II"). For the readers, the game

[4]The condition makes checking for at least one reader profitable.

then essentially turns into a public good game and the volume of scrutiny does not depend on the author's payoff parameters.[5] Low values for G and τ and high values for W', β and k reduce the individual probability of a reader to check a published result. Hence, if most publications can generally be trusted, checking costs are high and the readers' gains from refuting the article are limited, there might be a good chance for a cheating author to escape unscathed. When G increases (assuming that all other variables are held constant), eventually all readers want to check a publication with probability 1, that is when $G \geq W' \cdot (1-\tau)^{n-1} + \frac{k}{(1-\beta)\cdot\tau}$, a parameter constellation we ruled out by assumption. Note that for the parameter set between these two extremes (all readers want to check or no reader wants to check), there are many more asymmetric equilibria, where readers differ in their individual probability to examine a finding.

For $(1-\tau)^n \leq \frac{g}{B'+g}$, an author is kept indifferent between cheating and not cheating and both actions occur with positive probability. Hence we have a semi-separating equilibrium. For the audience, a publication is generally informative about the outcome of the research project, and an observed publication leaves any reader more confident that the author indeed has managed to obtain a success. Unlike in the second pooling equilibrium, positive probabilities for playing *check* now let the author abstain from cheating with positive probability. Therefore this equilibrium resembles more to that of the canonical inspection game, in which cheating is not a dominant strategy for the (potential) perpetrator. In particular, the author chooses to cheat more often, as k, W' and β increase and as G and τ decrease. The readers only respond to the author's payoff variables and check a publication more frequently as B' increases and g decreases. Notice that also for the case of $p \in (0,1)$, there are other asymmetric equilibria where readers check publications with dissimilar probabilities. Figure 2 illustrates the three different equilibria for varying values of G and B'.

From the author's perspective there exists a clear ordering regarding the three different types of equilibria. His expected payoff is highest in the first pooling equilibrium, second highest in the second pooling equilibrium and lowest in the semi-separating equilibrium. This ordering directly corresponds to the respective probability of being caught, which is highest in the semi-separating equilibrium and 0 in the first pooling equilibrium.

We are mostly interested in how audience size affects the share of debunked fraudulent research, as well as the total volume of fraudulent research. We obtain the counterintuitive result, that the volume of both, fraud in general and undetected fraud, will be weakly increasing in the number of readers. This is true for both kinds of equilibria, pooling equilibria and semi-separating equilibria.

Proposition 2. *In the deception game, considering symmetric equilibria, an increase in the number of readers from n to $n+1$ affects the equilibria as follows:*

[5] To be more precise, the game gets the structure of a volunteer's dilemma (Diekmann, 1985), where the public good would be provided with certainty, if there was only one potential contributor.

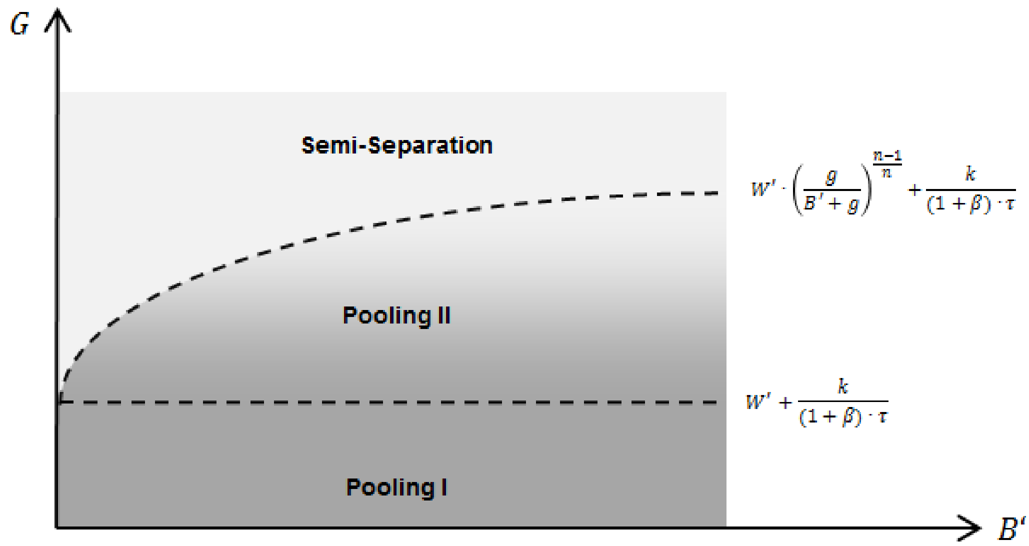

Figure 2: Resulting equilibria for different realizations of G and B'. Brighter shades refer to a higher overall probability of fraud detection, given that a publication has been released.

1. *For* $G < W' + \frac{k}{(1-\beta)\cdot\tau}$*: The absolute volume of fraud remains at* $(1-\beta)$ *and the volume of undetected fraud remains at 0.*

2. *For* $G \geq W' + \frac{k}{(1-\beta)\cdot\tau}$ *and* $(1-\tau)^n < \frac{g}{B'+g}$*: The absolute volume of fraud remains at* $(1-\beta)$ *and the level of undetected fraud increases to* $1 - \left(\frac{1}{W'} \cdot \left(G - \frac{k}{(1-\beta)\cdot\tau}\right)\right)^{\frac{n+1}{n}}$.

3. *For* $G \geq W' + \frac{k}{(1-\beta)\cdot\tau}$ *and* $(1-\tau)^n \geq \frac{g}{B'+g}$*: The absolute absolute volume of fraud increases to* $\frac{\beta}{1-\beta} \cdot \frac{k}{\tau\cdot\left(G - W'\cdot\left(\frac{g}{B'+g}\right)^{\frac{n}{n+1}}\right)-k}$ *and the volume of undetected fraud increases to* $\frac{\beta}{1-\beta} \cdot \frac{k}{\tau\cdot\left(G - W'\cdot\left(\frac{g}{B'+g}\right)^{\frac{n}{n+1}}\right)-k} \cdot \left(1 - \frac{g}{B'+g}\right)$.

4. *The parameter set for which "Pooling II"exists weakly decreases, and the parameter set for which "Semi-Separation"exists weakly increases.*

Proof: See Appendix.

Despite the fact that we have a larger supply of readers and therefore potentially a higher level of scutiny, the de facto level of checking decreases due to free riding behavior. This implies, no matter how large the readership is, the overall probability of fraud detection, will never exceed τ. As n becomes larger, the volume of misconduct remains unchanged or even increases. These results contradict common sense beliefs about the academic publication process and shows that the notion of a self-correcting scientific community may not be justified. However, caution should be exercised. A larger audience size might also imply different values for all other (payoff) parameters (see also section 4.3 in Kiri et al. (2015)). In particular, it is reasonable to assume that

B' and G are higher for larger audiences, thus encouraging the readers' scrutiny and deterring fraudulent behavior. Therefore we can only conclude that a large number of readers *alone* is not a sufficient condition for a high quality level of scientific publications.

With a higher level of n, we also observe a shift in the occuring equilibria. The set of parameter values for which the second pooling equilibrium emerges, grows at the expense of the set of parameter values that imply the semi-separating equilibrium. The validity of the first pooling equilibrium is not affected by a higher n.

What may come as a surprise is that τ does not affect the overall probability to detect flawed articles *within* the different equilibria. A lower τ is always compensated by a higher q_i, leaving the overall detection probabilities unaffected. Instead a higher value of τ leads to a downward shift of the threshold functions, depicted in Figure 2. As a consequence the set of parameter values that result in the first pooling equilibrium shrinks, whereas the set of parameter values that imply the semi-separating equilibrium will grow. It is furthermore interesting to see, that W (in contrast to W') has no influence on neither of the players' behavior.

3.2 Erroneous Research: The Delusion Game

In this section, we disregard the possibility to deceive the scientific community. Instead we make the assumption that the author is righteous, but might unwittingly produce a result that is not replicable or at least less potent than originally claimed. The reason for this can either be the author's individual negligence, or else, the result is simply a false positive one. We therefore adjust the presented game in the following way. At stage 1, nature determines the author's *observed* output level as well as the *actual* output level. With probability β, the author experiences and observes a true success. With probability α, the author erroneously observes a success, when in fact a failure has been produced, henceforth referred to as "spurious success". Then, with counter probability $1 - \beta - \alpha$, the author rightly observes a failure.[6]

To the readers, neither the observed, nor the actual output level is known. At stage 2, the author chooses his binary level of care (*no care* or *care*), and p now refers to the probability of not applying care. Investing care refers to any action that helps to rule out non-replicable or oversold results. For example, to avoid individual errors, he could consult colleagues to clarify whether a statistical method has been applied correctly or he double-checks all the data that has been processed by his student assistants. Or else, the author increases the sample size, to improve the robustness of his findings. Investing care comes at a cost $c > 0$. Should the author decide to invest care, he identifies a spurious success with certainty. If A initially observes a success, but refrains from investing care, we will refer to this as an "unsure success". At stage 3, he then decides to publish or not to publish the result, where *pub S*, *No Pub S*, *pub F* and

[6]For simplicity, we rule out type II errors, where the author observes a failure, though the result is actually a success.

no pub F denote the possible actions, in case care has been applied, whereas *pub Y* and *no pub Y* refer to the respective actions, when the author prefers not to apply care. If an article is released, the game enters stage 4, in which the audience can again decide to check or not to check the result and will find errors with individual probability τ.

The players' payoffs are determined analogously to the deception game, though other values for these parameters seem reasonable now (for example it is plausible that g, the author's utility loss in case a wrong result gets detected, is generally milder). We now assume that $B' < 0$ and $B' > -g$, i.e. a reader who knows about the article's flaws will prefer to not present it to the scientific community, and should the flaws be detected, his utility loss is larger than when his error remains unseen by the readers. We assume furthermore that the author is sufficiently optimistic about the validity of his observed success, such that he prefers to publish an unsure success (i.e when he did not apply care), rather than to abstain completely from a publication:

$$\mu_A^1 \cdot B + \left(1 - \mu_A^1\right) \cdot \left((1 - q_i \cdot \tau)^n \cdot B' + (1 - (1 - q_i \cdot \tau)^n) \cdot (-g)\right) \geq 0 \tag{4}$$

and $\mu_A^1 = \frac{\beta}{\beta+\alpha}$ denotes the author's updated probability of a true success, after having observed a success (true or spurious). Moreover, like in the deception game, we assume that not all readers strictly prefer to check. The following game tree illustrates the course of action, again under the simplified assumption of a single reader.

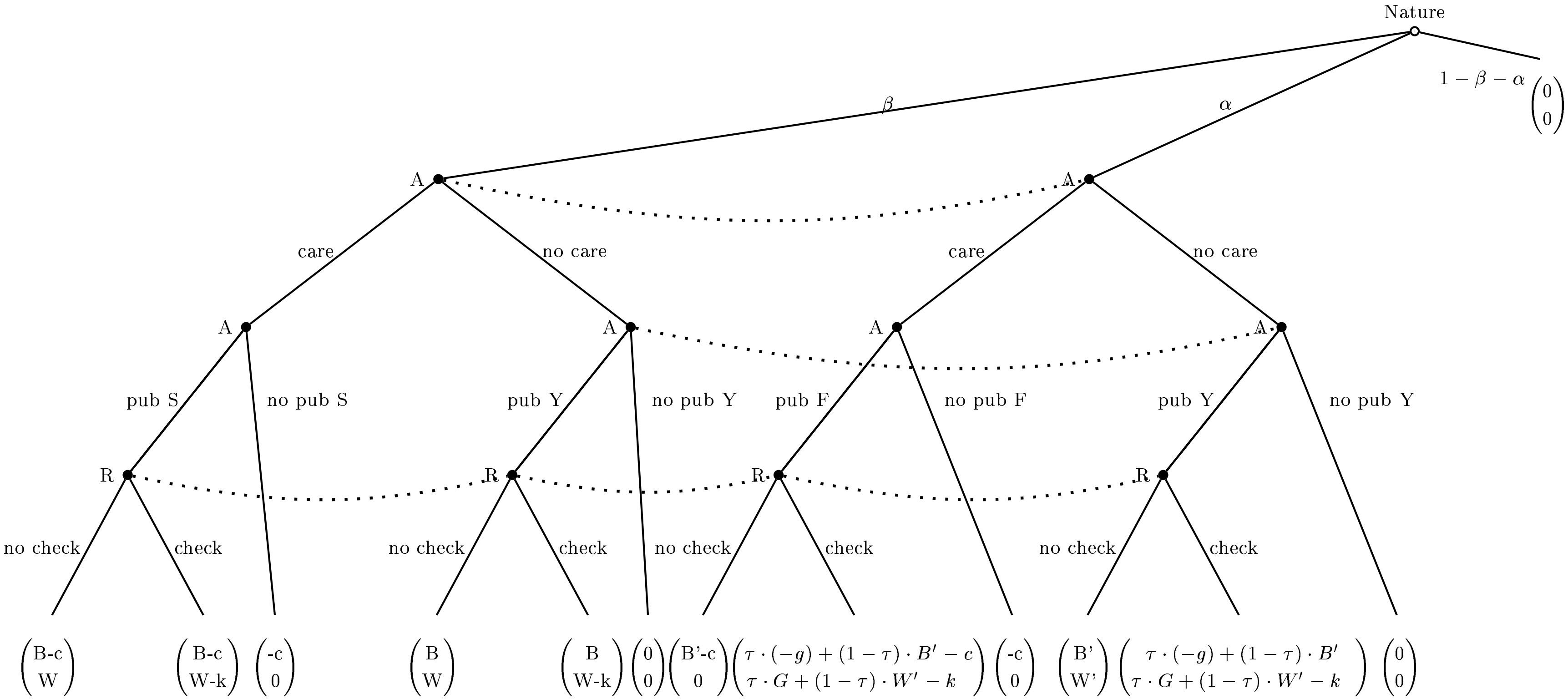

Figure 3: The delusion game under the simplified assumption of $n = 1$ reader

The game's equilibria are characterized by (a) the author's care level, (b) the author's publication decision in case he applies care and identifies a success, (c) the author's publication decision in case he applies care and identifies a failure, (d) the author's publication decision in case he applies no care, (e) the action chosen by any reader in case an article gets published, (f) the author's posterior belief μ_A^1 on the likelihood of true success upon observing a success, (g) the author's posterior belief μ_A^2 about the success after care has been applied, (h) the readers' posterior belief μ_R about the pubication's soundness if a publication is released.

Proposition 3. *For the delusion game, every parameter constellation yields exactly one symmetric equilibrium, such that*

1. *For* $B' < \frac{-c\cdot(\alpha+\beta)}{\alpha}$*:* $p = 0$, $q_i = 0$, $\mu_A^1 = \frac{\beta}{\beta+\alpha}$, $\mu_A^2 = 1$, $\mu_R = 1$. *The probability that a faulty article gets published is* 0. *We call this equilibrium "Separation".*

2. *For* $B' \geq \frac{-c\cdot(\alpha+\beta)}{\alpha}$ *and* $G < W' + \frac{k}{\left(1-\frac{\beta}{\beta+\alpha}\right)\cdot\tau}$*:* $p = 1$, $q_i = 0$, $\mu_A^1 = \frac{\beta}{\beta+\alpha}$, $\mu_A^2 = 1$, $\mu_R = \frac{\beta}{\beta+\alpha}$. *The probability that a faulty article gets published is* α *and the probability that a faulty article gets revealed (if published) is 0. We call this equilibrium "Pooling I".*

3. *For* $(1-\tau)^n > \frac{(-c+g)\cdot\alpha-c\cdot\beta}{\alpha\cdot(B'+g)}$ *and* $G < W'+\frac{k}{\left(1-\frac{\beta}{\beta+\alpha}\right)\cdot\tau}$*:* $p = 1$, $q_i = \left(1 - \sqrt[n-1]{\frac{1}{W'}\cdot\left(G-\frac{k}{\left(1-\frac{\beta}{\beta+\alpha}\right)\cdot\tau}\right)}\right)\cdot\frac{1}{\tau}$, $\mu_A^1 = \frac{\beta}{\beta+\alpha}$, $\mu_A^2 = 1$, $\mu_R = \frac{\beta}{\beta+\alpha}$. *The probability that a faulty article gets published is* α *and the probability that a faulty article gets revealed (if published) is* $1-(1-q_i\cdot\tau)^n = 1-\left(\frac{1}{W'}\cdot\left(G-\frac{k}{\left(1-\frac{\beta}{\beta+\alpha}\right)\cdot\tau}\right)\right)^{\frac{n}{n-1}}$. *We call this equilibrium "Pooling II".*

4. *For* $B' \geq \frac{-c\cdot(\alpha+\beta)}{\alpha}$, $(1-\tau)^n \leq \frac{(-c+g)\cdot\alpha-c\cdot\beta}{\alpha\cdot(B'+g)}$ *and* $G \geq W'+\frac{k}{\left(1-\frac{\beta}{\beta+\alpha}\right)\cdot\tau}$*:* $p = \frac{\beta}{\alpha}\cdot\frac{k}{\tau\cdot\left(G-\left(\frac{(-c+g)\cdot\alpha-c\cdot\beta}{\alpha\cdot(B'+g)}\right)^{\frac{n-1}{n}}\cdot W'\right)-k}$, $q_i = \left(1-\sqrt[n]{\frac{(-c+g)\cdot\alpha-c\cdot\beta}{\alpha\cdot(B'+g)}}\right)\cdot\frac{1}{\tau}$, $\mu_A^1 = \frac{\beta}{\beta+\alpha}$, $\mu_A^2 = 1$, $\mu_R = \frac{\beta}{\beta+\alpha\cdot p}$.
The probability that a faulty article gets published is $\alpha\cdot p$ *and the probability that a faulty article gets detected (if published) is* $1-(1-q_i\cdot\tau)^n = 1-\frac{(-c+g)\cdot\alpha-c\cdot\beta}{\alpha\cdot(B'+g)}$. *We call this equilibrium "Semi-Separation".*

In any of the equilibria, the author will always publish a sure success and an unsure success, but will never publish a sure failure.

Proof: See Appendix.

In contrast to the deception game, we obtain a full separation equilibrium in which no false results are published and therefore no reader ever wants to check a publication. This equilibrium occurs, whenever the cost of investing care are sufficiently small as compared to the cost

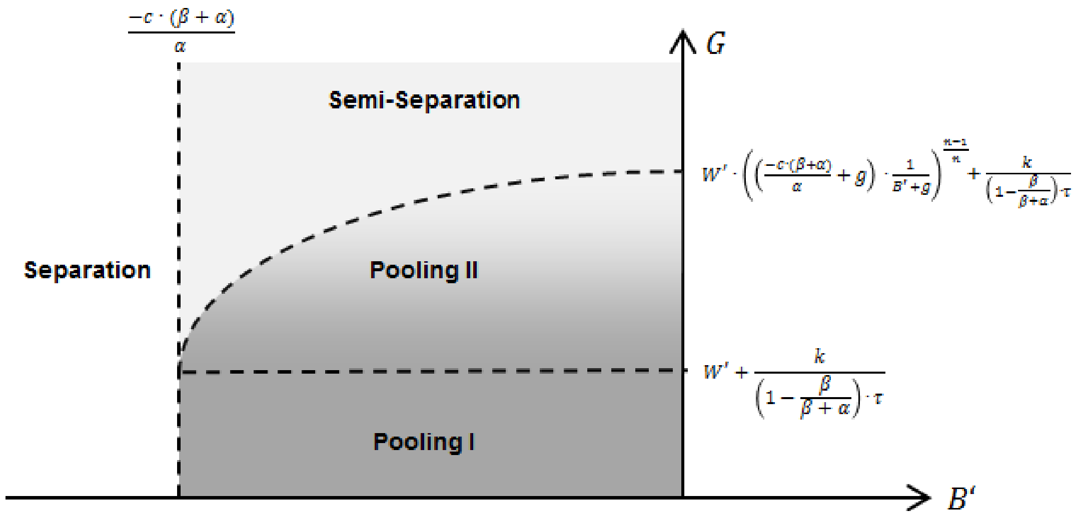

Figure 4: Resulting equilibria for different realizations of G and B'. Brighter shades refer to a higher overall probability of error detection.

of mistakenly publishing an unsound finding (even if not detected), that is when $B' < \frac{-c\cdot(\alpha+\beta)}{\alpha}$. If this condition fails to hold however, one of the remaining equilibria, all of which are qualitatively comparable to the deception game, will occur. Figure 4 illustrates all possible symmetric equilibria of the delusion game. The different (payoff) parameters have a similar effect on the occurance of equilibria as in the deception game and the above reasoning applies mutatis mutandis. Most importantly, Proposition 2 also applies for the delusion game, such that a larger readership entails a lower overall level of error detection in “Pooling II”and “Semi-Separation”equilibria, and a lower level of care in the semi-separating equilibrium.

4 Extensions and Applications

4.1 The Priority Principle in Science

As mentioned earlier, the notion that a reader’s (private) benefit from successfully debunking fraudulent research is independent from the number of other (successful) readers, could be considered as unrealistic. In academic research, priority matters (e.g. Dasgupta and David, 1994). If more than one reader writes an article about his findings, there is always the danger that a fellow reader will outpace her and own scrutiny will be worthless in hindsight. For the deception game, we therefore include this competition among readers and assume that a reader’s expected gain $E(G)$ is a function of the (expected) number of successful readers.[7] Since more general results are not obtainable, we analyze the case of $n = 2$. The expected gain from finding misbehavior equals

[7]For the delusion game, qualitatively similar results to those presented here can be obtained.

$$E(G) = \frac{G}{1 + (n-1) \cdot q_i \cdot \tau} \Leftrightarrow \frac{G}{1 + q_i \cdot \tau}. \tag{5}$$

Evidently, we assume that any of the successful readers gets only a fraction of the reputational gain. This assumption allows for two natural interpretations. Either any reader must manage to be the first of all successful peers to reveal the fraud, and only then she will get the entire gain. Or else we could assume that all successful readers will write a joint article, but the reputational gain must be shared equally among all successful peers. We compare the resulting equilibria to those obtained for the original game. In line with Kiri et al. (2015), we come to the conclusion that the contest among readers further weakens the average quality of research articles. More precisely, for a given G, the number of fraudulent articles and the share of overlooked fraudulent research will both weakly increase compared to the non-competitive setting.

Proposition 4. *In the deception game with $n = 2$ and competition among readers, considering symmetric equilibria*

1. *"Pooling II"emerges for a broader range of parameter values, while "Semi-Separation"emerges for a lower range of parameter values*
2. *the volume of undetetcted fraud is unaffected in the "Pooling I"and increases in "Pooling II"*
3. *the volume of overall and undetetcted fraud increases in "Semi-Separation"*

as compared to the non-competitive setting.

Proof: See Appendix.

It is most convenient to anaylze the results by reference to Figure 2: The parameter set for which the first pooling equilibrium arises remains unaffected and no fraudulent piece of research will ever be revealed. In the second pooling equilibrium, readers will check publications with individual probability $\widehat{q_i} = \frac{\frac{k}{1-\beta} - \sqrt{4 \cdot \tau \cdot W' \cdot \left(\tau \cdot W' + \frac{k}{1-\beta} - \tau \cdot G\right) + \left(\frac{k}{1-\beta}\right)^2}}{2 \cdot \tau^2 \cdot W'}$ which is lower than in the non-competitive setting. Therefore also the overall detection proability will be smaller. The set of parameter values, for which the second pooling equilibrium emerges will expand at the expense of values for which the semi-separating equilibrium will come up. Here the individual checking probability will remain the same as in the non-competitive setting, such that the author is still indifferent between fraud and compliance. The author will increase his volume of cheating, as through competition, the readers' expected gain becomes smaller. The reader's indifference can only be regained by commiting fraud more often. The adjusted level of cheating then equals $\widehat{p} = \frac{\beta}{1-\beta} \cdot \frac{k}{\tau \cdot \left(\frac{G}{1+\widehat{q}_i \cdot \tau} - W' \cdot \left(\frac{g}{B'+g}\right)^{\frac{n-1}{n}}\right) - k}$.

4.2 Strategic Audience Choice

A rational (and malevolent) author will take into consideration that for larger audience sizes the free rider problem is more pronounced. Therefore, all other parameters held constant, he will always weakly prefer a huge audience over a small one. In fact, as we will show in this section, the author can even have incentives to *strategically induce* the free rider problem, i.e. to take measures that are suited to increase the number of interested readers. The most obvious way to attract a higher number of readers to a scientific article is by offering a more interesting or spectacular result. If we allow for the possibility of more than two output levels, then this can have quite severe implications for the level of undetected fraud.

In the following we slightly modify the deception game at stage 1 and allow for a third outcome L, that represents a landmark result which is generally suited to arouse the interest of a larger audience. Such a breakthrough occurs with probability β^L, whereas a success occurs with probability β^S. At stage 2, if A experiences a failure, he can now choose between three options. Besides (*no pub* F) and (*pub* $\widetilde{S}$), he can also choose to oversell the failure as a landmark result (*pub* $\widetilde{L}$).[8] We will refer to the different transgression levels as "mild cheating"and "heavy cheating".

The audience consists of two sub-audiences, x and y and $n_x + n_y = n$. For a member of audience x, both output levels are perceived as equally valuable, and we have $W_x'^L = W_x'^S$ and $G_x'^L = G_x'^S$. In the absence of further checking readers, a member of audience x wants to check an article of type Y if

$$G_x^Y \geq W_x'^Y + \frac{k}{(1 - \beta^L - \beta^S) \cdot \tau} \tag{6}$$

is satisfied for both output levels. We assume hat this codition always holds. For a member of audience y, only L-results are of interest. Therefore we assume that

$$G_y^Y \geq W_y'^Y + \frac{k}{(1 - \beta^L - \beta^S) \cdot \tau} \tag{7}$$

is only satisfied for $Y = L$, but is not satisfied for $Y = S$. For simplicity we furthermore assume that L-results are perceived as equally by both audiences and we have $W_x'^L = W_y'^S$ and $G_x^L = G_y^S$. For the author's payoff parameters it is straightforward to assume that $B'^L \geq B'^S$ and $g^L \geq g^S$. Alltogether, our assumptions reflect the fact that some results are only of interest for a small (specialized) audience and that the author's reward is weakly higher if he succeeds to produce a result that is of more general interest.

In the following proposition we do not fully characterize all possible equilibria. Instead we content ourselves with showing the following results:

Proposition 5. *In the deception game with two transgression levels, "mild cheating" and "heavy cheating" and a heterogeneous composition of readers, there exist equilibria such that*

[8] We exclude the possibility of overselling a success as a landmark result.

1. *the author, if failing, cheats mildly with certainty*
2. *the author, if failing, cheats heavily with certainty*
3. *the author, if failing, cheats heavily or does not publish an article, both with positive probability.*

Should the author play heavy cheating with certainty, his risk of getting caught is always lower compared to playing mild cheating.

Proof: See Appendix.

From a perspective of optimal incentive design, these results are somewhat discouraging. By commiting a more severe offense (heavy cheating), the perpetrator can make himself better off and will actually *reduce* the probability of getting caught. Still an author might prefer a milder transgression level, when a more severe punishment for heavy cheating offsets the reduced likelihood of getting caught.

The above results are also interesting in the light of the findings of Furman et al. (2012) and their discussion in L&Z. For the case of biomedicine, the former authors find that it is mostly highly influential research that is retracted after publication and retractions occur relatively rarely in low-profile research. L&Z speculate that this finding can be explained with a reader's low reward of refuting "incremental"research, which is therefore not scrutinized at all. In line with the above findings, our alternative explanation for this phenomenon would be that a researcher does only commit fraud, when doing so earns him a high-profile publication and therefore attracts many readers. We therefore might not see much low-profile fraudulent research, not because it remains undetected, but because it rarely exists.

5 Discussion

In this section we want to briefly discuss the above findings. First, it is worthwhile to highlight the differences of our model to those of L&Z and Kiri et al. (2015). While L&Z mainly focus on different types of research (incremental vs. radical) and their respective odds of being fraudulently produced (and to be revealed as such), we disregard this distinction of research types and instead focus on the scientific community and its role in debunking deficient publications. Kiri et al. (2015) concentrate on the author's motivation to invest costly effort that positively affects the chances to produce high-quality research when the resulting article possibly undergoes a check by a single colleague. In an extension of their model, they introduce a second reader who can also check a publication's validity. Similar to us, they find that the overall probability to debunk low quality research can decrease by introducing an additional reader. However, the mechanism at work is completely different from ours, since their results are solely driven by the readers' quest for priority. Our main contribution is therefore to show that a volunteer's dilemma exists among members of the scientific community and to analyze how this dilemma in turn influences the author's willingness to cheat or to apply an inadequate level of diligence.

Our central finding is certainly Proposition 2 in which we show that an increasing number of readers is possibly detrimental to the average quality of scientific publications. This is clearly a counter-intuitive finding. What can positively affect the volume of detected flawed research though, is a high level of G (or equivalently a low level of W'). This shows that for the scientific community to be self-correcting, it is crucial that (at least a few) readers have some (ideological) distance to the author's presented findings. Our result herefore suggests that a fair share of devil's advocates", i.e. readers who are commited to a different theory or paradigm, is certainly helpful for reducing the volume of flawed publications. In an updated version of this work, we plan to address the issue of reader heterogeneity more explicitely. Certainly our central result can be criticized for several reasons. First, as already stated, the payoff parameters of all players are certainly not independent from the number of readers. Second our model does not explicitely address the role of follow-up research. More interesting findings (those with many readers) are more likely to spur future research activities, which could be helpful for refuting unsound articles. Third, the readers individual probability of finding errors might be not independent from each other, but (negatively) correlated. We also do not incorporate the process of peer-review in our model. In an updated version of this paper, we plan to analylze the role of an editor, who can check an article before being published.

There are some avenues for future research. First, it is not entirely clear, to which degree our results hold more generally, when we allow for richer action spaces, e.g. a continuum of checking levels or cheating levels. Second, as a possible extension, one could explicitely analyze the competition among authors, who compete for scarce journal space (similar to Gall Maniandis, 2015). The implications might be different from those of L&Z, who model a harsher publish or perish paradigm simply by a higher individual publication benefit. Third, our finding that deception possibilities increase with group size might be relevant in contexts other than academic publishing.[9]

6 Conclusion

We have presented a model of the scientific approval process when scrutinizing scientific publications is individually costly and causes a positive externality for the whole scientific community. In the model's basic version, an author can decide to publish a fraudulent article, should a research project turn out as a failure. Contrary to the intuitive view that a higher number of readers should more effectively deter authors from behaving fraudulently and also should increase the number of revealed cases of fraud, we find that the contrary might be true, de-

[9] Think of a politician who wants to cheat a large electorate, or an agent who wants to deceive a collective of principals.

pending on parameter size. The effect is due to the readers' individual free riding behavior that in turn affects the authors willingness to cheat. Therefore our model challenges the notion of self-correcting science. In an adjusted version of the model, we have analyzed the case of erroneous research, where a flawed publication is the result of lacking diligence, instead of deliberate fraud. Likewise, increasing readership size might be detrimental rather than helpful to reduce and uncover deficient publications. When incorporating the competition among the authors for priority, the level of (undetected) defective research might further increase. If we explicitely consider the possibility of two transgressions levels (mild and severe misconduct), it turns out that an author who decides for the more severe transgression level can actually reduce his risk of getting caught, because the free rider problem is more pronounced for the case of severe misconduct.

References

Andreozzi, L. (2004). Rewarding policemen increases crime: Another surprising result from the inspection game. *Public Choice 121*(1/2), 69–82.

Azoulay, P., J. Furman, J. Krieger, and F. Murray (2015). Retractions. *Review of Economics and Statistics 97*(5), 1118–1136.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature 533*(7604), 452–454.

Becker, G. (1968). Crime and punishment: An economic approach. *Journal of Political Economy 76*(2), 169–217.

Begley, C. and M. Lee (2012). Drug development: Raise standards for preclinical cancer research. *Nature 483*(7391), 531–533.

Bettis, R. (2012). The search for asterisks: Compromised statistical tests and flawed theories. *Strategic Management Journal 33*(1), 108–113.

Bobtcheff, C., B. J., and T. Mariotti (2017). Researcher's dilemma. *The Review of Economic Studies 84*(3), 969–1014.

Brodeur, A., M. Sangnier, and Y. Zylberberg (2016). Star wars: The empirics strike back. *American Economic Journal: Applied Economics 8*(1), 1–32.

Bruns, S. et al. (2017). Errors and biases in reported significance levels: Evidence from innovation research. *Working paper presented at DRUID 2017, available at http://conference.druid.dk/acc_papers/nuefo9ll2m6v0e0xca1qamg58nsm.pdf*.

Camerer, C. et al. (2016). Evaluating replicability of laboratory experiments in economics. *Science 351*(6280), 1433–1436.

Dasgupta, P. and P. David (1994). Toward a new economics of science. *Research Policy 23*(5), 487–521.

Di Tillio, A., M. Ottaviani, and P. Sørensen (2017). Persuasion bias in science: Can economics help? *The Economic Journal 127*, F266–F304.

Diekmann, A. (1985). Volunteer's dilemma. *Journal of Conflict Resolution 29*(4), 605 – 610.

Felgenhauer, M. and E. Schulte (2014). Strategic private experimentation. *American Economic Journal: Microeconomics 6*(4), 74–105.

Furman, J., K. Jensen, and F. Murray (2012). Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine. *Research Policy 41*(2), 1747–1759.

Gall, T. and Z. Maniadis (2015). Evaluating solutions to the problem of false positives. *University of Southampton Discussion Paper in Economics and Econometrics 1504*.

Goldfarb, B. and A. King (2016). Scientific apophenia in strategic management research: Significance tests & mistaken inference. *Strategic Management Journal 37*(1), 167–176.

Henry, E. and M. Ottaviani (2017). Research and the approval process: The organization of persuasion. *CEPR Discussion Paper No. DP11939*.

Herndon, T., M. Ash, and R. Pollin (2014). Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff. *Cambridge Journal of Economics 38*(2), 257–279.

Hwang, W. et al. (2005). Patient-specific embryonic stem cells derived from human SCNT blastocysts. *Science 308*(5729), 1777–1783.

Ioannidis, J. (2005). Why most published research findings are false. *PLoS Medicine 2*(8), e124.

Ioannidis, J. (2012). Why science is not necessarily self-correcting. *Perspectives on Psychological Science 7*(6), 645–654.

Kiri, B., N. Lacetera, and L. Zirulia (2015). Above a swamp: A theory of high-quality scientific production. *NBER Working Paper*.

Lacetera, N. and L. Zirulia (2011). The economics of scientific misconduct. *The Journal of Law, Economics, & Organization 27*(1), 568–603.

LaFollette, M. (1992). *Stealing into Print: Fraud, Plagarism, and Misconduct in Scientific Publishing*. University of California Press.

McElreath, R. and P. Smaldino (2015). Replication, communication, and the population dynamics of scientific discovery. *PLoS One 10*(8), e0136088.

Merton, R. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. The University of Chicago Press.

Necker, S. (2014). Scientific misbehavior in economics. *Research Policy 43*(10), 1747–1759.

Nissen, S., T. Magidson, K. Gross, and C. Bergstrom (2016). Publication bias and the canonization of false facts. *eLife 5*, e21451.

Reinhart, C. and K. Rogoff (2010). Growth in a time of debt. *American Economic Review: Papers & Proceedings 100*(2), 573–578.

Simmons, J., L. Nelson, and U. Simonsohn (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science 22*(11), 1359–1366.

Tsebelis, G. (1989). The abuse of probability in political analysis: The Robinson Crusoe fallacy. *American Political Science Review 83*(1), 77–91.

Tsebelis, G. (1990). Penalty has no impact on crime: A game-theoretic analysis. *Rationality and Society 2*(3), 255–286.

Wagenmakers, E., R. Wetzels, D. Borsboom, and H. van der Maas (2011). Why psychologists must change the way they analyze their data: The case of PSI: Comment on BEM (2011). *Journal of Personality and Social Psychology 100*(3), 426 432.

Wible, J. (1998). *The Economics of Science: Methodology and Epistemology as if Economics Really Mattered*. Routledge.

Appendix: Proofs

Proof of Proposition 1:
We start with showing that for all possible equilibria, the author will always publish an article, if he obtained a success. Publishing a successful project is superior to not publishing if

$$(1-(1-q_i\cdot\tau)^n)\cdot B+(1-q_i\cdot\tau)^n\cdot B\geq 0\Leftrightarrow B\geq 0. \tag{8}$$

This is true by assumption.

The condition that makes any reader prefer to not check a publication is

$$\begin{aligned}\mu\cdot W+(1-\mu)\cdot\Big(\tau\cdot G+(1-\tau)\cdot(1-q_i\cdot\tau)^{n-1}\cdot W'\Big)-k<\\ \mu\cdot W+(1-\mu)\cdot(1-q_i\cdot\tau)^{n-1}\cdot W'\end{aligned} \tag{9}$$

and $\mu=P(S|article)=\frac{P(article|S)\cdot P(S)}{P(article|S)\cdot P(S)+P(article|F)\cdot P(F)}=\frac{\beta}{\beta+p\cdot(1-\beta)}$.

The author's condition for cheating ($pub\ \widetilde{S}$) to be rational and to set $p=1$ is

$$(1-(1-q_i\cdot\tau)^n)\cdot(-g)+(1-q_i\cdot\tau)^n\cdot B'>0\Leftrightarrow(1-q_i\cdot\tau)^n>\frac{g}{B'+g}. \tag{10}$$

For "Pooling II"to exist, it must be that (10) is strictly satisfied. This implies that $\mu=\beta$. Since the author will always decide to publish a paper, the readers' updated posterior will be identical to the prior. The publication decision is not informative with respect to the research project's outcome (success or failure). Furthermore condition (9) must hold with equality. We then get

$$q_i=\left(1-\sqrt[n-1]{\frac{1}{W'}\cdot\left(G-\frac{k}{(1-\beta)\cdot\tau}\right)}\right)\cdot\frac{1}{\tau}. \tag{11}$$

To obtain $q_i\in(0,1)$, the following conditions must be satisfied: We have

$$q_i>0\Leftrightarrow G>W'+\frac{k}{(1-\beta)\cdot\tau} \tag{12}$$

and

$$q_i<1\Leftrightarrow G<W'\cdot(1-\tau)^{n-1}+\frac{k}{(1-\beta)\cdot\tau}. \tag{13}$$

Condition (13) is identical to condition (3) and true by assumption.

Substituting (11) into (10) yields

$$\sqrt[n-1]{\frac{1}{W'}\cdot\left(G-\frac{k}{(1-\beta)\cdot\tau}\right)}>\sqrt[n]{\frac{g}{B'+g}}\Leftrightarrow G<W'\cdot\left(\frac{g}{B'+g}\right)^{\frac{n-1}{n}}+\frac{k}{(1-\beta)\cdot\tau}. \tag{14}$$

The above inequality is implied by (3) if

$$\left(\frac{g}{B'+g}\right)^{\frac{n-1}{n}} < (1-\tau)^{n-1} \Leftrightarrow \frac{g}{B'+g} < (1-\tau)^{n}. \tag{15}$$

It is easy to see that conditions (12) and (15) can be satisfied together for any $\tau \in (0,1)$.

For "Pooling I" to exist, condition (10) must be satisfied with $q_i = 0$, such that the condition degenerates to $B' > 0$, a condition that is always true. Then $\mu = \beta$ and condition (9) yields

$$G < W' + \frac{k}{(1-\beta)\cdot\tau}. \tag{16}$$

One can readily see, that parameters that meet this condition can be easily found.

We proceed with "Semi-Separation". Inequality (10) must hold with equality and we obtain

$$\Leftrightarrow q_i = \left(1 - \sqrt[n]{\frac{g}{B'+g}}\right)\cdot\frac{1}{\tau}. \tag{17}$$

Also inequality (9) must hold with equality and yields

$$\tau\cdot\left(G - (1-q_i\cdot\tau)^{n-1}\cdot W'\right) - \frac{k}{1-\mu} = 0. \tag{18}$$

We can solve for p and obtain

$$\begin{aligned} p &= \frac{\beta}{1-\beta}\cdot\frac{k}{\tau\cdot\left(G - W'\cdot(1-q_i\cdot\tau)^{n-1}\right) - k} \\ &\Leftrightarrow \frac{\beta}{1-\beta}\cdot\frac{k}{\tau\cdot\left(G - W'\cdot\left(\frac{g}{B'+g}\right)^{\frac{n-1}{n}}\right) - k}. \end{aligned} \tag{19}$$

Next we derive the conditions, for which $q_i, p \in [0,1]$. For q_i we obtain

$$q_i \geq 0 \Leftrightarrow 1 \geq \frac{g}{B'+g} \Leftrightarrow B' \geq 0 \tag{20}$$

and

$$q \leq 1 \Leftrightarrow (1-\tau)^n \leq \frac{g}{B'+g} \tag{21}$$

and the first condition is always true. We have furthermore

$$\begin{aligned} p \geq 0 \Leftrightarrow \frac{\beta}{1-\beta}\cdot\frac{k}{\tau\cdot\left(G - W'\cdot(1-q_i\cdot\tau)^{n-1}\right) - k} \geq 0 \\ \Leftrightarrow G \geq W'\cdot(1-q_i\cdot\tau)^{n-1} + \frac{k}{\tau} \end{aligned} \tag{22}$$

and

$$p \leq 1 \Leftrightarrow \frac{\beta}{1-\beta} \cdot \frac{k}{\tau \cdot \left(G - W' \cdot (1 - q_i \cdot \tau)^{n-1}\right) - k} \leq 1$$
$$\Leftrightarrow G \geq W' \cdot (1 - q_i \cdot \tau)^{n-1} + \frac{k}{(1-\beta) \cdot \tau}. \tag{23}$$

Inequality (22) is less restrictive than inequality (23) and is therefore not binding. Plugging in (17) into (23) yields

$$G \geq W' \cdot \left(\frac{g}{B' + g}\right)^{\frac{n-1}{n}} + \frac{k}{(1-\beta) \cdot \tau}. \tag{24}$$

Referring to condition (3), it is easy to see that condition (24) holds, whenever conditions (12) and (21) are satisfied. □

Proof of Proposition 2:
In the first pooling equilibrium n readers prefer to not check a publication. It is straightforward to see that if n readers prefer to remain idle, this is also true for $n+1$ readers, since the validity of inequality (16) remains unaffected by the additional reader. The volume of fraudulent research remains $(1-\beta)$ and no bogus article will ever get revealed.

In the second pooling equilibrium, according to equation (9), the overall probability of detection decreases in n whenever

$$1 - \left(1 - \left(1 - \sqrt[n-1]{\frac{1}{W'} \cdot \left(G - \frac{k}{(1-\beta) \cdot \tau}\right)}\right)\right)^{n} > 1 - \left(1 - \left(1 - \sqrt[n]{\frac{1}{W'} \cdot \left(G - \frac{k}{(1-\beta) \cdot \tau}\right)}\right)\right)^{n+1}$$
$$\Leftrightarrow \left(\frac{1}{W'} \cdot \left(G - \frac{k}{(1-\beta) \cdot \tau}\right)\right)^{\frac{n+1}{n}} > \left(\frac{1}{W'} \cdot \left(G - \frac{k}{(1-\beta) \cdot \tau}\right)\right)^{\frac{n}{n-1}}. \tag{25}$$

The base of both sides must be in the closed unit intervall (otherwise q_i could not be $\in (0,1)$), and therefore the above inequality is equivalent to $\frac{n}{n-1} > \frac{n+1}{n} \Leftrightarrow n^2 > n^2 - 1$, which is obviously true. Since the overall probability of getting caught is lower for $n+1$ readers, the author will find cheating optimal, if he already did so with n readers.

When the parameter constellation for n readers imply the existence of a semi-separating equilibrium, we have to distinguish two cases: In case 1, if $(1-\tau)^{n+1} \geq \frac{g}{B'+g}$, a semi-separating equilibrium will also occur for $n+1$ readers. Then, referring to equation (19), the overall volume of fraud is higher for $n+1$ readers when

$$\frac{\beta}{1-\beta} \cdot \frac{k}{\tau \cdot \left(G - W \cdot \left(\frac{g}{B'+g}\right)^{\frac{n-1}{n}}\right) - k} < \frac{\beta}{1-\beta} \cdot \frac{k}{\tau \cdot \left(G - W \cdot \left(\frac{g}{B'+g}\right)^{\frac{n}{n+1}}\right) - k}$$
$$\Leftrightarrow \left(\frac{g}{B'+g}\right)^{\frac{n}{n+1}} < \left(\frac{g}{B'+g}\right)^{\frac{n-1}{n}} \Leftrightarrow n^2 - 1 < n^2. \tag{26}$$

Making use of equation (17), we see that after publication has been released, the fraction of articles that get scrutinized is not affected by n since

$$
\begin{aligned}
1-\left(1-\left(1-\sqrt[n]{\frac{g}{B'+g}}\right)\right)^{n} &= 1-\left(1-\left(1-\sqrt[n+1]{\frac{g}{B'+g}}\right)\right)^{n+1} \\
\Leftrightarrow 1-\frac{g}{B'+g} &= 1-\frac{g}{B'+g}.
\end{aligned} \tag{27}
$$

Since the overall volume of fraud increases and the share of debunked fraudulent articles remains constant, the absolute volume of undetected fraudulent articles is higher for $n+1$ readers than for n readers.

In case 2, if $(1-\tau)^{n+1} \geq \frac{g}{B'+g}$, a pooling equilibrium will emerge for $n+1$ readers. Hence, A now strictly prefers cheating and is not kept indifferent between cheating and not publishing any longer. Since A's payoff only depends on the overall likelihood of getting caught, we know that this likelihood is smaller in the pooling equilibrium than in the semi-separating equilibrium.

It is obvious that the parameter set for which "Pooling II"is an equilibrium weakly expands (and the set for which the semi-separating equilibrium exists weakly decreases), since $(1-\tau)^{n+1} \leq (1-\tau)^{n}$. $\square$

Proof of Proposition 3:

We start with showing that for all equilibria, after applying care, the author will always publish a sure success and will never publish a sure failure. If the author does not apply care, he will always publish the unsure success.

Upon investing care, the author prefers to publish the project's outcome whenever

$$
\begin{aligned}
&\mu_A^2 \cdot (1-(1-q_i\cdot\tau)^n)\cdot B + (1-q_i\cdot\tau)^n\cdot B + \\
&(1-\mu_A^2)\cdot((1-(1-q_i\cdot\tau)^n)\cdot(-g) + (1-q_i\cdot\tau)^n\cdot B') > 0
\end{aligned} \tag{28}
$$

and $\mu_A^2 = P(true\ success|care:\ success) = \frac{P(care:\ observed\ success|true\ success)\cdot P(true\ success)}{P(care:\ observed\ success)} = 1.$ If, after the uncertainty is revealed, the author identifies a success, the above inequality reduces to $B > 0$, which is true by assumption. If instead the author learns that his success was spurious, the inequality reduces to $((1-(1-q_i\cdot\tau)^n)\cdot(-g) + (1-q_i\cdot\tau)^n\cdot B') > 0$, which is never true, since B' and $(-g)$ are both negative. The result that an unsure success is always published is established in condition (4), and by assumption, B is large enough to make the condition hold.

The author will set $p = 0$ and always invests care if

$$
\begin{aligned}
&\mu_A^1\cdot B + (1-\mu_A^1)\cdot 0 - c > \\
\mu_A^1\cdot B + (1-\mu_A^1)\cdot &((1-(1-q_i\cdot\tau)^n)\cdot(-g) + (1-q_i\cdot\tau)^n\cdot B') \\
\Leftrightarrow &\frac{-c\cdot(\alpha+\beta)}{\alpha} > (1-q_i\cdot\tau)^n\cdot(B'+g) - g.
\end{aligned} \tag{29}
$$

Furthermore, any reader will abstain from checking a publication and sets $q_i = 0$ if

$$\mu_R \cdot W + (1-\mu_R) \cdot \left(\tau \cdot G + (1-\tau) \cdot (1-q_i \cdot \tau)^{n-1} \cdot W'\right) - k < \mu_R \cdot W + (1-\mu_R) \cdot (1-q_i \cdot \tau)^{n-1} \cdot W' \quad (30)$$

and $\mu_R = P(true\ success|article) = \frac{P(article|true\ success) \cdot P(true\ success)}{P(article)} = \frac{\beta}{\beta + \alpha \cdot p}$.

For "Seperation"to exist, it must be that inequalities (29) and (30) both strictly hold for $p = 0$ and $q_i = 0$. Then $\mu_R = 1$ and condition (30) reduces to $W - k < W$, which is always true. For $q_i = 0$ condition (29) can be simplified to

$$B' < \frac{-c \cdot (\alpha + \beta)}{\alpha}. \quad (31)$$

Therefore condition (31) alone is sufficient for the postulated equilibrium to exist.

Next we prove the existence of the second pooling equilibrium. First condition (29) must hold with reversed operator, such that "no care"yields the author a weakly higher utility and $p = 1$. This implies that $\mu_R = \frac{\beta}{\beta+\alpha}$. From condition (29) we can then conclude that

$$(1 - q_i \cdot \tau)^n \geq \frac{(-c+g) \cdot \alpha - c \cdot \beta}{\alpha \cdot (B' + g)}. \quad (32)$$

Readers mix between checking and not checking and condition (30) holds with equality. We then obtain

$$q_i = \left(1 - \sqrt[n-1]{\frac{1}{W'} \cdot \left(G - \frac{k}{\left(1 - \frac{\beta}{\beta+\alpha}\right) \cdot \tau}\right)}\right) \cdot \frac{1}{\tau}. \quad (33)$$

The conditions that ensure that $q \in (0,1)$ yield

$$q_i > 0 \Leftrightarrow G > W' + \frac{k}{\left(1 - \frac{\beta}{\beta+\alpha}\right) \cdot \tau}, \quad (34)$$

as well as

$$q_i < 1 \Leftrightarrow G < W' \cdot (1-\tau)^{n-1} + \frac{k}{\left(1 - \frac{\beta}{\beta+\alpha}\right) \cdot \tau}. \quad (35)$$

This last condition is true by assumption.

Substituting (33) in (32) yields

$$G < W' \cdot \left(\frac{(-c+g) \cdot \alpha - c \cdot \beta}{\alpha \cdot (B' + g)}\right)^{\frac{n-1}{n}} + \frac{k}{\left(1 - \frac{\beta}{\beta+\alpha}\right) \cdot \tau}. \quad (36)$$

The above inequality is implied by (35) if

$$\left(\frac{(-c+g) \cdot \alpha - c \cdot \beta}{\alpha \cdot (B' + g)}\right)^{\frac{n-1}{n}} < (1-\tau)^{n-1} \Leftrightarrow \frac{(-c+g) \cdot \alpha - c \cdot \beta}{\alpha \cdot (B' + g)} < (1-\tau)^n \quad (37)$$

One can easily verify that inequalities (34) and (37) can simultaneously hold true for any $\tau \in (0,1)$.

We proceed with the first pooling equilibrium. The readers will not check any publication ($q_i = 0$), when (31) holds true, again with $\mu_R = \frac{\beta}{\beta+\alpha}$. Then the inequality can be transformed to

$$G < W' + \frac{k}{\left(1 - \frac{\beta}{\beta+\alpha}\right) \cdot \tau}. \tag{38}$$

Referring to condition (29), the author prefers not to invest care and to set $p = 1$ if

$$B' \geq \frac{-c \cdot (\alpha+\beta)}{\alpha}. \tag{39}$$

It is straightforward to see that conditions (38) and (39) can hold simultaneously for a non-empty set of parameter values.

Finally we proove the existence of the semi-separating equilibrium. The readers set q_i such that A is indifferent between "no care"and "careand condition (29) must hold with equality and yields

$$q_i = \left(1 - \sqrt[n]{\frac{(-c+g)\cdot\alpha - c\cdot\beta}{\alpha\cdot(B'+g)}}\right) \cdot \frac{1}{\tau}. \tag{40}$$

Likewise, he author makes all readers indifferent between "check"and "no check"and (30) holds with equality:

$$\tau \cdot \left(G - (1 - q_i \cdot \tau)^{n-1} \cdot W'\right) - \frac{k}{1-\mu_R} = 0 \tag{41}$$

and $\mu_R = \frac{\beta}{\beta+\alpha\cdot p}$. We solve for p and obtain

$$p = \frac{\beta}{\alpha} \cdot \frac{k}{\tau \cdot \left(G - (1 - q_i\cdot\tau)^{n-1} \cdot W'\right) - k} \Leftrightarrow \frac{\beta}{\alpha} \cdot \frac{k}{\tau \cdot \left(G - \left(\frac{(-c+g)\cdot\alpha - c\cdot\beta}{\alpha\cdot(B'+g)}\right)^{\frac{n-1}{n}} \cdot W'\right) - k}. \tag{42}$$

We derive the conditions, for which $q_i, p \in [0,1]$. For q_i we obtain

$$q_i \geq 0 \Leftrightarrow B' \geq \frac{-c\cdot(\alpha+\beta)}{\alpha} \tag{43}$$

and

$$q \leq 1 \Leftrightarrow (1-\tau)^n \leq \frac{(-c+g)\cdot\alpha - c\cdot\beta}{\alpha\cdot(B'+g)}. \tag{44}$$

We have furthermore

$$p \geq 0 \Leftrightarrow G \geq W' \cdot \left(\frac{(-c+g)\cdot\alpha - c\cdot\beta}{\alpha\cdot(B'+g)}\right)^{\frac{n-1}{n}} + \frac{k}{\tau} \tag{45}$$

and

$$\begin{aligned} p \leq 1 &\Leftrightarrow G \geq W' \cdot (1 - q_i\cdot\tau)^{n-1} + \frac{k}{\alpha\cdot\tau} \\ &\Leftrightarrow G \geq W' \cdot \left(\frac{(-c+g)\cdot\alpha - c\cdot\beta}{\alpha\cdot(B'+g)}\right)^{\frac{n-1}{n}} + \frac{k}{\alpha\cdot\tau}. \end{aligned} \tag{46}$$

We derive the conditions, for which $q_i, p \in [0,1]$. For q_i we obtain

$$q_i \geq 0 \Leftrightarrow B' \geq \frac{-c \cdot (\alpha + \beta)}{\alpha} \tag{47}$$

and

$$q_i \leq 1 \Leftrightarrow (1-\tau)^n \leq \frac{(-c+g) \cdot \alpha - c \cdot \beta}{\alpha \cdot (B' + g)}. \tag{48}$$

We have furthermore

$$p \geq 0 \Leftrightarrow G \geq W' \cdot \left(\frac{(-c+g) \cdot \alpha - c \cdot \beta}{\alpha \cdot (B'+g)}\right)^{\frac{n-1}{n}} + \frac{k}{\tau} \tag{49}$$

and

$$p \leq 1 \Leftrightarrow G \geq W' \cdot (1 - q_i \cdot \tau)^{n-1} + \frac{k}{\alpha \cdot \tau}. \tag{50}$$

Inequality (49) is less restrictive than inequality (51) and therefore not binding. Plugging in (40) into (51) yields

$$G \geq W' \cdot \left(\frac{(-c+g) \cdot \alpha - c \cdot \beta}{\alpha \cdot (B'+g)}\right)^{\frac{n-1}{n}} + \frac{k}{\alpha \cdot \tau}. \tag{51}$$

Referring to condition (38), which is true by assumption, it is easy to see, that condition (51) holds, whenever (34), (49) and (51) are satisfied. One can readily see that parameter realizations which meet all conditions do exist.

Like in the deception game, the parameter sets for each equilibrium constitute a partition of the whole parameter space and are mutually exclusive. □

Proof of Proposition 4:
First we show that whenever the first pooling equilibrium will be obtained in the non-competitive setting, we will also obtain this equilibrium in the competitive setting. The revised form of condition (9), assuming that $n = 2$, yields

$$\begin{aligned} &\mu \cdot W + (1-\mu) \cdot (1 - \widehat{q}_i \cdot \tau) \cdot W' > \\ &\mu \cdot W + (1-\mu) \cdot \left(\tau \cdot \frac{G}{1 + \widehat{q}_i \cdot \tau} + (1-\tau) \cdot (1 - \widehat{q}_i \cdot \tau) \cdot W'\right) - k \end{aligned}. \tag{52}$$

The revisded form of inequality (10) equals

$$\left(1 - (1 - \widehat{q}_i \cdot \tau)^2\right) \cdot (-g) + (1 - \widehat{q}_i \cdot \tau)^2 \cdot B' > 0 \Leftrightarrow (1 - \widehat{q}_i \cdot \tau)^2 > \frac{g}{B' + g}. \tag{53}$$

It is obvious, that inequality (52) is equal to (9) for $\widehat{q}_i = 0$, and therefore the conditions that make the first pooling equilibrium appear, are identical to those of the original game.

For the second pooling equilibrium to exist, condition (52) must hold with equality and yields

$$\begin{aligned} &\Leftrightarrow \tau \cdot \frac{G}{1 + \widehat{q}_i \cdot \tau} - \tau \cdot (1 - \widehat{q}_i \cdot \tau) \cdot W' - \frac{k}{1-\mu} = 0 \\ &\Leftrightarrow \widehat{q}_i = \left(1 - \frac{1}{W'} \cdot \left(\frac{G}{1 + \widehat{q}_i \cdot \tau} - \frac{k}{(1-\mu) \cdot \tau}\right)\right) \cdot \frac{1}{\tau} \end{aligned} \tag{54}$$

and $\mu = \beta$. Solving for $\widehat{q}_i$ yields two potential solutions, viz.

$$\widehat{q}_i = \frac{\frac{k}{1-\beta} - \sqrt{4 \cdot \tau \cdot W' \cdot \left(\tau \cdot W' + \frac{k}{1-\beta} - \tau \cdot G\right) + \left(\frac{k}{1-\beta}\right)^2}}{2 \cdot \tau^2 \cdot W'} \tag{55}$$

and

$$\widehat{q}_i = \frac{\frac{k}{1-\beta} + \sqrt{4 \cdot \tau \cdot W' \cdot \left(\tau \cdot W' + \frac{k}{1-\beta} - \tau \cdot G\right) + \left(\frac{k}{1-\beta}\right)^2}}{2 \cdot \tau^2 \cdot W'} \tag{56}$$

Only the former solution yields values in $(0, 1)$. We have

$$\begin{aligned} \widehat{q}_i &> 0 \\ \Leftrightarrow 4 \cdot \tau \cdot W' \cdot \left(\tau \cdot W' + \frac{k}{1-\beta} - \tau \cdot G\right) + \left(\frac{k}{1-\beta}\right)^2 &> \left(\frac{k}{1-\beta}\right)^2 \\ \Leftrightarrow G &> W' + \frac{k}{(1-\beta) \cdot \tau}. \end{aligned} \tag{57}$$

This condition is identical to (9). Likewise, to obtain $\widehat{q}_i < 1$, it must be that

$$\begin{aligned} \widehat{q}_i &< 1 \\ \Leftrightarrow 4 \cdot \tau \cdot W' \cdot \left(\tau \cdot W' + \frac{k}{1-\beta} - \tau \cdot G\right) + \left(\frac{k}{1-\beta}\right)^2 &< \\ 4 \cdot \tau^4 \cdot W'^2 - 4 \cdot \tau^2 \cdot W' \cdot \frac{k}{1-\beta} + \left(\frac{k}{1-\beta}\right)^2 & \\ \Leftrightarrow G < W' \cdot (1 - \tau^2) + \frac{k}{1-\beta} + \frac{k}{(1-\beta) \cdot \tau} & \\ \Leftrightarrow G < \left(W' \cdot (1 - \tau) + \frac{k}{(1-\beta) \cdot \tau}\right) \cdot (1 + \widehat{q}_i \cdot \tau). & \end{aligned} \tag{58}$$

The second possible solution of $\widehat{q}_i$ fails to deliver meaningful results, as values larger than 0 are not obtainable, which is easy to verify with an adjusted version of condition (57).

To show that $q_i > \widehat{q}_i$, we make use of equation (11) and (54) obtain

$$\begin{aligned} \left(1 - \frac{1}{W'} \cdot \left(G - \frac{k}{(1-\beta) \cdot \tau}\right)\right) \cdot \frac{1}{\tau} > \left(1 - \frac{1}{W'} \cdot \left(\frac{G}{1 + \widehat{q}_i \cdot \tau} - \frac{k}{(1-\beta) \cdot \tau}\right)\right) \cdot \frac{1}{\tau} \\ \Leftrightarrow G > \frac{G}{1 + \widehat{q}_i \cdot \tau}. \end{aligned} \tag{59}$$

The fact that the *individual* detection probability is lower for Pooling II in the competitive setting, readily implies that also the *overall* detection probability must be lower. Hence, if cheating is rational in the original game (condition (10) holds), it must be that condition (53) is satisfied, and the equilibrium does exist.

We proceed with the semi-separating equilibrium. Condition (53) must hold with equality, to make the author indifferent. Therefore we get

$$\widehat{q}_i = \left(1 - \sqrt{\frac{g}{B' + g}}\right) \cdot \frac{1}{\tau} \tag{60}$$

and conditions (20) and (21)guarantee that $\widehat{q}_i \in [0, 1]$.

Also condition (52) must hold with equality. Solving for $\widehat{p}$ we yield

$$\widehat{p} = \frac{\beta}{1-\beta} \cdot \frac{k}{\tau \cdot \left(\frac{G}{1+\widehat{q}_i \cdot \tau} - W' \cdot \left(\frac{g}{B'+g} \right)^{\frac{n-1}{n}} \right) - k}. \tag{61}$$

Analogous to condition (23) we obtain

$$\Leftrightarrow G \geq \left(W' \cdot (1 - \widehat{q}_i \cdot \tau) + \frac{k}{(1-\beta) \cdot \tau} \right) \cdot (1 + \widehat{q}_i \cdot \tau). \tag{62}$$

Substituting (60) into the above condition we yield

$$\Leftrightarrow G \geq \left(W' \cdot \sqrt{\frac{g}{B'+g}} + \frac{k}{(1-\beta) \cdot \tau} \right) \cdot \left(2 - \sqrt{\frac{g}{B'+g}} \right). \tag{63}$$

The existence of the equilibrium is still guaranteed, as $\widehat{q}_i$ can become arbitrarily close to 0, when B' approaches zero. Then condition (63) becomes identical to (23), which holds if $(1-\tau)^n \leq \frac{g}{B'+g}$.

We can easily conclude that $\widehat{p} > p$, since

$$\begin{aligned} \frac{\beta}{1-\beta} \cdot \frac{k}{\tau \cdot \left(\frac{G}{1+\widehat{q}_i \cdot \tau} - W' \cdot \left(\frac{g}{B'+g} \right)^{\frac{1}{2}} \right) - k} &> \\ \frac{\beta}{1-\beta} \cdot \frac{k}{\tau \cdot \left(G - W' \cdot \left(\frac{g}{B'+g} \right)^{\frac{1}{2}} \right) - k} & \\ \Leftrightarrow G &> \frac{G}{1+\widehat{q}_i \cdot \tau}. \end{aligned} \tag{64}$$

The parameter set for which the semi-separating equilibrium (the second pooling equilibrium) arises gets smaller (larger) in the competitive setting due to inequalities (58) and (62). □

Proof of Proposition 5:
Applying the same reasoning from prior proofs, if the author produces a success or a landmark result, he will always choose to publish a corresponding article.

A will prefer *pub* $\widetilde{S}$ to *no pub* if

$$\left(1 - \left(1 - q_x^S \cdot \tau\right)^{n_x}\right) \cdot \left(-g^S\right) + \left(1 - q_x^S \cdot \tau\right)^{n_x} \cdot B'^S > 0 \Leftrightarrow \left(1 - q_x^S \cdot \tau\right)^{n_x} > \frac{g^S}{B'^S + g^S} \tag{65}$$

and q_x^S is the individual probability to check a S-publication for a member of audience x.

Likewise, A will prefer *pub* $\widetilde{L}$ to *no pub* if

$$\left(1 - \left(1 - q_i^L \cdot \tau\right)^{n_x+n_y}\right) \cdot \left(-g^L\right) + \left(1 - q_x^L \cdot \tau\right)^{n_x+n_y} \cdot B'^L > 0 \Leftrightarrow \left(1 - q_i^L \cdot \tau\right)^{n_x+n_y} > \frac{g^L}{B'^L + g^L} \tag{66}$$

and $q_i^L = q_x^L = q_y^L$ is true, due to our assumption that $G_x^L = G_y^L$ and $W_x'^L = W_y'^L$.

Moreover, A will prefer mild cheating ($pub\ \widetilde{S}$) over heavy cheating ($pub\ \widetilde{L}$) if

$$\begin{aligned}\left(1-\left(1-q_x^S\cdot\tau\right)^{n_x}\right)\cdot\left(-g^S\right)+\left(1-q_x^S\cdot\tau\right)^{n_x}\cdot B'^S >\\ \left(1-\left(1-q_i^L\cdot\tau\right)^{n_x+n_y}\right)\cdot\left(-g^L\right)+\left(1-q_i^L\cdot\tau\right)^{n_x+n_y}\cdot B'^L\\ \Leftrightarrow g^L-g^S>\left(1-q_i^L\cdot\tau\right)^{n_x+n_y}\cdot\left(B'^L-g^L\right)-\left(1-q_x^S\cdot\tau\right)^{n_x}\cdot\left(B'^S-g^S\right).\end{aligned} \tag{67}$$

Any R of audience a will check a publication of type Y if

$$\begin{aligned}\mu^Y\cdot W^Y+\left(1-\mu^Y\right)\cdot\left(\tau\cdot G_a^Y+(1-\tau)\cdot\left(1-q_a^Y\cdot\tau\right)^{n_x+n_y-1}\cdot W_x'^S\right)-k>\\ \mu^Y\cdot W_a^Y+\left(1--\mu^Y\right)\cdot\left(1-q_a^Y\cdot\tau\right)^{n_x+n_y-1}\cdot W_x'^S\end{aligned} \tag{68}$$

and $n_y=0$, should the result be of type S. Furthermore, $\mu_L=\frac{\beta^L}{\beta^L+(1-\beta^L-\beta^S)\cdot p^L}$ and $\mu_S=\frac{\beta^S}{\beta^S+(1-\beta^S-\beta^L)\cdot p^S}$ and p^L and p^S denote the respective probability for playing heavy and mild cheating.

First we want to prove, that heavy cheating can occur in pooling and semi-searating equilibria. Assume for the moment that $g^L=g^S$ and $B'^L=B'^S$ and $\beta^L=\beta^S$. Assume furthermore that (65) and (68) both hold and the reasoning that this is possible is similar to former proofs. Then, referring to inequality (15), it is clear that the author would never be deterred from cheating mildly, as long as $\frac{g^S}{B'^S+g^S}<(1-\tau)^{n_x-1}$ is satisfied. The individual probability of detection would be

$$q_x^S=\left(1-\sqrt[n_x-1]{\frac{1}{W_x'^S}\cdot\left(G_x^S-\frac{k}{\left(1-\frac{\beta^L}{1-\beta^S}\right)\cdot\tau}\right)}\right)\cdot\frac{1}{\tau}, \tag{69}$$

according to equation (11). Likewise, should the author decide to cheat heavily, his probability of detection would be

$$q_i^L=\left(1-\sqrt[n_x+n_y-1]{\frac{1}{W_i'^L}\cdot\left(G_i^L-\frac{k}{\left(1-\frac{\beta^S}{1-\beta^L}\right)\cdot\tau}\right)}\right)\cdot\frac{1}{\tau}, \tag{70}$$

These two equations would deliver identical values, if $n_y=0$. As soon as $n_y\geq 1$, the individual checking probability will be lower for heavy cheating, and according to condition (25), the overall probability of detection is lower for higher n. Since the punishment is identical for both forms of cheating, heavy cheating must strictly dominate mild cheating.

The existence of pooling equilibrium in which A will always cheat mildly is guaranteed by the following reasoning: Assume once more that that conditions (65) and (68) both hold. Assume furthermore that $g^L>g^S$. Then, although the overall probability of detection is lower for heavy cheating, the author still prefers mild cheating, because the higher chances of not getting detected are offset by the more severe punishment in the case of heavy cheating and (67) will hold.

Finally, there must be a semi-separating equilibrium, in which the author randomizes between $pub\ \widetilde{L}$ and *no pub*. Here, with analogous reasoning from prior proofs, condititions (66) and (68) must hold with equality and yield the author an expected payoff of zero. Randomizing between $pub\ \widetilde{S}$ and *no pub* would also yield a payoff of zero, but since inequality (67) must hold strictly, the author prefers to play heavy cheating with positive probability. □