Incentives for Science

By CHRISTIAN ROESSLER[1] AND PHILIPP D. KOELLINGER[2, 3, 4]

The academic peer review system is central to scientific progress, but it has well-known flaws. Symptoms are substantial delays in getting new research published and, most seriously, the uncertain quality of published works due to non-transparent refereeing criteria and effort, as well as poor incentives to replicate. This paper reviews empirical indicators of these issues and conceptually characterizes an efficient decentralized market for refereeing.

The challenge is not merely to balance the demand and supply in refereeing services. Efficiency entails, statically, that (1) the right works and properties of those works are reviewed, (2) the right experts perform the review, and (3) the right effort is expended on the review. Furthermore, efficiency entails, dynamically, that the system evolves toward cheaper and more accurate reviews. We introduce an anonymous consensus-based protocol, combined with a Vickrey-style auction, that can satisfy these demands.

[1] Department of Economics, California State University – East Bay, Hayward, CA, USA
[2] Department of Economics, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
[3] DeSci Labs AG, Bäch, Switzerland
[4] DeSci Foundation, Geneva, Switzerland

I.   Introduction

Systematic peer-review emerged as the dominant quality assurance mechanism in science during the 1970s, mainly in response to increasing

demands for US governmental agencies to justify funding for scientific research.[1] Today, peer review is omnipresent in the scientific enterprise and shapes scientists' careers, incentives, and research fund allocations. Despite its widespread use, the current peer review system is often criticized. For example, Richard Smith, MD, former editor of the British Medical Journal, has described the peer review process as "ineffective, largely a lottery, anti-innovatory, slow, expensive, wasteful of scientific time, inefficient, easily abused, prone to bias, and unable to detect fraud and irrelevant science".[2]

There is evidence to back up such opinions. Science faces a widely acknowledged replication crisis,[3–6] an increasing number of retractions,[7] numerous discoveries of "fake" research that passed through the peer review process and involves authors from the most prestigious research institutions,[8,9] and indications that scientific progress has slowed over the last few decades.[10,11] Several studies have also shown that the discrepancy between preprints and published peer-reviewed articles is very small, suggesting that peer review adds little to the overall value of the scholarly record.[12–14]

The social costs of forged, irreproducible, and low-quality research are high: It wastes substantial resources, misinforms policy, misdirects future scientific efforts onto unproductive paths, and generally undermines the public trust in science. A hypothetical mechanism that would reliably and efficiently evaluate research works promises enormous social benefits. Naturally, there are great challenges in designing such a mechanism. How do we assess the accuracy of claims made by researchers with highly specialized skills? How do we assess the value of knowledge that may be several steps removed from potential applications and that, by the very nature of the research enterprise, is at best a plausible temporary guess about the elusive truth?

It may indeed be an impossible undertaking to try to predict the longevity and ultimate usefulness of new insights. Realistically, we can verify whether research meets currently acceptable standards of scientific inquiry. However, much is accomplished by doing so. If the process by which knowledge was won is ascertained, then one can judge both the likelihood that the knowledge is accurate and the sensitivity to the employed methods.

This paper explores design flaws in the current peer-review process and how to realign incentives in order to efficiently determine relevant properties of scientific research. The proposed mechanism adapts the idea of decentralized public ledgers, familiar from cryptocurrencies, to a market-based validation of research outputs.

Finally, we formulate testable theoretical claims about an ideal mechanism. Efficiency in academic refereeing has static and dynamic components. At a given time, the market needs to identify an equilibrium price of reviews that matches the right referees to the right tasks and incentivizes optimal effort. Over time, the system should also enable innovation that causes the price of reviews to fall, so that more research is assessed.

We start by discussing inherent problems of a peer review system that have to be addressed (II). Next, we frame peer review as a general mechanism (III) and model a specific candidate mechanism (IV). Finally, we discuss efficiency properties (V). The mechanism is the basis for the incentive layer that is currently under development for DeSci Nodes, a decentralized publishing and peer review platform.

II. Issues in designing an efficient peer review system

Researchers in the US, UK, and China spent over 100 million hours on peer reviews in 2020, equivalent to over 15,000 years of labor. The estimated cost of these researchers' time spent on peer review is equivalent to a 3 billion USD donation of researchers to publishers, contributing to the high profit margins (30-40%) of commercial publishers that monetize the academics' willingness to perform peer review work for free.[23]

Because peer review services are donated, goodwill needs to outweigh the costs for scientists. Conventions have emerged (to different extents in different disciplines) that reduce the costs of refereeing in practice, making it more palatable to scientists to provide the service without compensation. Timing is flexible, and long review times are acceptable, so that the report can be a low-priority task for referees that they perform when the opportunity cost of time is low. Work scope is also flexible, and incomplete reviews are acceptable, so that referees are able to avoid some tasks that are relatively difficult or time-intensive for them to perform.

Post-publication monitoring by the broader scientific community could partially substitute for high-quality peer-review services at the time of publication. Indeed, some scientists engage in post-publication peer-review online (e.g. PubPeer, Data Colada, ResearchGate, X, Reddit). Several prominent fake science scandals published in reputable journals came to light this way,[24–28] but whistleblowers have also been sued in court (e.g. for defamation), creating a strong negative incentive.[24] A bias against publishing replication studies further impairs post-publication monitoring (i.e. repeat study bias), and the few replication studies that do get published typically report results contrary to the original findings (introducing a "gotcha bias").[29] Consequently, peer evaluation relies primarily on pre-publication monitoring, despite its apparent flaws.

There is a high cost to science from slow and imperfect reviews, in terms of immediacy and currency of research output, as well as reliability. Two general strategies to help fix these issues are paying referees and disaggregating the review work.

Paying referees. Peer reviews are typically unpaid, with few exceptions (e.g. *American Economic Review, ResearchHub*[20]), as well as "invisible." Since there is no systematic record of reviews performed, or their quality, providing this service cannot play an important role in scientists' career advancement. The "black box" approach to peer review implies a lack of referee availability and accountability. In contrast, open peer review tends to increase the quality of peer review reports,[18] while not affecting referees' willingness to review or turn-around times.[19]

The leading reasons for academics to do peer review despite the lack of specific incentives is a perception that "it's part of the job" (41%) and a sense of reciprocity (i.e. "do my fair share / reciprocate for reviews of my work" 35%).[21] Without explicit incentives, the result is congestion:[22] Authors face long waiting times to get their work reviewed and published.

It is hardly surprising that 75% of journal editors say that "finding referees and getting them to accept review invitations" is the hardest part of their job.[21] This contributes to an uneven distribution of peer review work, where 10% of reviewers are responsible for 50% of all reviews.[21] An obvious solution to the insufficient supply of (timely and quality) refereeing is offering payments, especially since authors appear to value faster turnaround enough so that there is probably a willingness to pay. However, payment needs to be decoupled from the outcome of the review – while still incentivizing referees to perform the work thoroughly.

Disaggregating review work. A mechanism needs, at minimum, the ability to clearly distinguish between different types of outcomes so that rewards or penalties can be conditioned on them. However, studies have shown that inter-rater agreements between reviewers' recommendations for publication are typically only around 30%.[15] One reason for the low incidence of agreement among reviewers is that referees often focus on different aspects of a study and vary in their expectations and beliefs.

Structuring peer review via specific questions that each referee is asked to answer was shown to increase inter-rater agreement[16] and has been called for to improve the reliability of quality control.[17] Thus, an important dimension for improving the validation process of scientific outputs is to disaggregate the review process into specific claims that can be independently evaluated while steering the process away from subjective judgements towards objective criteria. Disaggregation also allows referees to specialize by selectively performing tasks they are both competent and efficient at.

III. Peer review as a mechanism

Conceptually, the peer review process can be viewed as a mechanism to incentivize "good science." The "principal" here is the scientific community represented, in the closest approximation, by academic societies or publishers and their journals. The process should elicit genuine effort from scientists, the "agents," in applying accepted scientific methods accurately. There is a presumption that the scientist possesses the necessary skill to do so. Then, the challenge is to incentivize the costly effort required to exercise the skill. Principals rely on referees as intermediaries to extract pre-publication signals and may also verify outcomes post-publication. If the scientist does not have the skill, then the process concludes there is "no effort," and the negative incentives attached to this should discourage participation from such

6

unqualified scientists. We face a "double incentive problem" in that referees also have to exert unobservable effort.

Figure 1 captures this setting, encoding the actions of the main actors (i.e., authors, referees, community) as low or high-effort choices. The mechanism has three stages where an effort choice is made: first by the author, who applies scientific methods (or not); then by referees / editors, who carefully evaluate (or not); and finally by the community, which replicates published findings (or not).

Figure 1 – Science as a mechanism
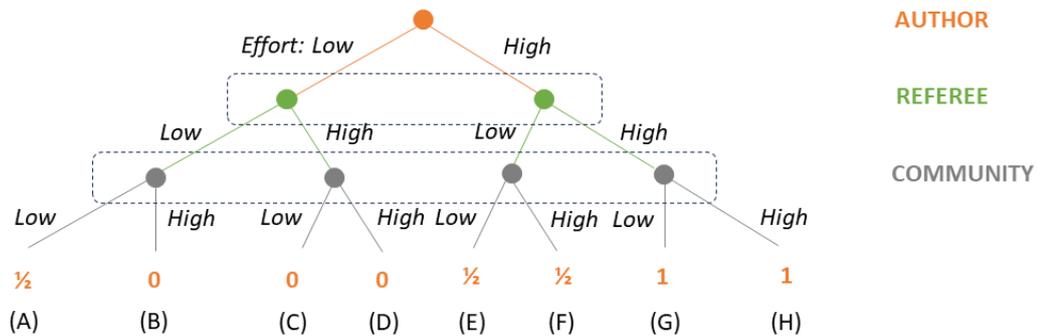
Authors want to publish papers, and it's their payoffs that are shown at the end of each branch: certain publication yields 1, certain rejection yields 0, and an uncertain decision yields ½.

If referees put in high effort, they will correctly distinguish between works with low author effort (rejected) and high author effort (accepted). If referees put in low effort, they might make right or wrong recommendations, so that publication is uncertain regardless of author effort. However, if author effort is low and the scientific community ultimately attempts (and thus fails) to validate the paper, i.e. if the community's effort is high in that scenario, the author is "caught," and payoff is zero.

An advantage of pre-publication monitoring is that it allows high-quality journals to emerge that enable efficient access to quality science. An advantage of post-publication monitoring is that it can create efficiencies in matching referees to research, since the scientists who conduct replications or in-depth post-publication reviews self-select based on their interest, availability, and effort cost. One of the most important roles that post-publication review could play is to induce journals to incentivize referees, given that the reputation cost of discovered fraud or fatal error falls not only on authors but also on journals and their editors. This would help induce high effort at the refereeing stage, which produces the most robust incentives for the authors to produce good science.

Generally, Figure 1 illustrates how low effort in pre-publication and post-publication review will yield low author effort, and therefore systematic flaws in the scientific record. Outcomes (E) and (F) are Type 1 publication errors where the referee process failed and good science sometimes gets mistakenly rejected. Outcomes (A) and (B) are Type 2 publication errors where bad science may get published because the refereeing process failed. In outcome (B), the community can rectify these errors ex-post. Outcomes (C)-(D) and (G)-(H) contain no errors because the referees performed their tasks with high effort.

The specific mechanism we are proposing aims to reduce Type 1 and Type 2 publication errors by incentivizing referees.

IV.   Candidate solution

A core concept in our marketplace for scientific evaluation services are "attestations[i]" that structure the peer-review process by breaking it down into specific, verifiable claims about research properties (e.g. "results supported by the presented evidence," "data / code available," "methods described in sufficient detail"). We refer to referees of attestations as "validators." [ii]

There is not a single market for reviews, but rather a variety of attestation services, each priced separately to balance demand and supply. Formally, an attestation is a list of criteria that define a property of the research object, as well as a consensus rule (i.e. how many validators need to check and what percentage of validators need to agree that the criteria are satisfied in order to verify the property). The list of criteria linked to an attestation is published to the system by the originator of the attestation (e.g. a journal).

Here is an overview of the process. At the bidding stage, validators submit a minimum price to perform the evaluation, and they stake at-risk tokens. The $\mu \cdot n$ bids with the highest stakes are considered (where n is the number of required evaluators, and μ is a fixed multiple); out of these, the n bidders with the lowest price demands are chosen to perform the requested evaluation service.

At the next stage, authors accept and pay a price that reflects the lowest price rejected at the initial stage (i.e. the lowest price outside the n successful bids). This is the market price for a specific attestation, which is expected to fluctuate over time. From each of the n successful bidders, a payment equal to the highest stake placed by one of the unsuccessful bidders is collected. The attestation is awarded (or denied) by consensus of a qualified majority. Stakes are returned

---

[i] See https://codex.desci.com/data-layer-definition/entities/attestation for the technical specifications of attestations.

[ii] A "validator" need not be a human "referee," but could simply be an algorithm. Indeed, there is competitive pressure over time to automate validation, and to create attestations capable of automation, thereby lowering the cost, speed, and reliability of attestation reviews.

to validators who are aligned with the consensus. A more detailed discussion follows, starting with the key objects and their properties.

Research objects. We start with a set Q of properties that a research object may possess, and we index these properties by k. When we need to refer to the number of properties, we denote it by $\#Q$. A research object (e.g. a paper, data, code, or a combination of those) $p \in P$ is a vector of length $\#Q$ with $p_k \in \{0,1\}$, where $p_k = 1$ if the object has the property k, and $p_k = 0$ otherwise. Every research object has a corresponding author responsible for requesting actions on it, such as claiming attestations and publishing in a journal.

Attestations. A journal creates an attestation for property k by specifying criteria that describe k and a standard that determines consensus on whether a paper has the property. The standard has two components: (1) a number n_k of validation protocols that need to be run on the research object that claims the attestation, and (2) a number $m_k < n_k$ of validation protocols that need to agree on whether the research object has that property to establish consensus one way or the other. A function $a_k : P \rightarrow \{0,1\}$ tracks which research objects have received attestations for property k.

Validation protocols. A set V_k of validators of property k creates unique, privately observed validation protocols v_{ik}, where i indexes validators. The validation protocol $v_{ik} : P \rightarrow \{0,1\}$ is a function that assigns to p the value $v_{ik}(p) = 1$ if, according to i's algorithm for confirming the property k, $p_k = 1$. Otherwise, $v_{ik}(p) = 0$. Validator i also stakes a token amount σ_{ik} on the validation protocol as collateral and specifies a minimum fee π_{ik} for the validation service.

Next, we look at the details of the process.

Validator selection. The author of research object p may claim an attestation for property k, i.e. assert that $p_k = 1$. The attestation claim triggers a bidding process in which validators are initially ranked by their stakes σ_{ik}, from highest to lowest. Among the cheapest $\mu \cdot n_k$ bidders, the n_k bidders with the lowest minimum fees π_{ik} are selected. Those selected bidders form the set S_k. Each selected validator receives the Vickrey auction price, i.e. the lowest price π_k^* that was acceptable to unsuccessful bidders in $V_k \backslash S_k$.

Note that $\mu \geq 1$ is a fixed parameter that determines how sensitive the selection of validators is to their price bid vs. their stake. The closer μ is to 1, the more stake matters relatively.

The amount each selected validator has at risk is equal to the highest stake σ_k^* that was acceptable to unsuccessful bidders in $V_k \backslash S_k$.

Consensus. Each selected validator runs their protocol. If $\sum_{i \in S_k} v_{ik}(p) \geq m_k$, then there is consensus to grant the attestation. The attestation state updates to $a_k(p) = 1$. Any validators who submitted a negative finding, $v_{ik}(p) = 0$ lose the at-risk funds σ_k^*.

If $\sum_{i \in S_k} v_{ik}(p) < n_k - m_k$, then there is consensus to deny the attestation. The attestation state updates to $a_k(p) = 0$. Any validators who submitted a positive finding, $v_{ik}(p) = 1$ lose the at-risk funds σ_k^*. If $n_k - m_k \leq \sum_{i \in S_k} v_{ik}(p) < m_k$, no consensus was reached, the attestation is not granted, but no funds are lost. The attestation state remains $a_k(p) = 0$.

To recap: An attestation is created, defined, and made publicly observable. Anyone can choose to serve as a potential validator and perform validation services by some (manual or algorithmic) privately known protocol. Validators

specify a minimum price they must be guaranteed for their work and also put up a stake (i.e. the amount they're willing to lose if, by consensus of a sufficiently large majority of other validators, they have misjudged the attestation claim). We now have a pool of potential validators.

Consensus is defined in the original attestation itself as the percentage of agreement between validators that is required for a definite conclusion. Reaching a consensus puts validator funds at risk. We would expect that originators of attestations define them narrowly enough to be able to set relatively high thresholds for consensus. Then, validators lose their at-risk funds only if they disagree with a large majority, i.e. when it is readily apparent that they have not performed the work diligently or competently. If consensus is not reached, no stakes are lost, since it is then quite possible that a validator could have differed from the majority despite having made a good-faith effort.

As an example, suppose three validators are needed (this number is also specified in the attestation itself). The mechanism then picks a multiple of 3 (say, 30) by sorting the pool of validators by their stakes, which express the bidders' confidence in their judgment. Among those 30 bidders who can perform the validation with confidence, the mechanism next looks at the price demands and picks the three cheapest bidders.

The fee that is actually shown to the attestation claimant and paid to each validator is determined by the lowest price bid that was not successful. Disentangling bids from payments in this way makes it optimal for each validator to bid their actual cost of performing the work to maximize the odds of being selected whenever it is profitable. Technically, this is a Vickrey auction. The benefit is that the cost of the most efficient validators – among those who also signal quality by putting sufficiently high stakes down – is revealed. Over time, the system is designed to generate the cheapest validations

possible, subject to maintaining a certain quality. Thus, the price mechanism incentivizes technological progress that leads to lower costs for a given level of quality.

V. Efficiency Defined

In this section, we formulate notions of static and dynamic efficiency for the peer review problem, and we break these concepts down into subsidiary claims that need to be true for the efficiency property to hold. Static efficiency concerns how effectively the mechanism coordinates resources at any given moment, while dynamic efficiency addresses how the mechanism evolves over time to reduce costs and improve reliability.

A. Static Efficiency

Let P be the set of papers (or research objects) that can be submitted for review. For each paper $p \in P$, let $V(p) \geq 0$ denote the social (or net) value of *accurately* validating its properties (i.e., the benefit to the scientific community of having reliable information about p).

Let R be the set of potential reviewers, each reviewer $r \in R$ having a cost function $c_r(p)$ for providing an accurate evaluation of p. This cost includes time, effort, and opportunity costs.

A mechanism M is a mapping from the environment $(P, R, \ldots)$ to an outcome ω, specifying which papers get reviewed, which reviewers perform the reviews, and how much reviewers are paid (if at all). The mechanism can also impose incentives (e.g., using stakes) so that reviewers exert the appropriate effort.

A mechanism M is said to be statically efficient if it satisfies:

1. **Participation Criterion**:

Only papers p for which $V(p) \geq Cost(p)$ are induced to request evaluation. Here, $Cost(p)$ refers to the total resources (payments plus any opportunity cost or stake forfeitures) needed to evaluate p. Formally, let $\omega \in \Omega$ be an outcome in which a subset of papers $P' \subseteq P$ is chosen for review. Then

For all $p \in P'$: $V(p) \geq Cost(p)$; for all $p \notin P'$: $V(p) < Cost(p)$

2. **Matching Criterion**:

Among the reviewers R, those assigned to paper p must be the ones minimizing the sum of individual costs (including effort) while achieving the desired accuracy. Let R_p denote the set of reviewers actually assigned to p. Then:

$$R_p = argmin_{S \subseteq R \ s.t. \ |S|=n} \sum_{r \in S} c_r(p)$$

subject to a given review protocol that requires, for example, n reviewers for paper p.

3. **Effort Criterion**:

Each reviewer assigned to paper p exerts optimal effort e_r^* to ensure an accurate evaluation, given the incentives. In other words, for each assigned $r \in R_p$,

$$e_r^* = argmax_{e \geq 0}\{U_r(e; M)\},$$

where U_r denotes reviewer r's expected utility, incorporating (i) any payment for reviewing, (ii) stake forfeiture risk if misreporting, (iii) the cost of exerting effort eee. Optimal effort implies that any avoidable Type I or Type II errors are minimized given the mechanism's design.

When these conditions (participation, matching, effort) are jointly satisfied in equilibrium, static efficiency is achieved: the system directs scarce refereeing resources to those papers that benefit most from validation, assigns the "best-suited" reviewers to each paper, and elicits the correct amount of diligence from them.

Peer review is not a one-shot game. It unfolds repeatedly over time as new papers arrive, new reviewers enter, and validation technologies improve. Let $t = 1,2,\dots$ index discrete time periods. For each period t:

- The set of papers P_t arrives (or remains from prior periods).
- A set of reviewers R_t is available, each with cost functions $c_r^t(p)$ that may evolve over time (e.g., through improved software tools or learning-by-doing).
- The mechanism M_t is executed to produce outcome ω^t.

A mechanism $M^\infty = \{M^1, M^2, \dots\}$ defined over an infinite or at least extended time horizon is said to be dynamically efficient if:

1. **Convergence to Lower Costs**:

   If a new technology or method emerges that reduces the **marginal cost** of accurate review for some subset of papers $\Delta \subseteq P_t$, then the mechanism eventually allocates those reviews to the new technology. Formally, if there exists $r' \in R_t$, with $c_{r'}^t(p) < c_r^t(p)$ for some $p \in \Delta$, then for sufficiently large τ,

   $$r' \in argmin_{S \subseteq R_t} \sum_{r \in S} c_r^\tau(p), \text{ for all } p \in \Delta.$$

   In other words, the mechanism should incentivize and enable adoption of superior, cheaper reviewing methods as soon as they become available.

2. **Learning-by-Doing**:

   Reviewers who prove more accurate over multiple rounds of validation (i.e., rarely forfeit stakes and provide reliably correct evaluations) should see their effective cost or demanded compensation π_r^t (and/or stake σ_r^t) adjust to reflect their skill. That is, high-skill reviewers who demonstrate repeated correctness can lower their bid π_r^t and/or raise their stake σ_r^t to outcompete lower-skill reviewers. This dynamic fosters competition and ultimately leads to better accuracy at lower prices.

3. **Flexible Market Expansion**:

   Over time, if the system gains efficiency (via cost reductions), it should become worthwhile to validate a wider range of papers. Let $P_{(t,validated)}$ be the set of papers validated at time t. If costs are decreasing and everything else stays constant, the set $P_{(t,validated)}$ should expand over time to include papers that previously had insufficient value-to-cost ratios. Formally, if $V(p)$ remains stable and c_r^t or the equilibrium price for validations decreases enough, then

   $$p \in P_{(t,validated)}, \text{ where previously } p \notin P_{(t,validated)}.$$

   When these conditions hold, the mechanism is said to be dynamically efficient: it converges toward ever-more cost-effective and accurate reviewing, while expanding the set of research outputs that receive formal validation.

VI. Efficiency of the Candidate Mechanism

To see why this mechanism can achieve static efficiency, note first that authors will only seek attestation for a paper p if they expect to gain at least as much as the total cost of validation. When the mechanism reveals the market-

clearing price, the author of p pays that price only if their valuation $V(p)$ exceeds the cost. Consequently, the set of papers that receive formal attestation comprises precisely those for which validation is privately justified. If we define $V(p)$ as the social value of a publication, then efficiency under the participation criterion will not generally be achieved. However, it is entirely unclear how this social value would be determined in the case of a research object with unknown impact. If it is not measurable, then there seems to be little or no hope of designing a mechanism that targets it, and optimal participation from a private standpoint appears to be the highest efficiency standard that is attainable.

Next, among the reviewers who bid, the mechanism selects those with the largest stakes and then, among this subset, those with the lowest price demands. This dual criterion combines cost-effectiveness (low π_r) with confidence or competence (high σ_r), thereby ensuring that the reviewers who are most capable of delivering accurate judgments at minimal cost are assigned to each paper. By design, the Vickrey properties of the auction encourage truthful revelation of price: each reviewer knows that, if selected, they will be paid an amount determined by the losing bids. Overstating π_r reduces the probability of selection, while understating it risks covering less than the actual cost of diligence or failing to account for the stake penalty. Hence, a truthful cost signal emerges, aligning with the standard argument for Vickrey auctions in single-dimensional mechanisms.

Turning to effort or diligence, the consensus-based stake forfeiture rule incentivizes reviewers to make correct evaluations. Each reviewer who disagrees with a large majority stands to lose the collateral σ_r. Under risk neutrality and in the absence of collusion, a reviewer finds it optimal to invest whatever amount of effort is necessary to form an accurate judgment. Because misreporting or negligence has a higher expected cost (forfeiting the stake) than benefit, the equilibrium outcome is one of adequate effort, in line with the backward-induction logic that each reviewer's best response is truthful

evaluation. Collectively, these features ensure that (i) only the right papers are validated (with the most appropriate reading of "right"), (ii) reviewers with the lowest net cost and highest reliability are assigned to each task, and (iii) each selected reviewer exerts the correct level of effort to avoid mistakes. All three conditions characterize a statically efficient allocation of refereeing resources.

The mechanism also exhibits properties consistent with dynamic efficiency. As it is repeated over time, new reviewers - equipped with cost-saving technologies or advanced capabilities - can enter the market. Because the payment for each validation is determined by the next-lowest losing bidder's price, newcomers with genuinely lower costs will, on average, be able to underbid incumbents. Over multiple rounds, the equilibrium price for a given type of attestation will decrease to reflect cost-reducing innovations. Furthermore, reviewers who consistently align with consensus can credibly offer higher stakes and lower fees in subsequent attestations, given that their probability of forfeiting the stake is low. This reputation dynamic gradually marginalizes less accurate or more expensive reviewers, raising the system's overall accuracy while lowering average costs.

A further implication is that, as prices fall with time, more papers will find it economically justifiable to request attestation. If $V(p)$ was previously below the old, higher equilibrium cost, but remains above the new, lower cost, then such a paper becomes newly eligible for cost-effective validation. This is precisely what one expects in a dynamically efficient mechanism: the set of validated research objects expands in response to falling prices, while the market-driven pressure on reviewers fosters continuous improvements in productivity and reliability.

In conclusion, if we assume risk-neutral reviewers, independent decision-making, and sufficiently objective criteria for consensus, the proposed attestation-and-stake mechanism attains static efficiency by selecting the right

papers, the right reviewers, and the right level of diligence in any given period. Over time, it also moves toward dynamic efficiency by enabling the adoption of cheaper technologies, rewarding reliable reviewers through reputational gains, and enlarging the scope of papers whose validation is worthwhile. Exceptions arise where collusion, subjective criteria, or high entry barriers undermine the truth-telling and cost-reduction dynamics central to this argument.

VII.   References

1.   Baldwin, M. Scientific Autonomy, Public Accountability, and the Rise of "Peer Review" in the Cold War United States. *Isis* **109**, 538–558 (2018).
2.   Smith, R. The peer review drugs don't work. *Times Higher Education (THE)* https://www.timeshighereducation.com/content/the-peer-review-drugs-dont-work.
3.   Ioannidis, J. P. A. Why most published research findings are false. *PLoS Med.* **2**, e124 (2005).
4.   Higginson, A. D. & Munafò, M. R. Current Incentives for Scientists Lead to Underpowered Studies with Erroneous Conclusions. *PLoS Biol.* **14**, e2000995 (2016).
5.   Begley, C. G. & Ellis, L. M. Raise standards for preclinical cancer research. *Nature* **483**, 531–533 (2012).
6.   Camerer, C. F. *et al.* Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nat Hum Behav* **2**, 637–644 (2018).
7.   Van Noorden, R. More than 10,000 research papers were retracted in 2023 - a new record. *Nature* **624**, 479–481 (2023).
8.   Verfaellie, M. & McGwin, J. The case of Diederik Stapel. *American Psychological Association* https://www.apa.org/science/about/psa/2011/12/diederik-stapel (2011).
9.   Piller, C. Blots on a field? A neuroscience image sleuth finds signs of fabrication in scores of Alzheimer's articles, threatening a reigning theory of the disease. *Science* **377**, 358–363 (2022).
10. Bloom, N., Jones, C. I., Van Reenen, J. & Webb, M. Are Ideas Getting Harder to Find? *Am. Econ. Rev.* **110**, 1104–1144 (2020).
11. Park, M., Leahey, E. & Funk, R. J. Papers and patents are becoming less disruptive over time. *Nature* **613**, 138–144 (2023).

12. Nelson, L. *et al.* Robustness of evidence reported in preprints during peer review. *Lancet Glob Health* **10**, e1684–e1687 (2022).

13. Klein, M., Broadwell, P., Farb, S. E. & Grappone, T. Comparing Published Scientific Journal Articles to Their Pre-print Versions. in *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries* 153–162 (Association for Computing Machinery, New York, NY, USA, 2016).

14. Carneiro, C. F. D. *et al.* Comparing quality of reporting between preprints and peer-reviewed articles in the biomedical literature. *Res Integr Peer Rev* **5**, 16 (2020).

15. Bornmann, L., Mutz, R. & Daniel, H.-D. A reliability-generalization study of journal peer reviews: A multilevel meta-analysis of inter-rater reliability and its determinants. *PLoS One* **5**, e14331 (2010).

16. Malički, M. & Mehmani, B. Structured peer review: pilot results from 23 Elsevier journals. *PeerJ* **12**, e17514 (2024).

17. Malički, M. Structure peer review to make it more robust. *Nature Publishing Group UK* http://dx.doi.org/10.1038/d41586-024-01101-9 (2024) doi:10.1038/d41586-024-01101-9.

18. Bruce, R., Chauvin, A., Trinquart, L., Ravaud, P. & Boutron, I. Impact of interventions to improve the quality of peer review of biomedical journals: a systematic review and meta-analysis. *BMC Med.* **14**, 85 (2016).

19. Bravo, G., Grimaldo, F., López-Iñesta, E., Mehmani, B. & Squazzoni, F. The effect of publishing peer review reports on referee behavior in five scholarly journals. *Nat. Commun.* **10**, 322 (2019).

20. ResearchHub. *ResearchHub* https://www.researchhub.com/.

21. Global State of Peer Review. *Clarivate* https://clarivate.com/lp/global-state-of-peer-review-report/ (2022).

22. Hardin, R. & Cullity, G. The Free Rider Problem. *Stanford Encyclopedia of Philosophy* https://plato.stanford.edu/entrieS/free-rider/ (2003).

23. Aczel, B., Szaszi, B. & Holcombe, A. O. A billion-dollar donation: estimating the cost of researchers' time spent on peer review. *Res Integr Peer Rev* **6**, 14 (2021).

24. Warren, M. Honesty researcher facing fraud concerns sues Harvard and accusers for $25 million. *Science* (2023).

25. Enserink, M. Mysterious whistleblower identified in new social psychology scandal. *Science* (2012).

26. Butler, D. Leading plant biologist found to have committed misconduct. *Nature* (2015) doi:10.1038/nature.2015.17958.

27. Kolata, G. He promised to restore damaged hearts. Harvard says his lab fabricated research. *The New York Times* (2018).

28. Archie, A. Stanford president resigns after fallout from falsified data in his research. *NPR* (2023).

29. Berinsky, A. J., Druckman, J. N. & Yamamoto, T. Publication biases in replication studies. *Polit. Anal.* **29**, 370–384 (2021).

32. Elson, M. Pay researchers to spot errors in published papers. *Nature Publishing Group UK* http://dx.doi.org/10.1038/d41586-024-01465-y (2024) doi:10.1038/d41586-024-01465-y.