

ARTIFICIAL INTELLIGENCE AS A RESEARCH INSTRUMENT AND AN EDITORIAL BURDEN: SCIENTIFIC JUDGMENT, FORESTRY SCIENTOMETRICS, AND ACADEMIC INCENTIVES

CHRIS J. CIESZEWSKI

Quantfor Co., Watkinsville, GA, USA

ABSTRACT. The increasing production of fake manuscripts with generative artificial intelligence (AI) is a genuine concern that leads editors to call for special countermeasures. In this opinion piece, I argue against targeting AI for these abuses, and I support this position with two levels of argument, micro and macro. First, AI is like other scientific tools and should be used just as they are. Second, the real reason for the escalating pressure to publish is a system that rewards the number of publications over their quality.

At the first, micro level, I examine examples of individual research tasks, competence, verification, replicability, and a defensible contribution distinguishing accountable assistance from delegated authorship. I use for it an example of an author-directed forestry scientometrics workflow on the Stanford/Ioannidis career database (the Stanford list). The provided examples show what guided exploration produces when used as an analytical tool, similar in nature to, for example, stepwise regressions, nonlinear mixed models, and even general applied mathematics and mathematical software, and what depends on documented author decisions. Along with the examples, I give some opinionated arguments that the evolutionary aspect of AI's game-changing empowerment is greater but no different in principle from other past advances, such as the emergence of mainframe computers, desktop computers and portable calculators, the Internet, software availability and other breakthroughs in scientific developments.

At the second, macro level, I consider the general predicament of publication systems and academic institutional requirements that shape and condition individual practices and values. I provide examples of an institution's public pay records, which track the volume of output by authorship position but show no appreciation of the impact of single-author papers, i.e., of citations per single-author paper. Awarding countable output more reliably than ambitious inquiry unavoidably promotes cheaper and faster manuscript production, resulting in the production of more lower-quality manuscripts pretending to be worthwhile research just to get published. The implied conclusion is that in this context AI is just the topping on the cake, while the reward system is the cake, and the observed trend towards fake publications is merely accelerated and facilitated by AI, not created or motivated by it.

Keywords: artificial intelligence; peer review; forestry scientometrics; authorship position; representation index; academic incentives; publish or perish.

1 INTRODUCTION

A manuscript assembled with AI assistance can arrive fluent, fully referenced, richly illustrated, and empty, and the cost of discovering that it is empty falls on the reviewers and editors, whose time was already the scarcest resource in the system. Weiskittel (2026), in this Issue, documents the arithmetic of Russell and Weiskittel (2026): a demand-to-supply index of forestry publication output against the indexed US forestry researcher

workforce that rose from about 1.0:1 in 2002 to 4.5:1 in 2024, with a projection of 9.4:1 by 2032, extending his earlier assessment of reviewer availability, incentives, and alternative review systems (Weiskittel, 2015). Earlier tools, deemed difficult before they became “mouse-click” operations that produced complex results cheaply, were more conspicuous. Stepwise regression outputs, permutation procedures, and elaborate nonlinear mixed-

effects models could be more easily seen in a manuscript and questioned.

Yet, the question an editor must answer is what to combat? The instrument cannot be the target, because the same instrument helps the editor read, check, and recalculate, and because the errors it produces are often more visible than an erroneous variance buried in a hand-built model. In this paper, I argue against targeting AI for the abuses, and I support this position with two levels of argument, micro and macro. At the first, micro level, AI is like other scientific tools before it and should be used just as they are, which the examples in Section 2 show. At the second, macro level, the real reason for the escalating pressure to publish is a system that rewards the number of publications over their quality, which Sections 3 and 4 examine. I don't discuss here any of the obvious manuscript issues, such as a manuscript containing fabricated data (i.e., falsification) or claims its author knows to be unsupported (i.e., misconduct), which is a different problem than AI itself.

Various sophisticated tools, including AI with them, can be used to create a multitude of dazzling objects in a package that pretends to be original research, and each, used as intended, lets a knowledgeable investigator reach deeper than without them. AI that corrects grammar and style while preserving the author's meaning is a service to editors and a leveler for authors who write in a second language or without institutional editing support. The author must still read every sentence, and must guide selecting a population, constructing a measure, or interpreting a fitted model. The boundary is whether assistance clarifies an argument the author sets forth or supplies an argument the author does not understand or intend.

Any research production should be characterized by the following four characteristics, three of which make a manuscript reviewable, and the fourth makes it worth reviewing:

1. Competence: the author can defend all the research aspects.
2. Verification: clarity of methods, selection and processing of data, and problem tractability.
3. Replicability: the data sources, tools used for different tasks (e.g., SAS used to compute, Mathematica to derive), and applied checks.
4. Contribution: a relevant result that the author can explain and defend and the readers can benefit from.

I see AI's empowerment as greater in degree, but no different in principle, from the arrival of computers, portable calculators, the Internet, and readily available

advanced software. Each was resented, each was misused, and each became an ordinary instrument of inquiry (e.g., my elementary school did not allow slide rules, later schools did not allow simple calculators, and during my first M.Sc. program I was not allowed to check out a 300-step programmable calculator I needed for my thesis, and none of my schools ever allowed spelling checkers). The new capacity lets us do more, faster, and easier, including more of what we ought not to do; deciding which is which remains the author's onus.

My arguments run at two levels and this paper is built accordingly. Section 2 is Exhibit A, for the micro level: an author-directed forestry scientometrics workflow, organized around what the guided exploration made learnable and what it could not decide. Section 3 is Exhibit B, for the macro level: a further exploratory question that the first exhibit raised, namely which features of a publication record one institution's public pay records track. Section 4 is the argument those exhibits serve, stated as my view and marked as such: responsible individual practice is necessary and not sufficient, because what an increase in productive capacity is used for depends on what the system rewards. Section 5 is the editorial response. The numbers related to both exhibits, with their tables and figures, are in Appendices A and B.

2 EXHIBIT A: A FORESTRY QUESTION EXPLORED WITH AI

2.1 Questions worth asking

Suppose we want to ask questions about the following: is a country with many listed forestry scientists strongly represented, or does it simply have many scientists? How much does a citation ranking depend on credit for coauthored papers? What does the database count when papers start in 1960 but citations in 1996? Which AI errors are found by arithmetic checks and which require reading? Can the list tell us whether forestry is under-represented at a particular school? These questions can be answered, including the last one, whose answer is that the list cannot tell us. Answering them requires more than asking AI to produce an impressive report.

I explored the questions with Claude Code under my direction, using the Stanford/Ioannidis career database (Ioannidis, 2025; Ioannidis et al., 2016, 2019, 2020). Appendix A gives the releases, cohort, definitions, calculations, and recorded checks. Appendix C contains the files (Cieszewski, 2026). The population had to fit the forestry question, i.e., primary Forestry, or primary Ecology with secondary Forestry (Archambault et al., 2011). The composite score of the list already includes authorship information. The rankings I discuss are my orderings of the extracted profiles by citation fields, not

that score. I would not have attempted this work by hand. Honestly, had I known how much time it could take to do it with AI, I would not have bothered with it either. The exploration took weeks, but I have no comparison with doing it any other way. My hope is that there is value in showing what various questions can make the tool do.

2.2 What the denominator and the attribution rule decide

Head count answers a question about size. I wanted to know whether forestry's share of a country's listed scientists exceeded its share of the whole list. The representation index (RI), Eq. (1) in Appendix A, is the commonly used location quotient (Bureau of Labor Statistics, 2024). It gives the United States 0.85, Canada 3.11, Finland 11.69, Sweden 3.35, China 0.28, and the United Kingdom 0.29. The United States has the most listed forestry scientists but falls below the world share by this measure. Fig. A1A makes the difference visible. A better graph of the original count would not have answered the new question. Choosing the denominator did.

The same applies to credit for publications. AI can sort any column, but I must decide what the resulting ranking means. Among the 854 profiles, the total-citation ranking has Spearman correlations of +0.55 with the single-or-first-author ranking and -0.03 with the single-author ranking. The median absolute rank change is 247 positions, and 57% of profiles move more than 200 positions. Country rankings also change with the authorship basis (Fig. A1B). The patterns suggest different publishing cultures. In Portugal, New Zealand, and Poland, single-author papers are cited more than the country's papers overall, while in the United Kingdom, the Netherlands, and France they are cited far less than the highly cited multi-author work. Is one ranking therefore the correct measure of contribution? No. Important work is done jointly, and authorship practices differ. The useful result is that the dependence on the credit rule can now be examined. Neither an RI nor a citation rank measures national quality or an individual's scientific worth.

2.3 What a successful calculation can miss

The most instructive discrepancy survived correct arithmetic. Subtracting single-or-first-authored papers from all papers produced negative residuals. The original log treated them as curation noise and "absorbed" them. The later rebuild found the explanation in the source key: np counts papers from 1960, citations start in 1996, and no publication window is stated for the authorship paper counts. There are 52 records with pre-1960 papers, representing 15 profiles, of which one starting in

1956 has np = 7, cpsf = 8, and npsfl = 10. The subtraction agreed exactly with the source but compared different time windows. Appendix A gives the sensitivity checks. Understanding the variable, not making the arithmetic agree once more, resolved the anomaly.

Three other examples show the same difference between correct computation and a valid result. (1) The first population was the one I asked for, not the one I meant: my first mapping, Forestry or Ecology in either subfield slot, was correctly extracted, and S1 reported its 5:1 balance in favor of ecology. My later instruction to require Forestry in both inclusion paths fixed the question I meant to ask. (2) A sentence described Northern Europe as an impact-intensity leader. However, its citations per paper were 39.5, below South America's 63.0, Western Europe's 53.7, Northern America's 48.0, Australia and New Zealand's 43.1, and Central Europe's 40.5. A reading intended to find errors compared the description with the measure. (3) The rebuild found an old count-share image in the figure folder, and an earlier draft of this paper mistook it for the report figure. The report actually embeds the correct nine-rate chart. Recalculating the table could not catch a file name read as the artifact; only opening the embedded picture could. The evidence and checks are in Table A5.

The tool also questioned my assumption that the University of Georgia was Georgia's largest university, and it was right. My request to add John Nash returned the hydrologist J. E. Nash. A search of the full files across the five releases, from 186,177 to 230,333 profiles, found no profile matching the game theorist. It did find a J. F. Nash in dermatology, so the matching name gave another incorrect result. These are reasons to use AI critically, not reasons to pretend that either it or its author is always the source of error. Arithmetic checks determine whether an operation was performed as specified. Reading is needed to determine whether the specification and the resulting claim answer the question.

2.4 What I take from the exercise

The final analysis required author direction in this project. I chose the disciplinary boundary, the authorship comparisons, and the interpretations in the original work, and the denominator in the later rebuild. The window explanation also came in the rebuild. The deposited record does not show what every AI would do without direction, and a log written by the tool is not an independent transcript. It does show which original decisions the author had to defend. Appendix A gives the detailed numbers so that readers can examine them without treating the calculations as the argument of this opinion piece.

There is also a question that this exhibit cannot settle. Science-Metrix subfields classify a publication record and do not provide a list of a school's faculty. Forestry-school scientists appear under plant biology, agronomy, fisheries, parasitology, and tourism as well as Forestry (Table A6). Thus, the list cannot show institutional under-representation, regardless of how accurately AI counts the labels. To me, this is the value of guided exploration: it produces useful answers and exposes questions for which the available data supply no defensible answer. The author must recognize both before asking an editor to spend a reviewer's time on the result.

3 EXHIBIT B: WHAT ONE INSTITUTION'S PAY RECORDS TRACK

Which features of a publication record go with institutional rewards? The State of Georgia's public pay records let me explore that question in an afternoon (Georgia Department of Audits and Accounts, 2025). Of 321 nominal University of Georgia profiles, 197 matched the records by name, and 154 persons had usable observed pay. In the three authorship classes, i.e., all positions, last author, and single author, Spearman correlations with standardized pay are +0.38, +0.37, and +0.08 for papers. They are +0.36, +0.44, and +0.04 for citations, and +0.11, +0.22, and -0.04 for citations per paper. The order is clear. Pay goes about equally with all papers and last-author papers, most with citations to last-author papers, and least with single-author papers and citations. Citations per paper have a clear association only for last-author papers, while the weak rank correlation in all positions is not significant. The weakest association of all, slightly negative, is with citations per single-author paper, i.e., with the most cited original work of authors whose authorship is not in doubt. Publication volume has an elasticity of 0.24 to 0.27 across the regression specifications, and the last-author share of papers correlates with pay at +0.20. Appendix B gives the sample flow, screening and standardization, the full correlations (Table B2), the model estimates (Table B3), and the scatter plots (Fig. B2). The largest pay difference in these records is at the top (Fig. B1): depending on the college, deans were paid 1.2 to 2.3 times the mean pay of the listed scientists of their college. A dean does not have to produce the papers, citations, and grants by which the college and its faculty are evaluated. The last-author share is significant in the pooled models but not on observed pay only, while citations per last-author paper are positive under both estimators of model F, also on observed pay only. I do not call any of these variables scientific quality.

Two limitations apply to this argument. (1) College determines the standardization denominator, but it was

first inferred from the same subfield labels shown to be unreliable in Table A6. That inference was wrong for 32 of the 169 persons then in the analysis, and one payroll match was a different person. Every person's college was therefore checked against the university's public web pages, since labels are not faculty lists, as Exhibit A also shows. The results above use the checked colleges (Appendix B), although joint appointments and indirect sources may still leave uncertainty about the appropriate college. The associations hold without the 50,000-dollar pay floor and without the rule that removes partial exit years (Appendix B). (2) This is a cross-sectional comparison between persons, with career output through 2024 compared with pay in 2021 to 2025. It cannot determine cause and effect, identify pay criteria, or test whether the volume and citation associations differ. Academic rank is the obvious omitted variable. It tracks career output, the last-author share, and pay, so the volume and last-author associations could arise from promotion alone. Rank varies little in this sample, because 133 of the 154 persons held the title Professor in every observed year, but it was not controlled. A significant coefficient beside a nonsignificant coefficient does not test this difference, and gross pay is not a salary rate. In this selected sample, volume has a consistent positive association with pay, and citations per paper have one only for last-author papers. That motivates my question about what institutions reward. It does not settle it. Does assessment distinguish countable output from valuable contribution, and what happens when producing that output becomes cheaper?

4 THE REWARD SYSTEM AND THE TOOL

What follows is my view, and I mark it as such. The two exhibits motivate it. They do not prove it.

Consider the asymmetry that Exhibit B displays. The initial analysis that found an association between pay and publication volume took an afternoon. Further verification would require verified appointments, salary rates, an unselected faculty population, and a longitudinal design, i.e., a study that nobody is rewarded for doing, on a question that no department counts. Initial analysis has become cheap while verification remains costly. That asymmetry is general, and it is precisely the condition under which a system that rewards publications will use the cheap half. When producing one more publishable-looking output brings a more dependable reward than improving the depth or reliability of an existing study, lowering the cost of production makes the additional output more attractive. When checking that output still requires substantial expert attention, part of the cost is simply transferred to reviewers and readers. Reviewers do this work unpaid, while the production of

manuscripts is paid in careers and salaries. AI assistance can therefore help an individual study, as Exhibit A shows, while increasing pressure on the publication system, as the assumed AI scenario in Weiskittel (2026) suggests. The conditions matter. If the effort saved is reinvested in verification, better questions, or clearer explanation, the effect is beneficial. If it is reinvested in additional submissions, the effect is a burden. Which reinvestment predominates is not a property of the tool. It is a property of what the surrounding incentives buy.

This process has been modeled and requires neither dishonesty nor a lack of talent. Smaldino and McElreath show that selection based on publication output favors methods that produce more publishable results, without anyone intending to cut corners. The drift persists as long as the reward stays attached to the count (Smaldino and McElreath, 2016, 2023). Edwards and Roy describe the same hypercompetition and optimization of measures as researchers examining perverse incentives (Edwards and Roy, 2017). A system does not have to be populated by compulsive writers without talent to produce rivers of papers. It only needs to reward practices that yield readily counted outputs more reliably than slower and harder inquiry. The practices are easy to recognize: prolific production, optimization of whatever measure happens to be counted, analyses whose complexity serves no scientific purpose, and presentation substituted for explanation. None of these practices requires AI, but AI makes all of them cheaper.

The people who do the counting are also part of this process. In Exhibit B, pay is associated with the number of papers and citations in all positions and in the last-author position. These are the features of a publication record that an administrator who does not publish can relate to and count, whereas citations per paper, the feature that comes closest to what only a reader can judge, have a weak association with pay in all positions, a clear one for last-author papers, and none for single-author papers. The rewards run in a telling order: about equally with all papers and last-author papers, most with citations to last-author papers, whose authors often lead or manage groups, and least with single-author work, with the least of all for the most cited single-author papers, i.e., for the best original work of authors whose authorship is not in doubt. I hold the view, stated here as a premise and not as a result of this paper, that the growing administrative and managerial layer of universities has weakened the culture of knowledge (Ginsberg, 2011). When research is evaluated by people who count rather than read, it becomes research about the number of papers rather than their content. The highest rewards go to those who demonstrate accomplishment by counting the publications of others. The same records show how large these rewards

are (Fig. B1). Deans were paid up to 2.3 times the mean pay of the most cited scientists of their colleges.

The pay premium is itself the pathology: a dean does not have to produce the papers, citations, and grants by which the college is evaluated, yet is paid more than those who do. This pay model comes from work settings that cannot be compared with a university. In a machine shop the supervisor earns the most because the supervisor knows the work best, shows how it is done, and holds it to its standards. The university is the opposite. A dean is the least knowledgeable person in every field but their own, which is irrelevant to the others, and cannot make other professors better in their fields or hold their work to its standards. In my view, the premium follows administrative power, not knowledge or contribution. The same layer that counts the faculty's output also decides the faculty's raises, and I do not think that this power to reward loyalty and discourage objection goes unused. Some of the most meritocratic universities in the world fill the office from the professors, for a fixed term, in rotation. At ETH Zurich the president appoints the head of a department from its full and associate professors, at the department's request, for two or three years and at most eight in total, with 15,000 Swiss francs a year added to the salary (ETH Zurich, 2015, 2024). In Baden-Württemberg the faculty council elects the dean from its full-time professors for four years, and the dean stays a professor (Land Baden-Württemberg, 2005). A dean of the Jagiellonian University serves four years and at most two terms in a row (Jagiellonian University, 2019), and a Cambridge head of department is one of its teaching and research officers, appointed for up to five years at a time (University of Cambridge, 2025).

Exhibit B is consistent with that premise in one institution's records. It does not establish it anywhere. The publish-or-perish requirement of so many papers per year, imposed on faculty by that layer, is exactly the motivation that promotes the mass production of easy and expedient publications by any means available. AI is simply the newest and most efficient of these means. It is the topping, not the cake.

This is why the editorial resentment, understandable as it is, is aimed at the wrong object when it is aimed at the tool. The quality of submissions is likely to decrease, not because a tool exists, but because the system ignores the quality and rewards only quantity. There are no rewards for difficult, valuable inquiry prevailing over easily counted scores, and the tool merely makes such scoring even cheaper, which further promotes it. Data-based papers that excel in glamorous graphics, massive exploration, and incomprehensibly complicated analysis of an ad hoc question are the natural product of that price fall. These papers will find journals because many journals include data-based papers in their scope, which

reflects a difference in purpose rather than journal quality. The published scope of this journal states that using data is not inferior to theoretical approaches, but using data alone is insufficient reason for publication. Its filter is a well-posed mathematical or computational question and an argument the reader can follow. AI does not change that filter. Incentives do, and the filter is the defense. Editors who resent the flood should aim at the incentive that produces it and at the authors who cannot defend their claims, not at the tool. In the hands of someone who can, the tool extends exactly what a journal like this one exists to publish. Be cautious about what you resent; do not throw the exploratory computation out with the costume science.

5 WHAT AN EDITOR CAN ASK FOR

The response that follows from Sections 1 to 4 is more demanding than an AI label and less demanding than a ban. Table 1 expands the four characteristics of Section 1, i.e., competence, verification, replicability, and contribution, into conditions an editor can look for in a manuscript before review. Verification asks whether what is presented is authentic. Replicability asks whether a reader has the information needed to do the same kind of work, by the same kind of methods, with their own data or in their own environment. A statement that AI was used does not provide that information.

All four conditions require work by the author before submission and do not require the reviewer to reconstruct the project. In practice, the verification and replicability conditions can be met with a one-page methods note submitted with the manuscript: the data sources and their releases; the software used for each task, for example, SAS to compute or Mathematica to derive; the frozen inputs; and, for each numbered claim, the check that recovered it.

Disclosure of AI use is a different matter, and I do not think that it can be required. As a category, the use of AI is not binary, unequivocal, or even mildly meaningful. It ranges from coincidental use, such as an unknowingly AI-enhanced Google search, through marginal use, such as a spelling or grammar check, and through help with code, analysis, and drafting, to plain abuse in a manuscript that only pretends to be research. Where a given use falls on that range can be truly judged only in the mind of the beholder. A meaningful disclosure requirement would either have to cover every spelling check or rely on each author's own judgment of what is worth disclosing, and it would burden the honest authors who comply while doing nothing about those who would not. I am therefore ambivalent about disclosure. It is probably best treated as a matter of the author's

Table 1: The four characteristics of research production (Section 1) as conditions an editor can look for in a manuscript before review.

Condition	What the editor should find in the manuscript or its package
Competence	The author can defend every aspect of the research. Every selection condition, denominator, and category label is defined once, with its meaning in this study, and the limits of inference are stated beside each claim
Verification	Clear methods, selection and processing of data, and a tractable problem, so that what is presented can be checked as authentic: saved unchanged inputs and executable scripts that produce every numbered claim and figure; the measure and confirming check for each comparison; no certification of a narrative by the generator that wrote it
Replicability	The information needed to do the same kind of work by the same methods with your own data or in your own environment: the data sources and their releases; the software or tool used for each task, stated as a plain fact (e.g., SAS used to compute, Mathematica to derive); the settings and checks applied. A statement that AI was used does not provide this information
Contribution	A relevant result that the author can explain and defend and the readers can benefit from: a question the author can state, and what the analysis lets readers understand that makes it worth publishing, whether its purpose concerns theory, methods, explanation, or practical use

All conditions are required. A manuscript that meets competence, verification, and replicability but not contribution is trustworthy within its stated limits and still not worth publishing. A manuscript that claims a contribution without the other three has transferred the burden of establishing its basis to the reviewer. None of the four requires additional unpaid checking by reviewers. All four require work by the author before submission. Disclosure of AI use is not one of them (Section 5).

discretion, or possibly encouraged, as in the proposed Guidelines of Weiskittel (2026), but not required. It is also worth asking editors what they will do with such a disclosure. And, once on the subject, whether they ask different authors what technical writing or editorial services they used, and whether any paid services were involved in addressing the problems of a manuscript.

There are three consequences. First, an editor may decline a manuscript that fails the contribution condition without establishing anything about the other three. A manuscript that fails competence, verification, or replicability may also be declined without alleging misconduct, because burden and insufficiency are sufficient grounds. Second, an AI label is a method statement, not evidence of merit or of wrongdoing. The questions remain whether there is a contribution, competent inquiry, verifiable evidence, and enough information to replicate the work. Weiskittel (2026) proposes guidelines for this

journal that separate the internal EIC Report from the Author Decision Letter and advise against uploading manuscript content to AI tools that store or transmit it externally without confirmed zero-data-retention agreements. Those guidelines are mostly the reviewer's half of the same arrangement. This paper is the author's half, and like the Guidelines, its conditions are what an editor could ask of authors, offered for the same discussion that the Guidelines invite. Third, repeated AI review is not independent review, because similar systems checking an account they already know repeat the same assumptions. Agreement among generated reports does not provide independent support. My safeguard in Exhibit A was to state what each check examined, i.e., a selection condition, a raw cell, a recalculated rank, a claim of direction, or an identity, and to name the checks intended to find errors as such. The arguments that research measures should be used with expert assessment and attention to content (DORA, 2012; Hicks et al., 2015) apply equally to the tool and the indicator.

6 CONCLUSIONS

AI-assisted research and AI-driven publication inflation are different uses of the same increase in the ability to produce. Neither use is an inevitable result of that increase. Exhibit A shows how this ability can be applied to a question that the author defines, examines, and takes responsibility for. The representation index separates disciplinary representation from head count, and the ranking comparison shows dependence on the credit rule. The negative residuals were explained by understanding what the source variables counted, not by making the arithmetic agree. The original work also includes a clearly written false sentence and a correct extraction of the wrong population. In the later rebuild, an unused figure file was mistaken for the report figure, an error that recalculation could not detect. It is because some of these errors hide so well that the editorial burden is real. Exhibit B asks the further question of whether institutional rewards separate publication volume from scientific contribution, but it does not provide an answer. In one institution's public records, pay tracked the number of papers and of citations, in all positions and in the last-author position alike, and its association with citations per paper was weak in all positions, clear for last-author papers, and absent for single-author papers. The association with pay was about equal for all papers and last-author papers, largest for citations to last-author papers, and smallest for single-author work, and the weakest association of all was with the citations per single-author paper, i.e., with the best original work of authors whose authorship is not in doubt. The deans who receive these counts, who do not have to produce

what the counts measure, were paid up to 2.3 times as much as the most cited scientists of their colleges.

Where outputs that are easy to count receive more reliable rewards than the difficult work of explanation and verification, cheaper manuscript production increases an existing problem. AI will be used to multiply plausible papers because that is what the system buys. Editors are therefore right to resist submissions that transfer the burden of establishing their scientific basis to the reviewers. However, AI use alone does not identify these submissions, and misuse can be recognized only from the submitted work. The appropriate response is neither automatic resentment nor automatic acceptance. It requires a contribution the author can defend and author responsibility under the conditions of competence, verification, and replicability. It also requires examination of the incentives that determine how an increased ability to produce is used. These requirements apply to research generally and are not special exceptions for AI users. A journal that applies them does, within its own jurisdiction, the opposite of what the reward system does in the rest of the academy.

ACKNOWLEDGMENTS

I used Claude Code (Anthropic) and Codex (OpenAI), and their programs, to conduct successive analysis and produce the results under my guidance. Conversely, I used Grok and Gemini as reviewers for contrary positions, verification, refuting, and testing the breaking strength of the produced trends and interpretations. In this arrangement, I was the implementer of the analyses and the super-reviewer who judged the reviews and decided what to accept. The cumulative cost of all agents for these analyses was less than \$500, and my time was about 160 hours. The handoff records one model change; otherwise the original record did not record model versions, which are reported as unavailable rather than supplied from dynamic memory records. The later rebuild, recalculation, output checks, and revisions of this version also used Codex (OpenAI) and Claude Code (Anthropic). The model identifiers were taken from session metadata and are listed in the README of Appendix C. The assistance included code, recalculation, figure generation, drafting from my written specifications, checking, and the look-up of each listed scientist's college on public university pages (Appendix B). Every interpretation, definition, and selection of examples is mine, and I reviewed and approved this disclosure. Before submission I commissioned critical reports on successive drafts from AI systems and answered them in response files deposited with Appendix C. Those reports are author-requested critiques, not the journal peer review, and nothing in this paper is a statement about the

journal's handling of it. Aaron Weiskittel's correspondence of March and April 2026 helped develop Section 1, but he has no responsibility for Section 4. My inclusion as a listed profile, my University of Georgia affiliation, and my MCFNS editorial role are disclosed interests. The handling rule is that an editor with no interest in the exhibits decides on this manuscript and chooses its referees, and that I have no access to the editorial file.

DATA AND CODE AVAILABILITY

Appendix A gives the numerical details of the forestry exhibit, and Appendix B, included in this paper, gives the full pay analysis of Section 3. It contains the sample flow, screening, standardization, modeled retiree values, complete model tables, dean comparison, and limitations. Appendix C is the reproducibility package, published with this paper as one archive (Cieszewski, 2026). It contains the unchanged cohort extract and the whole-database country-by-subfield-by-year file. It also contains the field-name mapping to the columns of the original releases, i.e., versions 3, 5, 6, 7, and 8 of Elsevier Data Repository record 10.17632/btchxktyw. The package also contains the original outputs and the record, i.e., S1, the final report, the handoff, and the running log. Further contents are the original figure files with hashes, the pay workbook, the checked college list with its sources, scripts that recompute every numerical table cell and figure, and annual representation indices. The package also includes tables of citation-integrity indicators, database coverage comparisons, and country results that this paper does not reproduce. Every file in the package is accompanied by a SHA-256 hash that can be used to check the file. Appendices D and E list every program verbatim in two separate files. The raw career tables are the public Elsevier releases and are not redistributed.

REFERENCES

- Archambault, E., Beauchesne, O.H., Caruso, J., 2011. Towards a multilingual, comprehensive and open scientific journal ontology. In: Noyons, E., Ngulube, P., Leta, J. (Eds.), *Proceedings of the 13th International Conference of the International Society for Scientometrics and Informetrics*, Vol. 1, Durban, pp. 66–77. <https://science-metrix.com/towards-a-multilingual-comprehensive-and-open-scientific-journal-ontology/>.
- Bureau of Labor Statistics, 2024. QCEW location quotient details. *Quarterly Census of Employment and Wages*, U.S. Department of Labor, Washington, DC. <https://www.bls.gov/cew/about-data/location-quotients-explained.htm>.
- Cieszewski, C.J., 2026. Stanford/Ioannidis forestry representation project record: broad report (supplementary file S1), final report, running project log, and handoff. Author-directed AI-assisted working documents and source extracts, published with this paper as Appendix C (see Data and code availability). <https://mcfns.com/index.php/Journal/article/view/18.3>.
- DORA, 2012. San Francisco Declaration on Research Assessment. Declaration developed at the American Society for Cell Biology meeting, San Francisco. <https://sfedora.org/read/>.
- Edwards, M.A., Roy, S., 2017. Academic research in the 21st century: maintaining scientific integrity in a climate of perverse incentives and hypercompetition. *Environmental Engineering Science* 34, 51–61. <https://doi.org/10.1089/ees.2016.0223>.
- ETH Zurich, 2015. Richtlinien betreffend Entschädigung für Mitglieder der Departementsleitungen, RSETHZ 207. *Rechtssammlung ETH Zurich*, Zurich. <https://rechtssammlung.sp.ethz.ch/Dokumente/207.pdf>.
- ETH Zurich, 2024. Verordnung über die Organisation der Eidgenössischen Technischen Hochschule Zürich (Organisationsverordnung ETH Zürich) vom 21. November 2024, RSETHZ 201.021, Art. 80. *Rechtssammlung ETH Zurich*, Zurich. <https://rechtssammlung.sp.ethz.ch/Dokumente/201.021.pdf>.
- Georgia Department of Audits and Accounts, 2025. Open Georgia: salaries and travel reimbursements, University of Georgia, fiscal years 2015 to 2025. Public records portal. <https://open.ga.gov/>.
- Ginsberg, B., 2011. *The Fall of the Faculty: The Rise of the All-Administrative University and Why It Matters*. Oxford University Press, New York. <https://academic.oup.com/book/40915>.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., Rafols, I., 2015. Bibliometrics: The Leiden Manifesto for research metrics. *Nature* 520, 429–431. <https://doi.org/10.1038/520429a>.
- Hirsch, J.E., 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences USA* 102, 16569–16572. <https://doi.org/10.1073/pnas.0507655102>.
- Ioannidis, J.P.A., 2025. August 2025 data-update for "Updated science-wide author databases of standardized citation indicators". Elsevier Data Repository, version 8, published 19 September 2025; career data through 2024. Earlier releases used: version 3 (19 October 2021), version 5 (3 November 2022), version

- 6 (4 October 2023), version 7 (16 September 2024). <https://doi.org/10.17632/btchxktzyw.8>.
- Ioannidis, J.P.A., Klavans, R., Boyack, K.W., 2016. Multiple citation indicators and their composite across scientific disciplines. *PLoS Biology* 14, e1002501. <https://doi.org/10.1371/journal.pbio.1002501>.
- Ioannidis, J.P.A., Baas, J., Klavans, R., Boyack, K.W., 2019. A standardized citation metrics author database annotated for scientific field. *PLoS Biology* 17, e3000384. <https://doi.org/10.1371/journal.pbio.3000384>.
- Ioannidis, J.P.A., Boyack, K.W., Baas, J., 2020. Updated science-wide author databases of standardized citation indicators. *PLoS Biology* 18, e3000918. <https://doi.org/10.1371/journal.pbio.3000918>.
- Jagiellonian University, 2019. Statut Uniwersytetu Jagiellońskiego (Statute of the Jagiellonian University), adopted by the Senate on 29 May 2019, consolidated text, §§ 93–94. Kraków. <https://bip.uj.edu.pl/dokumenty/statut>.
- Land Baden-Württemberg, 2005. Gesetz über die Hochschulen in Baden-Württemberg (Landeshochschulgesetz, LHG), as amended, § 24 Dekanin, Dekan. Stuttgart. <https://www.landesrecht-bw.de/>.
- Russell, M.B., Weiskittel, A.R., 2026. Improving peer review capacity and integrity in forestry journals. *Forest Ecosystems* 16, 100455. <https://doi.org/10.1016/j.fecs.2026.100455>.
- Schreiber, M., 2008. To share the fame in a fair way, hm modifies h for multi-authored manuscripts. *New Journal of Physics* 10, 040201. <https://doi.org/10.1088/1367-2630/10/4/040201>.
- Smaldino, P.E., McElreath, R., 2016. The natural selection of bad science. *Royal Society Open Science* 3, 160384. <https://doi.org/10.1098/rsos.160384>.
- Smaldino, P.E., McElreath, R., 2023. Correction to "The natural selection of bad science". *Royal Society Open Science* 10(9), 231026. <https://doi.org/10.1098/rsos.231026>.
- University of Cambridge, 2025. Statutes and Ordinances of the University of Cambridge, Chapter IX, Departments and Heads of Departments, Regulation 2. Cambridge. <https://www.admin.cam.ac.uk/univ/so/>.
- Weiskittel, A.R., 2015. Evaluating traditional peer-review processes and their alternatives: an opinionated discussion. *Math. Comput. For. Nat.-Res. Sci.* 7(2), 81–92.
- Weiskittel, A.R., 2026. Artificial Intelligence and Peer Review Capacity in Forestry: A Decade of Evolution and an Introduction to Proposed MCFNS Guidelines for AI-Assisted Review. *Math. Comput. For. Nat.-Res. Sci.* 18(1), 23–35.

APPENDIX A: THE FORESTRY SCIENTOMETRICS EXHIBIT IN NUMBERS

A.1 The question, the data, and the record

My first exhibit is a project that I would not have attempted without AI, because its question was easy to state but tedious to answer. How well are highly cited forestry researchers represented in the Stanford/Ioannidis career-long database of standardized citation indicators (Ioannidis, 2025; Ioannidis et al., 2016, 2019, 2020)? How much does the answer change when publication and citation records are separated by authorship position? Table A1 lists the five releases used, versions 3, 5, 6, 7, and 8 of Elsevier Data Repository record 10.17632/btchxktzyw, with career data through 2020 to 2024. The list selects authors by a composite score and by subfield, and the composite indicator already includes information on authorship. When I criticize an indiscriminated ranking below, I mean ordering the extracted authors by total citations, not the database's own score.

Table A1: Citation database releases used, listed author profiles, and credited career publications, by data year and nested authorship set.

Data year	Version (published)	Listed profiles	Single papers	Single-or-first papers	All papers
2020	3 (19 Oct 2021)	709	5,777	19,985	73,804
2021	5 (3 Nov 2022)	735	5,975	21,062	80,137
2022	6 (4 Oct 2023)	769	6,262	22,459	87,418
2023	7 (16 Sep 2024)	819	6,579	24,015	96,216
2024	8 (19 Sep 2025)	854	6,850	25,125	102,096

Data year: last year of career data in the release. Version and publication date: Elsevier Data Repository record 10.17632/btchxktzyw.N, where N is the version. Version 5 replaced version 4 for the 2021 data year. Listed profiles: author-release records in the cohort requiring Forestry (primary Forestry, or primary Ecology with secondary Forestry). Single papers: sum of nps (single-authored papers, no window stated); single-or-first papers: sum of cpsf; all papers: sum of np (papers 1960 to the data year). These are career counts summed over profiles in each release, rather than unique publications produced in that year. The panel cannot separate entry, exit, reclassification, affiliation change, and database revision. Source: cohort extract; checked against auth6.json (Appendix C).

I did the work with Claude Code (Anthropic) under my direction, and later I had it rebuilt and independently recomputed, as described in the Acknowledgments. Three kinds of record were retained: (1) files produced during the work, i.e., the broad report (supplementary file S1), the final report two days later, their source extracts and outputs, and a running log kept during the work, with my instructions quoted or paraphrased; (2) documents written later, i.e., a handoff four days after the final report and the build notes of the rebuild (the latter are not in Appendix C); and (3) the scripts of Appendix C, which recompute this paper's numerical results from the unchanged cohort extract and country totals, the pay workbook and the checked college list for Appendix B, and the raw releases for the Nash search (Cieszewski, 2026). Since the interaction transcripts were not archived, the available log is an account written by the tool under my direction and is not an independent recording. I also have no comparison to offer for time, cost, or quality. This is a case described after the fact, not a comparison with an analyst working without AI. What I can say is that the exploration described below took weeks, that I would not have done it by hand, and that its value came from the questions I asked of it. That is an account of one project, not a measured general effect.

The cohort follows the population fixed by my later, corrective instruction during the original work and contains the 3,271 author-release records whose primary Science-Metrix subfield is Forestry, plus the 615 whose primary subfield is Ecology and secondary subfield is Forestry. These give 3,886 in all, from a broad extraction of 23,374 records matching my first mapping, Forestry or Ecology in either subfield position (Archambault et al., 2011). The observation unit is the author-release record, i.e., the career record in a given release. It is neither a scientist nor a year's production, because a scientist appears in several releases with most of the same history. In addition, a coauthored paper is included in the records of every listed coauthor. The 2024 data year contains 854 records under 852 distinct names, with unresolved profile identities for the two repeated names. Excluding these names changes the reported rank correlations by no more than 0.01. Retaining one record per repeated name changes Germany's representation index from 1.01 to 0.98.

A.2 Representation index versus head count

The first thing AI produced was what anybody would produce first: a ranking of countries by counts, of the broad cohort and then of the final cohort. The ranking by primary-Forestry profiles, with the United States on top at 239

in the 2024 data year, came in the later rebuild. Is a count an answer to the question of whether forestry is well represented among highly cited scientists? I decided it was not, and in the rebuild I asked a different question instead: what share of a country's listed scientists are forestry scientists, compared with the world share? The comparison can be computed using the whole-database counts available for each country and subfield. The representation index of country c is

$$RI_c = \frac{F_c/F}{A_c/A} = \frac{F_c/A_c}{F/A} \quad , \quad (1)$$

where:

- RI_c is the representation index of country c , a dimensionless ratio equal to unity when the country holds the same share of the listed Forestry scientists as of all listed scientists;
- F_c is the number of listed profiles with primary subfield Forestry in country c ;
- F is the number of listed profiles with primary subfield Forestry in all countries ($F = 733$ in the 2024 data year);
- A_c is the number of listed profiles of all primary subfields in country c ; and
- A is the number of listed profiles of all primary subfields in all countries, including the 252 records without a country ($A = 230,333$ in the 2024 data year).

Eq. (1) is nothing more than a location quotient, used in labor statistics to compare an area's industry share with the national share (Bureau of Labor Statistics, 2024). Only its application to this database is new here. Table A2 and Fig. A1A show the result. The United States, first by count, has an RI of 0.85 because its share of the whole list is larger still. Canada has 95 profiles and an RI of 3.11. Finland's 57 give 11.69 and Sweden's 37 give 3.35. China has 11 profiles and 0.28, and the United Kingdom 19 and 0.29. The useful result is not yet another ranking. Absolute numbers and disciplinary representation within the list answer different questions, so the ranking by one must not be described as the ranking by the other. A country with many listed scientists in medicine can have a low forestry RI even if it has many forestry scientists. Changing the denominator changed the question being answered, which could not be done by improving the graph.

Panel B of Fig. A1 compares citations per single-authored paper with citations per paper in all positions for the same 25 countries. The rates for countries with few profiles are based on only a few records. The panel suggests different publishing cultures. Single-author papers are cited more than the country's papers in all positions in Portugal (61.0 against 35.9 citations per paper), New Zealand (42.3 against 29.8), and Poland (30.3 against 23.1). The opposite holds where multi-author work is highly cited: in the Netherlands (37.2 against 65.7), France (31.4 against 52.1), and Denmark (40.2 against 53.4), and, on a large base, in the United Kingdom (19.9 against 39.3). Both rates are low in Turkey, Slovakia, China, and Japan, and both are high in Brazil (60.8 against 64.3). In most countries, including the United States (46.0 against 52.5), single-author papers are cited somewhat less than the average paper. These are pooled author-release records of a selected list, and a few profiles can dominate a small country: one profile holds 37% of Portugal's single-author citations and 55% of Brazil's.

The formula also has a limitation when the index is based on few profiles, because the index changes when a profile enters, leaves, or is reclassified. Across the five data years Slovakia ranges from 3.23 to 14.68 and Slovenia from 8.02 to 10.14, whereas Finland ranges from 11.69 to 13.44 and the United States from 0.83 to 0.87. The annual values are in Appendix C. The pooled column of Table A2 summarizes the same five years and provides neither validation nor an uncertainty interval. Neither version tests the causal effects of any forestry policy.

A.3 The attribution rule and the ranking

The second question was what a total-citation ranking of the 854 profiles measures. AI could rank on any field, but I had to decide which fields to rank on and what to call them. The release provides single-authored papers and their citations (nps, ncs), single-or-first-authored (cpsf, ncsf), and single-first-or-last-authored (npsfl, ncsfl). At my instruction, the final report called the middle set "senior," although the field does not support that name, so this paper names the sets by what they count. Table A3 divides the pooled cohort into the mutually exclusive single, first, and other positions. These account for 7%, 18%, and 74% of the papers and 6%, 21%, and 73% of the citations, with 37.2, 49.6, and 43.6 citations per paper against 44.2 overall. Most credited work is in the other-coauthor category, which shows the extent of joint work in the records without measuring the contributions of non-first coauthors.

Table A2: Representation index of primary-Forestry scientists by country, relative to all listed scientists, for the 2024 data year and pooled 2020–2024 data.

Country	Forestry count	All listed count	Forestry share (%)	All listed share (%)	RI 2024	RI pooled
United States	239	87,859	32.61	38.14	0.85	0.85
Canada	95	9,589	12.96	4.16	3.11	3.30
Finland	57	1,532	7.78	0.67	11.69	12.34
Australia	42	8,025	5.73	3.48	1.64	1.64
Germany	39	12,165	5.32	5.28	1.01	0.99
Sweden	37	3,471	5.05	1.51	3.35	3.25
France	19	7,285	2.59	3.16	0.82	0.89
United Kingdom	19	20,573	2.59	8.93	0.29	0.34
Japan	16	8,730	2.18	3.79	0.58	0.46
Switzerland	12	3,659	1.64	1.59	1.03	1.32
Spain	12	3,816	1.64	1.66	0.99	1.06
New Zealand	12	1,167	1.64	0.51	3.23	2.29
China	11	12,374	1.50	5.37	0.28	0.27
Italy	11	6,839	1.50	2.97	0.51	0.47
Austria	10	1,472	1.36	0.64	2.13	1.96
Netherlands	9	4,755	1.23	2.06	0.59	0.62
Denmark	8	2,207	1.09	0.96	1.14	0.94
Poland	8	1,412	1.09	0.61	1.78	1.16
Slovenia	7	217	0.95	0.09	10.14	9.25
Portugal	6	711	0.82	0.31	2.65	2.82
Turkey	6	1,308	0.82	0.57	1.44	1.20
Czechia	5	588	0.68	0.26	2.67	1.74
Slovakia	5	107	0.68	0.05	14.68	6.63
South Africa	5	897	0.68	0.39	1.75	0.98
Brazil	4	1,195	0.55	0.52	1.05	0.70

Forestry count F_c : listed profiles with primary subfield Forestry in country c . All listed count A_c : listed profiles of every primary subfield in country c . Shares are F_c/F and A_c/A , with world totals $F = 733$ and $A = 230,333$. A includes 252 records with unknown country; no primary-Forestry record has an unknown country. $RI = (F_c/F)/(A_c/A)$, Eq. (1); $RI = 1$ means equal shares. RI pooled: the same ratio computed from the five releases 2020 to 2024 together, giving larger releases more weight. Top 25 countries by F_c ; ties ordered by ISO code. Brazil, India, Indonesia, and Norway tie at four profiles for the 25th place; ISO order keeps Brazil and leaves out India, Indonesia, and Norway. Indonesia's RI of 22.05 would be the highest in the table. Annual values for every country and year are in `annual_ri.csv` (Appendix C). Source: `world_subfield_year.csv`, checked against the Forestry extract.

Table A3: Pooled cohort (3,886 author-release records, 2020–2024) divided into mutually exclusive authorship positions.

Position	Papers (%)	Citations (%)	Citations / paper
Single	7	6	37.2
First	18	21	49.6
Other	74	73	43.6
All	100	100	44.2

Single: nps papers and ncs citations. First: first author of a multi-authored paper, cpsf minus nps papers and ncsf minus ncs citations. Other: every non-first coauthored position, np minus cpsf papers and nc minus ncsf citations. Papers (%) and citations (%) are shares of the pooled totals. Separately rounded shares may not sum to 100. Citations per paper: class citations divided by class papers. Window limitation (Section A.4): np counts papers from 1960 to the data year, while no window is stated for nps and cpsf. Thus, the other-paper residual is approximate for the 52 records with a first publication before 1960. It is negative for five of them (one profile, Adler). Without those 52 records, paper shares are 7.0 / 18.4 / 74.6% and citations per paper are 37.3 / 50.0 / 43.7. Excluding only the five negative-residual records gives 7.1 / 18.5 / 74.4% and 37.1 / 49.6 / 43.6 (full cohort 7.2 / 18.5 / 74.4%). Source: extract; `authorship.json` and final report Table 9.

Ranking the 854 records by total citations gives a Spearman correlation of +0.55 with the single-or-first-author citation ranking and -0.03 with the single-author ranking. The top-25 overlaps are 14 and 5 of 25. Between the two bases, the median absolute change is 247 positions, and 57% of the profiles move by more than 200 positions. Several details affect the interpretation. The pooled coefficient is approximately zero, while the within-group coefficients

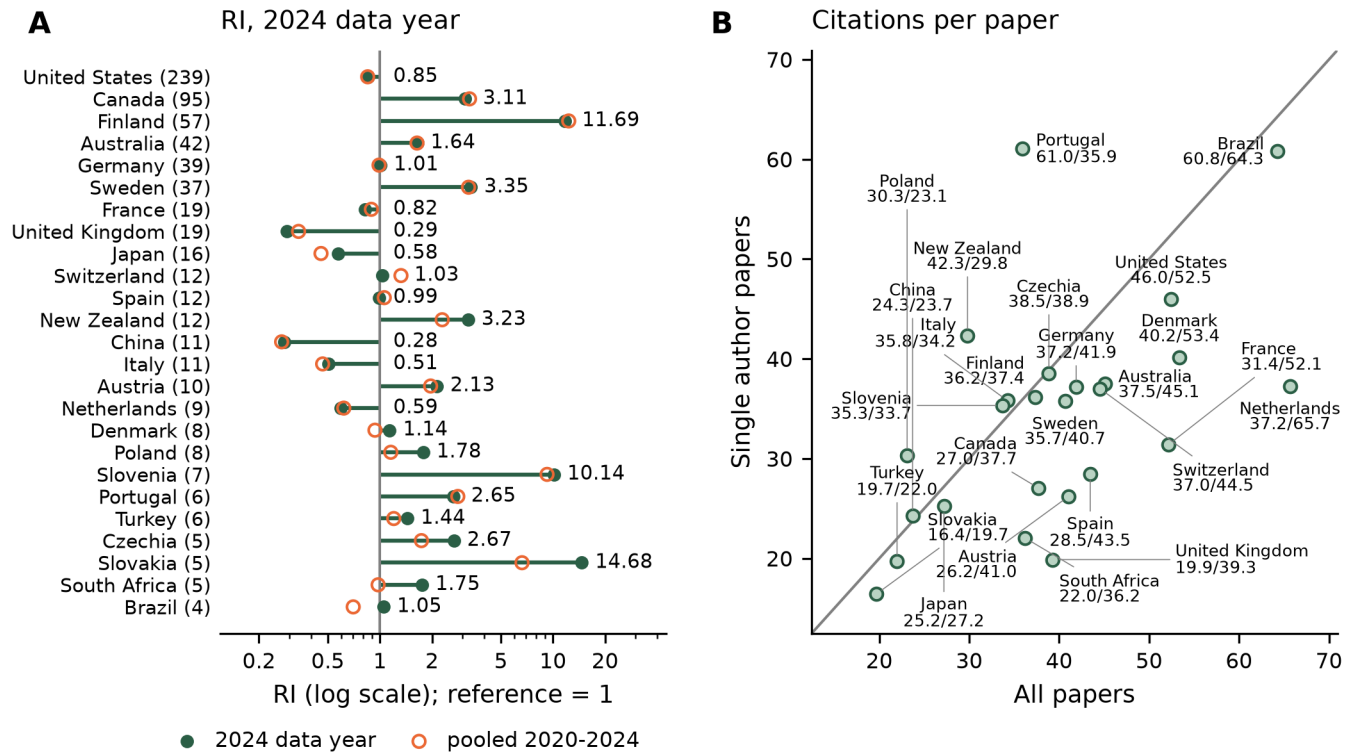


Figure A1: Publishing base and impact for the 25 countries with the most listed primary-Forestry scientists in the 2024 data year (count after the name; ties ordered by ISO code). Brazil, India, Indonesia, and Norway tie at four profiles for the 25th place; ISO order keeps Brazil and leaves out India, Indonesia, and Norway. Indonesia's RI of 22.05 would be the highest in panel A. A: representation index $RI_c = (F_c/F)/(A_c/A)$, Eq. (1). F_c , A_c : listed profiles with primary subfield Forestry and with all subfields in country c . $F = 733$ and $A = 230,333$: world totals. The vertical line marks $RI_c = 1$, equal shares of listed Forestry scientists and listed scientists overall. Filled green markers and labels: 2024 data year. Hollow orange markers: ratio pooled over the five releases 2020 to 2024. The axis is logarithmic. A high index shows specialization within the list and does not measure national quality. Countries with few profiles change most between the two bases. B: citations per single-authored paper (ncs / np) against citations per paper in all positions (nc / np) for the same countries, pooled author-release records 2020 to 2024. Labels give the country and the two rates as single/all. The gray line marks equality. Rates for countries with few profiles are based on few records, and the panel suggests different publishing cultures but does not explain them. Data: Table A2 (A) and Appendix C (B); source: world_subfield_year.csv, the extract, statistics.json; build/make_figures.py.

are negative (-0.25 in the Forestry group and -0.15 in the Ecology group). Thus, neither subgroup shows a positive association and the pooled value describes neither subgroup. The fields excluding self-citations give +0.55 and +0.01. In addition, 45 profiles have no single-authored citations and 37 have no single-authored papers. Restricting the comparison to the 809 profiles with single-authored citations gives +0.06. Again, the useful result is not that one ranking is the correct one. These results show substantial dependence on the credit rule and allow that dependence to be examined, but they do not identify a correct ranking. The single-author ranking does not necessarily measure individual contribution, because authorship practices differ between fields and generations, and important work is done jointly. The profiles entered the list partly through the citation quantities being compared, so these correlations describe the selected list and need not apply to all forestry researchers. Fig. A1B shows similar dependence at the country level, where citations per single-authored paper and per paper in all positions rank the 25 countries differently. The greatest scatter is among countries with few profiles. The database's composite score already includes authorship information, so differences from a total-citation ranking are expected and do not imply a problem with that score.

A.4 Computational success is not measurement validity

The third result came from examining another question. Subtracting cpsf from np to obtain the other-position papers gave negative residuals in five records. The original running log treated the negative residuals as curation noise and "absorbed" them. My earlier draft gave the same kind of explanation, a source inconsistency in one profile. The later rebuild found the window explanation in the release key sheet, reproduced in Table A4. The field np counts papers published from 1960 to the data year (np6024), while nc counts citations received from 1996 to the data year. Since no time window is stated for nps, cpsf, and npsfl, authorship counts can exceed np for profiles with papers before 1960. The cohort contains 52 such records (15 profiles with a first publication before 1960). The field npsfl exceeds np in 20 of them, and cpsf exceeds np in five. All five belong to one profile with a first publication in 1956 (np = 7, cpsf = 8, and npsfl = 10 in every release). The citation fields are nested in every record. Therefore, the division into positions shows no inconsistency for the 3,834 records without pre-1960 papers and differs by a few papers for the rest. Dropping the 52 records changes the paper shares to 7.0%, 18.4%, and 74.6%. Note that the arithmetic agreed with the source both before and after the correction. The problem was resolved not by making the numbers agree but by understanding what the source variables counted. This is exactly the difference between correct computation and valid measurement that an editor cannot see from a fluent page.

Table A4: Fields used from the career release: names in the extract and release (version 8), definitions in the release key sheet, and time windows.

Extract name	Release column (v8)	Definition (release key)	Window
np	np6024	number of papers	papers 1960 to the data year
nc	nc9624	total citations	citations 1996 to the data year
nps / ncs	nps / ncs	single-authored papers / their citations	no paper or citation window stated
cpsf / ncsf	cpsf / ncsf	single-or-first-authored papers / their citations	as nps / ncs
npsfl / ncsfl	npsfl / ncsfl	single, first, or last-authored papers / their citations	as nps / ncs
h / hm	h24 / hm24	Hirsch index (Hirsch, 2005) / coauthorship-adjusted hm index (Schreiber, 2008)	as of the end of the data year
self%	self%	self-citation share of total citations	no window stated
cprat	cprat	total citations divided by distinct citing papers	no window stated
np cited	np6024 cited9624	papers cited at least once	papers 1960 to the data year
np_rw	np6024_rw	papers marked as retractions in the Retraction Watch database	papers 1960 to the data year; versions 7 and 8 only
sm-subfield-1 / -2;	same	first and second Science-Metrix subfield;	profile of the data year
sm-field		top-ranked field	
firstyr / lastyr	same	year of first / most recent publication	

Release column names end with the data-year suffix (6024 = 1960 to 2024, 9624 = 1996 to 2024 in version 8; earlier versions use their own year). Citation columns other than the _rw and _d columns are provided with and without self-citations (suffix ns). This paper uses the all-citations columns unless stated otherwise. The extract removes the year suffixes from column names to combine the five releases. The different windows for np and the authorship paper counts explain the negative residuals in Section A.4. Source: Key sheet of each release; Appendix C README.

A.5 What the checks caught, and what only a reader could

Table A5 lists the recorded episodes of the work, the evidence for each, the class of failure, and the check that caught it. Three examples show how well some mistakes in AI-assisted work hide. First, the broad cohort was a correct extraction of the population I asked for, not the one I meant. My first mapping required Forestry or Ecology in either subfield slot. S1 reported the 5:1 balance in favor of ecology. My later instruction that both inclusion paths require Forestry fixed the question. Second, a clearly written but false sentence in the first report draft called Northern America and Northern Europe the leaders on impact intensity. The regional citations per paper in the same report put Northern Europe (39.5) near the bottom, below South America (63.0), Western Europe (53.7), Northern America (48.0), Australia and New Zealand (43.1), and Central Europe (40.5). The numbers had been checked, but the sentence survived until a reading intended to find errors compared the description with its measure. The log says it was fixed without quoting a corrected sentence; the final report still contains a claim contradicted by its own region figure. Third, the rebuild notes found an old count-share image, ff4_group.png, in the figure folder, and an

earlier draft of this paper mistook it for the final report figure. The report actually embeds `ff_ratio_group.png`, the correct nine-rate chart. The mistake was a file name read as the artifact; no arithmetic check opens the embedded image. On the other side of the ledger, AI questioned my assumption that the University of Georgia was Georgia’s largest university, and it was right. My request to add John Nash returned the hydrologist J. E. Nash, and the handoff records my correction of the identity. A full-file search of all five releases (186,177 to 230,333 profiles) found no profile matching the game theorist but found a J. F. Nash in dermatology. This was a second wrong identity that a name match would have happily accepted. The interpretation errors were found by reading the claims, checking the identities, and looking at the embedded image, while recalculation caught none of them. A reviewer who receives the fluent page receives none of that record unless the author supplies it.

Table A5: Recorded episodes of the work, the evidence in the record for each, the class of failure, and the lesson.

Episode	Evidence in the record (Appendix C)	Class	Lesson
Broad Forestry-or-Ecology cohort	S1 Section 2 (population definition); the author’s first either-slot mapping and later NAMING instruction in the log; S1 reports ecology dominance; 23,374 versus 3,886 records (Section A.1)	Revised research objective (author)	A correct extraction can answer the question asked rather than the one intended
Northern Europe as intensity leader	First-draft sentence reported in log; final report Section 4.2 and its region figure; regional citations per paper (South America 63.0, Western Europe 53.7, Northern Europe 39.5); log entry "Adversarial verify workflow" says fixed but quotes no corrected sentence (Section A.5)	Interpretation error (tool)	Agreement of table values does not verify written comparisons
Nash identity	Returned profile J. E. Nash (environmental engineering); intended J. F. Nash (game theory); author correction in handoff; full-file surname search of all five releases, <code>nash_search.json</code> (Section A.5)	AI implementation error (identity matching)	A name match does not verify identity
Qualitative-oriented label	Classification rule in S1 (absence of a methods subfield); resulting 94% label	Interpretation error (concept)	A rule can run correctly without measuring the stated concept
USDA and United States Unknown counts	USDA affiliation sequence 32 / 73 / 36 in S1 Table 8a; Unknown counts 28 to 110 in Table 7	Source data problem, unresolved	Sudden changes require source checking before explanation
Old figure file	Final report Figure 4 uses <code>ff_ratio_group.png</code> (nine rates); <code>ff4_group.png</code> (count shares) is not embedded in the final report; original <code>build_docx6.py</code> and figure hashes in Appendix C (Section A.5)	Interpretation error (later rebuild)	Open the embedded image: an earlier draft of this paper mistook an unused file for the report figure
Scholar raw counts	Log entry on reading only the first page; author-supplied counts 211 and 361; <code>dedup_scholar.py</code> and <code>dedup_bettinger.py</code> ; Bettinger estimate about 271 in the log versus about 300 in the final report	AI implementation error, then conflicting reported estimates	A written counting rule does not reconcile conflicting estimates
Georgia Tech enrollment	Author assumption in the log; contrary AI evidence; assumption withdrawn	Interpretation error (author), corrected by the tool	Oversight runs in both directions
Top-20 to top-25 display	Instruction in the log; final report Table 5 (scientist and publication counts by country)	Presentation preference	Neither party made an error

Class labels distinguish AI implementation error, revised research objective, source data problem, interpretation error (tool, author, or later rebuild), and presentation preference. Log entries are identified by their original headings in the running log (Appendix C, record folder). S1 is the broad report.

A.6 What the exhibit does and does not show

The final analysis required author direction in this project, with its disciplinary boundary, comparison rules, measure definitions, and interpretations depending on author decisions. Of the denominator comparison, authorship dependence, and time-window problem, only the authorship dependence is documented in the deposited record as a calculation requested by the author. The denominator comparison and window explanation came in the later rebuild. That is a claim about this record. It is not a claim that no AI could produce such an analysis, which this case did not test, and it is not an estimate of how much the assistance saved me. The exhibit also cannot show whether forestry

is under-represented at any institution, because the Science-Metrix labels classify journals and do not provide school faculty lists. Forestry-school scientists at my university appear under plant biology, agronomy, fisheries, parasitology, and tourism as well as forestry (Table A6, used again in Section 3). What the exhibit does show is an exploration under competent direction that produced results worth having, together with a record of what it took to trust them.

Table A6: University of Georgia scientists on the 2024 list whose forestry-school or related affiliation is not shown by the Science-Metrix subfield labels.

Scientist	Primary subfield in the database	Why a forestry reader would expect a different label
Bettinger	Forestry (secondary: Ecology)	Operations research and GIS; the secondary label does not describe the work
Cieszewski	Forestry	Biometrics and modeling; there is no biometrics label
Peterson	Ecology (secondary: Forestry)	Forest disturbance ecology in the forestry-anchored union
Teskey	Plant Biology and Botany	Tree physiology, Warnell (retired)
Markewitz	Agronomy and Agriculture	Forest soils, Warnell
Grossman	Fisheries	Stream ecology, Warnell (emeritus)
Yabsley	Mycology and Parasitology	Wildlife disease, Warnell
Woosnam	Sport, Leisure and Tourism	Nature-based tourism, Warnell
Helfman	Fisheries	Odum School of Ecology (emeritus); grouped with Warnell in the running log until the affiliation check

Warnell: Warnell School of Forestry and Natural Resources; Odum: Odum School of Ecology; both at the University of Georgia. Primary subfield: sm-subfield-1 of the 2024 data year. This classification is based on journals and does not provide a school faculty list. The database has no department field. Surnames and initials were matched to the school directory and checked by hand. Source: extract; running log; final report Tables 12 and 13.

APPENDIX B: PUBLIC PAY RECORDS OF UNIVERSITY OF GEORGIA SCIENTISTS ON THE 2024 CITATION LIST: A DESCRIPTIVE EXERCISE

B.1 Status

This appendix reports an analysis included in the Discussion of an earlier draft and separated in response to author-commissioned AI critiques. It compares public fiscal-year pay with career bibliometric measures for University of Georgia (UGA) scientists on the 2024 data year of the Stanford/Ioannidis list. It also compares deans with the listed scientists of their colleges. The main text uses these results as an illustration, not as a causal test. The analysis describes the data and suggests hypotheses for further study. Colleges were first inferred from subfields, the inference shown to be unreliable in Table A6 in Appendix A, and then checked person by person against public university pages (Section B.3). The outcome is recorded fiscal-year gross pay, which is not a salary rate, and 204 of the 818 person-years are modeled rather than observed. The cross-sectional design cannot identify the criteria or causal processes that determine pay. Observed pay is the primary basis of every figure and the active-only model columns. Modeled retiree values are reported separately.

B.2 Data and sample flow

The State of Georgia publishes the name, title, and fiscal-year gross pay of every University System employee on the Open Georgia portal, open.ga.gov (Georgia Department of Audits and Accounts, 2025). I matched the 321 nominal UGA profiles of the 2024 data year by surname and given name, ignoring diacritics, to the UGA pay records for fiscal years 2021 to 2025. This gave 197 matched profiles and 124 unmatched. One name match was a different person, a university employee with the same surname and first name and a non-academic title, and is excluded (Section B.3). Three of the unmatched profiles belong to universities in the country of Georgia, so the nominal UGA count includes them but the pay results do not. The 124 unmatched profiles were searched in an additional extract for fiscal years 2015 to 2020, and 26 were found. The transcribed historical list contains 27 rows, with 26 matching a listed profile after transcription. Rows without a listed profile are excluded. Of the matching profiles, 26 had a college found and were included in the panel. The 98 profiles unmatched in either window were not verified individually. Failure to find a payroll record does not by itself show that a person has retired, died, or left, and the list gives last-known affiliations. Table B1 shows the successive selection steps. Two workbook sheets were prepared before the match ignoring diacritics: Scientists x FY, with 190 persons, and Not found in OpenGA, with 131 profiles. These sheets

predate seven matches, which had no college before the check and now have one, and are retained unchanged so the work can be traced. The panel sheet (Std panel) is the source to use.

Table B1: Successive sample selection from the 321 nominal UGA profiles of the 2024 data year, including three unmatched profiles at universities in the country of Georgia (Stanford/Ioannidis career database, version 8) to each figure and model sample. FY = state fiscal year (July to June). A usable person-year is an observed screened year (Section B.3) or a modeled year. Observed persons have at least one observed screened year. Source: workbook sheets Std panel and Scientists long; build/redu_standardized.py; build/compute_ugapay.py.

Step	Persons (or persons / person-years)	Use
Listed UGA profiles in the 2024 data year	321	all
Matched to FY2021 to 2025 pay records (diacritic-insensitive name match)	197	continue
Not matched in FY2021 to 2025	124	searched in FY2015 to 2020
found in FY2015 to 2020 records	26	continue
carried into the panel (college found)	26	continue
not found in either window	98	excluded, unverified
Persons in the assembled panel	223	197 + 26
Persons with no usable person-year (every year administrative, temporary or retiree title, below the pay floor, or a partial exit year, or the one payroll match that was a different person)	43	excluded
Persons with at least one usable person-year	180	panel
with at least one observed screened year (observed persons)	154	primary
with modeled years only (pre-window retirees)	26	secondary
observed persons who also have modeled years (retired inside the window)	27	counted among observed
Person-years: observed / modeled / total	614 / 204 / 818	
Persons / person-years in models E	180 / 818	pooled
Persons / person-years in models E, observed pay only	154 / 614	primary
Persons / person-years in models F (at least one single-author and one last-author paper)	167 / 759	pooled

Table B2 gives Spearman correlations between observed standardized pay and papers, citations, and citations per paper in the three authorship classes. Sections B.3 and B.4 provide the definitions and model estimates.

Table B2: Spearman rank correlations of standardized pay with career papers, citations, and citations per paper for University of Georgia scientists on the 2024 list. The three authorship classes are all positions, last author, and single author. Only persons with observed pay are included.

Feature of the publication record	rho	p	n
Papers, all positions (np)	+0.38	<0.001	154
Papers as last author (np_last)	+0.37	<0.001	154
Papers as single author (nps)	+0.08	0.30	154
Citations, all positions (nc)	+0.36	<0.001	154
Citations to last-author papers (nc_last)	+0.44	<0.001	154
Citations to single-author papers (ncs)	+0.04	0.62	154
Citations per paper, all positions (nc / np)	+0.11	0.16	154
Citations per paper, last author (nc_last / np_last)	+0.22	0.007	153
Citations per paper, single author (ncs / nps)	-0.04	0.67	141

Pay: mean standardized fiscal-year gross pay over each person's observed screened years, fiscal 2021 to 2025. Standardization divides pay by the five-year mean for observed listed scientists of the same college and multiplies it by their university-wide mean (185,676 dollars). Values remain in dollars and are rescaled to a common college mean. Persons: listed profiles matched to public pay records with at least one non-administrative, non-temporary year above 50,000 dollars (Appendix B, Table B1). The value of n is smaller where an authorship class has no paper. Authorship classes come from the version 8 career table (papers 1960 to 2024, citations 1996 to 2024). All positions: np, nc. Last author: np_last = npsf minus cpsf and nc_last = ncsf minus ncsf. These are papers on which the person is last but not first or sole author, and their citations. Single author: nps, ncs. Citations per paper of a class are class citations divided by class papers. Regressions of log pay on log papers, with year and career-age controls, give elasticity 0.24 to 0.27 in every specification (Appendix B, Table B3). Source: Open Georgia pay records; build/make_pay_figures.py and build/compute_ugapay.py.

B.3 Screening, college assignment, outcome, and standardization

Screening. A person-year enters the observed base only if its recorded title and pay meet the following conditions. Administrative titles are excluded: dean, associate or assistant dean, department chair or head, director, provost, and chief academic affairs officer. Temporary, retiree, former, limited-term, affiliate, occasional, and part-time classes are also excluded. Gross pay must be at least 50,000 dollars and at least 80% of the previous screened year. A lower year is treated as a partial exit year and excluded. Appointment fraction, contract length, and duties are not observed. Thus, screening approximates a regular full-time academic appointment but does not verify that status, and the earlier full-time label is withdrawn. Persons with administrative titles in some years contribute only their non-administrative years. In the panel, 43 persons had no usable year. Their most frequent titles were TEMPORARY FACULTY RETIREE (49 person-years), DEPT CHAIR/HEAD AC (41 person-years), DEAN AC (15 person-years), ASSOC/ASST DEAN AC (12 person-years).

College assignment. The citation database has no department field, so colleges were first inferred from subfields, with Warnell and Odum corrections, the same inference criticized in the main text (Table A6 in Appendix A). Because college sets the standardization denominator and the dean comparison, every person was then looked up with AI assistance on public university pages, i.e., department and college pages, profiles, news releases and CVs, or in a publication affiliation. Of 224 persons checked, 202 were confirmed and 22 were probable because of joint appointments or indirect or older sources. The inference was wrong for 55 persons, including 32 of the 169 then in the analysis, for example psychology-labelled scientists moving to Family and Consumer Sciences and plant scientists to CAES. The false payroll match and the retiree row without a profile on the citation list were excluded. School of Computing faculty are counted with Franklin. Three small units keep their own group, each with one usable person: the journalism college (Grady), the medical partnership (Medicine), and the CCRC research center. The file college_checked.csv in Appendix C gives each person's college, source, and a short quotation. Colleges and usable persons: Franklin 62, CAES 32, Education 16, Engineering 15, Vet Med 14, Odum 11, Warnell 8, Public Health 6, SPIA 4, FACS 3, Pharmacy 3, Terry 3, CCRC 1, Grady 1, Medicine 1.

Outcome and standardization. The outcome is recorded fiscal-year gross pay. Values were standardized for differences in college pay levels using $\text{pay_std} = \text{pay} / F_c * U$, where F_c is the mean gross pay of observed screened person-years in college c over FY2021 to 2025. $U = 185,676$ dollars is the mean of the same person-years over all colleges. Modeled values enter neither F_c nor U . In the 3 groups with one usable person (CCRC, Grady, Medicine), F_c is that person's own mean over the observed screened years, so their mean standardized observed pay equals U . This transformation removes only differences in college means and does not make positions, contracts, or labor markets comparable. It leaves every within-college ratio unchanged because dean and scientist pay in the same college are multiplied by the same factor. Thus, the dean pay ratios in Figure B1 are raw within-college ratios. Standardization affects only the pooling of scientists across colleges.

Modeled retiree pay. Scientists who retired within the window (27 persons) and before it (26 persons) were retained by projecting their last screened pay forward at the yearly raise of the observed screened scientists of the same college. For fiscal 2016 to 2020, the median observed raise came from the historical extract. The raise was zero for fiscal 2021, and the college-year mean raise, computed with the checked colleges, was used for fiscal 2022 to 2025. These values estimate earnings under assumed continued employment and are not recorded university pay. Retirement was inferred from a missing screened record or a retiree title and was not verified. In every figure, observed persons are plotted using only the mean of their observed screened years. Thus, the 27 in-window retirees appear as observed persons, and their modeled years are not used. The 26 pre-window retirees appear as hollow triangles at the mean of their modeled FY2021 to 2025 values. Five-year means use the years available for each person (1 to 5).

Definitions. Authorship counts come from the version 8 career table in three classes. All positions: np = papers 1960 to 2024 and nc = citations 1996 to 2024. Last author: np_last = papers on which the person is last but not first or sole author ($npsfl$ minus $cpsf$), and nc_last = their citations ($ncsfl$ minus $ncsf$). Single author: nps and ncs . `build/check_authorship_fields.py` recomputes np_last , nc_last , nps , and ncs from the raw version 8 release for all 321 profiles and finds no difference. Shares are np_last / np , nc_last / nc , nps / np and ncs / nc . Citations per paper of a class are class citations divided by class papers. If the class has no paper, this value is undefined and the person is omitted from that panel. The figures add 1 to last-author and single-author counts to show zeros on the logarithmic axis. Log citations per single-author and per last-author paper are undefined for zero papers or zero citations, so those person-years are omitted from regression model F (six person-years have single-author papers but no single-author citations). Lines are fitted by ordinary least squares through observed persons only, using the plotted scales (log-log where both axes are logarithmic).

Deans and the listed scientists of their colleges, standardized pay

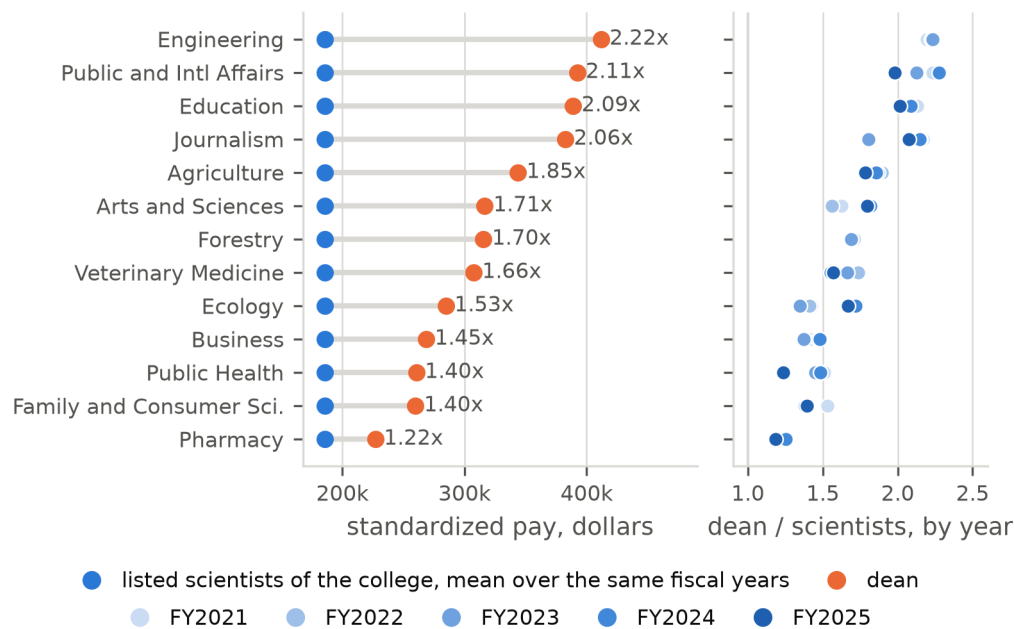


Figure B1: Deans and the listed scientists of their colleges on the standardized pay scale, college-years without a dean transition only. Left: for each college, the mean standardized pay of its observed screened listed scientists over the same clean fiscal years as its dean (blue) and of its dean (orange), joined by a gray bar and labelled with the ratio dean / scientists. In this panel, pay is standardized over those same years, so the scientists of every college are at $U = 185,676$ dollars. Right: the same ratio for each clean fiscal year, FY2021 (lightest blue) to FY2025 (darkest blue); vertical line at ratio 1. Because standardization multiplies a dean and the scientists of one college by the same factor, the ratios equal the raw within-college ratios. Colleges: Engineering, Public and International Affairs (SPIA), Education, Journalism (Grady), Agriculture (CAES), Arts and Sciences (Franklin), Forestry (Warnell), Veterinary Medicine, Ecology (Odum), Business (Terry), Public Health, Family and Consumer Sciences (FACS), Pharmacy. Source: workbook sheet Std dean vs sci; build/make_pay_figures.py.

B.4 Results

The pay-record comparison (Georgia Department of Audits and Accounts, 2025) gives the following results. Of the nominal University of Georgia profiles, 197 matched pay records by name, and 154 had at least one observed year meeting the screening rules. These 154 persons, with pay averaged over their observed years, provide the basis for the observed-pay results reported here, while pooled models also include modeled retiree pay. Table B1 gives the sample flow, and Sections B.3 to B.5 give screening criteria, modeled values, model tables, the dean comparison, and limitations. I standardized each pay value for differences in college pay levels, dividing by the mean for observed listed scientists of the same college over the study window and multiplying by the university-wide mean for that group. Thus, the values remain in dollars and are rescaled to a common college mean. Table B2 gives results for all positions, last author, and single author, where last author means papers on which the person is last but not first or sole author. Fig. B2 gives the scatter plots. Standardized pay goes about equally with papers in all positions (Spearman $\rho = +0.38$, $p < 0.001$, $n = 154$) and last-author papers ($+0.37$, $p < 0.001$), while single-author papers have a weak association ($+0.08$, $p = 0.30$). The elasticity is 0.24 to 0.27 across the regression specifications with year and career-age controls. For citations, the association is strongest for last-author papers ($+0.44$), followed by all positions ($+0.36$) and single-author papers ($+0.04$). Citations per paper have a positive association for last-author papers ($+0.22$, $p = 0.007$), a weak and nonsignificant one in all positions ($+0.11$, $p = 0.16$), and none for single-author papers (-0.04 , $p = 0.67$). The last-author share of papers correlates with pay ($+0.20$, $p = 0.01$) and is significant in the pooled models, but not on observed pay only. Citations per last-author paper are positive under both estimators of model F, also on observed pay only (Table B3 and Section B.5).

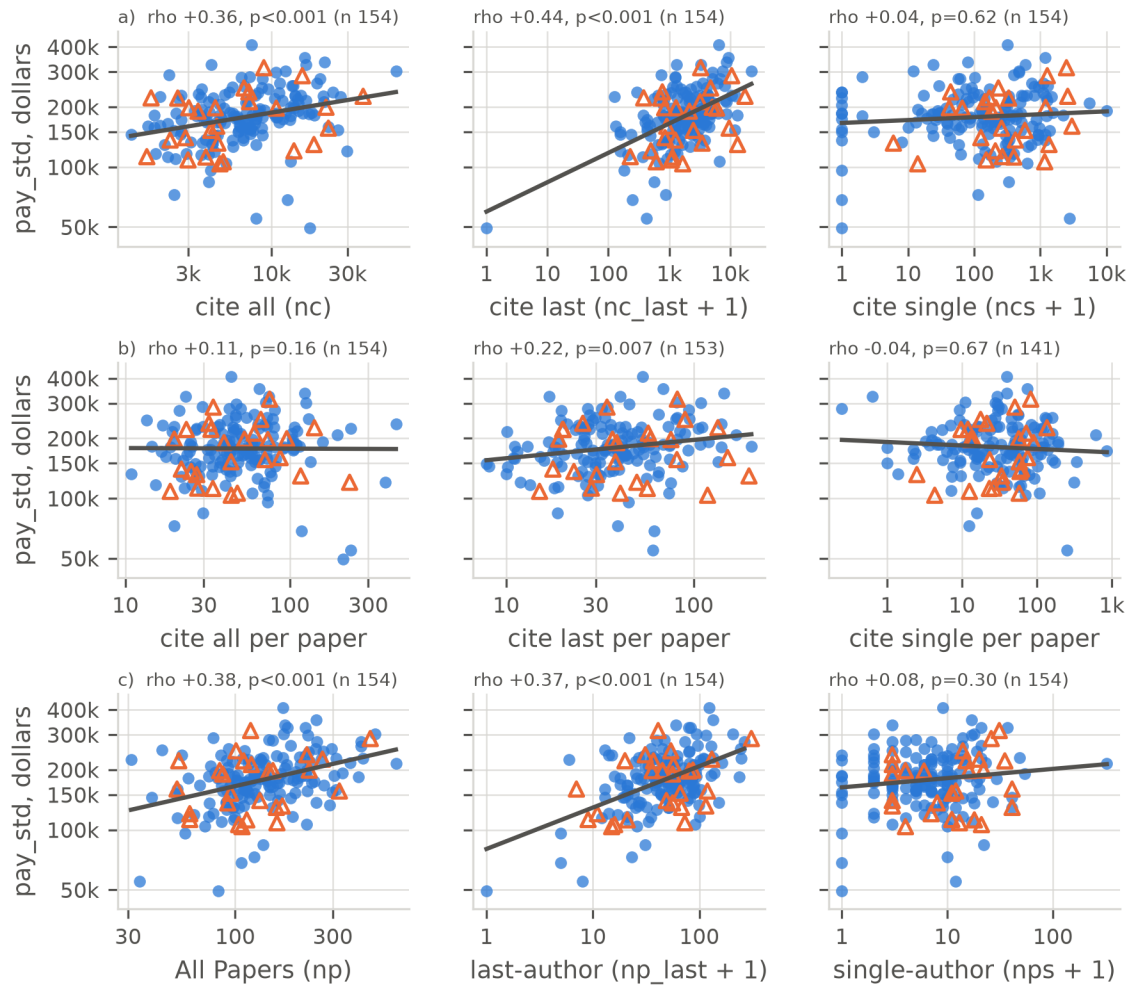


Figure B2: Standardized gross pay of UGA scientists on the 2024 list, by authorship class. Columns from left to right: all positions, last author, single author, as defined in Section B.3. The three rows, a to c: standardized pay (pay_std, dollars), averaged over the person's observed screened fiscal years 2021 to 2025. Pay is compared with citations per author (a), citations per paper (b, omitting persons with no paper in the class), and papers (c). Last-author and single-author counts have 1 added to show zeros on the logarithmic axes. Blue dots: 154 observed persons. Orange hollow triangles: 26 pre-window retirees with modeled pay only. Line: least-squares fit through observed persons. Text above each panel: Spearman rho, p-value, and n for observed persons. Source: workbook sheets Std panel and Authorship indices (321); build/make_pay_figures.py.

Figure B2 compares standardized pay with career citations per author (row a), citations per paper (row b), and papers (row c). Columns are the three authorship classes of Section B.3: all positions, last author, and single author. Figure B1 compares deans with the listed scientists of their colleges. For observed persons, Spearman correlations of pay with citations per author were $\rho = +0.36$ ($p < 0.001$, $n = 154$) for all positions, $\rho = +0.44$ ($p < 0.001$, $n = 154$) for last author, and $\rho = +0.04$ ($p = 0.62$, $n = 154$) for single author. With citations per paper, they were $\rho = +0.11$ ($p = 0.16$, $n = 154$) for all positions, $\rho = +0.22$ ($p = 0.007$, $n = 153$) for last author, and $\rho = -0.04$ ($p = 0.67$, $n = 141$) for single author. With papers, they were $\rho = +0.38$ ($p < 0.001$, $n = 154$) for all positions, $\rho = +0.37$ ($p < 0.001$, $n = 154$) for last author, and $\rho = +0.08$ ($p = 0.30$, $n = 154$) for single author. The largest coefficient change is 0.039 when the persons with modeled pay only are included as well (values in pay_figures.json, Appendix C).

Sensitivity. The associations hold without the 50,000-dollar pay floor, without the rule that removes partial exit years, and without either rule (Table B4). Without both rules, the elasticity rises to 0.30 to 0.35. The one large

Table B3: Complete estimates for models of log standardized pay: coefficient (standard error; p-value). Dependent variable: natural logarithm of pay_std (Section B.3) for each person-year. Model E uses log np, log citations per paper (all positions), log career age, and the four authorship shares. Model F uses log np, log career age, log citations per single-author paper, and log citations per last-author paper. It includes person-years with at least one paper in each class. Every model includes fiscal-year effects, with fiscal 2021 as the reference. Person RI: linear mixed model with a random intercept per person, fitted by restricted maximum likelihood. OLS cl.: ordinary least squares with standard errors clustered by person. Observed: observed screened person-years only (primary). Pooled: observed and modeled person-years together. Coefficients on logged predictors are elasticities. Coefficients on shares give the change in log pay from share 0 to share 1. Year coefficients give log differences from fiscal 2021. Sources: Open Georgia pay records and the version 8 career table; build/compute_ugapay.py (statsmodels).

Term	E observed, person RI	E observed, OLS cl.	E pooled, person RI	E pooled, OLS cl.	F pooled, person RI	F pooled, OLS cl.
Intercept	+10.514 (0.393; p<0.001)	+10.616 (0.340; p<0.001)	+10.619 (0.369; p<0.001)	+10.630 (0.327; p<0.001)	+10.951 (0.316; p<0.001)	+10.764 (0.308; p<0.001)
Fiscal 2022 (vs 2021)	+0.055 (0.009; p<0.001)	+0.055 (0.016; p<0.001)	+0.056 (0.007; p<0.001)	+0.045 (0.010; p<0.001)	+0.058 (0.007; p<0.001)	+0.053 (0.009; p<0.001)
Fiscal 2023 (vs 2021)	+0.075 (0.009; p<0.001)	+0.075 (0.017; p<0.001)	+0.074 (0.007; p<0.001)	+0.062 (0.012; p<0.001)	+0.076 (0.007; p<0.001)	+0.065 (0.011; p<0.001)
Fiscal 2024 (vs 2021)	+0.098 (0.009; p<0.001)	+0.101 (0.018; p<0.001)	+0.094 (0.007; p<0.001)	+0.085 (0.012; p<0.001)	+0.094 (0.007; p<0.001)	+0.087 (0.012; p<0.001)
Fiscal 2025 (vs 2021)	+0.135 (0.009; p<0.001)	+0.158 (0.020; p<0.001)	+0.128 (0.007; p<0.001)	+0.131 (0.013; p<0.001)	+0.128 (0.007; p<0.001)	+0.132 (0.013; p<0.001)
log np (papers, all positions)	+0.272 (0.054; p<0.001)	+0.242 (0.044; p<0.001)	+0.264 (0.049; p<0.001)	+0.262 (0.042; p<0.001)	+0.237 (0.042; p<0.001)	+0.258 (0.045; p<0.001)
log citations per paper (all positions)	+0.084 (0.039; p=0.030)	+0.066 (0.044; p=0.137)	+0.074 (0.035; p=0.033)	+0.054 (0.042; p=0.192)		
log career age (2025 minus first publication year)	-0.119 (0.088; p=0.173)	-0.087 (0.075; p=0.243)	-0.131 (0.078; p=0.094)	-0.113 (0.070; p=0.106)	-0.122 (0.080; p=0.125)	-0.085 (0.075; p=0.259)
Single-author share of papers (nps / np)	+0.213 (0.381; p=0.576)	+0.214 (0.344; p=0.535)	+0.358 (0.331; p=0.279)	+0.400 (0.291; p=0.169)		
Single-author share of citations (ncs / nc)	+0.357 (0.317; p=0.261)	+0.272 (0.262; p=0.299)	+0.202 (0.297; p=0.497)	+0.093 (0.256; p=0.716)		
Last-author share of papers (np_last / np)	+0.356 (0.297; p=0.231)	+0.253 (0.332; p=0.445)	+0.569 (0.268; p=0.034)	+0.615 (0.295; p=0.037)		
Last-author share of citations (nc_last / nc)	+0.276 (0.254; p=0.278)	+0.368 (0.248; p=0.137)	+0.051 (0.229; p=0.824)	+0.011 (0.246; p=0.963)		
log citations per single-author paper					-0.000 (0.017; p=0.990)	+0.005 (0.016; p=0.767)
log citations per last-author paper					+0.094 (0.034; p=0.006)	+0.072 (0.033; p=0.029)
Person-years / persons	614 / 154	614 / 154	818 / 180	818 / 180	759 / 167	759 / 167

change is the SPIA ratio in fiscal 2021, which reaches 3.25 without the floor. One partial year below 50,000 dollars lowers the mean in a college with three listed scientists, which is why the floor is used.

Dean premia. In each fiscal year the university carried 17 to 20 academic deans and 52 to 58 associate or assistant deans. Comparing each dean with the observed screened listed scientists of the same college, in college-years without a dean transition, the ratio of dean pay to mean scientist pay ranged from 1.18 to 2.28: SPIA 1.98 to 2.28; Engineering 2.19 to 2.23; Grady 1.81 to 2.17; Education 2.01 to 2.13; CAES 1.78 to 1.90; Franklin 1.56 to 1.82; Vet Med 1.55 to 1.74; Odum 1.35 to 1.72; Warnell 1.69 to 1.71; FACS 1.38 to 1.53; Public Health 1.23 to 1.51; Terry 1.37 to 1.48; Pharmacy 1.18 to 1.25. The comparison is with listed scientists, not with the faculty of the college; the scientist means use observed screened person-years only (in Figure B1 over the same fiscal years as the dean); and differences between colleges may reflect market structure or joint appointments. In raw dollars, the dean of the forestry school (Warnell) was paid 224,573 dollars in fiscal 2021 against a mean of 131,398 for the 6 listed Warnell scientists with observed pay (1.71 times), 229,573 dollars in fiscal 2022 against a mean of 135,847 for the 6 listed Warnell scientists with observed pay (1.69 times), 235,823 dollars in fiscal 2023 against a mean of 139,618 for the 5 listed Warnell scientists with observed pay (1.69 times); fiscal 2024 and 2025 are dean-transition years and are excluded. The listed scientists are not named in this appendix.

Table B4: Sensitivity to the pay floor and partial-exit-year rule, with checked colleges in every variant. The whole pay pipeline is rerun for each variant.

Measure	Base	No floor	No exit rule	Neither
Observed persons	154	157	154	157
Papers rho, all positions	+0.38	+0.35	+0.37	+0.35
Papers rho, last author	+0.37	+0.35	+0.37	+0.35
Papers rho, single author	+0.08	+0.05	+0.06	+0.02
Citations per paper rho, all positions	+0.11	+0.13	+0.13	+0.13
Citations per paper rho, last author	+0.22	+0.22	+0.23	+0.21
Citations per paper rho, single author	-0.04	-0.02	-0.02	-0.02
Elasticity range	0.24 to 0.27	0.27 to 0.33	0.24 to 0.27	0.30 to 0.35
Model F on observed pay, citations per last-author paper, person random intercept	+0.125 (p < 0.001)	+0.142 (p = 0.02)	+0.126 (p < 0.001)	+0.140 (p = 0.01)
Model F on observed pay, citations per last-author paper, OLS clustered	+0.116 (p < 0.001)	+0.122 (p < 0.001)	+0.116 (p < 0.001)	+0.122 (p < 0.001)
Dean ratio range	1.18 to 2.28	1.18 to 3.25	1.18 to 2.32	1.18 to 3.25

Note. Columns give the screening variants. Base: 50,000-dollar floor and 80% exit rule. No floor retains the exit rule. No exit rule retains the floor. The Neither variant removes both rules. The rho rows are Spearman correlations of mean standardized observed pay with papers and with citations per paper, one row per authorship class. The elasticity range is the range of the log np coefficient over the six Table B3 specifications. The two model F rows give the coefficient on log citations per last-author paper, with its p-value, separately for the person random-intercept estimator and OLS with errors clustered by person. The dean range covers clean annual within-college ratios. Source: build/pay_sensitivity.json; build/pay_sensitivity.py.

B.5 Limitations and the hypothesis

(1) College was checked for every person, but joint appointments and the 22 probable cases may still sit in the less relevant college, which would change the standardization denominator and the dean comparison. (2) The sample contains listed profiles with matched pay rather than UGA faculty, and unmatched profiles were not verified. (3) Gross pay is not a salary rate, and screening only approximates a regular appointment. (4) Modeled retiree pay is included in the pooled models. Observed-pay results are the observed columns of Table B3, the blue markers of Figure B2, and the dean comparison of Figure B1. (5) Career records through 2024 are compared with pay in fiscal 2021 to 2025. This cross-sectional comparison between persons cannot determine cause and effect. (6) Last-author position has not been validated here as a measure of leading the work reported in a paper. (7) Nonsignificant coefficients do not show that an effect is absent. (8) Academic rank was not used as a variable. Rank tracks career output, the last-author share, and pay, so the volume and last-author associations could arise from promotion alone. Of the 154 persons with observed pay, 133 held the title Professor in every observed year, 15 held the title Associate or Assistant Professor in at least one observed year, and 6 held research titles.

The results show a positive association between publication volume and standardized pay across the reported specifications. The elasticity of log np is between 0.24 and 0.27, including the observed-pay models. In the rank correlations, papers were associated about equally with pay in all positions and the last-author class, while citations and citations per paper were strongest in the last-author class. The single-author class had the weakest associations, and the lowest value (-0.04, p = 0.67) was for citations per single-author paper. The differences between these correlations were not tested. Associations with authorship shares and citation-intensity measures depend on the predictor, sample, and estimator. For last-author share of papers, the observed-pay coefficients are +0.36 and +0.25, with p = 0.23 and 0.45. The pooled models give p = 0.03 and 0.04. In model F, citations per last-author paper have coefficients +0.094 and +0.072, with p = 0.006 and p = 0.03. This is an association with citation intensity. Fitted to observed pay only (564 person-years of 141 persons), model F gives +0.125 and +0.116 for citations per last-author paper (p < 0.001 and p < 0.001), and about zero for citations per single-author paper (p = 0.60 and 0.90). These initial associations do not identify pay criteria or show a preference for group leadership over per-paper impact. Testing hypotheses about pay criteria requires verified appointments, salary rates, an unselected faculty population, and a longitudinal design, and no such test is made here. The results are consistent with the hypothesis that institutional rewards give more weight to volume and to last-author work than to the per-paper impact of single-author work, but other hypotheses are also possible.

APPENDIX C: REPRODUCIBILITY PACKAGE AND PROGRAM LISTINGS (APPENDICES D AND E)

Appendix C is the archive `Cieszewski2026_AI_AppendixC_package.zip` published with this paper (Cieszewski, 2026). It contains the unchanged cohort extract, `StanfordList_Forestry_Ecology_career_2020-2024.xlsx`, with 23,374 author-release records from versions 3, 5, 6, 7, and 8 of Elsevier Data Repository record 10.17632/btchxktzyw. It also contains the whole-database country-by-subfield-by-year file, `world_subfield_year.csv`, and the original outputs used for verification: `auth6.json`, `authorship.json`, `integrity.json`, and `refs.json`. The record includes the text of the broad report (S1), the final report, the handoff, and a running log kept during the work, with instructions quoted or paraphrased. Original figure files are included with their hashes, among them `ff4_group.png`, the unused count-share plot that an earlier draft of this paper mistook for the report figure (Section A.5). The pay workbook of Appendix B, the checked college list `college_checked.csv` with its sources, and the original, unedited analysis programs are also included. Programs that recompute every numerical table cell and figure of this paper are `compute_normalization.py`, `revision_checks.py`, `nash_search.py`, `make_figures.py`, `manuscript_content.py`, `build_tex.py`, and `verify_numbers.py`. Those for Appendix B are `redo_standardized.py`, `compute_ugapay.py`, `make_pay_figures.py`, `pay_sensitivity.py`, and `appendix_a.py`, with `paynames.py` as their shared name key. Generated outputs are `statistics.json`, `normalization.json`, `annual_ri.csv`, `revision_checks.json`, `nash_search.json`, `rank_vectors.json`, `ugapay.json`, `pay_figures.json`, `pay_sensitivity.json`, and `verify_output.txt`. SHA-256 hashes are provided for every file. The README gives the field-name mapping and selection conditions.

Appendix D (`Cieszewski2026_AI_AppendixD_programs.pdf`) lists the verification programs verbatim. Appendix E (`Cieszewski2026_AI_AppendixE_original_pipeline.pdf`) lists the 86 original analysis programs verbatim. Both are separate files because of their length. The run order is `compute_normalization.py`, `revision_checks.py`, `nash_search.py` (requires the raw releases), `make_figures.py`, `redo_standardized.py`, `compute_ugapay.py`, `make_pay_figures.py`, `pay_sensitivity.py`, and `appendix_a.py`. The source text of this paper is not included, so `build_tex.py` and `verify_numbers.py` cannot be rerun from the package. The output of the final check, `verify_output.txt`, ends with RESULT PASS.