Does Peer Review Penalize Scientific Risk Taking? Evidence from NIH Grant Renewals*

Pierre Azoulay
MIT and NBER
100 Main Street
E62-487
Cambridge, MA 02142

Wesley H. Greenblatt
Boston Children's Hospital
and Harvard Medical School
300 Longwood Ave.
Boston, MA 02115

February 11, 2025

Abstract

Scientific projects that carry a high degree of risk may be more likely to lead to breakthroughs yet also face challenges in winning the support necessary to be carried out. We analyze the determinants of renewal for more than 100,000 R01 grants from the National Institutes of Health between 1980 and 2015. We use four distinct proxies to measure risk taking: extreme tail outcomes, disruptiveness, pivoting from an investigator's prior work, and standing out from the crowd in one's field. After carefully controlling for investigator, grant, and institution characteristics, we measure the association between risk taking and grant renewal. Across each of these measures, we find that risky grants are renewed at markedly lower rates than less risky ones. We also provide evidence that the magnitude of the risk penalty is magnified for more novel areas of research and novice investigators, consistent with the academic community's perception that current scientific institutions do not motivate exploratory research adequately.

Keywords: scientific risk, scientific productivity, peer review, NIH, government funding.
JEL Classification: O32, O38, H51

* We thank Alessandro Bonatti, Matt Clancy, Danielle Li, and Scott Stern, as well as seminar participants at the Max Planck Institute for Innovation and Competition, NBER Summer Institute, and NBER Productivity Seminar for useful feedback.

1. Introduction

No funder of science has ever prided itself on supporting "low risk, low reward" projects, yet few ideas elicit as much agreement as the claim that peer review punishes risk taking. These laments appear in the popular press, in U.S. Congressional testimony, errant tweets, or the pages of *Nature*, and are authored by successful investigators and leading science policy makers (Kolata 2009, Kornberg 2007, Nielson 2022, Sutter 2022, Woolston 2014). Prominent scientists have penned editorials satirizing the behavior of risk-averse grant funders, and ponder whether Nobel Prize winners could receive funding for their seminal discoveries in the modern funding climate (Fields 2014, Petsko 2012). While economists would caution that a benevolent social planner might want to tame individual scientists' risk appetites, the proposition that current institutions do not adequately reward risky research strikes practicing scientists as self-evidently true.

Yet, for all the colorful anecdotes, blue-ribbon panels, and editorials, there is surprisingly little evidence to support the claim that peer review punishes risk taking, and less evidence still that society as a whole would be better served if funders could clamp down on their conservatism. Azoulay et al. (2011) compare two grant funding mechanisms—the Howard Hughes Medical Institute Investigatorships and National Institutes of Health (NIH) R01 grants—and provide evidence that the former encourages more risk taking than the latter, but this does not necessarily imply that peer review at NIH punishes risk taking. Other scholars have examined how novelty shapes grant funding outcomes (Ayoubi et al. 2021, Packalen and Bhattacharya 2020, Veugelers et al. 2022). However, novelty is distinct from risk taking: investigators can pursue novel research which isn't risky, and risky research which isn't novel. What is widely agreed upon is that grant funders' tolerance for risk has direct implications for what research is undertaken in the first place. In a survey of investigators supported by *Fast Grants*, 78% stated they would make significant changes to their research program if they were not constrained by the demands of grant funding agencies, including pursuing more ambitious research programs, pivoting to new topics, and testing hypotheses others see as unlikely to succeed (Collison et al. 2021).

More generally, the choice of what scientists choose to study is fundamental to determining the rate and direction of innovation. An assumption typically made by models of the innovation process is that exploratory projects are less likely to bear fruit than projects that merely seek incremental advances, but with more upside if they are indeed successful (Manso 2011). This is the sense in which science funders always claim to crave "high risk, high reward" projects even though the trade-off between risk and reward in the market for scientific ideas is less obvious than in financial markets (Nielsen and Qiu 2022). While both risky exploratory research and incremental lines of inquiry are valuable from a societal point of view, a first-order concern in contemporary science policy is that funding programs tilt the mix of projects they support too far in the direction of incrementalism. Moreover, the societal costs of undersupplying risky research may be especially prominent for early-stage, basic research, which may lead to substantial knowledge spillovers (Azoulay and Li 2022).

What causes individual scientists to select risky projects? Rather than being solely determined by an individual's preference for risk taking, this choice is shaped in important ways by the incentives, organization, and supporting institutions that surround a researcher. Tolerance for failure or lack thereof (Tian and Wang 2014), time horizons (Lerner and Wulf 2007), team structure (Wu et al. 2019), and the structure of financial incentives (Graff Zivin and Lyons 2020) shape how much risk innovators choose to bear when selecting projects. The choice of research approach is made more challenging by the difficulty in predicting the ultimate benefits if the project is successful; this is especially the case for early stage and highly uncertain projects (Rosenberg 1995).

In this paper, we study how a specific institution, grant peer review at NIH, shapes scientific risk taking. In grant peer review, scientific proposals are evaluated by a panel of experts, with the resulting scores having a major influence on the likelihood of funding. Scientists who are awarded funding use these resources to conduct the investigations they committed to in their proposal. At the end of the grant cycle, a funded scientist has the option of extending the research program by applying for an additional cycle of support. To study the relationship between peer review and risk taking, one would ideally have access to all proposals (regardless of funding status), reviewer scores, and grant outcomes in the form of publications and citations that can be traced to a particular funding stream.[1] With access limited to publicly available data, we focus on funded applications, and ask whether risk taking is associated with the likelihood that a grant is renewed competitively for an additional cycle. Since we are studying the impact of risk taking conditional on a successful initial grant application, the estimates we present should be construed as a lower bound on the risk taking penalty that would be observed if we had access to the full set of grant applications, included those that were not supported. We also link each grant cycle with the publications they produced, leveraging the requirement that authors acknowledge their NIH grant support in their published work.

R01 grants are the "bread and butter" of academic biomedical research, and a requirement for scientists who aspire to establish and make viable their laboratory over the long run. We study how risk taking is penalized or rewarded in 103,164 NIH-funded R01 and R01-equivalent grant cycles between 1980 and 2015. After controlling for a comprehensive set of covariates for investigator demographics, institution quality, and grant characteristics, we estimate the relationship between risk taking and whether the grant was competitively renewed. Scientific risk taking, much like scientific creativity, is a nebulous concept. There is no consensus on what constitutes risk in science nor how it should be measured (Althaus 2005, Franzoni and Stephan 2023). Since it is unlikely that a single measure can fully represent the phenomenon, we develop four distinct measures of risk taking, each of which captures a different aspect of the risks taken

[1] These data are typically not made available to researchers. NIH employees have access to funded and unfunded grant applications and average scores; even these insiders do not have access to the full distribution of scores at the proposal level, as these scores are destroyed in order to protect reviewer confidentiality.

or borne by scientists, and examine if our findings depend on the particular measure used. All four measures rely on the realized outcomes of the initial grant cycle to characterize the risk profile of subsequent cycles. A maintained assumption in our analysis is therefore that risk taking is imprinted on a particular scientific trajectory at the initial stage, and largely persists across future cycles.

One way to conceptualize risk is through variability in outcomes—projects that produce either groundbreaking breakthroughs or complete failures (Azoulay et al. 2011). Another perspective links risk to disruption, distinguishing research that challenges the status quo from work that consolidates existing knowledge (Funk and Owen-Smith 2017). A third approach considers risk in terms of intellectual movement, where scientists pivot away from their prior work to explore new directions, emphasizing that risk stems not just from the project itself but from who undertakes it. Finally, risk may manifest in intellectual distinctiveness, as researchers choose projects that diverge sharply from those of their peers, making their successes—or failures—particularly visible to the scientific community.

Across each of these measures, we find that risk taking is penalized. When comparing grant cycles in the top and bottom decile of risk taking, grants with greater risk taking have a 9.5% lower renewal rate (20.5% decline) when measuring risk taking using extreme tail outcomes, an 11.1% lower renewal rate (24.4% decline) when measuring risk taking by its disruptiveness, a 6.9% lower renewal rate (15.2% decline) when measuring risk taking by an investigator pivoting from her prior research, and a 4.0% lower renewal rate (9.5% decline) when measuring risk taking by standing out from what other investigators are studying. In contrast to our measures of risk taking, novelty is associated with *higher* grant renewal rates in our data. At the very least, this finding buttresses our claim that novelty and risk taking are different concepts that it is best not to conflate.

Of course, it may well be that the penalty applied to riskier projects is appropriate, given the real-world constraints faced by funders. For public funding at least, the committees and officials doling out funding to individual proposals are accountable to political paymasters, and may feel pressure to present evidence of accumulating and steady success, rather than a series of failed projects occasionally punctuated by an extraordinary outcome (Clancy 2023). We present two separate pieces of evidence consistent with the claim that funders penalize risk even beyond what a benevolent social planner might wish: First, the gradient in the renewal-risk taking relationship appears steeper for novel proposals, relative to less novel ones. Second, for renewed grant cycles, those we measure as taking on more risk lead to elevated rates of citations of the renewed grants' publications in patents, relative to the renewed cycles we measure as taking on less risk.

Our study goes beyond reaffirming the widely held belief that peer review favors conservative research over high-risk, potentially higher-reward projects. By offering systematic evidence of this bias, we take the conversation one step further, entertaining the possibility—as

ridiculous as it may sound to practicing scientists— that the heightened scrutiny placed on risky projects is not entirely misplaced. While society may benefit from imposing a higher bar on risky endeavors—filtering out failures that are not necessary milestones on the way to promising breakthroughs, our findings suggest that the penalty may be excessive, particularly in novel research domains where bold exploration is most likely to pay off. Our results also imply that peer review, far from being an immovable pillar of the scientific process, is a powerful lever for policymakers to shape the kinds of challenges researchers pursue—and it can be reformed to better balance the trade-off between risk and societal benefit, steering innovation toward the breakthroughs that matter most.

Prior studies have explored how aspects of a research proposal influence the evaluation and funding of scientific projects and publications. This includes novelty (Ayoubi et al. 2021, Boudreau et al. 2016, Packalen and Bhattacharya 2020, Teplitskiy et al. 2022, Veugelers et al. 2022), interdisciplinarity (Banal-Estañol et al. 2019, Bromham et al. 2016), and the interplay between novelty and feasibility (Krieger and Nanda 2022, Lane et al. 2022a). Our study builds on this literature in several ways.

First, we introduce an eclectic set of risk-taking measures, each capturing a different dimension of risk, and explicitly distinguish between risk and novelty in the analysis. Second, rather than characterizing an investigator's prior research trajectory to predict future funding—an approach common in previous studies—we focus on risk measures directly tied to a specific grant proposal and the publications emerging from it. This distinction is important because researchers often manage multiple projects simultaneously, each carrying different levels of risk. An investigator's past work may not accurately reflect the level of risk in a particular grant proposal, and in some cases, scientists may deliberately avoid proposing high-risk ideas to funders they suspect might penalize bold experimentation, instead pursuing such projects through alternative funding sources.

Our closest intellectual antecedents are Boudreau et al. (2016), who experimentally assign grant proposals to peer evaluators, and Carson et al. (2023), who manipulate the distribution of reviewer scores to assess how peers (rather than funders) value disagreement among reviewers. Like ours, these studies exploit proposal-level variation to uncover the determinants of funding success. Unlike ours, the studies control the treatment of interest in a framed experimental design (List and Metcalfe 2014)—an unambiguous methodological strength. However, these studies do not measure proposal risk, nor can they replicate the high stakes of NIH grant funding competitions. Our approach, by contrast, relies on observational data. While this limits our ability to identify causal effects, it allows us to examine risk-taking as the primary treatment of interest in a real-world setting of substantial policy interest—namely, the entire universe of R01 NIH grants over an extended time period.

The remainder of the paper is structured as follows. Section 2 outlines our argument for the central role of risk taking in scientific research and examines how, or whether, funding peer

review influences the level of risk accepted by investigators, evaluators, and funders. Section 3 introduces the setting, data, and addresses measurement challenges. In Section 4, we describe our empirical strategy and provide descriptive statistics. Section 5 presents the results of our analysis. Finally, Section 6 concludes by discussing the broader implications of our findings.

2. Scientific risk taking and peer review

There is no consensus on what constitutes risk in science (Althaus 2005, Franzoni and Stephan 2023). Following how the term is frequently used in scientific vernacular, we use risk to mean uncertainty surrounding the probability that a scientific project can be successfully brought to fruition. Failure, however, pervades the scientific research process. While some degree of risk is inherent in nearly all innovation projects, a central challenge facing projects with high levels of risk is both that they are less likely to succeed, and that conditional on success, it may be more difficult to predict their ultimate impact (Rosenberg 1995).[2]

As Franzoni and Stephan (2023) explain, risk in science is a multilayered concept. Different types of risk in science include technical or scientific risk stemming from nature's answer to the research question being asked, execution risk in being able to successfully carry out the planned experiments, competitive risk from being "scooped" by other investigators studying the same or similar question (Hill and Stein 2024), and risk in how a successfully completed project will be perceived by peers or advance the investigator's career.

To illustrate how projects may entail different aspects of risk, consider a few examples. The efforts to synthesize insulin using recombinant DNA technology in the late 1970s was perceived to carry limited scientific risk but significant execution and competitive risks. While the foundational work by Cohen and Boyer (Cohen et al. 1973) had already established the basic techniques for gene splicing, challenges in gene sequencing, protein expression, and proper folding presented substantial execution hurdles. Genentech and UCSF led the intense competition to overcome these challenges first and this scientific contest not only transformed diabetes treatment but also heralded the biotechnology era (Hall 1987). In contrast, early efforts to develop cancer immunotherapy entailed very substantial scientific risk, as it was perceived to be very challenging to develop treatments targeting cancer cells but not normal tissue, but they faced arguably lower levels of competitive risk (Littman 2015). Similarly, designing processes to encourage handwashing may carry little technical risk, but entailed career risk for Ignaz Semmelweis, whose empirical observations conflicted with the established scientific and medical opinions of the mid-19th century (Nuland 2004).

Why would individual scientists or society seek out research projects that involve risk? One common justification is the assumption that the probability of success and the impact of

[2] Of course, the returns or impact of innovative projects are uncertain as well, but we find it conceptually useful to not conflate risk and return in what follows.

scientific results are often inversely related: higher-impact discoveries tend to come with lower chances of success. However, unlike in finance, this relationship is probably not systematic, as there are celebrated cases where both high value and high probability of success coexist. For instance, the Human Genome Project had a very high likelihood of success when it was launched given previous advances in sequencing, and its outcomes were anticipated to be highly impactful for both human health and future scientific progress. Additionally, the value of failure in science weakens the analogy to financial risk and return. Failed experiments often generate insights that increase the expected returns of future research, though individual scientists or firms may struggle to capture that long-term value (Frankel et al. 2023). Despite these nuances—and the limited empirical evidence of the risk/return tradeoff in scientific research—this study will maintain the assumption of a positive correlation between scientific risk and the expected importance of the results.

A scientist's choice of projects incorporates their subjective assessment of risk with their judgement of the value (to their own career and to society) of the project's potential impact. Why would scientists ever propose projects whose impact is *ex ante* not commensurate with their perceived probability of success? One possible explanation is that researchers are simply overconfident, holding optimistic beliefs about their projects' probability of success or the value of their potential results (Gan et al. 2012, Sanchez and Dunning 2023). But a countervailing influence is that of funding institutions. Long gone is the era of the "gentleman scientist" who could, like Charles Darwin, fund his own investigations. Undertaking scientific research almost invariably requires harnessing the support of a financial backer, whether the source of funds is a government agency, a philanthropic foundation, or the private sector. While researcher optimism suggests that science funders might justifiably limit risky projects to align with the goals of a benevolent social planner, the topics scientists ultimately pursue are shaped by the risk tolerance and priorities of their funding sources.

In academic research, peer review is the primary mechanism for allocating research funds. This raises an important question: does peer review bias funders toward or away from risky projects? A reasonable starting point is that funders, by supporting a portfolio of projects with largely independent probabilities of success, are better positioned to absorb risk than individual scientists (Arrow and Lind 1970). Yet, despite the frequent rhetorical emphasis on funders' tolerance for failure and enthusiasm for "high risk/high reward" projects, there are several reasons to suspect that risk aversion may permeate the peer review process: from the level of individual investigators seeking funds, to the level of peer reviewer evaluating proposals, to review committees, and even to the funding institutions themselves (Azoulay and Li 2022, Franzoni et al. 2022).

Peer review may shape the degree of risk investigators choose to pursue through several mechanisms. Many academic labs rely on external grants to fund their operations. An inability to win external funding may result in a lab's closure, laying off dependent personnel (Tham et al. 2024), and loss of the investigator's job and potentially career (Ruben 2017). To the extent

investigators *perceive* grant peer review to punish risk taking, this will lead them to instead propose safer, more incremental projects (Kent 2018, Langer 2012, Nielson 2022). Additionally, the intense competition for funds that is the byproduct of a system in which the vast majority of proposals will not receive support may push investigators to propose projects with relatively quick pay-offs (Alberts et al. 2014, Fang and Casadevall 2015). In short, peer review may shape investigators' incentives for risk taking even in the absence of any actual bias against risk for the evaluators and funders.

The penalization of risk taking may also occur at the level of peer reviewers. Individuals may have a psychological bias against the uncertainty that characterizes risky research compared to more incremental projects (Fox and Tversky 1995, Mueller et al. 2012). In line with this, evaluators of research proposals have been found to give lower scores to proposals with greater novelty (Boudreau et al. 2016) and to emphasize feasibility (Krieger and Nanda 2022, Lane et al. 2022a). The bias against novelty may also lead to delayed recognition of the underlying potential of a scientific work (Wang et al. 2017).

Peer review is conventionally undertaken in a committee. Bias against risk taking can occur due to both committee processes and committee membership. Proposals with greater risk taking may have greater variance in reviewer assessment of their quality (Langer 2012). Committee processes that emphasize consensus and uniformly high evaluations across its members to be funded may penalize risk taking (Carson et al. 2023).[3] Similarly, given the wider variation in evaluations of risky projects, processes where each member independently fully evaluates a proposal prior to the committee meeting may be more likely to result in identifying a judge who enthusiastically supports the project compared to processes where one or two committee members have primary responsibility for evaluating each proposal and then present their recommendation to the committee. Exposure to negative assessments of other committee members prior to the final judgment my lead evaluators to revise their own assessment downwards more so than exposure to positive assessments leads them to revise their assessments upwards, biasing against projects with more disagreement (Lane et al. 2022b).

Finally, some aspects of the potential for peer review to penalize risk taking emerge at the level of the funder and innovation ecosystem. Funders can influence risk taking through how they structure committee procedures and membership, set expectations for preliminary evidence, and establish explicit evaluation criteria. Pressure to show political paymasters that administrators' are effective stewards of money can also tilt the selection of projects towards more conventional or incremental approaches (Lorsch 2015).

Of course, many funders think about risk at the level of their entire portfolio of projects, rather than on a grant-by-grant basis. Funders sometimes manage riskier projects using distinct

[3] For instance, a process where a proposal's overall score is simply the average of each judge's evaluation will emphasize consensus more than having scores aggregated after dropping the lowest judge score(s) or a process which emphasizes having a few judges with very high levels of support (Gius 2024).

mechanisms or organizational models, with explicitly different evaluation criteria, while applying a more conservative filter in other parts of their project portfolio (Azoulay et al. 2019a).[4] While there are portions of the NIH portfolio where risk taking is not the goal, such as P01 grants, some clinical trials, or support for research cores where there is more of a focus on operational excellence, we focus on R01 grants, the primary NIH funding mechanisms for biomedical research aimed at extending the scientific frontier and a setting where an excessive penalty for risk taking might plausibly slow scientific progress.

Since beliefs regarding a project's probability of success and its potential impact are unobservable, under what conditions might it be feasible to show that peer review *excessively* punishes risk taking? Under the maintained assumption that risk and return are positively correlated in the scientific setting, we argue that while a benevolent social planner might in some circumstances wish to discourage risk taking, it would seem perverse for her to do so when a project's expected impact is especially uncertain.

A useful way to identify cases where penalizing risk-taking is socially inefficient is to consider how informative prior assessments are about a project's potential value. When prior beliefs are less certain—meaning the initial expectations about a project's impact are more diffuse—punishing risk-taking is more likely to stifle valuable scientific exploration. This variation in the strength of prior beliefs offers a way to isolate instances where excessive conservatism in funding decisions is particularly costly. When are prior beliefs relatively more uncertain? One case is when a scientist is early in their career and has a limited track record, making it harder to form strong expectations about their projects' likely outcomes. Another occurs in highly novel research areas, where prior knowledge is sparse, and the value of success is more difficult to predict. Because novel research proposals carry greater option value, funders should, in theory, be more tolerant of risk in emerging fields than in well-established areas of inquiry. Certainly, a finding that risk taking is penalized more strongly for novice investigators or novel research areas appears difficult to reconcile with a reasonable model of optimal scientific risk taking.

Recognizing that the interaction between risk and uncertainty in research returns helps isolate cases of excessive penalization of risk-taking also underscores a crucial distinction: riskiness and novelty are not the same. It is easy to see why risk and novelty are often conflated: truly groundbreaking scientific breakthroughs frequently involve both. A revolutionary idea (high novelty) often carries a high chance of failure (high risk) simply because it ventures into the unknown. Conversely, research that is very safe (low risk) is often incremental and thus less likely to be novel. However, researchers can take risks without being novel, just as they can pursue novelty without taking large risks. Consider, for example, the notion that a bacteria (*H. pylori*)

[4] The NIH has created a number of distinct funding mechanisms with explicit mandates around high-risk, high-reward research, such as the Pioneer Award and Director's Transformative R01 Award. In 2021, the NIH funded 106 such awards with an expected accompanying outlay of $329 million over five years (NIH 2021). This represents only 1.4% by number of awards and 1.8% by funding amount of all R01-equivalent awards that year at the NIH (Lauer 2022).

causes stomach ulcers, the hypothesis proposed by Marshall and Warren (1984) that went against decades of medical dogma blaming stress and acid. Initially, their work seemed risky in terms of acceptance (many peers were skeptical), but the experiment to prove it was straightforward: they identified the bacteria in patients and Marshall even ingested it himself to show it caused gastritis. While bold (and a bit personally risky health-wise!), this was a relatively direct test of a novel hypothesis—not a massive, uncertain project (Blaser 2005). In contrast, Judah Folkman's early research on the role of angiogenesis in cancer faced extreme skepticism (Stephenson et al. 2013). Because his findings were viewed as highly uncertain and technically challenging to build upon, few scientists were willing to engage with them for over a decade—by which time follow-on work may no longer have been considered novel.

Further complicating the relationship between novelty and risk, scientists' careers depend on how their contributions are perceived by peers. As Merton (1973) notes, employment and promotion decisions reward originality. A researcher who consistently works on problems deemed insufficiently "cutting edge" may face significant career risk, even if their chosen projects are technically safe. Likewise, even a risk-averse funding agency has a mandate to pursue new ideas, which may lead it to favor "safe" forms of novelty—projects that appear innovative without being truly high-risk.

We view the study of scientific novelty as essential in its own right. However, research in this area should (i) carefully distinguish between novelty and risk-taking, (ii) develop empirical measures that do not conflate the two, and (iii) treat any correlation between them as an object of empirical interest rather than a taken-for-granted assumption. The present study endeavors to abide by these principles.

3. Setting, data, and measurement

The National Institutes of Health

The NIH plays a prominent role in funding biomedical research. In 2012, it comprised 27% of all biomedical research funding in the United States and nearly two-thirds of all public and philanthropic research funding (Moses et al. 2015). In 2021, the NIH spent $32 billion to support extramural research projects (Lauer 2022). While the NIH funds clinical research (including clinical trials, health services research, and behavioral studies), the bulk of their funding supports basic research aimed at advancing the frontier of scientific understanding. With research project grants having an average size of over $580,000 in 2021, NIH grants are a major source of funding for the researchers they support. Not only does this funding play an important role in supporting the purchase of materials and the salaries of lab personnel, many academic researchers are in "soft money" positions, funding their own salary out of awarded grants. The ability to successfully raise grant funding is an important metric upon which academic employment and promotion decisions are made.

Understanding the process the NIH uses to evaluate and fund grant applications can provide useful context for our empirical analysis. Briefly, the NIH is comprised of 27 separate institutes or centers (IC) which are typically organized around a set of diseases (e.g., National Cancer Institute) or organ systems (e.g., National Heart, Lung and Blood Institute). Each IC is responsible for funding research that is potentially relevant to its mission. Scientific evaluation of grant proposals occurs primarily in standing review committees known as study sections (e.g., "Cellular and Molecular Immunology" or "Atherosclerosis and Vascular Inflammation"). Each study section may review and score grants from multiple ICs, which are then funded by the ICs in order of their study section score until the IC exhausts its funds.[5] The amount of resources devoted to reviewing grant proposals—both from the NIH with staff to manage the process and from study section members—is substantial. In 2021, the NIH reviewed nearly thirty-eight thousand applications for R01-equivalent grants (Lauer 2022).

The NIH is an attractive setting for studying how peer review might respond to risk taking. First, professional norms lead most academic research to be published in peer-reviewed journals (Dasgupta and David 1994). When combined with rich bibliometrics in biomedicine, this provides a robust "paper trail" we can follow and characterize to measure the degree of grant risk taking. Second, as NIH research is a source of significant improvements in health and positive economic spillovers, understanding how the NIH may more effectively utilize its funds to support breakthrough innovation is of considerable policy interest (Azoulay et al. 2019b, Fleming et al. 2019).

Data

As we do not observe grant applications, we focus on whether an awarded grant was renewed. We rely on the NIH's Consolidated Grant Applicant File to identify all those researchers who received an NIH grant, 1980-2015. We start our analysis in 1980 as we are not able to reliably link publications to the NIH grants which supported them before this date, and end in 2015 to allow sufficient time for follow up to observe research outcomes and renewal status of grants awarded in that year. We limit our analysis to R01-equivalent grants.[6] These investigator-initiated grants are a major funding mechanism to support the work of principal investigators who have achieved career independence and run their own laboratory. We also exclude grant awards from special emphasis study sections, which may not have an option for grant renewal, and grants without associated publications for which many of our measures of risk taking would be undefined.

[5] Institute directors do have discretion to fund applicants out of order considering factors such as fit with the institute's mission, the overall portfolio of projects being supported, and their own evaluation of application quality.

[6] We consider R01-equivalent grants to be those with an NIH activity code of R01 or R37. We exclude several grant mechanisms considered by the NIH to be R01-equivalent but which may have distinct dynamics around risk taking. For instance, we exclude DP1, DP2 and DP5 which are part of the NIH's effort to fund high risk, high reward projects; R56 and RL1 which may not be applied for by the investigator; U01 in which NIH staff actively support the research; and R35 and RF1 which provide longer-term support to investigators.

We focus our analysis on the grant cycle, which corresponds to a single competitive review of a funding proposal. A cycle starts the year a grant proposal is funded and ends either at the conclusion of grant funding or when a proposal for grant renewal is competitively reviewed and funded. As grants may have multiple principal investigators and investigator characteristics are explicitly considered during renewal decisions, we analyze our sample at the investigator-grant-cycle level.[7] This gives a final sample of 103,164 investigator-grant-cycles from 63,101 grants and 37,222 investigators.

Our publication data is from *PubMed*, an online resource from the National Library of Medicine to support access to the biomedical research literature that indexes over 30,000 journals and over 34 million articles. We limit our analysis to original research publications; this excludes other types of publications such as review articles, letters, and editorials. For each grantee, we identify their career publications in *PubMed* using Author-ity (Torvik and Smalheiser 2021, Torvik and Smalheiser 2009, Torvik et al. 2005). Author-ity disambiguates authors using the insight that authors of the same name sharing coauthors, affiliations, keywords, and other such characteristics are likely, in fact, to be the same person. Author-ity has been shown to be highly accurate (Lerchenmueller and Sorenson 2016).

We use NIH RePORTER to link publications attributable to each NIH grant. This data relies on the grant having been listed in the publication's acknowledgement section as a funding source, which is required as a condition of accepting NIH funding. We consider publications attributable to a grant cycle if they are published up to one year after the last year of the grant cycle to allow for publication delays.[8] As investigators may have different degrees of risk taking across different projects and funding sources, in all cases, our measures of risk taking are calculated solely using publications during a grant cycle that acknowledge funding from the grant rather than also including publications which do not acknowledge funding from the grant.

To measure research topic and methodology, we capitalize on Medical Subject Headings (MeSH). MeSH is a hierarchical controlled vocabulary created and maintained by the National Library of Medicine used for indexing and searching the biomedical literature. In the 2019 edition of MeSH, there are over 29,000 terms, with an average of about 12 MeSH keywords associated with each *PubMed* publication. Crucially, MeSH are assigned by professional indexers. While it is difficult to assess accuracy as there is no single "correct" way to index articles, MeSH have generally been found to be consistent across indexers (Coletti and Bleich 2001, Funk and Reid 1983). In addition to MeSH associated with grant publications, we use the National Library of

[7] 138 (0.13%) grant cycles appear in our sample more than once as they have multiple principal investigators.
[8] Among publications acknowledging a focal NIH grant in our sample, 20.8% are published more than one year after the end of grant funding. These publications and the results they are based on will typically not have been available to study sections when they review a grant for competitive renewal and so do not inform our outcome or measures of interest.

Medicine Medical Text Indexer to assign MeSH to grant abstracts. Grant abstract text is available from the NIH RePORTER.

We supplement publication data with citations from Clarivate's *Web of Science*. All grants are adjusted to 2020 dollars using the Biomedical Research and Development Price Index.

Measuring risk taking

Given the different sources of risk in science, no single measure can fully capture the dynamics of risk-taking across the diverse range of projects in our data. We therefore employ four distinct measures, each reflecting a different aspect of scientific risk. Since these measures have varying strengths and weaknesses, considering them together provides a more comprehensive perspective than any single measure in isolation. Importantly, our analysis remains non-parametric, examining the full distribution of risk-taking levels. This ensures that our findings are not driven solely by conventional levels of risk-taking but also account for whether penalties for risk arise primarily in the tails of the distribution.

A key consideration in our measurement strategy is that we construct our risk measures using information from publications emerging from the initial grant cycle. Ideally, risk would be assessed directly from the content of the renewal proposal rather than inferred from past outcomes. This necessarily embeds an assumption: that a project's risk profile at the renewal stage largely reflects the level of risk taken in the initial funding period. While investigators may maintain a consistent approach to risk-taking across multiple grant cycles—whether due to inertia, specialization, or the constraints of their research agenda—this remains an inherent limitation of our method.

Access to full proposal texts, including those of unfunded applications, would allow for a more direct and precise assessment of risk at the time of renewal. However, given the constraints of publicly available data, our strategy represents a meaningful step forward in understanding how risk-taking influences funding outcomes. We aim to be transparent about these limitations so that readers can judge for themselves whether our approach is reasonable. In our view, the compromises we have made are a necessary price for progress in studying this important and otherwise difficult-to-analyze question. With these general considerations in place, we now turn to the construction of the four risk-taking metrics.

First, risk-taking may lead to greater variance in outcomes, increasing both the likelihood of a breakthrough and the likelihood of failure (Azoulay et al. 2011). We capture this dynamic with a measure we term "extreme tail outcomes," defined as the difference between the highest and lowest vintage-adjusted citation percentiles among publications from the initial grant cycle. We focus on long-run vintage-adjusted citation percentiles for several reasons. Relying solely on short-run citation data would introduce considerable noise, as the impact of scientific contributions—especially those employing unconventional approaches—often takes time to be fully recognized (Wang et al. 2017). Additionally, using percentiles enables meaningful comparisons

across grants awarded in different time periods, ensuring that differences in citation practices over time do not distort our measure.[9]

Research can be either consolidative, reinforcing earlier findings, or disruptive, challenging the status quo. Our second measure leverages this distinction with the disruption index proposed by Funk and Owen-Smith (2017), which has received increased attention (and validation) from the science of science community (Bornmann et al. 2020; Wu et al. 2019). This index classifies a paper as highly disruptive if future research cites it while largely ignoring its references, whereas highly consolidative research is characterized by future studies citing both the focal paper and its references. Our measure of "disruption" for a grant cycle is the highest percentile of the disruption index for the publications emerging out of the initial round of funding (calculated relative to all publications appearing in *PubMed* in the same year).

As a third measure, we examine the intellectual distance between a grant cycle's publications and the investigator's prior work. This approach recognizes that some aspects of risk stem not just from the project itself but also from who is conducting the research. Scientists who integrate their existing expertise with new knowledge—whether by addressing novel topics, adopting new techniques, or working with unfamiliar model organisms—are taking on risk. Such shifts may yield valuable discoveries but can also devalue prior expertise and invite skepticism from peer evaluators (Arts and Fleming 2018; Hill et al. 2021). We term this measure "pivoting." To compute it, we calculate the fraction of MeSH term pairs in a grant cycle's publications that are new relative to those in the investigator's prior work during the preceding five years.

Our final measure of risk taking captures intellectual distance from what other NIH-funded investigators are studying within the same scientific domain. This measure reflects the risks of pursuing contrarian research—projects that diverge from the mainstream of one's peers—thereby increasing career risk (Fang and Casadevall 2015). We term this measure "standing out" and operationalize it as the fraction of MeSH term pairs in a grant cycle's publications that appear exclusively in the focal investigator's work, meaning no other funded researchers in the same IC-study section-year use these term pairs in their grant proposal abstracts.[10,11]

---

[9] A limitation of this approach is that we infer failure from exceptionally low citation rates. While this is a reasonable proxy, truly failed projects are more likely to remain unpublished in a file drawer rather than being published in a minor outlet.

[10] The relevant sample size varies across each of these measures. Extreme tail outcomes is only defined for those grant cycles resulting in at least two publications indexed in Web of Science. Disruption is only defined for those grant cycles resulting in at least one publication that is cited at least once. Pivoting is defined only if there is at least one publication with MeSH terms in both the current grant cycle and the preceding five years. Standing out is defined only if there is one publication with MeSH terms in the current cycle and sufficient funded grant proposal abstracts in the same IC-study section-year are available.

[11] We compare an investigator's grant cycle publication MeSH to that for all available awarded grant proposal abstracts, including those that are not R01-equivalent grants, in the same IC-study section-year. Among the 103,164 investigator-grant-cycles in our sample, we have abstracts for other grants in the same IC-study section-year for 83,451. The mean number of abstracts in the comparison group for each investigator-grant-cycle is 99. We limit our measure to those investigator-grant-cycles with at least 20 other

Measuring novelty

To empirically distinguish novelty from risk taking, we construct a separate measure of novelty. Prior research has taken various approaches, including analyzing unusual or intellectually distant combinations of references (Shibayama et al. 2021, Uzzi et al. 2013, Wang et al. 2017), uncommon combinations of patent classes (Fleming 2001, Verhoeven et al. 2016), citation network centrality (Shibayama and Wang 2020), and the age of keywords or keyword combinations (Arts et al. 2024, Misha and Torvik 2016, Packalen and Bhattacharya 2020). Other studies have used domain-specific measures, such as novelty in chemical structures (Krieger et al. 2022, Rzhetsky et al. 2015). Following Boudreau et al. (2016), we define novelty based on new combinations of MeSH terms. Specifically, our measure captures the fraction of MeSH term pairs in a grant cycle's publications that were first used in any *PubMed*-indexed publication within the preceding three years.[12]

4. Research design and a descriptive look at the data

We observe the universe of NIH grants actually awarded, but unfortunately, do not observe grant applications which are not ultimately funded. Because of this, we empirically focus on NIH R01-equivalent grants that were awarded and measure whether or not the grant was renewed. This approach lets us reliably observe all investigators who were at risk of having their grant renewed. In this population, a grant may not be renewed either because the investigator applied for but did not receive a competitive renewal, or because they never applied for a renewal. Given the career importance of grant funding in maintaining a lab, we expect most investigators will apply for renewal.[13] Furthermore, if an investigator with risky research is dissuaded from applying for grant renewal due to their perception of a low likelihood of renewal and the time involved in applying, the implications for our research question are very similar to having applied and not been selected. As our sample evaluates the degree of risk taking among scientists all of whom have an R01-equivalent grant, this suggests these scientists will have similarly benefited from the act of applying for and receiving an R01-equivalent grant (Ayoubi et al. 2019, Jacob and Lefgren 2011).

Using this research design, the primary threat to unbiased estimation of the impact of risk taking on grant renewal is omitted variable bias, i.e., unmeasured factors correlated with both our proxies for risk taking and determinants of a grant's likelihood of renewal. To mitigate this concern,

abstracts in the same IC-study section-year. We prefer this approach to alternatives, as the comparison group is thicker for those with sufficient available data, facilitating our focus on the tail-measures of intellectual distance from other investigators. Similar results are obtained using other thresholds for the number of abstracts in the same IC-study section-year.

[12] 7.9% of MeSH pairs in our sample have three or fewer years between when they were first used in *PubMed* and their use in a grant cycle publication. Our findings are robust to using thresholds other than three years, including using MeSH pairs new to *PubMed* in the past year.

[13] Applications for competitive grant renewal, using the NIH's more expansive definition of R01-equivalency (see footnote 6) were funded at a rate of 34.5% in 2015 (Lauer 2016), close to the 35.9% renewal rate in our sample in the same year, and consistent with the assumption that most grantees do apply for renewal.

we incorporate a rich set of controls for investigator, institution, and grant cycle characteristics.[14]

We also consider controlling for publication and citation outcomes during the initial grant cycle, but the inclusion pf these controls warrants careful consideration. Study sections evaluating renewal applications assess an investigator's progress, often reflected in publications acknowledging the grant's support. Omitting research productivity and impact measures could introduce omitted variable bias, as investigator quality may correlate with both productivity and risk-taking. However, since the relationship between risk-taking and research output (quantity or quality) may vary in magnitude and direction, the bias introduced by exclusion is difficult to predict. At the same time, publication outcomes lie on the causal pathway between risk-taking and grant renewal, making their inclusion problematic. Controlling for them effectively conditions on a collider (Elwert and Winship 2014), potentially biasing estimates of risk-taking's effect on renewal. To address this tradeoff, we present results both with and without these productivity controls. Including them generally attenuates effect sizes toward zero, but our qualitative insights remain robust. Importantly, we do not claim to estimate the causal effect of risk-taking on funding success; rather, we document robust conditional correlations.

Methodological considerations

Our primary estimating equation relates the dependent variable of interest, $renewal_{ij,t(j)}$, for investigator *i,* grant *j*, and grant cycle *t(j)* to risk taking assessed for the preceding cycle of the same grant *j*, *t(j)-1*. Formally, we estimate variations of:

$$E\left[renewal_{ij,t(j)} = 1 \mid X_{ijt}\right] = \beta_0 + \beta_1 risk\_taking_{ij,t(j)-1} + X_{ij,t(j)}$$

where *renewal* is an indicator variable equal to 1 if grant *j* was renewed for the next cycle *t(j)*, and 0 otherwise. We include high dimensional vectors of controls for investigator and grant cycle characteristics $X_{ij,t(j)}$. This vector includes sets of indicator variables for investigator age, degree and gender; log investigator career citations, log prior investigator NIH funding, an indicator variable for no prior investigator NIH funding, and log institutional NIH total funding as measures of investigator and institutional quality; and sets of indicator variables for year of evaluation, initial cycle length, and IC-study section fixed effects. Select specifications also include controls based on measures of research quality and quantity during cycle *t(j)*, including log of mean journal impact factor, sets of indicator variables for the number of publications during the grant cycle (separately for those acknowledging and not acknowledging funding from the grant), and indicator

[14] The NIH instructs reviewers to scores each application based on five criteria: significance, innovation, approach, investigator, and environment. While the first three assess the importance, novelty, and feasibility of the proposed project itself, the last two evaluate investigator qualifications and ability of the institution to support the work, respectively.

variables for having publications in highly selective journals.[15,16] We use robust standard errors with two-way clustering at the investigator and grant level throughout. For β_1 to have a causal interpretation, the measures of $risk\_taking_{ij,t(j)-1}$ must be uncorrelated with the error term of the corresponding regression equation, after conditioning on the observable covariates included in X. In spite of the high dimensionality of X, this remains a very strong assumption, reinforcing the earlier caveat that our estimates correspond to robust conditional statistical associations rather than causal effects.

Our primary statistical model is a linear probability model (LPM). While LPM can produce fitted values outside the interval [0,1]—which makes it unsuitable for predictive purposes—our objective is to estimate the marginal effects from the conditional expectation function. In this context, the LPM is typically more appropriate (Angrist and Pischke 2008: pp. 94-106). Specifically, LPM offers a straightforward estimation of marginal effects and is computationally efficient when dealing with a high-dimensional vector of covariates and fixed effects, which is crucial for our analysis.

Descriptive Statistics

Tables 1a and 1b present descriptive statistics for investigators and grant cycles, respectively. Reflecting the intense competition for R01-equivalent grants, the sample consists of highly accomplished scientists. On average, investigators begin the grant cycle 18.7 years after completing their terminal degree, with 56 prior publications and 2,234 career citations. Approximately 5% of investigators in the sample have achieved elite status, defined as having won a Nobel Prize, Lasker Award, Howard Hughes Medical Institute Investigatorship, or membership in the Institute of Medicine or National Academy of Sciences.

The majority of an investigator's research output is not directly funded by the focal R01-equivalent grant. Among publications during the cycle period, 39.9% acknowledge the focal grant, while 59.6% of non-grant-funded publications still receive NIH support, either through other NIH grants to the investigator or coauthors. Overall, 45.2% of grant cycles were renewed, with renewal rates higher for subsequent cycles (51.5%) compared to the first cycle (40.4%).[17] Renewed grant cycles were associated with younger investigators (17.6 vs. 19.6 years of career experience), greater research productivity in cycle *t(j)* (8.1 vs. 7.0 publications and 663 vs. 422 long-run citations, limiting the comparison to articles directly acknowledging the focal grant), and a greater percentage of the investigator's research activity funded by the focal grant (46% vs. 38.2% of

[15] We use journal impact factor as a proxy for publication quality in grant cycle publications, as this metric is visible to peer review committees at the time of evaluation, whereas early citations provide a noisy signal of quality. Our results remain robust when instead measuring grant productivity using citations accrued by the time of renewal or, alternatively, long-run citations.

[16] For this measure we include those publications which were published in *Nature*, *Cell*, *Science*, and the *New England Journal of Medicine*.

[17] In contrast, the success rate of applicants for a new R01-equivalent grant is lower. In 2015, 16.1% of applications for a new R01-equivalent grant were funded (Lauer 2016).

publications). Figure 1 provides an histogram of the distribution of the number of publications per grant cycle, separately for renewed and terminated grants.

Table 1c presents descriptive statistics for the four measures of risk taking, as well as novelty. Although the mean level of risk taking is generally higher for terminated grants than for renewed ones, the corresponding magnitudes are modest. Figure 2 displays histograms for each of the four measures by grant cycle renewal status, providing initial evidence that renewal rates vary across different levels of risk taking. Finally, Table 1d presents correlations between measures of risk taking and novelty. The relatively low correlation between these measures highlights that each captures a distinct approach to risk taking, with novelty offering a contrast to the risk-taking metrics.

5. Results

Risk taking and grant renewal

Table 2 reports estimates of the impact of risk taking on the likelihood of grant renewal for each of the four measures. Overall, we observe a clear inverse relationship: high levels of risk are associated with lower renewal rates, while low levels of risk correlate with higher renewal rates across all measures. Not only are these effects statistically significant, but their magnitudes are also considerable. For example, when comparing grant cycles in the bottom versus top decile of the "extreme tail outcomes" metric in Table 2A, column 2, those in the bottom decile—reflecting low levels of risk taking—experience a 5.1% higher renewal rate (an 11.0% increase relative to intermediate deciles), whereas those in the top decile have a 4.4% lower renewal rate (a 9.5% decrease relative to intermediate deciles). Similar patterns emerge across the other risk-taking measures for the top and bottom deciles: disruption (8.9% higher [19.6% increase] vs. 2.2% lower [4.8% decrease]) in Table 2B, pivoting (4.1% higher [9.0% increase] vs. 2.8% lower [6.2% decrease]) in Table 2C, and standing out (2.7% higher [6.4% increase] vs. 1.3% lower [3.1% decrease]) in Table 2D.[18] These findings remain even after controlling for publication quantity and quality during the grant cycle, though in some cases, the magnitudes of the risk penalty shrink somewhat. Importantly, we find no evidence suggesting that the penalty for risk taking diminishes in later cycles (i.e., on the second or third competitive renewal for the same grant).

We further investigate the full distribution of risk-taking levels and their association with grant renewal, as depicted in Figures 3, 4, 5, and 6. In these specifications, separate indicator variables are included for each ventile of the risk measure. Consistent with Table 2, we find higher

[18] For extreme tail outcomes, pivoting, and standing out, percentiles are calculated relative to other grant cycles in the same evaluation year. For disruption, we emphasize work that is disruptive in an absolute sense and so use the percentile of the most disruptive paper, calculated relative to all publications in *PubMed* appearing in the same year.

renewal rates for lower levels of risk taking and lower renewal rates for higher levels. Notably, several measures show a nearly monotonic decrease in renewal rates as risk taking increases.[19]

Additionally, we examine how the penalty for risk taking varies across NIH Institutes and Centers. While both the magnitude of the penalty and the types of risk taking most penalized differ across Institutes, the pattern persists across the entire NIH. We also examine temporal changes in the penalty for risk taking. Given the increasing competition for R01-equivalent grants and evolving NIH policies, caution is needed when interpreting trends over time. Nevertheless, we find that the magnitude of the risk-taking penalty has remained relatively stable (Appendix Figure A2).

Novelty and risk taking

We next examine whether novelty confounds the relationship between risk taking and grant renewal. We estimate models similar to those shown in Figures 3 through 6, but with the addition of indicator variables for each ventile of the novelty measure. The results of this analysis are presented in Figure 7. Two key findings emerge from this analysis. First, in contrast to our measures of risk taking, we observe that novelty is rewarded. The magnitude of this reward is comparable to the penalty for risk taking. Second, the inclusion of novelty does not substantially reduce the observed penalty for risk taking. This provides empirical evidence that novelty and risk-taking measures may operate independently, supporting the notion that novelty, at least in some contexts, is not a valid proxy for risk taking. Appendix Figure A4 replicates this analysis by assessing the effect of novelty on grant renewal without incorporating measures of risk taking. As in Figure 7, we find that novelty is associated with a higher probability of grant renewal.

Is the risk penalty excessive?

As previously discussed, we propose that variation in the informativeness of prior beliefs about the potential value of research pursued in a grant may help identify cases where the penalty for risk taking is likely excessive from a social standpoint. To test this empirically, we examine the interaction between novelty and risk taking on the likelihood of grant renewal. We also analyze how career stage interacts with risk taking, as less established or younger scientists may lack a sufficient track record for evaluators to form an informative prior about the potential success of their investigations, compared to their senior counterparts.

Following NIH classifications, we focus on new investigators—those in their first R01-equivalent grant cycle—and early stage investigators—those within 10 years of their terminal

[19] An exception is with disruption, where there is a non-linearity in the association between disruption and grant renewal: grant cycles with the very highest degree of disruption have higher renewal rates than those with elevated but less extreme levels of disruption.

research degree or clinical training.[20] Additionally, we assess the risk-taking penalty for members of the National Academy of Sciences, whose rarefied status likely provides evaluators with more informative prior beliefs about the value and feasibility of their proposed research.[21]

Figure 8 summarizes these findings, with Appendix Tables A1 and A2 providing the full estimates. Several trends emerge. First, across most measures of risk taking, more novel research, new investigators, and early stage investigators face a higher penalty for risk taking, while members of the National Academy of Sciences experience a lower penalty.[22] Second, at higher levels of novelty, the penalty for risk taking increases for pivoting and standing out, but remains statistically indistinguishable from less novel research for the extreme tail outcomes and disruption measures (Appendix Table A2). Finally, some of these results are imprecisely estimated, particularly with regard to status, due to the small number of elite investigators.

We provide additional evidence suggesting that the penalty for risk-taking may be excessive. Specifically, we ask whether, among grant cycles that were renewed, those characterized by higher levels of risk-taking yield better outcomes than their lower-risk counterparts (Table 3, as well as Appendix Tables A3 and A4). To assess this, we examine the relationship between risk-taking and research relevance, measured by citations to renewed grant cycles' publications in patents (Marx and Fuegi 2020, 2022). Across all four measures, we find a positive association between risk-taking and impact on private-sector patenting—although for the "extreme tails" metric, this effect reverses when controlling for initial grant cycle productivity (column 5).[23] In line with Becker (1957), these results suggest that NIH peer review systematically disadvantages risky proposals, imposing a higher bar for funding despite their potential for significant private-sector impact.

[20] As we do not observe residency or fellowship training, we assume for determining early stage investigator status an average length of postgraduate clinical training of five years for researchers holding an M.D. or M.D./Ph.D.

[21] This measure also includes members of the National Academy of Medicine. We exclude any members who are also Howard Hughes Medical Institute Investigators who may have distinct dynamics and constraints for risk taking (Azoulay et al. 2011).

[22] New and early stage investigators may need to show some degree of pivoting from the research done during the training phase of their career to establish independence and a distinct research identity from their mentors.

[23] The Appendix (Table A4, Panel A) provides similar evidence after limiting the patents being considered to those linked to FDA-approved drugs (Durvasula et al. 2023). Because this outcome is more skewed, the corresponding analysis subsamples are smaller than in Table 3. We also examine the association between risk taking in the current grant cycle and publications and patents in the next grant cycle for those grant cycles that were renewed (Table A4, Panel B). We find these analyses less compelling since they essentially regress publication or citation outcomes in cycle *t(j)* on a transformation of publication/citation outcomes in cycle *t(j)-1*. Yet, as above, the results imply that risk taking is associated with superior outcomes in the future on the sample of renewed grants. An exception is observed for the pivoting measure, which is associated with fewer publications in the subsequent grant cycle, consistent with a decline in productivity when shifting to new research topics.

Robustness checks and ancillary results

We conduct several robustness checks to assess the sensitivity of our estimates to alternative assumptions and subsamples. First, we test alternative measures of risk taking and novelty (Appendix Table A5). This includes the difference between the maximum and median, median and minimum, winsorized maximum and minimum, and variance in vintage-adjusted citation percentiles for extreme tail outcomes (Appendix Figures A5 and A6); the mean rather than the maximum disruption index and a variation of the index following Bornmann et al. (2020) for disruption (Appendix Figures A7 and A8); fraction new single MeSH rather than MeSH pairs and cosine dissimilarity score for single MeSH terms after term frequency-inverse document frequency weighting for pivoting (Appendix Figures A9 and A10); fraction unique single MeSH rather than MeSH pairs and cosine dissimilarity score for single MeSH terms after term frequency-inverse document frequency weighting for standing out (Appendix Figures A11 and A12), and MeSH pair age as well as two measures—the fraction of bigrams which are new and the average number of new bigrams per publication—constructed from textual analysis of publication title and abstract compared to the corpus of publications in *OpenAlex* following Arts et al. (2024) for novelty (Appendix Figures A3 and A4). Emphasizing how pivoting and standing out are distinct from novelty, the findings with these measures are robust to exclusion of MeSH terms and MeSH term pairs which were first used in *PubMed* within the past decade. As risk taking measures may be undefined for some grant cycles, the sample used for analysis varies across each risk taking measure (see footnote 10 for details). Our results are robust to using a subsample for which all four risk taking measures are defined for all grant cycles (Appendix Table A6).

The penalty for pivoting might be smaller for those grants for which the investigator proposed—and peer review approved and funded—to pivot in the initially submitted planned research studies. To examine this, we identify the fraction of MeSH terms shared between the grant proposal abstract and an investigator's prior work and split the sample at the median for this measure. We find overall a similar magnitude penalty for pivoting across both groups (Appendix Table A7). In other words, grants which proposed a pivot and then investigators actually pivoted in the work carried out are still renewed at lower rates.

Author-ity has been shown to be highly accurate (Lerchenmueller and Sorenson 2016). It is likely the case, however, that its accuracy of publication disambiguation is higher for rare names within *PubMed*. To ensure our results are not driven by investigators with poorly disambiguated publications in Author-ity, we split the sample at the median name frequency based on frequency within the entire corpus of *PubMed* and find similar results across both groups (Appendix Table A8).

One reason investigators may choose not to renew their grant is that their primary research focus lies outside of the scope of what the NIH typically supports. For instance, some organic chemists, electrical engineers, or sociologists may have an isolated project at the interface of biomedicine within the context of a broader research agenda elsewhere. To ensure our results are

not driven by investigators with only a weak attachment to NIH funding, we calculated the fraction of career publications with any NIH funding and split the sample at the median of this measure (Appendix Table A9). We find a similar penalty for risk taking across both those with above and below the median career attachment to NIH funding.

Finally, we use a bounding technique proposed by Oster (2019) to gauge the sensitivity of our results to omitted variable bias. The intuition behind this approach is that the stability of the coefficient for risk taking when varying the set of control variables included in the model, scaled by the movement in R^2, provides information about the potential impact of unobserved covariates. To generate these bounds, the analyst must assume proportionality between the covariances of the outcomes with observed and unobserved covariates and posit a maximum value for R^2 if the regression could include all observed and unobserved covariates. Oster's technique generates δ, which can be interpreted as the degree of selection on unobservables relative to observables necessary to reduce the magnitude of the effect of the regressor of interest to zero. Appendix Table A10 reports the results of this exercise. For all four risk taking measures, δ is above one, the threshold recommended by Oster to suggest robustness to the influence of unobservable covariates.[24]

6. Discussion

Using a comprehensive dataset of awarded NIH R01-equivalent grants and multiple measures of risk-taking, we find that investigators who take greater risks experience lower renewal rates than their lower-risk counterparts. This provides systematic evidence that NIH peer review penalizes risk-taking. We also show that novelty functions differently from our risk-taking measures, cautioning against the practice of using the former as a proxy for the latter. Given the NIH's central role in U.S. biomedical research funding, these findings have significant policy implications for peer review and innovation funding.

A key strength of our study is its examination of risk-taking across the full distribution of observed values, capturing effects at the extremes. Not all scientific innovation contributes equally to progress, and the most radical advances often emerge from the highest-risk projects (Rzhetsky et al. 2015, Veugelers and Wang 2019, Wang et al. 2017). That peer review punishes risk-taking is particularly striking given mixed evidence on its ability to predict scientific productivity and impact (Cole et al. 1981, Danthi et al. 2014, Li and Agha 2015, Pier et al. 2018).[25]

[24] Some values for δ are negative. This implies that the correlation between unobservables and risk taking would need to be in the opposite direction as the correlation between observables and risk taking to make the corresponding results disappear.

[25] This parallels the dynamics seen for innovation in other settings, such as entrepreneurship. For instance, even among VC-backed startups, a handful companies that are outliers in their success account for the majority of VC profits and impact on the larger economy. Likewise, while investors may be able to identify

Our results cannot be given a causal interpretation if unobserved covariates drive both risk-taking and grant renewal. However, given the rich set of controls included in the statistical models, we believe this concern may be less significant than it appears. In fact, our estimates may understate the true penalty for risk-taking. Some of our measures depend on the existence of published work, meaning we may miss projects that failed to generate publications, potentially underestimating risk-related attrition. Additionally, because we observe only funded grant applications, any initial selection bias against risky projects would make our sample less venturesome than the full applicant pool. Investigators aware of the NIH's risk aversion may also self-select out of the funding process altogether, particularly for their high-risk projects (Langer 2012). Finally, young researchers training under NIH-funded mentors may internalize a preference for safer projects, reinforcing long-term conservatism in research agendas.

Each of our risk-taking measures captures different facets of risky research, each with its own strengths and imperfections. An investigator branching into a new research area, for instance, may be engaging in a promising recombination of ideas—or straying too far from their expertise, reducing their chances of success (Arts & Fleming 2018). Similarly, pursuing underexplored topics might signal a potential breakthrough or simply an unpromising research avenue. Given these differences, one could reasonably quibble with any single measure. Yet despite their idiosyncratic flaws and the fact that they are constructed in distinct ways—sometimes even drawing on different underlying data—the results remain remarkably consistent across measures. This consistency is reassuring: if each metric is flawed in its own way, much like Tolstoy's unhappy families, they are at least flawed differently. And yet, they tell the same story.

Both high-risk, high-reward and low-risk, incremental research play vital roles in scientific progress. However, there are strong reasons for the NIH to increase support for high-risk research. Established research fields naturally lend themselves to incremental work, while emerging areas require greater exploration and risk-taking. Instead, we find the opposite: the penalty for risk-taking is largest in more novel areas, suggesting a potential welfare loss. NIH risk aversion may also disproportionately affect research areas with limited private-sector incentives, where public support is most essential (Aghion et al. 2008). Governments are uniquely positioned to bear risks too large for private industry, especially given declining corporate investment in science (Arora et al. 2017, Fleming et al. 2019). NIH's size and influence mean its hesitancy to fund risky fundamental research is unlikely to be offset by other funding sources, potentially slowing technological progress (Babina et al. 2023). The same logic applies to NIH-funded clinical research, where industry has limited incentives to invest in high-risk studies with uncertain commercial returns (Greenblatt et al. 2023). While private philanthropies provide alternative funding, their priorities may not align with the greatest public needs (Murray 2012).[26]

---

those companies with little growth potential, they have only a very limited ability to discern those within their portfolio most likely to generate outsized returns (Kerr et al. 2014).

[26] Researchers who do not receive NIH funding may be able to continue their research with support from other funders. For instance, Chuban and Hackett (1990) find that many investigators are able to pursue

Recognizing these challenges, the NIH has created dedicated high-risk, high-reward funding mechanisms, such as the Pioneer Awards, Director's Transformative R01 Awards, and New Innovator Awards (NIH 2019). In 2021, these programs funded 106 awards totaling $329 million over five years (NIH 2021). However, such mechanisms represent only a small fraction of NIH's overall funding portfolio (see footnote 4). Several private science funders—including the Howard Hughes Medical Institute, Open Philanthropy Project, Fast Grants, Chan Zuckerberg Initiative, and Wellcome Trust—have also launched initiatives to support high-risk research.

While our study focuses on peer review in public funding, its implications extend to other contexts where high-risk innovation is evaluated. Rather than treating risk tolerance as an afterthought or relying on vague exhortations to evaluators, funders should proactively design processes that encourage the desired level of risk-taking. Dedicated funding mechanisms are an important step, but they must also address inherent biases against risk. For instance, as high-risk projects may have greater variation in evaluator assessments of quality, moving from a consensus-based to a champion-based approach or emphasizing enthusiastic support rather than mean score may facilitate the selection of riskier projects (Gius 2024). The composition of committees and panels should also be scrutinized. Expert evaluators may have more pronounced biases against novelty in their own field (Wang et al. 2017). Increasing diversity of expertise and traditions and considering an evaluator's prior risk-taking behavior when selecting reviewers may help mitigate risk penalization (Nicholson and Ioannidis 2012).

Finally, a fundamental challenge in motivating risk-taking in science is that, *ex post*, a failed but risky research project and a low-quality idea may generate similar outcomes—no publication, patent, or product. While fame and fortune can both act as powerful motivators for pursuing risky science, investigators may still hesitate if failed risky projects lead to job loss or lab closure. Without compromising high standards and accountability, motivating risk-taking may require simultaneous adjustments to multiple features of the funding environment, including feedback, time horizon, and job security (Manso 2011, Azoulay et al. 2011).

their proposed research even if they do not receive the focal grant. Similarly, Jacob and Lefgren (2011) find the marginal impact of receiving an NIH grant to be only about one additional publication over five years, in part because researchers can substitute to other funding sources. Yet, NIH's penalization of risk taking may influence the type of proposals investigators submit to these alternative sources. The preeminent position of NIH within the biomedical research funding ecosystem implies that its risk aversion could "contaminate" the behavior of other funders.

References

Aghion, Philippe, Mathias Dewatripont, and Jeremy C. Stein. 2008. "Academic freedom, private-sector focus, and the process of innovation." *RAND Journal of Economics*, 39(3):617-635.

Alberts, Bruce, Marc W. Kirschner, Shirley Tilghman, and Harold Varmus. 2014. "Rescuing US biomedical research from its systemic flaws." *Proceedings of the National Academy of Sciences*, 111(16):5773-5777.

Althaus, Catherine E. 2005. "A disciplinary perspective on the epistemological status of risk." *Risk Analysis*, 25(3):567-588.

Angrist, Joshua D., and Jorn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Arora, Ashish, Sharon Belenzon, and Andrea Patacconi. 2017. "The decline of science in corporate R&D." *Strategic Management Journal*, 39(1):3-32.

Arrow, Kenneth J., and Robert C. Lind. 1970. "Uncertainty and the Evaluation of Public Investment Decisions." *The American Economic Review*, 60(3): 364-378.

Arts, Sam, and Lee Fleming. 2018. "Paradise of Novelty—Or Loss of Human Capital? Exploring New Fields and Inventive Output." *Organization Science*, 29(6):1074-1092.

Arts, Sam, Nicola Melluso, Reinhild Veugelers. 2024. "Beyond citations: Measuring novel scientific ideas and their impact in publication text." Forthcoming, *Review of Economics and Statistics*.

Ayoubi, Charles, Michele Pezzoni, and Fabiana Visentin. 2021. "Does it Pay to do Novel Science? The Selectivity Patterns in Science Funding." *Science and Public Policy,* 48(5):635-648.

Ayoubi, Charles, Michele Pezzoni, and Fabiana Visentin. 2019. "The important thing is not to win, it is to take part: What if scientists benefit from participating in research grant competitions?" *Research Policy*, 48(1):84-97.

Azoulay, Pierre, Erica Fuchs, Anna P. Goldstein, and Michael Kearney. 2019a. "Funding Breakthrough Research: Promises and Challenges of the 'ARPA Model'." *Innovation Policy and the Economy*, 19:69-96.

Azoulay, Pierre, Joshua S. Graff Zivin, Danielle Li, and Bhaven Sampat. 2019b. "Public R&D and Private-Sector Patenting: Evidence from NIH Funding Rules." *The Review of Economic Studies,* 86(1):117-152.

Azoulay, Pierre, Joshua S. Graff Zivin, and Gustavo Manso. 2011. "Incentives and creativity: evidence from the academic life sciences." *RAND Journal of Economics*, 42(3):527-554.

Azoulay, Pierre and Danielle Li. 2022. "Scientific Grant Funding." Chapter 4 in Goolsbee, Austan and Benjamin Jones (Eds). *Innovation and Public Policy*, pp. 117-150. Chicago, IL: University of Chicago Press.

Babina, Tania, Alex Xi He, Sabrina T. Howell, Elisabeth Ruth Perlman, Joseph Staudt. 2023. "Cutting the Innovation Engine: How Federal Funding Shocks Affect University Patenting, Entrepreneurship, and Publications." *Quarterly Journal of Economics,* 138(2):895-954.

Banal-Estañol, Albert, Ines Macho-Stadler, and David Perez-Castrillo. 2019. "Evaluation in research funding agencies: Are structurally diverse teams biased against?" *Research Policy*, 48(7):1823-1840.

Becker, Gary S. 1957. *The Economics of Discrimination.* Chicago, IL: University of Chicago Press.

Blaser, Martin J. 2005. "An Endangered Species in the Stomach." *Scientific American,* 292(2):38-45.

Bornmann, Lutz, Sitaram Devarakonda, Alexander Tekles, and George Chacko. 2020. "Are disruption index indicators convergently valid? The comparison of several indicator variants with assessments by peers." *Quantitative Science Studies,* 1(3):1242-1259.

Boudreau, Kevin J., Eva C. Guinan, Karim R. Lakhani, and Christoph Riedl. 2016. "Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science." *Management Science*, 62(10):2765-2783.

Bromham, Lindell, Russell Dinnage, and Xia Hua. 2016. "Interdisciplinary research has consistently lower funding success." *Nature*, 534(7609):684-687.

Carson, Richard T., Joshua S. Graff Zivin, and Jeffrey G. Shrader. 2023. Choose Your Moments: Peer Review and Scientific Risk Taking. NBER Working Paper #31409.

Chubin, Daryl E., and Edward J. Hackett. 1990. *Peerless Science: Peer Review and U.S. Science Policy.* Albany, NY: SUNY Press.

Clancy, Matt. 2023. "Biases Against Risky Research." *New Things Under the Sun.* https://www.mattsclancy.substack.com/p/biases-against-risky-research

Cohen, Stanley N., Annie C.Y. Chang, Herbert W. Boyer, and Robert B. Helling. 1973. "Construction of Biologically Functional Bacterial Plasmids *In Vitro*." *Proceedings of the National Academy of Sciences*, 70(11), 3240-3244.

Cole, Stephen, Jonathan R. Cole, and Gary A. Simon. 1981. "Chance and Consensus in Peer Review." *Science*, 214(4523):881-886.

Coletti, Margaret H. and Howard L. Bleich. 2001. "Medical Subject Headings Used to Search the Biomedical Literature." *Journal of the American Medical Informatics Association*, 8(4):317-323.

Collison, Patrick, Tyler Cowen, and Patrick Hsu. 2021. "What We Learned Doing Fast Grants." *Future.* https://future.com/what-we-learned-doing-fast-grants/

Danthi, Narasimhan, Colin O. Wu, Peibei Shi, and Michael Lauer. 2014. "Percentile ranking and citation impact of a large cohort of National Heart, Lung, and Blood Institute-funded cardiovascular R01 grants." *Circulation Research*, 114(4):600-606.

Dasgupta, Partha, and Paul A. David. 1994. "Toward a new economics of science." *Research Policy*, 23:487-521.

Durvasula, Maya, C. Scott Hemphill, Lisa L. Ouellette, Bhaven Sampat, and Heidi L. Williams. 2023. "The NBER Orange Book Dataset: A user's guide." *Research Policy*, 52(7):104791.

Elwert, Felix, and Christopher Winship. 2014. "Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable." *Annual Review of Sociology*, 40:31-53.

Fang, Ferric C., and Arturo Casadevall. 2015. "Competitive science: Is competition ruining science?" *Infection and Immunity*, 83(4):1229-1233.

Fields, Stanley. 2014. "Would Fred Sanger get funded today?" *Genetics*, 197(2):435-439.

Fleming, Lee. 2001. "Recombinant Uncertainty in Technological Search." *Management Science*, 47(1):117-132.

Fleming, Lee, Hillary Greene, Guan-Cheng Li, Matt Marx, and Dennis Yao. 2019. "Government-funded research increasingly fuels innovation." *Science*, 364(6446):1139-1141.

Fox, Craig R., and Amos Tversky. 1995. "Ambiguity Aversion and Comparative Ignorance." *Quarterly Journal of Economics,* 110(3):585-603.

Frankel, Alexander P., Joshua L. Krieger, Danielle Li, and Dimitris Papanikolaou. 2023. "Evaluation and Learning in R&D Investment." NBER Working Paper #31290.

Franzoni, Chiara, and Paula Stephan. 2023. "Uncertainty and risk-taking in science: Meaning, measurement and management in peer review of research proposals." *Research Policy*, 52(3):104706.

Franzoni, Chiara, Paula Stephan, and Reinhilde Veugelers. 2022. "Funding Risky Research." *Entrepreneurship and Innovation Policy,* 1:103-133.

Funk, Mark E., and Carolyn Anne Reid. "Indexing consistency in MEDLINE." *Bulletin of the Medical Library Association*, 71(2):176-183.

Funk, Russell J., and Jason Owen-Smith. 2017. "A Dynamic Network Measure of Technological Change." *Management Science*, 63(3):791-817.

Gan, Hui K., Benoit You, Gregory R. Pond, and Eric X. Chen. 2012. "Assumptions of Expected Benefits in Randomized Phase III Trials Evaluating Systemic Treatments for Cancer." *Journal of the National Cancer Institute*, 104(8):590-598.

Gius, Luca. 2024. "Disagreement Predicts Startup Success: Evidence from Venture Competitions." Forthcoming, *Strategy Science*.

Graff Zivin, Joshua, and Elizabeth Lyons. 2021. "The Effects of Prize Structures on Innovative Performance." *AEA Papers and Proceedings*, 111:577-581.

Greenblatt, Wesley, Charu Gupta, and Jennifer Kao. 2023. "Drug Repurposing During the COVID-19 Pandemic: Lessons for Expediting Drug Development and Access." *Health Affairs*, 42(3):424-432.

Hall, Stphen S. 1987. *Invisible Frontiers: The Race to Synthesize a Human Gene.* New York: Atlantic Monthly Press.

Hill, Ryan and Carolyn Stein. 2024. "Race to the Bottom: Competition and Quality in Science." Forthcoming, *Quarterly Journal of Economics*.

Hill, Ryan, Yian Yin, Carolyn Stein, Dashun Wang, and Benjamin F. Jones. 2021. "Adaptability and the Pivot Penalty in Science." Working Paper, Northwestern University.

Jacob, Brian A. and Lars Lefgren. 2011. "The Impact of research grant funding on scientific productivity." *Journal of Public Economics,* 95(9-10):1168-1177.

Kerr, William R., Ramana Nanda, and Matthew Rhodes-Kropf. 2014. "Entrepreneurship as Experimentation." *Journal of Economic Perspectives,* 28(3):25-48.

Kolata, Gina. 2009. "Grant System Leads Cancer Researchers to Play it Safe." *New York Times*.

Kornberg, Roger D. 2007. "2006 Novel Laureates." U.S. Senate Committee on Commerce, Science, and Transportation.

Krieger, Joshua, Danielle Li, and Dimitris Papanikolaou. 2022. "Missing Novelty in Drug Development." *The Review of Financial Studies*, 35(2):636-679.

Krieger, Joshua, and Ramana Nanda. 2022. "Are Transformational Ideas Harder to Fund? Resource Allocation to R&D Projects at a Global Pharmaceutical Firm." Harvard Business School Working Paper No. 23-014.

Lane, Jacqueline N., Zoe Szajnfarber, Jason Crusan, Michael Menietti, and Karim R. Lakhani. 2022a. "Are Experts Blinded by Feasibility? Experimental Evidence from a NASA Robotics Challenge." Harvard Business School Working Paper No. 22-071.

Lane, Jacqueline N., Misha Teplitskiy, Gary Gray, Hardeep Ranu, Michael Menietti, Eva C. Guinan, and Karim R. Lakhani. 2022b. "Conservatism Gets Funded? A Field Experiment on the Role of Negative Information in Novel Project Evaluation." *Management Science*, 68(6):4478-4495.

Langer, James S. 2012. "Enabling Scientific Innovation." *Science*, 338(6104):171.

Lauer, Mike. 2022. "FY 2021 by the Numbers: Extramural Grant Investments in Research." National Institutes of Health Office of Extramural Research. https://nexus.od.nih.gov/all/2022/03/07/fy-2021-by-the-numbers-extramural-grant-investments-in-research/

Lauer, Mike. 2016. "Are Attempts at Renewal Successful?" National Institutes of Health Office of Extramural Research. https://nexus.od.nih.gov/all/2016/02/16/are-attempts-at-renewal-successful/

Lerchenmueller, Marc J. and Olav Sorenson. 2016. "Author Disambiguation in PubMed: Evidence on the Precision and Recall of Author-ity among NIH-Funded Scientists." *PLoS One,* 11(7):e0158731.

Lerner, Josh and Julie Wulf. 2007. "Innovation and Incentives: Evidence from Corporate R&D." *Review of Economics and Statistics*, 89(4):634-644.

Li, Danielle and Leila Agha. 2015. "Big names or big ideas: Do peer-review panels select the best science proposals?" *Science*, 348(6233):434-438.

List, John A. and Robert Metcalfe. 2014. "Field experiments in the developed world: an introduction." *Oxford Review of Economic Policy*, 30(4):585-596.

Littman, Dan R. 2015. "Release the Brakes on Cancer Immunotherapy." *Cell,* 162(6):1186-1190.

Lorsch, Jon R. 2015. "Maximizing the return on taxpayers' investments in fundamental biomedical research." *Molecular Biology of the Cell*, 26(9):1578-1582.

Manso, Gustavo. 2011. "Motivating Innovation." *Journal of Finance*, 66(5):1823-1860.

Marshall, Barry J. and J. Robin Warren. 1984. "Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration." *Lancet*, 323(8390): 1311-1315.

Marx, Matt, and Aaron Fuegi. 2020. "Reliance on science: Worldwide front-page patent citations to scientific articles." *Strategic Management Journal,* 41(9):1572-1594.

Marx, Matt, and Aaron Fuegi. 2022. "Reliance on science by inventors: Hybrid extraction of in-text patent-to-article citations." *Journal of Economics and Management Strategy,* 31(2):369-392.

Merton, Robert K. 1973. *The Sociology of Science: Theoretical and Empirical Investigation.* Chicago, IL: University of Chicago Press.

Mishra, Shubhanshu and Vetle I. Torvik. 2016. "Quantifying Conceptual Novelty and the Biomedical Literature." *Dlib Mag.*, 22(9-10).

Moses 3rd, Hamilton, David H. M. Matheson, Sarah Cairns-Smith, Benjamin P. George, Chase Palisch, and E. Ray Dorsey. 2015. "The anatomy of medical research: US and international comparisons." *Journal of the American Medical Association*, 313(2):174-89.

Mueller, Jennifer S., Shimul Melwani, and Jack A. Goncalo. 2012. "The Bias Against Creativity: Why People Desire but Reject Creative Ideas." *Psychological Science*, 23(1):13-17.

Murray, Fiona. 2012. "Evaluating the Role of Science Philanthropy in American Universities." *Innovation Policy and the Economy*, 13:23-59.

National Institutes of Health. 2021. "NIH supports 106 grants featuring high-risk, high-reward research." https://www.nih.gov/news-events/news-releases/nih-supports-106-grants-featuring-high-risk-high-reward-research. Press release.

National Institutes of Health. 2019. "Report of the ACD Working Group on High-Risk, High-Reward Research." Advisory Committee to the Director Working Group on High-Risk High-Reward Programs, National Institutes of Health.

Nielsen, Michael and Kanjun Qiu. 2022. "The trouble in comparing different approaches to science funding." The Science++ Project. https://scienceplusplus.org/trouble_with_rcts/

Nielsen, Michael [@michael_nielsen]. 2022. "(A brief afternoon rant.)" X. https://x.com/michael_nielsen/status/1559318292120027137.

Nicholson, Joshua M. and John P.A. Ioannidis. 2012. "Conform and be funded." *Nature*, 492(7427):34-36.

Nuland, Sherwin B. 2003. *The Doctors' Plauge: Germs, Childbed Fever, and the Strange Story of Ignác Semmelweis.* New York: Norton.

Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business and Economic Statistics*, 37(2):187-204.

Packalen, Mikko, and Jay Bhattacharya. 2020. "NIH funding and the pursuit of edge science." *Proceedings of the National Academy of Sciences*, 117(22):12011-12016.

Petsko, Gregory A. 2012. "Goodbye, Columbus." *Genome Biology*, 13(5):155.

Pier, Elizabeth L. Markus Brauer, Amarette Filut, Anna Kaatz, Joshua Raclaw, Mitchell J. Nathan, Cecilia E. Ford, and Molly Carnes. 2018. "Low agreement among reviewers evaluating the same NIH grant applications." *Proceedings of the National Academy of Sciences,* 115(2):2952-2957.

Rosenberg, Nathan. 1995. "Uncertainty and Technological Change." *Federal Reserve Bank of Boston Conference Proceedings*, 40(Jun):91-125.

Ruben, Adam. 2017. "Another tenure-track scientist bites the dust." *Science*, https://www.science.org/content/article/another-tenure-track-scientist-bites-dust.

Rzhetsky, Andrey, Jacob G. Foster, Ian T. Foster, and James A. Evans. 2015. "Choosing experiments to accelerate collective discovery." *Proceedings of the National Academy of Sciences*, 112(47):14569-74.

Sanchez, Carmen and David Duning. 2023. "Are experts overconfident?: An interdisciplinary review." *Research in Organizational Behavior*, 43(11):100195.

Shibayama, Sotaro, Deyun Yin, and Kuniko Matsumoto. 2021. "Measuring novelty in science with word embedding. *PLoS ONE*, 16(7):e0254034.

Shibayama, Sotaro and Jian Wang. 2020. "Measuring originality in science." *Scientometrics*, 122:409-427.

Sutter, Paul M. 2022. "Risk Aversion Is Ruining Science." *Undark*. https://undark.org/2022/04/27/risk-aversion-is-ruining-science/.

Stephenson, J.A., J.C. Goddard, O. Al-Taan, A.R. Dennison, and B. Morgan. 2013. "Tumour Angiogenesis: A Growth Area—From Hunter to Judah Folkman and Beyond." *Journal of Cancer Research,* 2013:895019.

Teplitskiy, Misha, Hao Peng, Andrea Blasco, and Karim R. Lakhani. 2022. "Is novel research worth doing? Evidence from peer review at 49 journals." *Proceedings of the National Academy of Sciences*, 119(47):e2118046119.

Tham, Wei Yang, Joseph Staudt, Elisabeth R. Perlman, Stephanie D. Cheng. 2024. "Scientific Talent Leaks Out of Funding Gaps." arXiv: 2402.07235.

Tian, Xuan and Tracy Yue Wang. 2014. "Tolerance for Failure and Corporate Innovation." *Review of Financial Studies*, 27(1):211-255.

Torvik, Vetle I. and Neil R. Smalheiser. 2021. Author-ity 2018—PubMed author name disambiguated dataset. University of Illinois at Urbana-Champaign.

Torvik, Vetle I. and Neil R. Smalheiser. 2009. "Author name disambiguation in MEDLINE." *ACM Transactions on Knowledge Discovery from Data*, 3(3):11.

Torvik, Vetle I, Marc Weeber, Don R. Swanson, and Neil R. Smalheiser. 2005. "A Probabilistic Similarity Metric for Medline Records: A Model for Author Name Disambiguation." *Journal of the American Society for Information Science and Technology*, 56(2):140-158.

Uzzi, Brian, Satyam Mukherjee Michael Stringer, and Ben Jones. 2013. "Atypical Combinations and Scientific Impact." *Science*, 342(6157):468-472.

Verhoeven, Dennis, Jurrien Bakker, and Reinhilde Veugelers. 2016. "Measuring technological novelty with patent-based indicators." *Research Policy*, 45(3):707-723.

Veugelers, Reinhilde, Jian Wang, and Paula Stephan. 2022. "Do Funding Agencies Select and Enable Risky Research: Evidence from ERC Using Novelty as a Proxy of Risk Taking." NBER Working Paper #30320.

Veugelers, Reinhilde and Jian Wang. 2019. "Scientific novelty and technological impact." *Research Policy*, 48:1362-1372.

Wang, Jian, Reinhilde Veugelers, and Paula Stephan. 2017. "Bias against novelty in science: A cautionary tale for users of bibliometric indicators." *Research Policy,* 46:1416-1436.

Woolston, Chris. 2014. "Sanger's legacy stirs up digital chatter." *Nature*, 511(9).

Wu, Lingfei, Dashun Wang, and James A. Evans. 2019. "Large teams develop and small teams disrupt science and technology." *Nature*, 566(7744):378-382.

Table 1a. Descriptive statistics: Investigator characteristics

	Mean	Median	Std. Dev.	Min.	Max.
Female	0.25	0	0.44	0	1
Degree year	1982.15	1982	12.26	1932	2012
Degree: PhD	0.68	1	0.47	0	1
Degree: MD	0.20	0	0.40	0	1
Degree: MD/PhD	0.10	0	0.31	0	1
Degree: Other	0.02	0	0.13	0	1
Nb of R01-equivalent grants	1.72	1	1.16	1	14
Nb of R01-equivalent grant cycles	3.05	2	2.78	1	37
Career NIH R01 funding ($2020)	10,268,954	6,953,402	10,515,349	133,632	134,399,904
Career NIH funding ($2020)	16,717,503	8,803,164	27,005,356	133,632	1,743,186,688
Career elite status	0.05	0	0.22	0	1

Note: Number of grants and cycles refers to only those grants that started 1980-2015. Career elite status identifies investigators who were ever won a Nobel Prize, Lasker Award, Howard Hughes Medical Institute Investigatorship, or membership in the Institute of Medicine or National Academy of Sciences. N=37,222 investigators.

Table 1b. Descriptive statistics: Grant cycle

	Renewed grant cycles					Not renewed grant cycles				
	Mean	Median	Std. Dev.	Min.	Max.	Mean	Median	Std. Dev.	Min.	Max.
Grant cycle number	1.94	1	1.28	1	10	1.69	1	1.15	1	11
First cycle of grant	0.51	1	0.50	0	1	0.62	1	0.49	0	1
Year started	1997.85	1999	9.16	1980	2015	2000.38	2002	9.82	1980	2015
Cycle length (years)	4.37	4	1.59	1	19	3.16	3	1.11	1	16
Time, degree to cycle start (years)	17.62	16	8.48	-22	61	19.57	18	9.42	-18	66
Investigator pubs. prior to cycle	53.35	36	57.52	0	1148	58.66	39	63.74	0	1196
Investigator citations prior to cycle	2,118.12	887	3,952.72	0	89,857	2,329.03	967	4,443.51	0	184,460
Publications acknowledging grant										
Nb of publications	8.06	6	7.87	1	184	7.01	5	7.34	1	164
Journal impact factor	6.28	5	4.00	0	72	5.36	5	3.51	0	43
Citations	662.91	355	1,146.86	1	6,4421	421.90	201	797.27	1	67,451
Highly selective journal	0.12	0	0.33	0	1	0.07	0	0.25	0	1
Publications not acknowledging grant										
Nb of publications	9.38	5	13.25	0	328	11.33	6	15.26	0	258
Journal impact factor	6.08	5	4.25	0	43	5.19	4	3.42	0	43
Citations	824.09	291	1,764.14	0	50,927	780.04	287	1,669.59	0	65,447
Highly selective journal	0.17	0	0.37	0	1	0.12	0	0.33	0	1
Fraction with any NIH funding	0.58	1	0.33	0	1	0.60	1	0.32	0	1

Note: Unit of analysis is at the investigator-grant-cycle level. N=103,164 (46,619 renewed and 56,545 not renewed investigator-grant-cycles). Grant cycle number minus one is the number of time the grant has previously been competitively renewed. Time from degree to cycle is measured from the last terminal degree earned by the investigator; 55 individuals (92 investigator-grant-cycles) earned an R01-equivalent grant prior to their last degree. This largely reflects investigators with multiple doctorates, frequently with large time gaps between them. Publications acknowledging a grant reflect those publications citing funding from the focal grant which were published up to one year after the conclusion of the grant cycle; publications not acknowledging the focal grant represents the investigator's other publications during the same time period. Citations are measured through 2020.

Table 1c. Descriptive statistics: Measures of risk taking, novelty, and impact

	Renewed grant cycles					Not renewed grant cycles				
	Mean	Median	Std. Dev.	Min.	Max.	Mean	Median	Std. Dev.	Min.	Max.
Extreme tail outcomes	46.69	47	23.80	0	100	47.40	48	23.79	0	100
Disruption	62.43	63	21.35	0	100	64.59	66	20.77	0	100
Pivoting	0.73	1	0.14	0	1	0.74	1	0.14	0	1
Standing out	0.81	1	0.10	0	1	0.81	1	0.10	0	1
Novelty	0.08	0	0.06	0	1	0.07	0	0.06	0	1
Patent citations	11.46	1	45.25	0	2,887	6.24	0	26.36	0	1,734

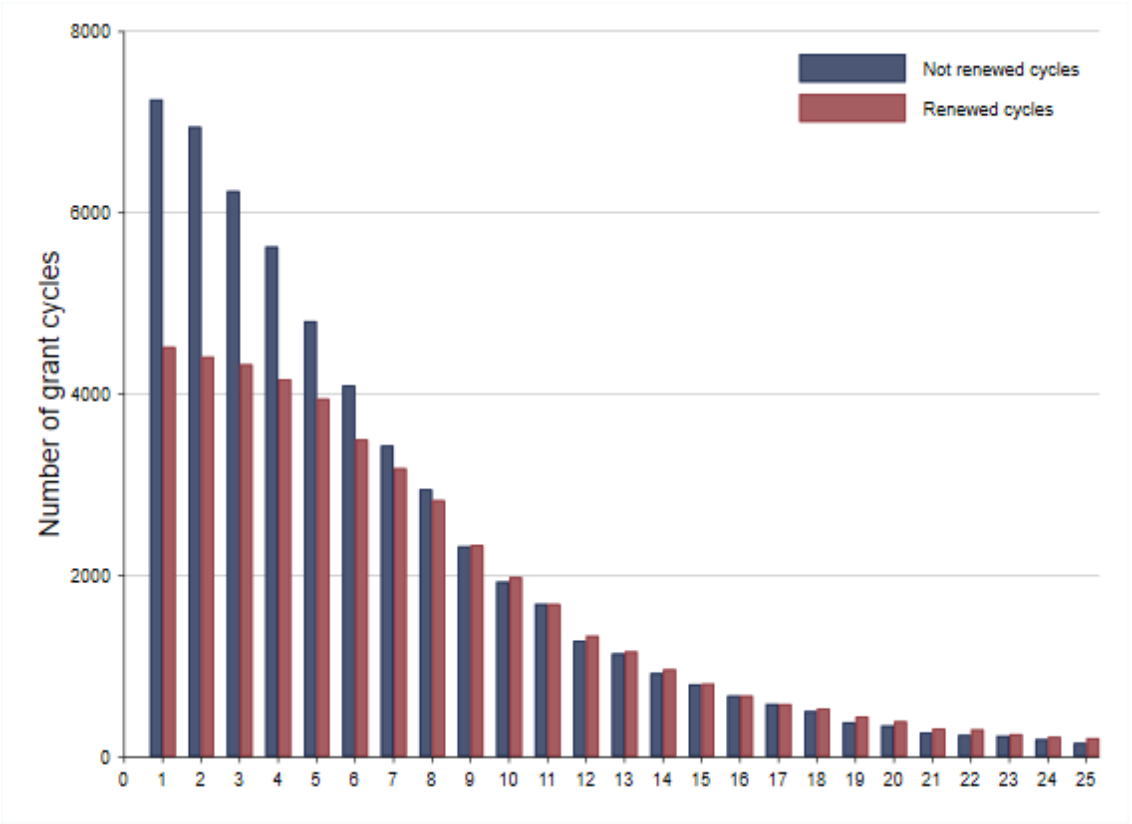
Note: Outcomes are measured at the investigator-grant-cycle level (N=103,164 investigator-grant-cycles). Measures are only defined for those cycles with qualifying associated publications (see footnote 11 for details).

Table 1d. Descriptive statistics: Correlations across measures of risk taking and novelty

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Extreme tail outcomes	1.000						
(2) Disruption	0.400	1.000					
(3) Pivoting	-0.071	0.024	1.000				
(4) Standing out	0.071	0.032	0.242	1.000			
(5) Novelty	-0.067	-0.030	0.319	0.158	1.000		
(6) Nb. of publications	0.509	0.352	-0.139	0.060	-0.044	1.000	
(7) Log citations	0.361	0.198	-0.025	0.065	0.085	0.603	1.000

Note: N=103,164 investigator-grant-cycles. Measures are only defined for those cycles with qualifying associated publications (see footnote 11 for details). Number of publications and citations are based solely on publications acknowledging funding from the focal grant during the grant cycle. Citations are measured through 2020.

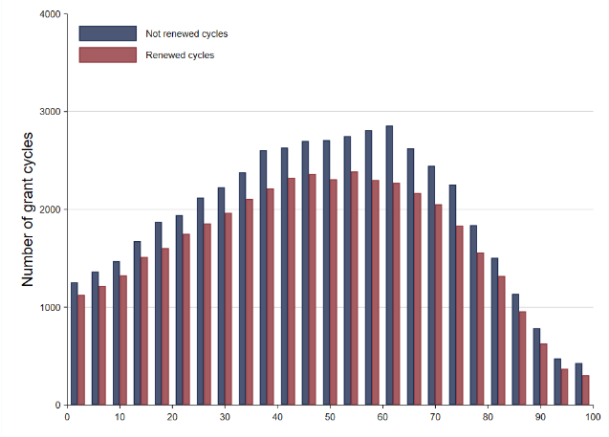
Figure 1. Distribution of grant cycle publications by renewal status



Note: Histogram of the number of publications acknowledging the grant within the grant cycle. Unit of analysis is at the investigator-grant-cycle level. Two thousand nine hundred fifty-two (2.9% of the sample) outliers with over 25 grant cycle publications are not shown.

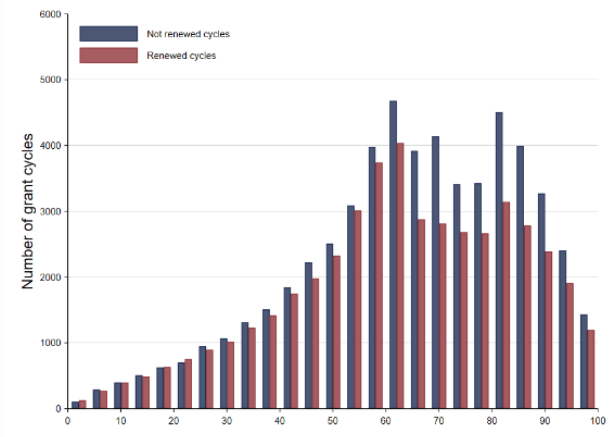
Figure 2. Distribution of risk-taking measures by renewal status

A. Extreme tail outcomes



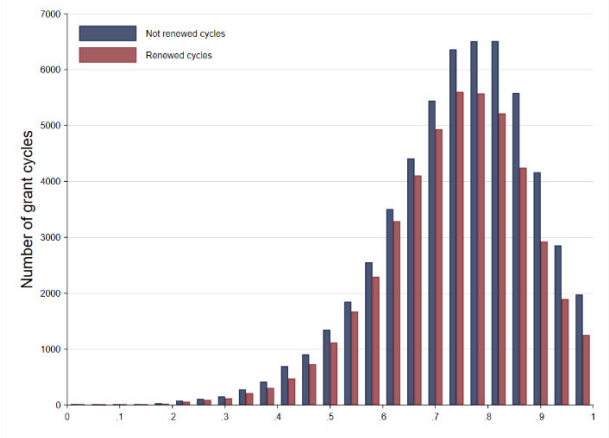
Note: Histogram of range between the maximum and minimum vintage-adjusted citation percentile for grant cycle publications. Unit of analysis is at the investigator-grant-cycle level.

B. Disruption



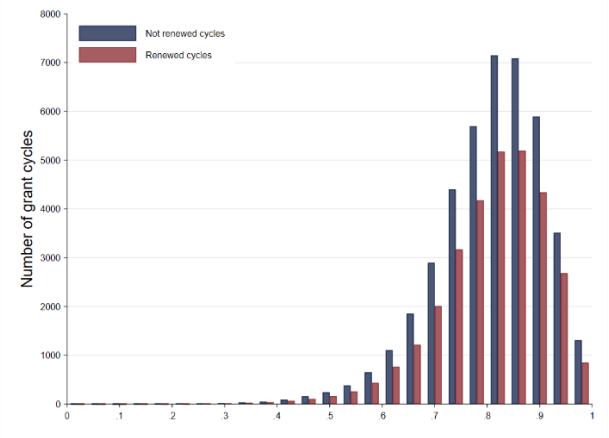
Note: Histogram of the maximum disruption index percentile among grant cycle publications. Unit of analysis is at the investigator-grant-cycle level.

C. Pivoting



Note: Histogram of the fraction of MeSH term pairs for grant cycle publications that were not used in the investigator's publications in the 5 years preceding the grant. Unit of analysis is at the investigator-grant-cycle level.

D. Standing out



Note: Histogram of the fraction of MeSH term pairs for grant cycle publications that were not used by funded grant proposal abstracts in the same NIH IC-study section-year. Unit of analysis is at the investigator-grant-cycle level.

Table 2. Effect of risk taking on grant renewal

A. Extreme tail outcomes

	All renewals			First renewal			Second and future renewals		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<10%ile, max-min citation percentile	0.054**	0.051**	0.022**	0.061**	0.055**	0.028**	0.064**	0.057**	0.018*
	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)	(0.006)	(0.008)	(0.008)	(0.008)
>90%ile, max-min citation percentile	-0.039**	-0.044**	-0.021**	-0.039**	-0.040**	-0.018**	-0.047**	-0.048**	-0.020**
	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Investigator demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grant year and length	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IC x study section	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Institutional quality	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Investigator quality	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Pub. quantity, quality during cycle	No	No	Yes	No	No	Yes	No	No	Yes
Mean of dependent variable	0.463	0.463	0.463	0.413	0.413	0.413	0.523	0.523	0.523
Std. Dev. of dependent variable	0.499	0.499	0.499	0.492	0.492	0.492	0.499	0.499	0.499
Effect bottom %ile group, in s.d. units	0.108	0.103	0.044	0.123	0.111	0.057	0.129	0.115	0.037
Effect top %ile group, in s.d. units	-0.078	-0.088	-0.041	-0.080	-0.082	-0.037	-0.093	-0.095	-0.040
Adjusted R^2	0.4685	0.5026	0.5175	0.5055	0.5267	0.5370	0.4572	0.4983	0.5166
Nb. of investigators	33,409	33,403	33,403	30,446	30,445	30,445	17,274	17,260	17,260
Nb. of investigator-grant-cycles	90,123	90,075	90,075	48,530	48,529	48,529	40,859	40,810	40,810

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Columns 1-3 includes the full sample, columns 4-6 limit the sample to the first cycle for each investigator-grant, and columns 7-9 limit it to cycles 2 and above for each investigator-grant. The independent variables are indicator variables for <10%ile and >90%ile for the difference between the maximum and minimum vintage-adjusted citation percentile of grant cycle publications, with the 10-90%iles as the omitted category. Columns 1, 4 and 7 include indicator variables for gender, investigator degree, investigator career age, year of grant evaluation, length of the grant cycle, and a set of 2,911 indicator variables corresponding to each NIH institute or center and study section. Columns 2, 5, and 8 include all of the previous controls, but also include log NIH funding to the investigator's institution, log investigator prior career citations, log investigator prior NIH grant funding, and an indicator for no prior NIH grant funding. Finally, columns 3, 6, and 9 include all previous controls, but also add log mean journal impact factor, indictor variables for grant and non-grant publications in one of four highly selective journals, and a set of 21 indicator variables separately for the number of publications during the grant cycle acknowledging and not acknowledging the grant. Robust standard errors were used, clustering at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

B. Disruption

	All renewals			First renewal			Second and future renewals		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<10%ile, maximum disruption index	0.091**	0.089**	0.054**	0.094**	0.089**	0.057**	0.080**	0.078**	0.037
	(0.012)	(0.012)	(0.012)	(0.014)	(0.013)	(0.013)	(0.024)	(0.023)	(0.023)
>90%ile, maximum disruption index	-0.019**	-0.022**	0.008†	-0.020**	-0.020**	0.008	-0.023**	-0.024**	0.009
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Investigator demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grant year and length	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IC x study section	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Institutional quality	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Investigator quality	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Pub. quantity, quality during cycle	No	No	Yes	No	No	Yes	No	No	Yes
Mean of dependent variable	0.454	0.454	0.454	0.405	0.405	0.405	0.519	0.519	0.519
Std. Dev. of dependent variable	0.498	0.498	0.498	0.491	0.491	0.491	0.500	0.500	0.500
Effect bottom %ile group, in s.d. units	0.183	0.179	0.108	0.192	0.180	0.116	0.159	0.156	0.075
Effect top %ile group, in s.d. units	-0.039	-0.044	0.016	-0.040	-0.041	0.016	-0.045	-0.048	0.017
Adjusted R^2	0.4481	0.4833	0.5017	0.4798	0.5028	0.5169	0.4427	0.4861	0.5074
Nb. of investigators	36,876	36,876	36,876	34,889	34,889	34,889	18,196	18,196	18,196
Nb. of investigator-grant-cycles	102,250	102,250	102,250	57,667	57,667	57,667	43,832	43,832	43,832

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Columns 1-3 includes the full sample, columns 4-6 limit the sample to the first cycle for each investigator-grant, and columns 7-9 limit it to cycles 2 and above for each investigator-grant. The independent variables are indicator variables for <10%ile and >90%ile for the maximum of the disruption index for grant cycle publications, with the 10-90%iles as the omitted category. Columns 1, 4 and 7 include indicator variables for gender, investigator degree, investigator career age, year of grant evaluation, length of the grant cycle, and a set of 2,911 indicator variables corresponding to each NIH institute or center and study section. Columns 2, 5, and 8 include all of the previous controls, but also include log NIH funding to the investigator's institution, log investigator prior career citations, log investigator prior NIH grant funding, and an indicator for no prior NIH grant funding. Finally, columns 3, 6, and 9 include all previous controls, but also add log mean journal impact factor, indictor variables for grant and non-grant publications in one of four highly selective journals, and a set of 21 indicator variables separately for the number of publications during the grant cycle acknowledging and not acknowledging the grant. Robust standard errors were used, clustering at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

C. Pivoting

	All renewals			First renewal			Second and future renewals		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<10%ile, fraction new MeSH pairs	0.051**	0.041**	0.045**	0.039**	0.034**	0.035**	0.048**	0.040**	0.044**
	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
>90%ile, fraction new MeSH pairs	-0.043**	-0.028**	-0.047**	-0.024**	-0.017**	-0.034**	-0.053**	-0.041**	-0.064**
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.009)	(0.008)	(0.008)
Investigator demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grant year and length	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IC x study section	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Institutional quality	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Investigator quality	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Pub. quantity, quality during cycle	No	No	Yes	No	No	Yes	No	No	Yes
Mean of dependent variable	0.455	0.455	0.455	0.406	0.406	0.406	0.519	0.519	0.519
Std. Dev. of dependent variable	0.498	0.498	0.498	0.491	0.491	0.491	0.500	0.500	0.500
Effect bottom %ile group, in s.d. units	0.101	0.083	0.091	0.080	0.069	0.072	0.096	0.081	0.088
Effect top %ile group, in s.d. units	-0.086	-0.056	-0.094	-0.048	-0.034	-0.069	-0.106	-0.082	-0.128
Adjusted R^2	0.4492	0.4844	0.5035	0.4799	0.5030	0.5179	0.4442	0.4881	0.5096
Nb. of investigators	36,543	36,538	36,538	34,496	34,496	34,496	18,127	18,114	18,114
Nb. of investigator-grant-cycles	101,387	101,336	101,336	57,007	57,006	57,006	43,628	43,576	43,576

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Columns 1-3 includes the full sample, columns 4-6 limit the sample to the first cycle for each investigator-grant, and columns 7-9 limit it to cycles 2 and above for each investigator-grant. The independent variables are indicator variables for <10%ile and >90%ile for the fraction of MeSH term pairs from grant cycle publications that were not used by the investigator in the 5 years preceding the grant cycle, with the 10-90%iles as the omitted category. Columns 1, 4 and 7 include indicator variables for gender, investigator degree, investigator career age, year of grant evaluation, length of the grant cycle, and a set of 2,911 indicator variables corresponding to each NIH institute or center and study section. Columns 2, 5, and 8 include all of the previous controls, but also include log NIH funding to the investigator's institution, log investigator prior career citations, log investigator prior NIH grant funding, and an indicator for no prior NIH grant funding. Finally, columns 3, 6, and 9 include all previous controls, but also add log mean journal impact factor, indictor variables for grant and non-grant publications in one of four highly selective journals, and a set of 21 indicator variables separately for the number of publications during the grant cycle acknowledging and not acknowledging the grant. Robust standard errors were used, clustering at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

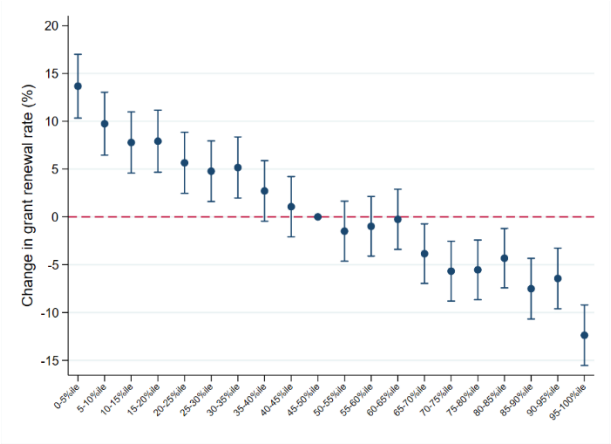
D. Standing out

	All renewals			First renewal			Second and future renewals		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<10%ile, fraction unique MeSH pairs	0.027**	0.027**	0.011*	0.031**	0.029**	0.014*	0.025**	0.021*	0.005
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)	(0.009)
>90%ile, fraction unique MeSH pairs	-0.020**	-0.013*	-0.019**	-0.011	-0.005	-0.014*	-0.030**	-0.024**	-0.025**
	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
Investigator demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grant year and length	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IC x study section	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Institutional quality	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Investigator quality	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Pub. quantity, quality during cycle	No	No	Yes	No	No	Yes	No	No	Yes
Mean of dependent variable	0.419	0.419	0.419	0.378	0.378	0.378	0.475	0.475	0.475
Std. Dev. of dependent variable	0.493	0.493	0.493	0.485	0.485	0.485	0.499	0.499	0.499
Effect bottom %ile group, in s.d. units	0.055	0.054	0.022	0.064	0.061	0.028	0.050	0.042	0.010
Effect top %ile group, in s.d. units	-0.040	-0.026	-0.038	-0.022	-0.011	-0.028	-0.059	-0.047	-0.051
Adjusted R^2	0.3976	0.4374	0.4586	0.4388	0.4633	0.4798	0.3965	0.4472	0.4713
Nb. of investigators	31,871	31,867	31,867	28,429	28,429	28,429	15,611	15,603	15,603
Nb. of investigator-grant-cycles	73,252	73,218	73,218	42,268	42,267	42,267	30,908	30,875	30,875

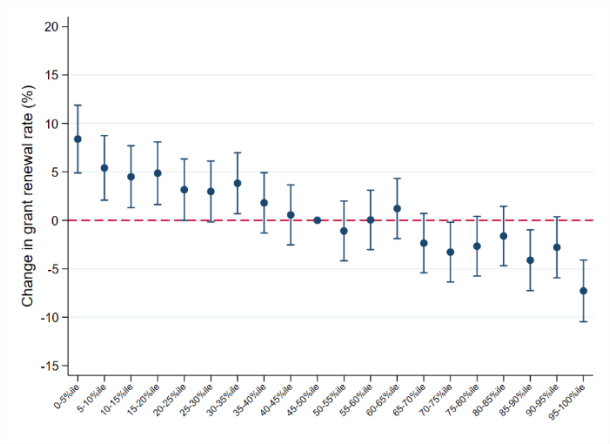
Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Columns 1-3 includes the full sample, columns 4-6 limit the sample to the first cycle for each investigator-grant, and columns 7-9 limit it to cycles 2 and above for each investigator-grant. The independent variables are indicator variables for <10%ile and >90%ile for the fraction of MeSH term pairs from grant cycle publications that were not used by funded grant proposal abstracts in the same NIH IC-study section-year, with the 10-90%iles as the omitted category. Columns 1, 4 and 7 include indicator variables for gender, investigator degree, investigator career age, year of grant evaluation, length of the grant cycle, and a set of 2,911 indicator variables corresponding to each NIH institute or center and study section. Columns 2, 5, and 8 include all of the previous controls, but also include log NIH funding to the investigator's institution, log investigator prior career citations, log investigator prior NIH grant funding, and an indicator for no prior NIH grant funding. Finally, columns 3, 6, and 9 include all previous controls, but also add log mean journal impact factor, indictor variables for grant and non-grant publications in one of four highly selective journals, and a set of 21 indicator variables separately for the number of publications during the grant cycle acknowledging and not acknowledging the grant. Robust standard errors were used, clustering at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

Figure 3. Effect of risk taking on grant renewal: Extreme tail outcomes

A. Base controls



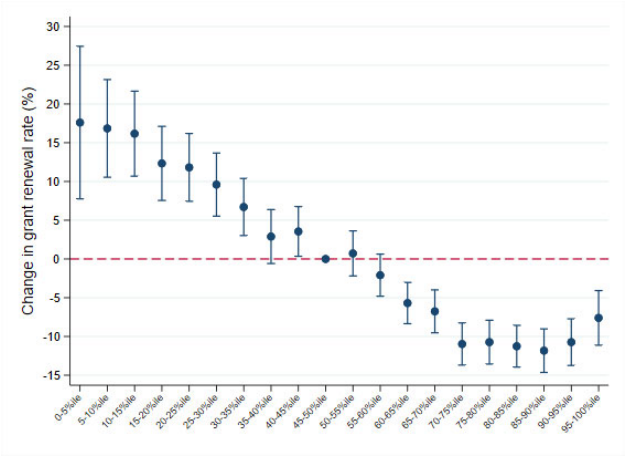
B. Controlling for grant publication quantity and quality



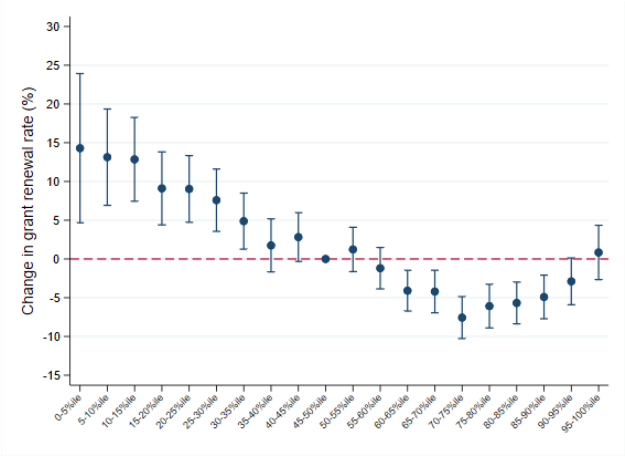
Note: The plots above present the coefficient estimates (and associated 95% confidence intervals) from a linear probability model where the dependent variable is an indicator variable for if the grant cycle was renewed. The unit of analysis is the investigator-grant-cycle. Panel A was modeled on Table 2, column 2, while panel B was modeled on Table 2, column 3. Each regression includes a set of 19 indicator variables for each ventile of the independent variable of interest, the difference between the maximum and minimum vintage-adjusted citation percentile among grant cycle publications. The omitted category is the 45-50%ile. Robust standard errors were used, clustered at the investigator and grant.

Figure 4. Effect of risk taking on grant renewal: Disruption

A. Base controls



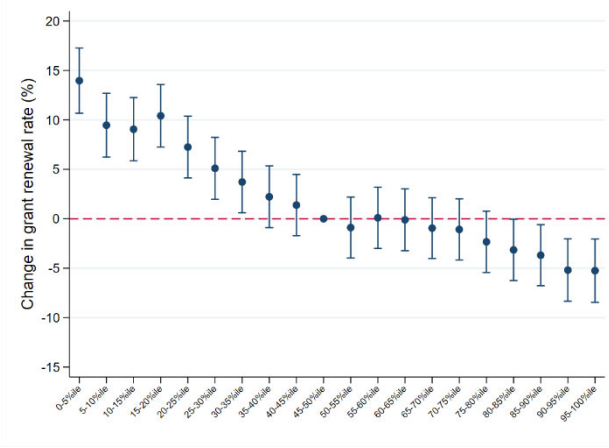
B. Controlling for grant publication quantity and quality



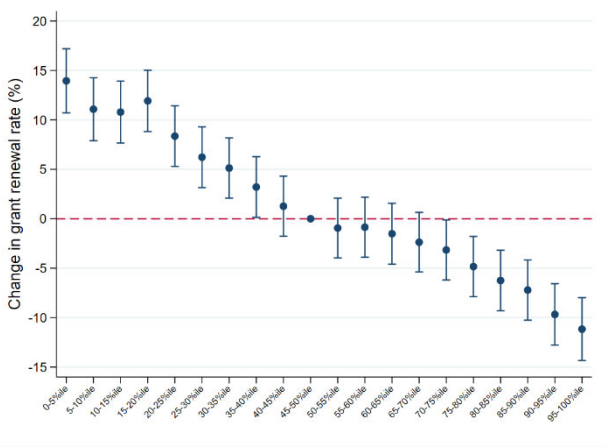
Note: The plots above present the coefficient estimates (and associated 95% confidence intervals) from a linear probability model where the dependent variable is an indicator variable for if the grant cycle was renewed. The unit of analysis is the investigator-grant-cycle. Panel A was modeled on Table 2, column 2, while panel B was modeled on Table 2, column 3. Each regression includes a set of 19 indicator variables for each ventile of the independent variable of interest, the maximum of the disruption index percentile for grant cycle publications. The omitted category is the 45-50%ile. Robust standard errors were used, clustered at the investigator and grant.

Figure 5. Effect of risk taking on grant renewal: Pivoting

A. Base controls



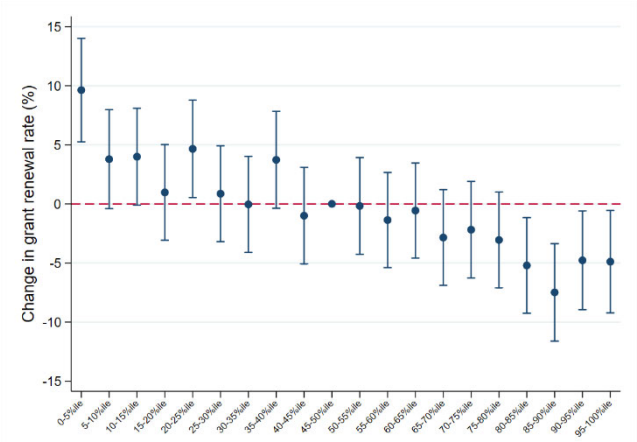
B. Controlling for grant publication quantity and quality



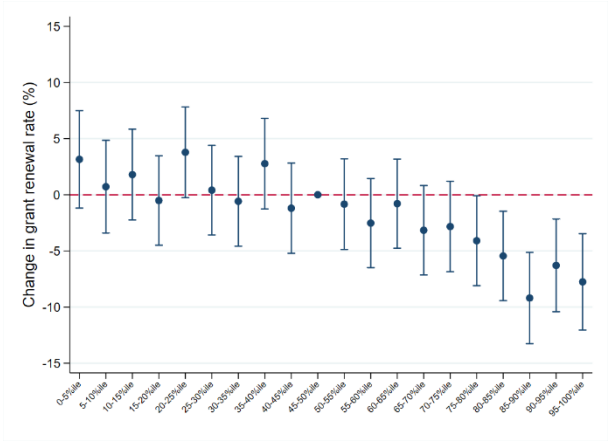
Note: The plots above present the coefficient estimates (and associated 95% confidence intervals) from a linear probability model where the dependent variable is an indicator variable for if the grant cycle was renewed. The unit of analysis is the investigator-grant-cycle. Panel A was modeled on Table 2, column 2, while panel B was modeled on Table 2, column 3. Each regression includes a set of 19 indicator variables for each ventile of the independent variable of interest, the fraction of MeSH term pairs from grant cycle publications that were not used by the investigator in the 5 years preceding the grant cycle. The omitted category is the 45-50%ile. Robust standard errors were used, clustered at the investigator and grant.

Figure 6. Effect of risk taking on grant renewal: Standing out

A. Base controls

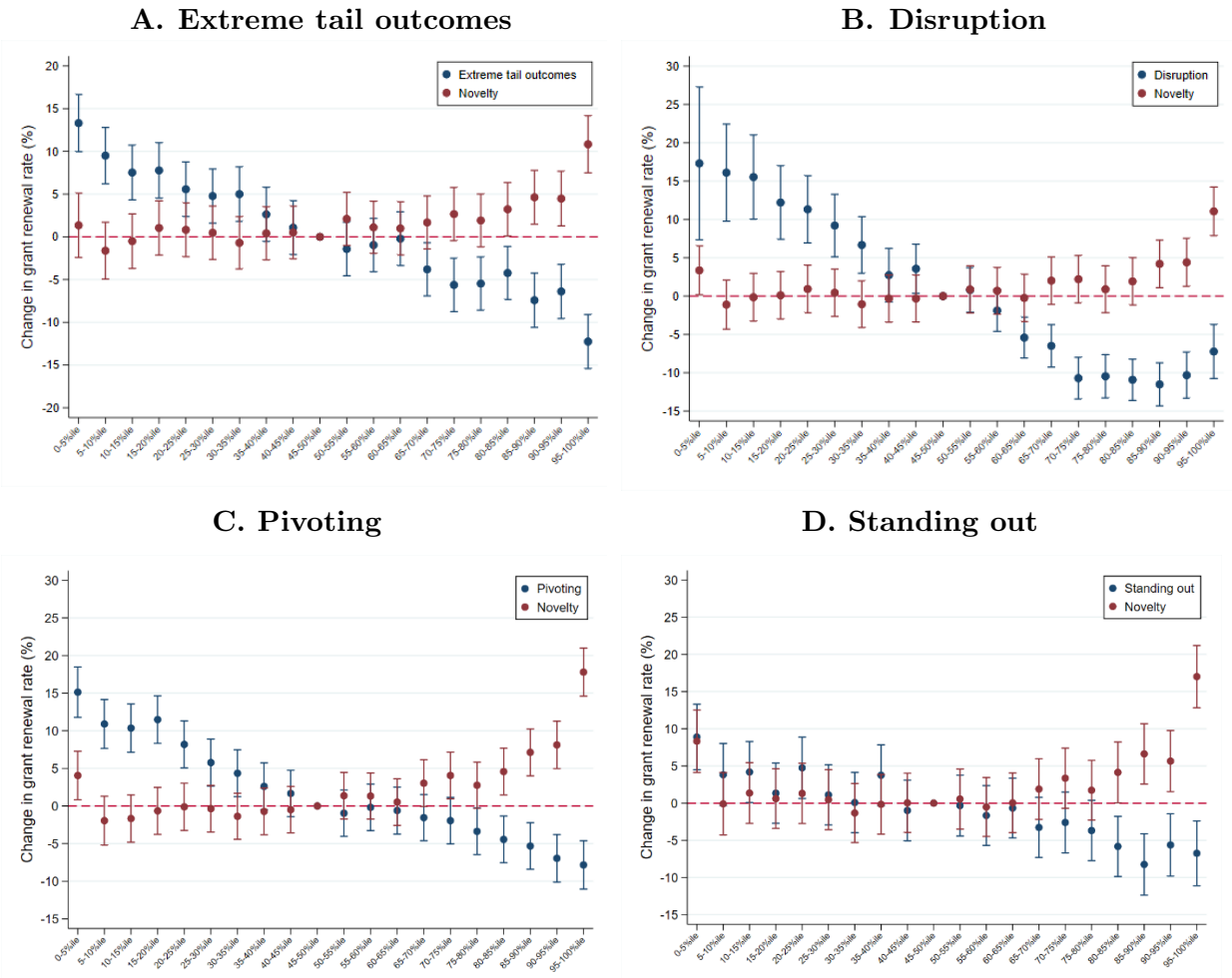


B. Controlling for grant publication quantity and quality



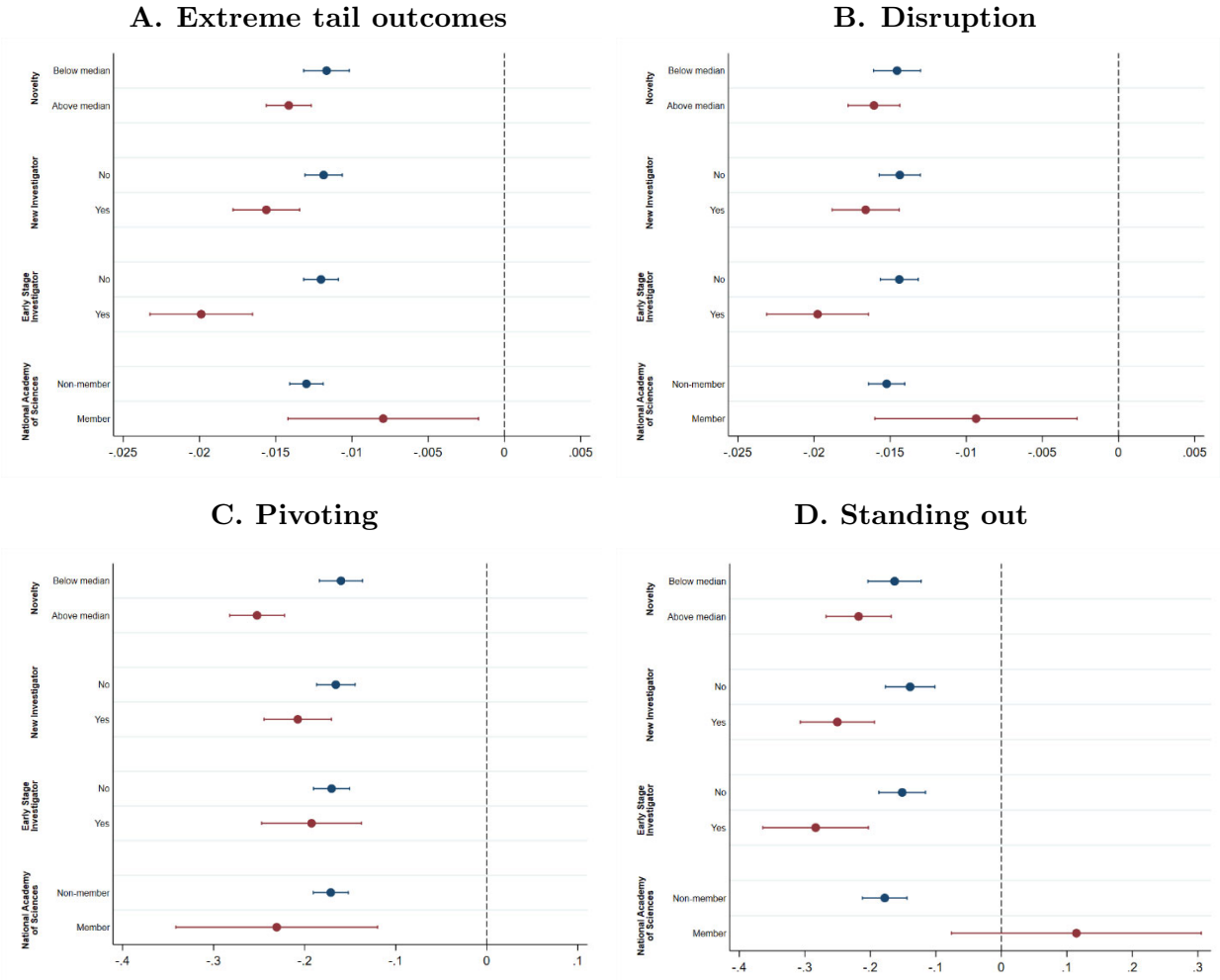
Note: The plots above present the coefficient estimates (and associated 95% confidence intervals) from a linear probability model where the dependent variable is an indicator variable for if the grant cycle was renewed. The unit of analysis is the investigator-grant-cycle. Panel A was modeled on Table 2, column 2, while panel B was modeled on Table 2, column 3. Each regression includes a set of 19 indicator variables for each ventile of the independent variable of interest, the fraction of MeSH term pairs from grant cycle publications that were not used by funded grant proposal abstracts in the same NIH IC-study section-year. The omitted category is the 45-50%ile. Robust standard errors were used, clustered at the investigator and grant.

Figure 7. Effect of risk taking and novelty on grant renewal



Note: The plots above present the coefficient estimates (and associated 95% confidence intervals) from a linear probability model where the dependent variable is an indicator variable for if the grant cycle was renewed modeled after Table 2, column 2. The unit of analysis is the investigator-grant-cycle. Each regression includes a set of 19 indicator variables for each ventile of the risk-taking measure, and 19 indicator variables for each ventile of novelty. The omitted category is the 45-50%ile. Robust standard errors were used, clustered at the investigator and grant.

Figure 8. Heterogeneity in the penalty for risk taking



Note: The plots above present the coefficient estimates (and associated 95% confidence intervals) from a linear probability model where the dependent variable is an indicator variable for if the grant cycle was renewed modeled after Table 2, column 2. The unit of analysis is the investigator-grant-cycle. Each estimate includes a risk-taking measure as a continuous variable; an indicator variable for being above median novelty, a new investigator, an early stage investigator or a member of the National Academy of Sciences as noted in the figure, as well as an interaction term between this indicator variable and the risk-taking measure. The extreme tail outcomes is scaled so a change of 1 in the regressor reflects a change of 10 in the long-run vintage-adjusted citation percentile spread, and disruption is scaled so a change of 1 in the regressor reflects a change of 10 in the maximum disruption index. Robust standard errors were used, clustered at the investigator and grant.

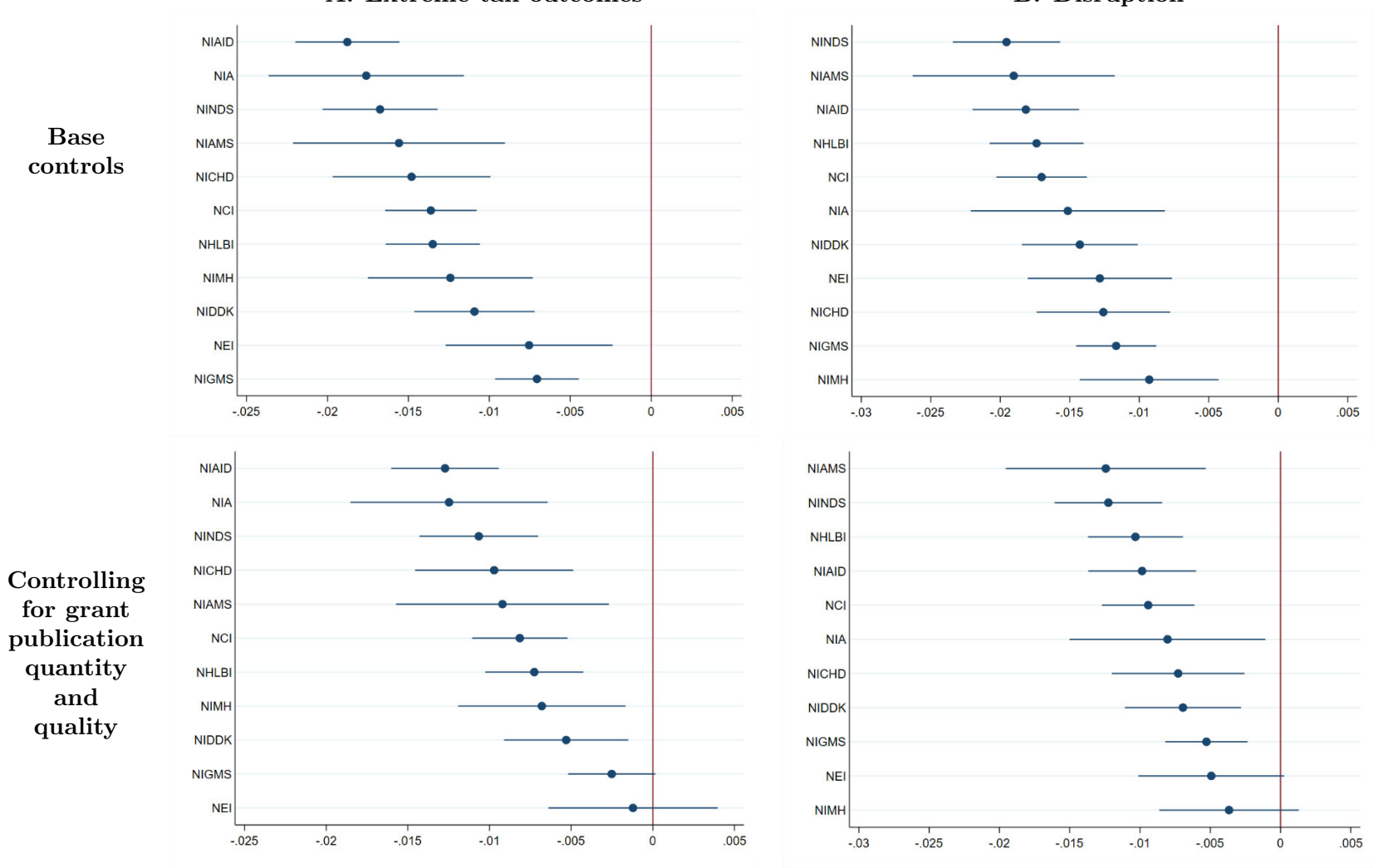
Table 3. Effect of risk taking on patents citing renewed grant cycles' publications

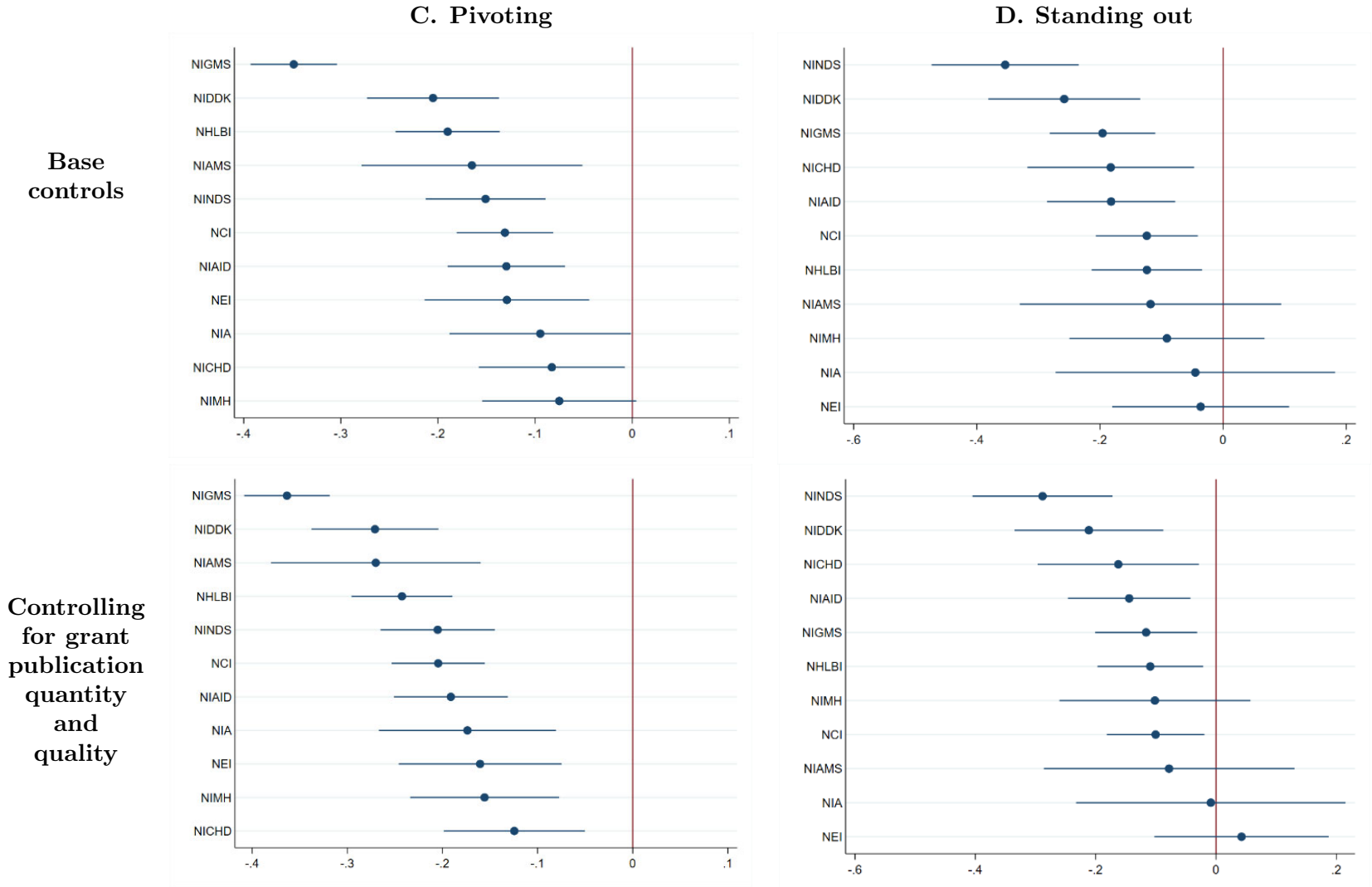
	Base controls				Grant output controls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<10%ile, Extreme tail outcomes	-0.2085**				0.4947**			
	(0.0694)				(0.0737)			
>90%ile, Extreme tail outcomes	0.3776**				-0.0053			
	(0.0539)				(0.0532)			
<10%ile, Disruption		-0.7184**				0.1323		
		(0.1543)				(0.2194)		
>90%ile, Disruption		0.7649**				0.4364**		
		(0.0589)				(0.0585)		
<10%ile, Pivoting			-0.4976**				-0.3804**	
			(0.0587)				(0.0559)	
>90%ile, Pivoting			0.3074**				0.5074**	
			(0.0910)				(0.0832)	
<10%ile, Standing out				-0.4137**				-0.1791**
				(0.0724)				(0.0687)
>90%ile, Standing out				0.1490				0.2257†
				(0.1096)				(0.1188)
Pseudo-R^2	0.3656	0.3787	0.3695	0.3422	0.4753	0.4847	0.4863	0.4510
Log pseudolikelihood	-552,146	-581,905	-588,151	-380,605	-456,677	-482,638	-479,159	-317,626
Nb. of investigators	16,883	18,338	18,208	16,141	16,883	18,338	18,208	16,141
Nb. of investigator-grant-cycles	40,657	45,164	44,820	30,225	40,657	45,164	44,820	30,225

Note: Estimates stem from a Poisson model where the unit of analysis is the investigator-grant-cycle, with the sample restricted to those investigator-grant-cycles which were renewed. The dependent variable is patent citations to grant cycle publications. Columns with base controls are very similar to Table 2, column 2, while those with grant output controls are very similar to Table 2, column 3. Each estimate includes a pair of indicator variables for being <10%ile and >90%ile for the included risk taking measure, with 10-90%ile as the omitted category, using extreme tail outcomes (columns 1 and 5), disruption (columns 2 and 6), pivoting (columns 3 and 7), and standing out (columns 4 and 8). Robust standard errors, double-clustered at the investigator and grant level. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

Appendix
Ancillary Results and Robustness Checks

Figure A1. Heterogeneity in the penalty for risk taking by NIH Institute

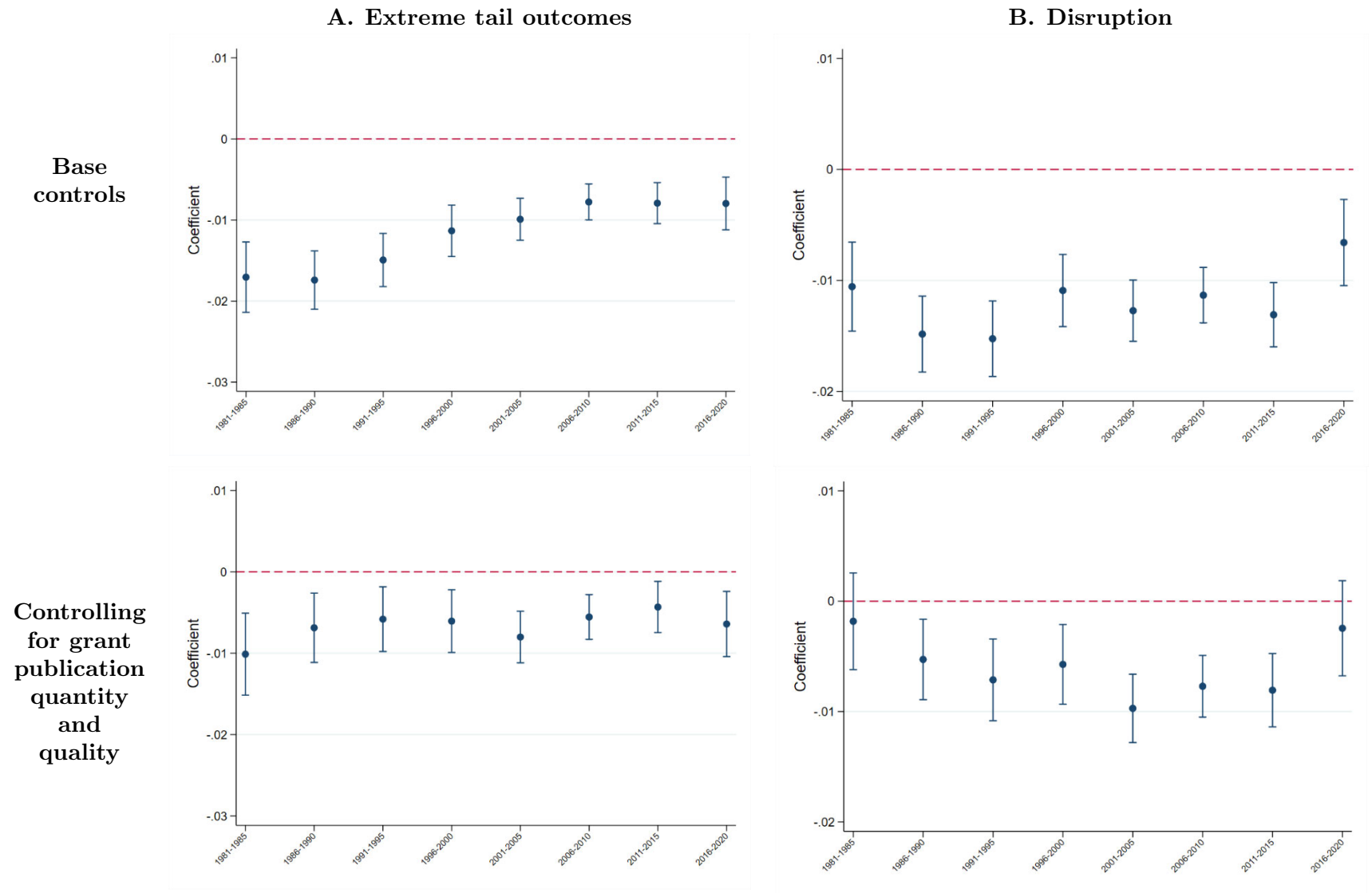


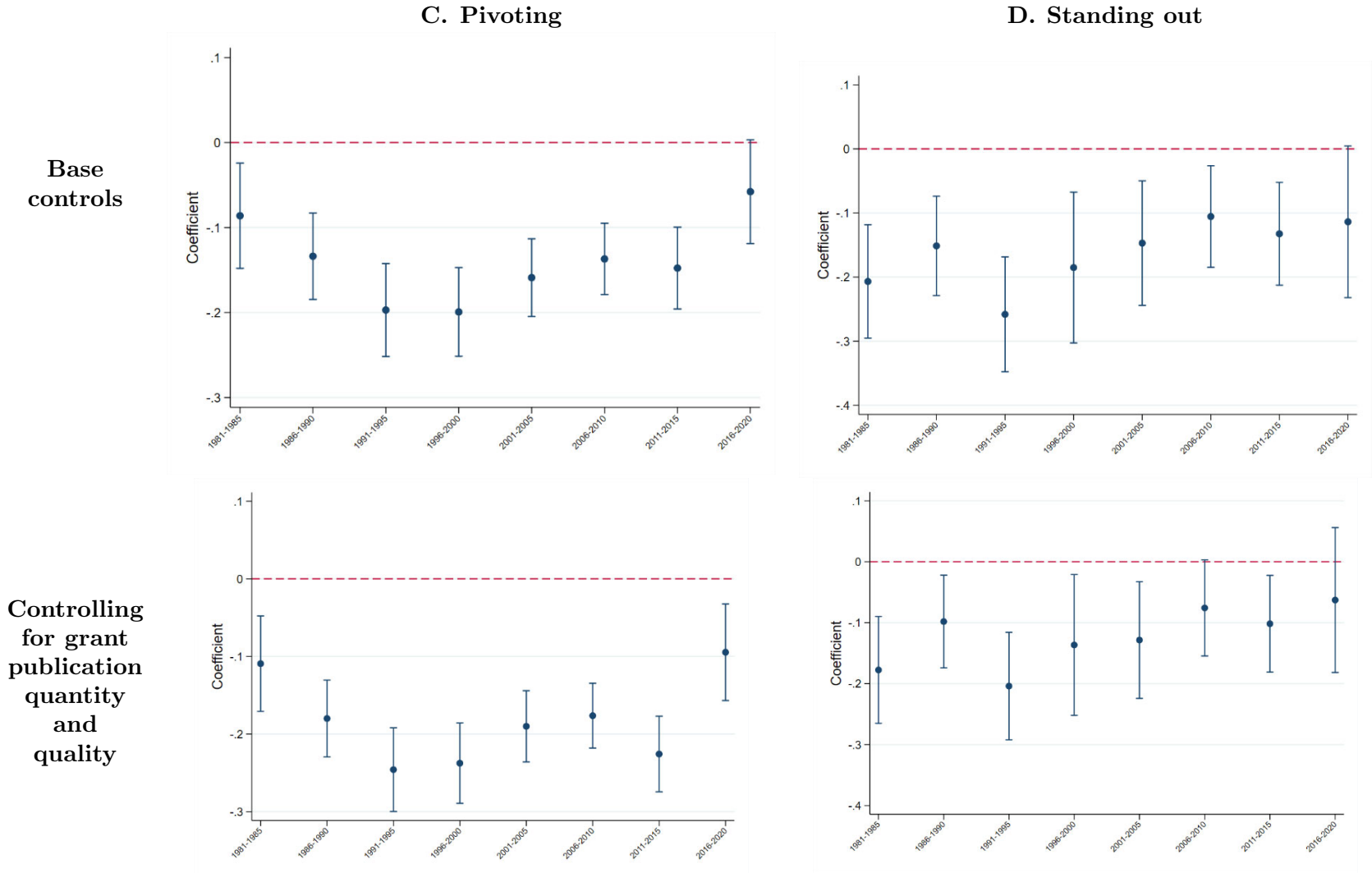


Note: The plots above present the coefficient estimates (and associated 95% confidence intervals) for the interaction between the risk taking measure and an indicator variable for each NIH Institute of Center from a linear probability model. The dependent variable is an indicator variable for if the

grant cycle was renewed. The unit of analysis is the investigator-grant-cycle. Each panel presents the results for a different risk taking measure included as a continuous variable: extreme tail outcomes scaled so a change of 1 in the regressor reflects a change of 10 in the long-run vintage-adjusted citation percentile spread (Panel A), disruption scaled so a change of 1 in the regressor reflects a change of 10 in the maximum disruption index percentile (Panel B), pivoting (Panel C), and standing out (Panel D). Row 1 was modeled on Table 2, column 2, while row 2 was modeled on Table 2, column 3. Coefficients for the 11 largest NIH institutes or centers, by number of R01-equivalent grant cycles, are displayed. Robust standard errors were used, clustered at the investigator and grant. National Cancer Institute – NCI; National Eye Institute – NEI; National Heart, Lung, and Blood Institute – NHLBI; National Institute on Aging – NIA; National Institute of Allergy and Infectious Diseases – NIAID; National Institute of Arthritis and Musculoskeletal and Skin Diseases – NIAMS; National Institute of Child Health and Human Development – NICHD; National Institute of Diabetes and Digestive and Kidney Diseases – NIDDK; National Institute of General Medical Sciences – NIGMS; National Institute of Mental Health – NIMH; National Institute of Neurological Disorders and Stroke – NINDS.

Figure A2. Heterogeneity in the penalty for risk taking over time



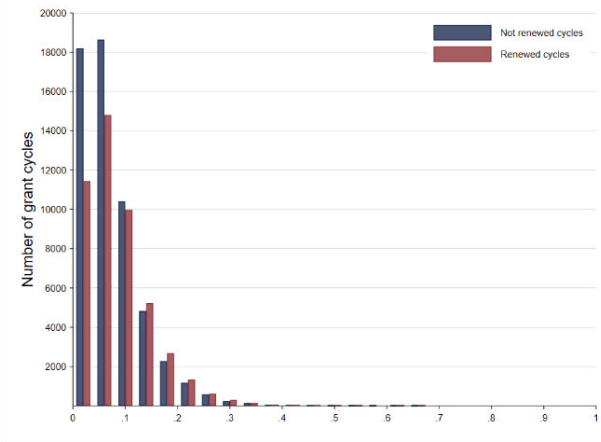


Note: The plots above present the coefficient estimates (and associated 95% confidence intervals) from a linear probability model where the dependent variable is an indicator variable for if the grant cycle was renewed. Each plotted coefficient reflects the results of a separate regression, with each panel presenting the results for a different risk taking measure included as a continuous variable: extreme tail outcomes scaled so a change of 1 in the regressor reflects a change of 10 in the long-run vintage-adjusted citation percentile spread (Panel A), disruption scaled so a change of 1 in the

regressor reflects a change of 10 in the maximum disruption index percentile (Panel B), pivoting (Panel C), and standing out (Panel D). The unit of analysis is the investigator-grant-cycle. Row 1 was modeled on Table 2, column 2, while row 2 was modeled on Table 2, column 3. Robust standard errors were used, clustered at the investigator and grant.

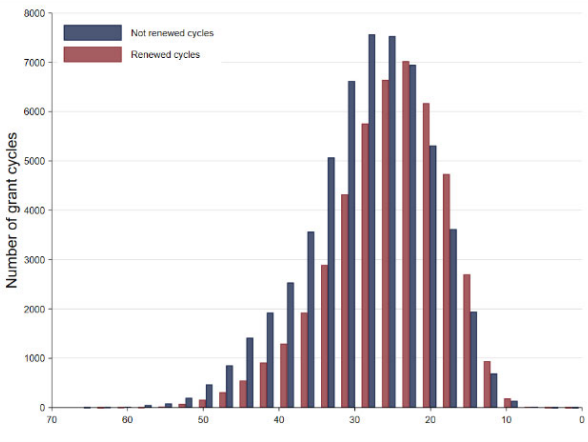
Figure A3. Distribution of measures of novelty

A. Fraction novel MeSH pairs



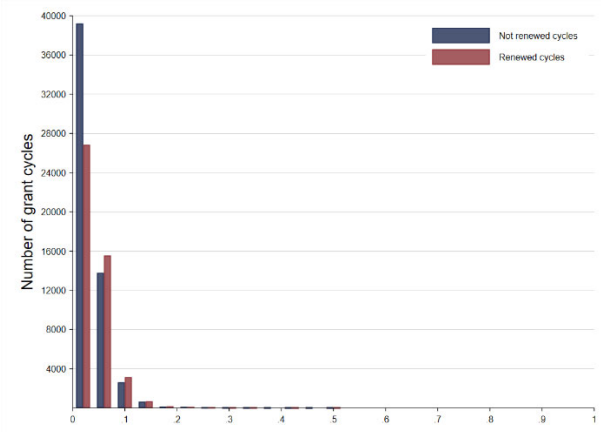
Note: Histogram of the fraction of MeSH pairs among grant cycle publications that were first used by any PubMed publication within the prior 3 years. Unit of analysis is at the investigator-grant-cycle level.

B. MeSH pair age, mean



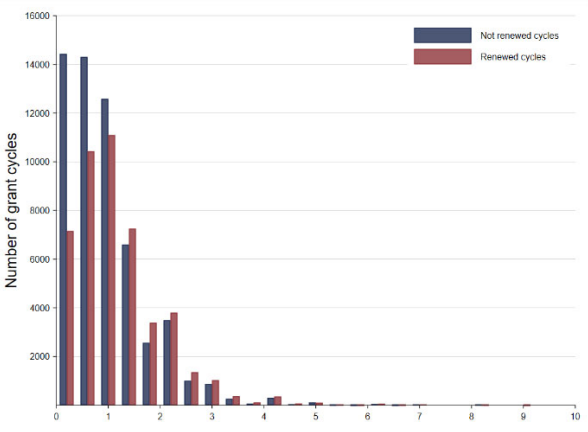
Note: Histogram of the mean MeSH pair age among all grant cycle publication MeSH pairs. Age is defined as the difference between the year of the focal publication and the year the MeSH pair was first used by any publication in PubMed. Unit of analysis is at the investigator-grant-cycle level.

C. Fraction new bigrams



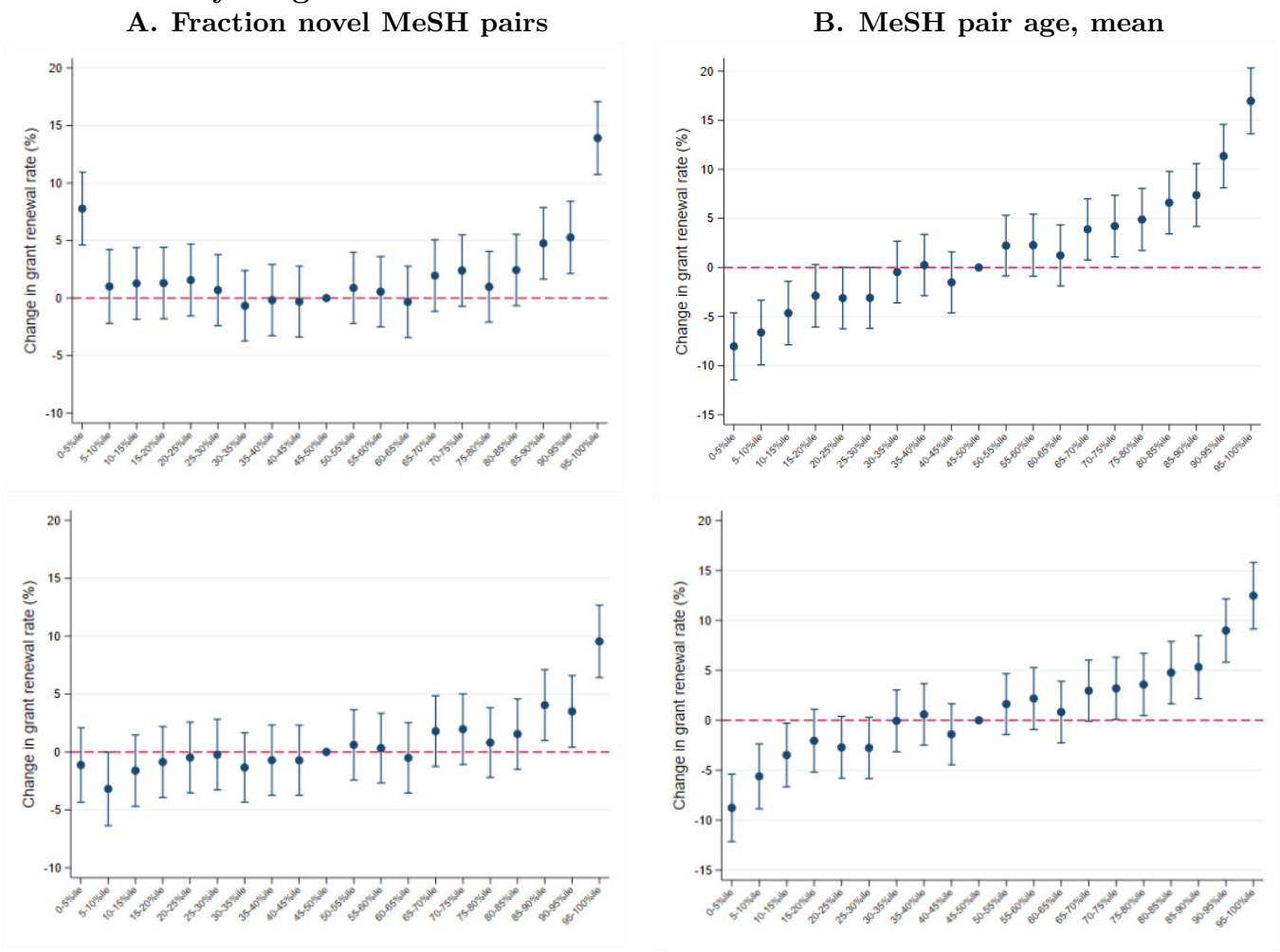
Note: Histogram of the fraction of abstract and title bigrams which are new to Microsoft Academic Graph among grant cycle publications. Unit of analysis is at the investigator-grant-cycle level.

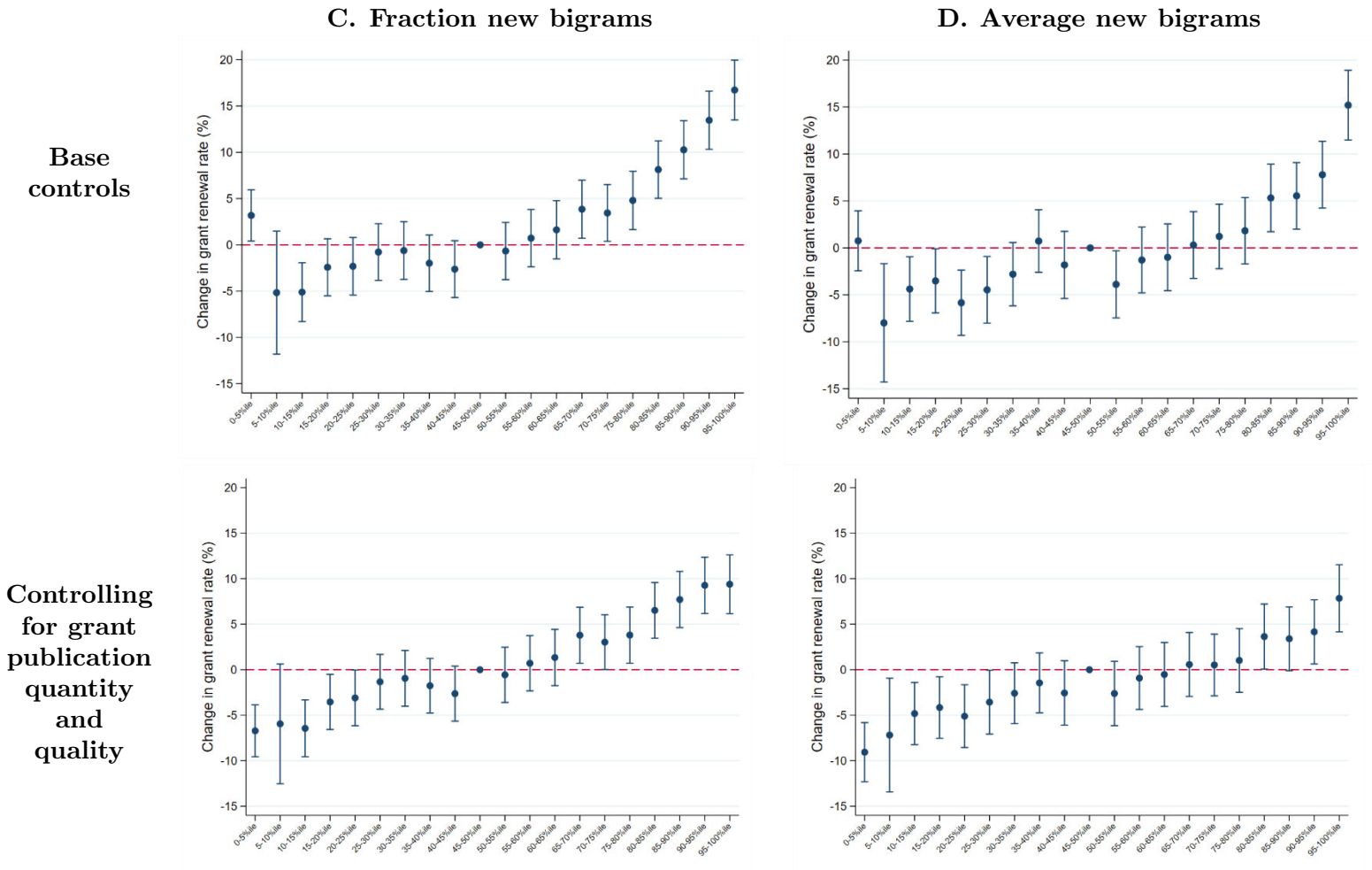
D. Average new bigrams



Note: Histogram of the mean number of abstract and title bigrams per publication which are new to Microsoft Academic Graph among grant cycle publications. Unit of analysis is at the investigator-grant-cycle level. Two outliers with more than 10 average new bigrams are excluded.

Figure A4. Effect of novelty on grant renewal

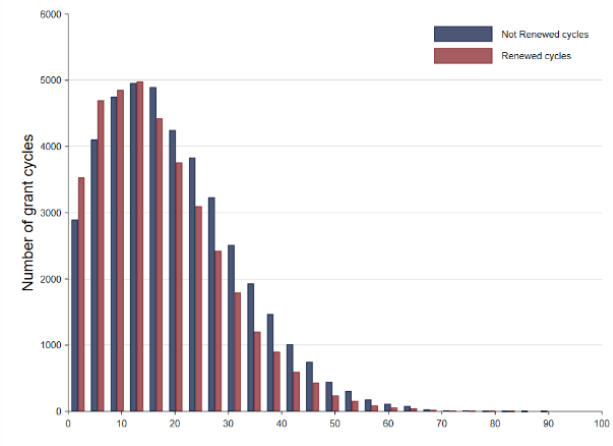




Note: The plots above present the coefficient estimates (and associated 95% confidence intervals) from a linear probability model where the dependent variable is an indicator variable for if the grant cycle was renewed. The unit of analysis is the investigator-grant-cycle. Row 1 was modeled on Table 2, column 2, while row 2 was modeled on Table 2, column 3. Each regression includes a set of 19 indicator variables for each ventile of the independent variable of interest, which is the fraction of MeSH pairs first used in PubMed within the prior 3 years (Panel A), the mean MeSH pair age (Panel B), fraction of abstract and title bigrams new to Microsoft Academic Graph (Panel C), and average number of new bigrams per publication (Panel D) for grant cycle publications. For ease of comparison to other figures, age (Panels A and B) is inverted so that higher percentiles represent grant cycles with younger mean ages. The omitted category is the 45-50%ile. Robust standard errors were used, clustered at the investigator and grant.

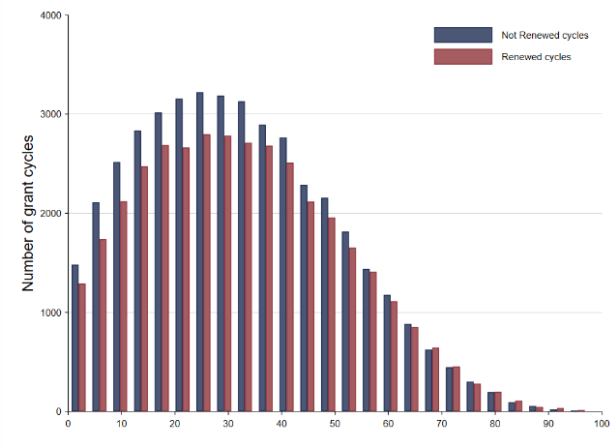
Figure A5. Distribution of alternative measures of extreme tail outcomes

C. Range, max-median citation percentile



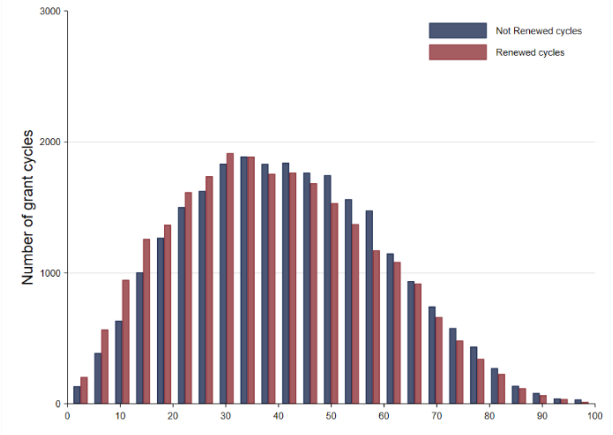
Note: Histogram of the range between the maximum and median vintage-adjusted citation percentile for grant cycle publications. Unit of analysis is at the investigator-grant-cycle level.

D. Range, median-min citation percentile



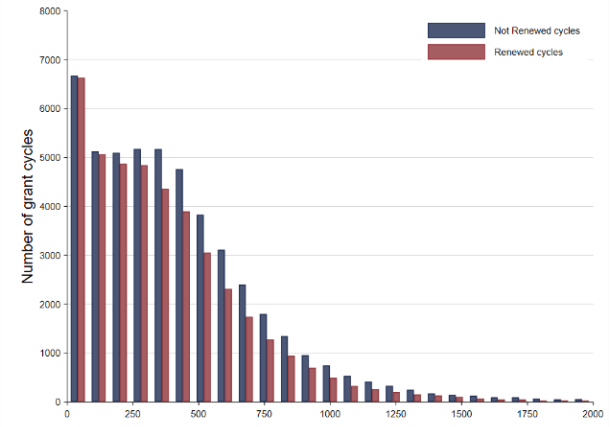
Note: Histogram of the range between the median and minimum vintage-adjusted citation percentile for grant cycle publications. Unit of analysis is at the investigator-grant-cycle level.

C. Range, winsorized max-min citation percentile



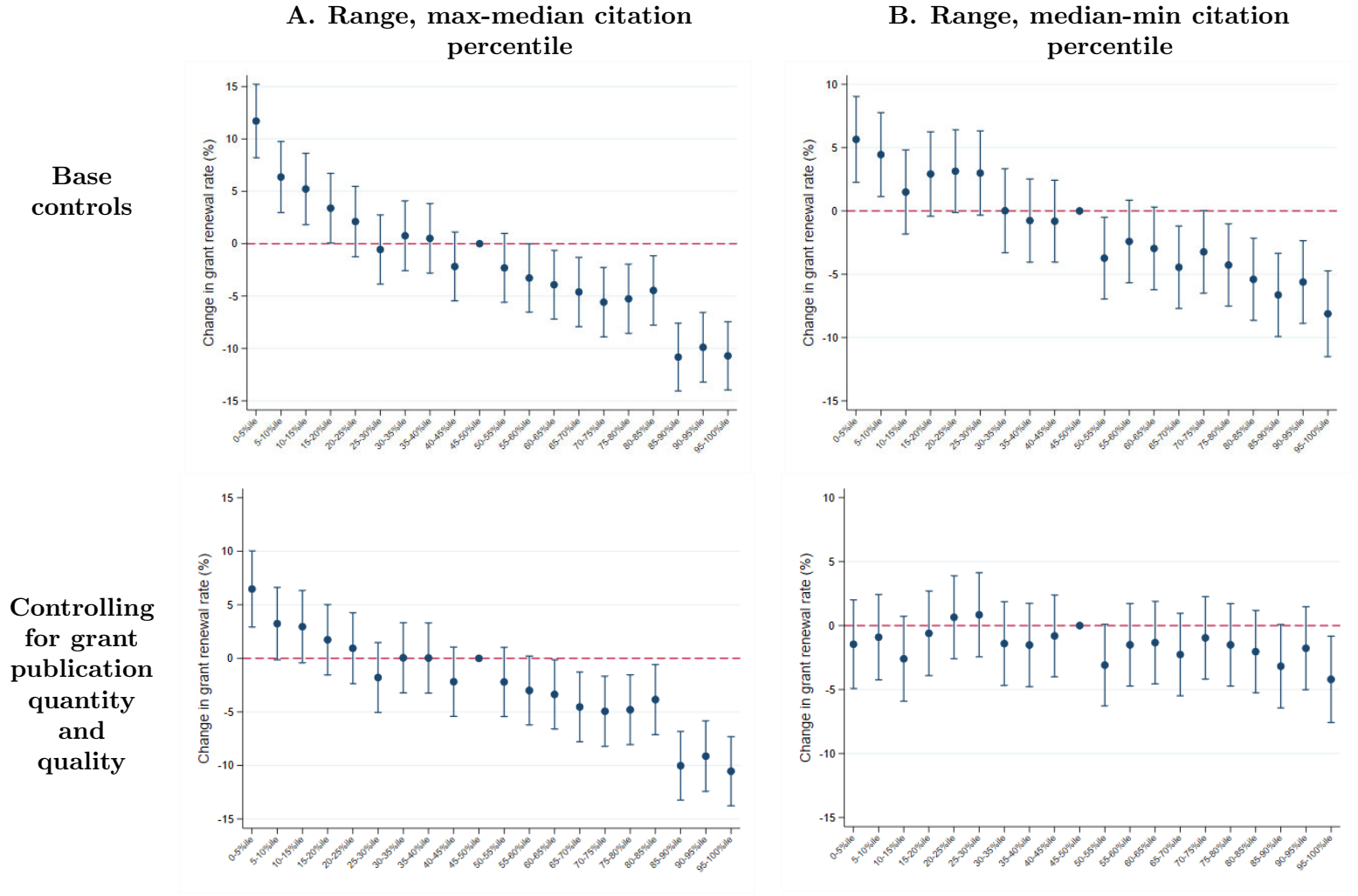
Note: Histogram of the range between the winsorized maximum and minimum vintage-adjusted citation percentile for grant cycle publications. Unit of analysis is at the investigator-grant-cycle level.

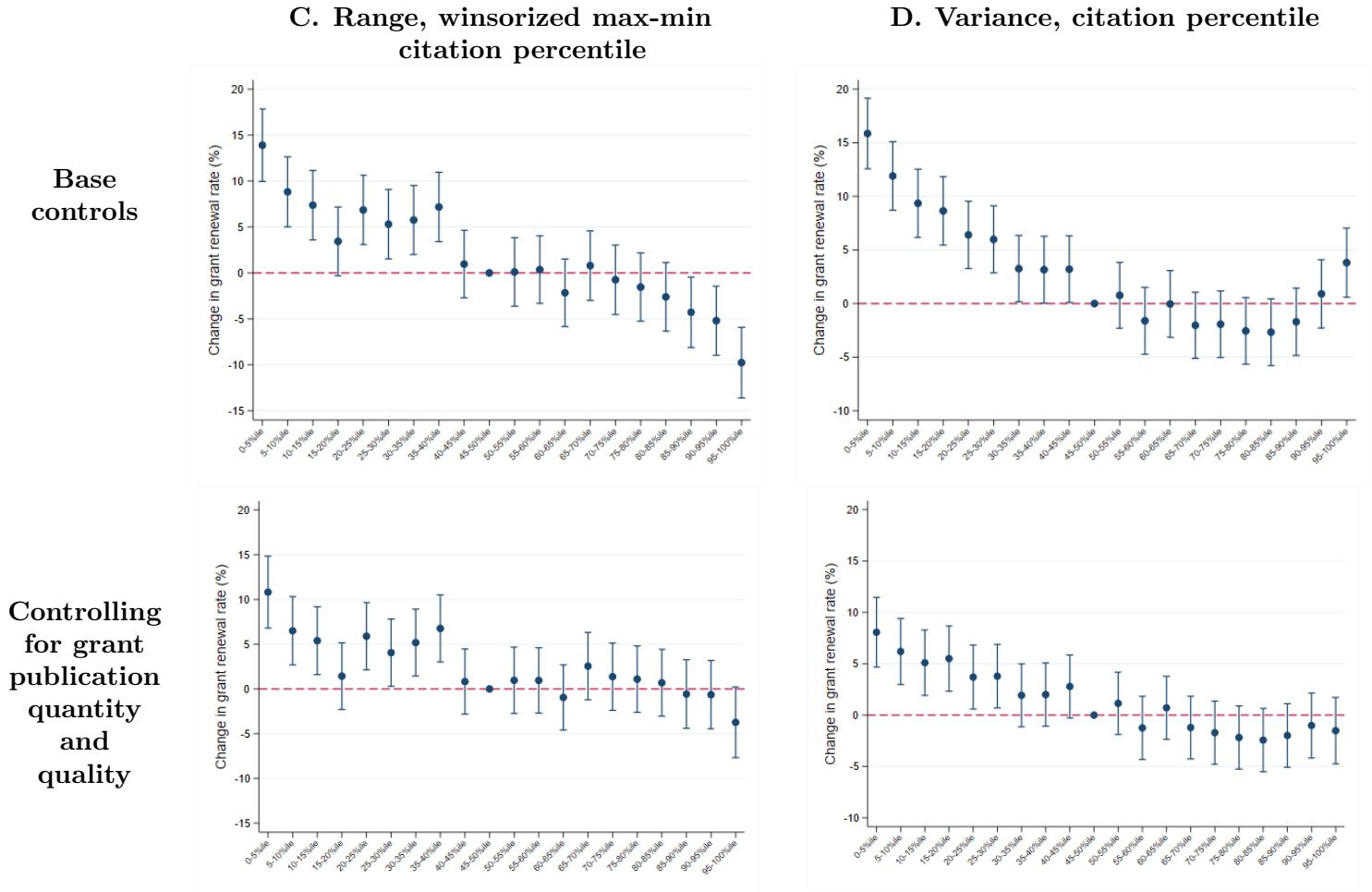
D. Variance, citation percentile



Note: Histogram of the variance in vintage-adjusted citation percentile for grant cycle publications. Unit of analysis is at the investigator-grant-cycle level. Four hundred seventy four outliers with a variance over 2,000 are not shown.

Figure A6. Effect of risk taking on grant renewal: Alternative measures of extreme tail outcomes

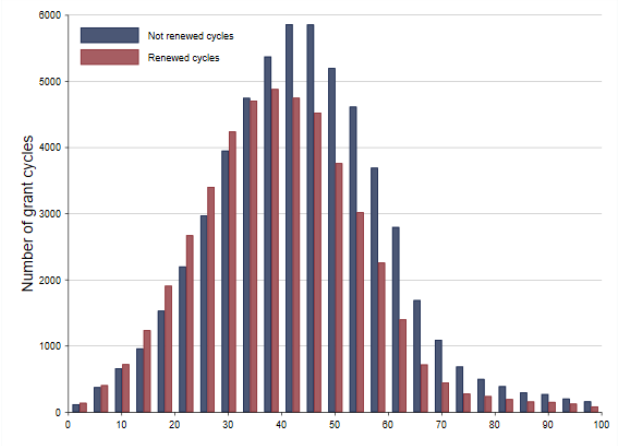




Note: The plots above present the coefficient estimates (and associated 95% confidence intervals) from a linear probability model where the dependent variable is an indicator variable for if the grant cycle was renewed. The unit of analysis is the investigator-grant-cycle. Row 1 was modeled on Table 2, column 2, while row 2 was modeled on Table 2, column 3. Each regression includes a set of 19 indicator variables for each ventile of the independent variable of interest: the difference between the maximum and median vintage-adjusted citation percentile (Panel A), difference between the median and minimum vintage-adjusted citation percentile (Panel B), difference between the winsorized maximum and minimum vintage-adjusted citation percentile (Panel C), and variance in vintage-adjusted citation percentile (Panel D) among grant cycle publications. The omitted category is the 45-50%ile. Robust standard errors were used, clustered at the investigator and grant.

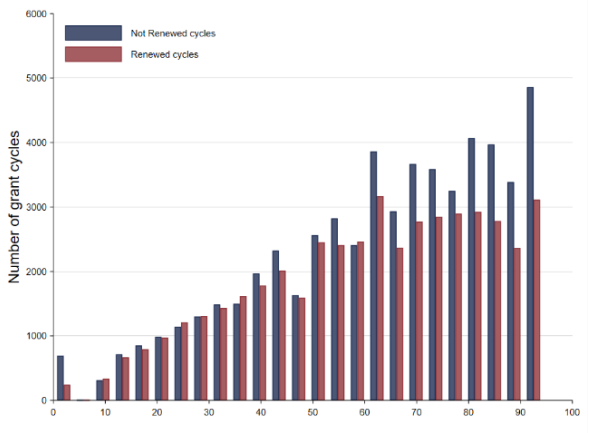
Figure A7. Distribution of alternative measures of disruption

A. Disruption index percentile, mean



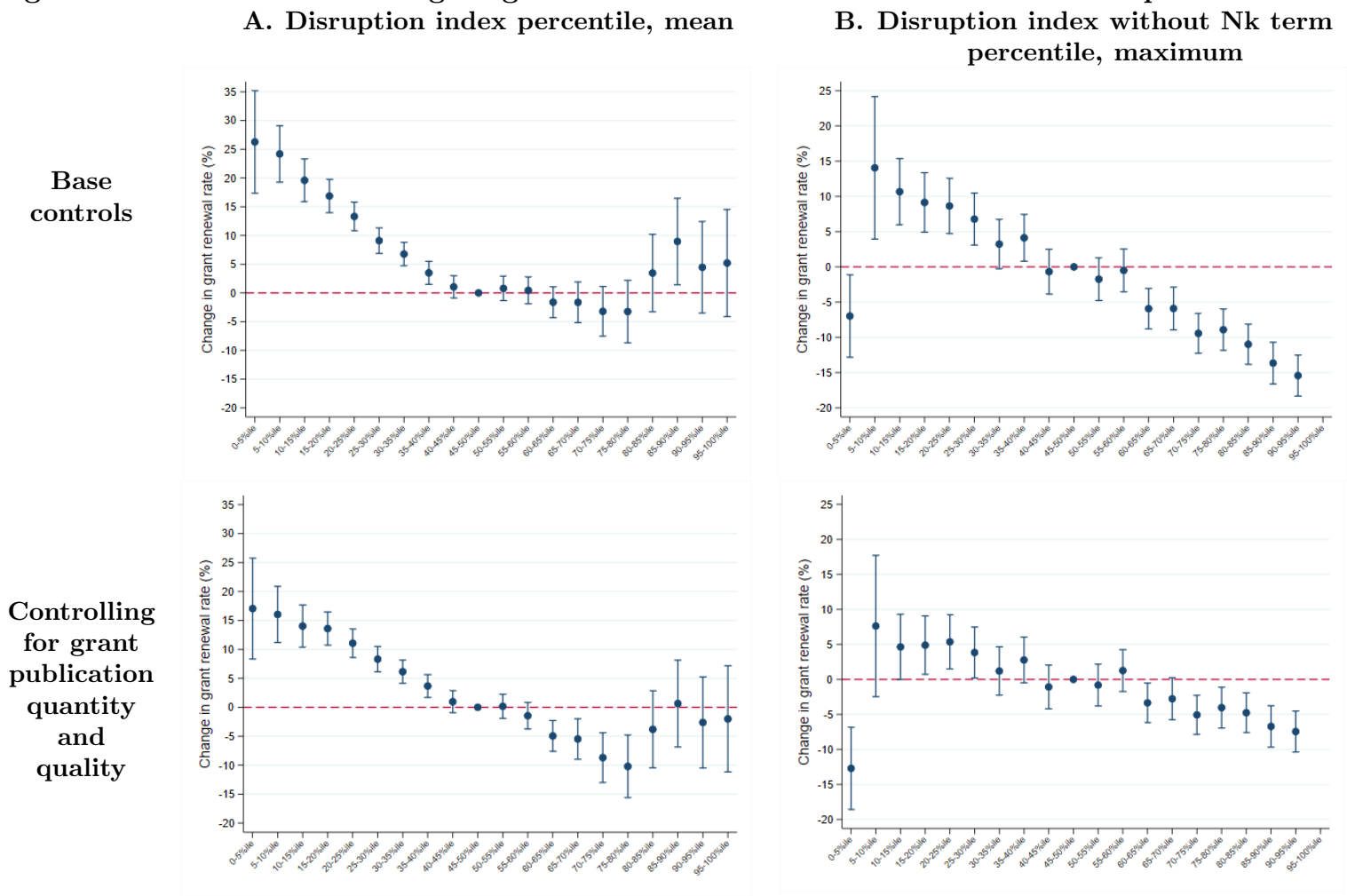
Note: Histogram of the mean disruption index percentile among grant cycle publications. Unit of analysis is at the investigator-grant-cycle level.

B. Disruption index without Nk term percentile, maximum



Note: Histogram of the maximum disruption index percentile among grant cycle publications. This alternative specification of the disruption index follows Bornmann et al. (2020) in excluding the Nk term, which adjusts the disruption index for citations to a focal paper's references that do not cite the focal paper itself. Unit of analysis is at the investigator-grant-cycle level.

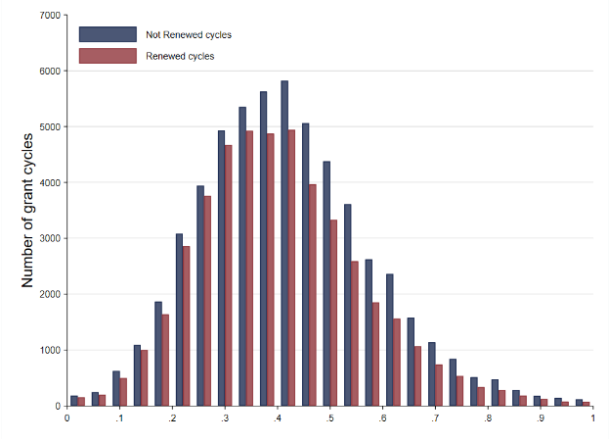
Figure A8. Effect of risk taking on grant renewal: Alternative measures of disruption



Note: The plots above present the coefficient estimates (and associated 95% confidence intervals) from a linear probability model where the dependent variable is an indicator variable for if the grant cycle was renewed. The unit of analysis is the investigator-grant-cycle. Row 1 was modeled on Table 2, column 2, while row 2 was modeled on Table 2, column 3. Each regression includes a set of 19 indicator variables for each ventile of the independent variable of interest, which is the mean disruption index percentile (Panel A) and maximum disruption index percentile following Bormann et al. (2020) in excluding the Nk term (Panel B) among grant cycle publications. The omitted category is the 45-50%ile. Robust standard errors were used, clustered at the investigator and grant.

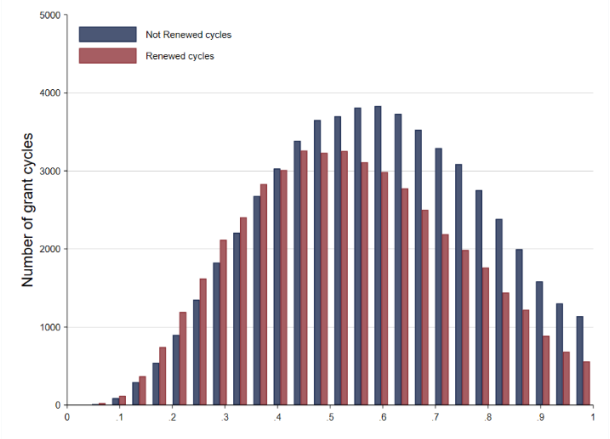
Figure A9. Distribution of alternative measures of pivoting

A. Fraction new MeSH vs. investigator's prior work



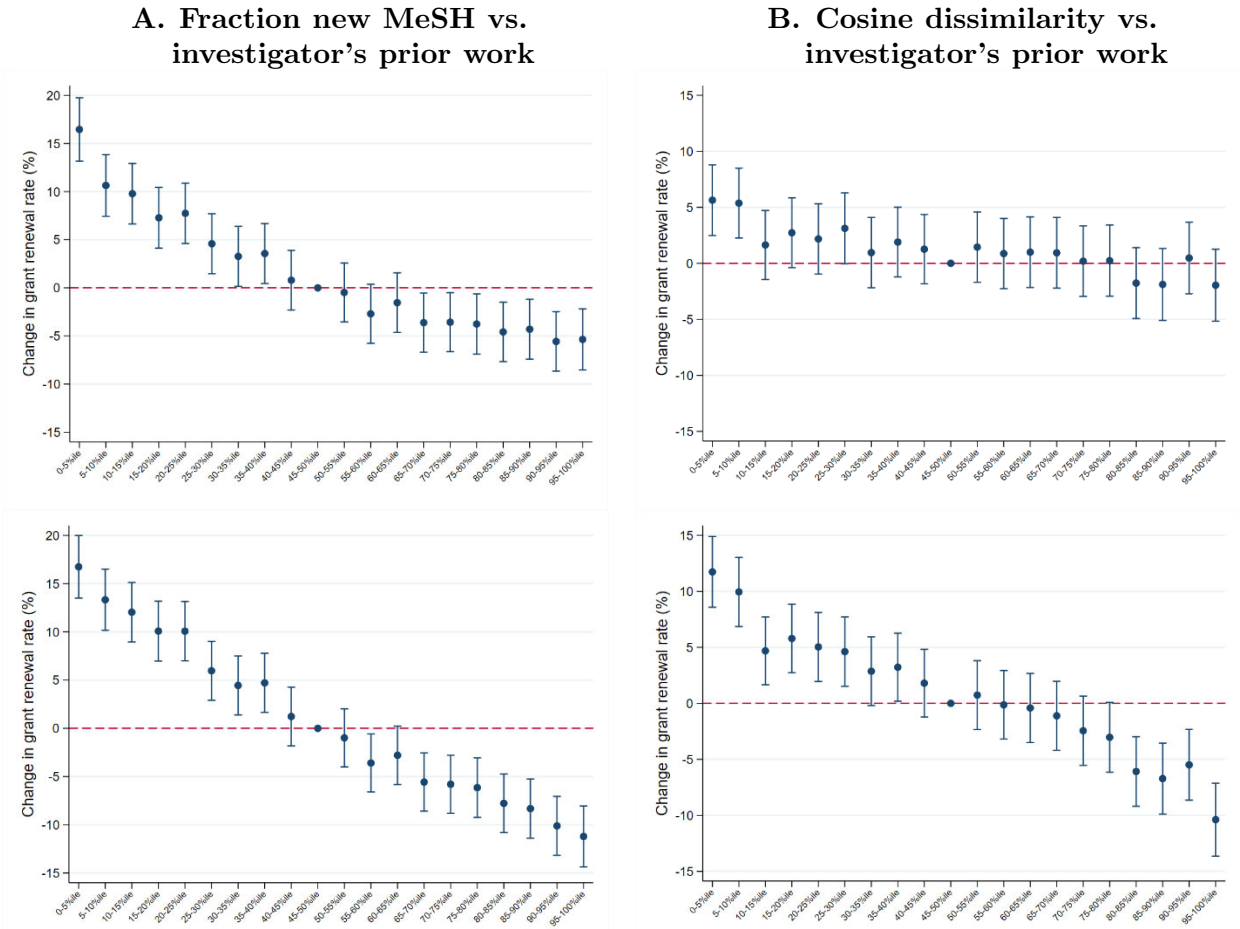
Note: Histogram of the fraction of MeSH terms from grant cycle publications that were not used in the MeSH for the investigator's publications in the 5 years preceding the grant cycle. Unit of analysis is at the investigator-grant-cycle level.

B. Cosine dissimilarity vs. investigator's prior work



Note: Histogram of the cosine dissimilarity between the MeSH of grant cycle publications and MeSH for the investigator's publications in the 5 years preceding the grant. MeSH were weighted using term-frequency-inverse document frequency. Unit of analysis is at the investigator-grant-cycle level.

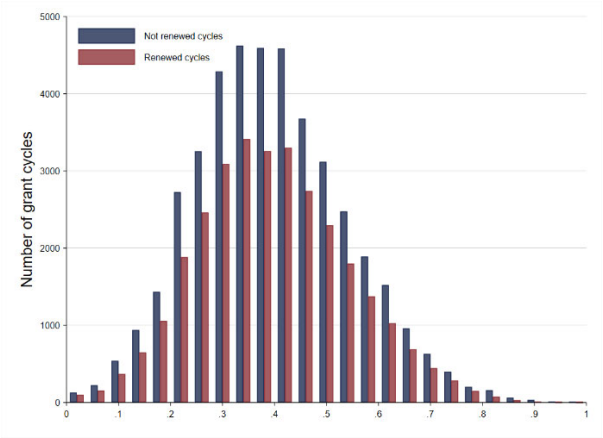
Figure A10. Effect of risk taking on grant renewal: Alternative measures of pivoting



Note: The plots above present the coefficient estimates (and associated 95% confidence intervals) from a linear probability model where the dependent variable is an indicator variable for if the grant cycle was renewed. The unit of analysis is the investigator-grant-cycle. Row 1 was modeled on Table 2, column 2, while row 2 was modeled on Table 2, column 3. Each regression includes a set of 19 indicator variables for each ventile of the independent variable of interest, which is the fraction of new MeSH (Panel A) and cosine dissimilarity with term frequency-inverse document frequency weighting (Panel B) comparing MeSH from grant cycle publications to those for the investigator in the 5 years preceding the grant cycle. The omitted category is the 45-50%ile. Robust standard errors were used, clustered at the investigator and grant.

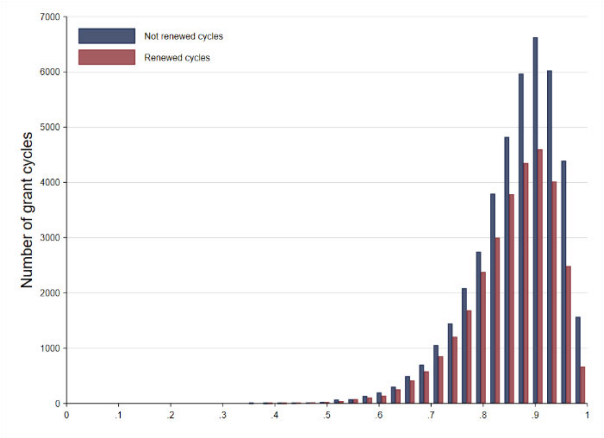
Figure A11. Distribution of alternative measures of standing out

A. Fraction unique MeSH vs. other investigators



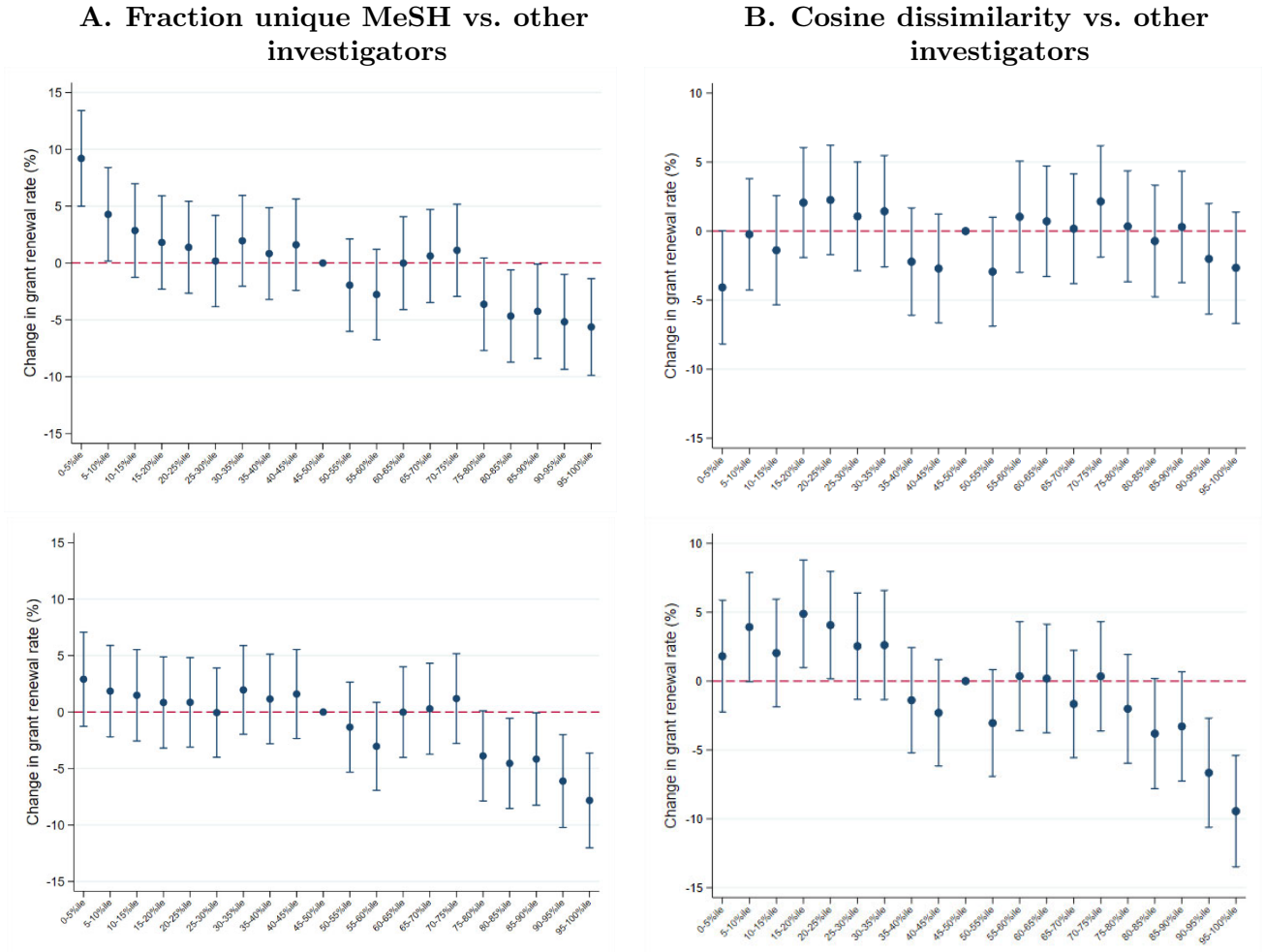
Note: Histogram of the fraction of MeSH terms for grant cycle publications that were not used by any funded grant proposal abstracts in the same NIH IC-study section-year. Unit of analysis is at the investigator-grant-cycle level.

B. Cosine dissimilarity vs. other investigators



Note: Histogram of the cosine dissimilarity between the MeSH for grant cycle publications and all funded grant proposal abstracts in the same NIH IC-study section-year. MeSH were weighted using term-frequency-inverse document frequency. Unit of analysis is at the investigator-grant-cycle level.

Figure A12. Effect of risk taking on grant renewal: Alternative measures of standing out



Note: The plots above present the coefficient estimates (and associated 95% confidence intervals) from a linear probability model where the dependent variable is an indicator variable for if the grant cycle was renewed. The unit of analysis is the investigator-grant-cycle. Row 1 was modeled on Table 2, column 2, while row 2 was modeled on Table 2, column 3. Each regression includes a set of 19 indicator variables for each ventile of the independent variable of interest, which is the fraction of unique MeSH (Panel A) and cosine dissimilarity with term frequency-inverse document frequency weighting (Panel B) comparing MeSH from grant cycle publications to those for funded grant proposal abstracts in the same NIH IC-study section-year. The omitted category is the 45-50%ile. Robust standard errors were used, clustered at the investigator and grant.

Table A1. Effect of novelty on the risk taking grant renewal penalty

A. Novelty above the 50%ile

	Base controls				Grant output controls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Novelty >50%ile	0.0254**	0.0211**	0.0922**	0.0578*	0.0284**	0.0191*	0.1200**	0.0683**
	(0.0057)	(0.0077)	(0.0142)	(0.0242)	(0.0056)	(0.0076)	(0.0139)	(0.0238)
Extreme tail outcomes	-0.0117**				-0.0058**			
	(0.0008)				(0.0009)			
Extreme tail outcomes × Novelty >50%ile	-0.0025*				-0.0025*			
	(0.0010)				(0.0010)			
Disruption		-0.0145**				-0.0078**		
		(0.0008)				(0.0008)		
Disruption × Novelty >50%ile		-0.0015				-0.0007		
		(0.0011)				(0.0011)		
Pivoting			-0.1601**				-0.2124**	
			(0.0121)				(0.0119)	
Pivoting × Novelty >50%ile			-0.0922**				-0.1178**	
			(0.0184)				(0.0181)	
Standing out				-0.1630**				-0.1254**
				(0.0207)				(0.0205)
Standing out × Novelty >50%ile				-0.0549†				-0.0620*
				(0.0294)				(0.0289)
Mean of dependent variable	0.464	0.455	0.455	0.419	0.464	0.455	0.455	0.419
Std. Dev. of dependent variable	0.499	0.498	0.498	0.493	0.499	0.498	0.498	0.493
Adjusted R^2	0.5043	0.4867	0.4860	0.4382	0.5182	0.5028	0.5063	0.4593
Nb. of investigators	33,348	36,799	36,538	31,867	33,348	36,799	36,538	31,867
Nb. of investigator-grant-cycles	89,960	102,035	101,336	73,218	89,960	102,035	101,336	73,218

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Columns with base controls are very similar to Table 2, column 2, while those with grant output controls are very similar to Table 2, column 3. Each estimate includes a risk-taking measure as a continuous variable: extreme tail outcomes scaled so a change of 1 in the regressor reflects a change of 10 in the long-run vintage-adjusted citation percentile spread (columns 1 and 5), disruption scaled so a change of 1 in the regressor reflects a change of 10 in the maximum disruption index percentile (columns 2 and 6), pivoting (columns 3 and 7), and standing out (columns 4 and 8). All estimates include an indicator variable for above median novelty as well as an interaction term between this indicator variable and the risk-taking measure. Robust standard errors were used, clustered at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

B. Novelty above the 75%ile

	Base controls				Grant output controls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Novelty >75%ile	0.0204**	0.0117	0.1016**	0.0980**	0.0233**	0.0095	0.1203**	0.0905**
	(0.0064)	(0.0088)	(0.0185)	(0.0306)	(0.0063)	(0.0087)	(0.0182)	(0.0300)
Extreme tail outcomes	-0.0126**				-0.0069**			
	(0.0006)				(0.0007)			
Extreme tail outcomes × Novelty >75%ile	-0.0004				-0.0007			
	(0.0012)				(0.0012)			
Disruption		-0.0153**				-0.0084**		
		(0.0007)				(0.0007)		
Disruption × Novelty >75%ile		0.0010				0.0014		
		(0.0013)				(0.0013)		
Pivoting			-0.1807**				-0.2305**	
			(0.0106)				(0.0106)	
Pivoting × Novelty >75%ile			-0.0882**				-0.1107**	
			(0.0233)				(0.0229)	
Standing out				-0.1709**				-0.1326**
				(0.0187)				(0.0184)
Standing out × Novelty >75%ile				-0.0876*				-0.0812*
				(0.0366)				(0.0359)
Mean of dependent variable	0.464	0.455	0.455	0.419	0.464	0.455	0.455	0.419
Std. Dev. of dependent variable	0.499	0.498	0.498	0.493	0.499	0.498	0.498	0.493
Adjusted R^2	0.5044	0.4868	0.4862	0.4385	0.5182	0.5028	0.5061	0.4594
Nb. of investigators	33,348	36,799	36,538	31,867	33,348	36,799	36,538	31,867
Nb. of investigator-grant-cycles	89,960	102,035	101,336	73,218	89,960	102,035	101,336	73,218

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Columns with base controls are very similar to Table 2, column 2, while those with grant output controls are very similar to Table 2, column 3. Each estimate includes a risk-taking measure as a continuous variable: extreme tail outcomes scaled so a change of 1 in the regressor reflects a change of 10 in the long-run vintage-adjusted citation percentile spread (columns 1 and 5), disruption scaled so a change of 1 in the regressor reflects a change of 10 in the maximum disruption index percentile (columns 2 and 6), pivoting (columns 3 and 7), and standing out (columns 4 and 8). All estimates include an indicator variable for above the 75%ile in novelty as well as an interaction term between this indicator variable and the risk-taking measure. Robust standard errors were used, clustered at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

C. Novelty above the 90%ile

	Base controls				Grant output controls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Novelty >90%ile	0.0206*	0.0167	0.1327**	0.1705**	0.0246**	0.0131	0.1359**	0.1368**
	(0.0090)	(0.0120)	(0.0284)	(0.0447)	(0.0089)	(0.0118)	(0.0279)	(0.0436)
Extreme tail outcomes	-0.0128**				-0.0072**			
	(0.0006)				(0.0007)			
Extreme tail outcomes × Novelty >90%ile	0.0018				0.0009			
	(0.0018)				(0.0017)			
Disruption		-0.0151**				-0.0083**		
		(0.0006)				(0.0007)		
Disruption × Novelty >90%ile		0.0024				0.0025		
		(0.0019)				(0.0018)		
Pivoting			-0.1855**				-0.2349**	
			(0.0101)				(0.0100)	
Pivoting × Novelty >90%ile			-0.1017**				-0.1136**	
			(0.0347)				(0.0342)	
Standing out				-0.1722**				-0.1339**
				(0.0178)				(0.0176)
Standing out × Novelty >90%ile				-0.1547**				-0.1259*
				(0.0527)				(0.0513)
Mean of dependent variable	0.464	0.455	0.455	0.419	0.464	0.455	0.455	0.419
Std. Dev. of dependent variable	0.499	0.498	0.498	0.493	0.499	0.498	0.498	0.493
Adjusted R^2	0.5044	0.4869	0.4863	0.4386	0.5182	0.5028	0.5059	0.4593
Nb. of investigators	33,348	36,799	36,538	31,867	33,348	36,799	36,538	31,867
Nb. of investigator-grant-cycles	89,960	102,035	101,336	73,218	89,960	102,035	101,336	73,218

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Columns with base controls are very similar to Table 2, column 2, while those with grant output controls are very similar to Table 2, column 3. Each estimate includes a risk-taking measure as a continuous variable: extreme tail outcomes scaled so a change of 1 in the regressor reflects a change of 10 in the long-run vintage-adjusted citation percentile spread (columns 1 and 5), disruption scaled so a change of 1 in the regressor reflects a change of 10 in the maximum disruption index percentile (columns 2 and 6), pivoting (columns 3 and 7), and standing out (columns 4 and 8). All estimates include an indicator variable for above the 90%ile in novelty as well as an interaction term between this indicator variable and the risk-taking measure. Robust standard errors were used, clustered at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

D. Novelty above the 95%ile

	Base controls				Grant output controls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Novelty >95%ile	0.0323**	0.0333*	0.1700**	0.2559**	0.0371**	0.0298*	0.1599**	0.2028**
	(0.0119)	(0.0152)	(0.0390)	(0.0607)	(0.0117)	(0.0148)	(0.0381)	(0.0591)
Extreme tail outcomes	-0.0127**				-0.0072**			
	(0.0006)				(0.0007)			
Extreme tail outcomes × Novelty >95%ile	0.0028				0.0014			
	(0.0025)				(0.0025)			
Disruption		-0.0149**				-0.0081**		
		(0.0006)				(0.0007)		
Disruption × Novelty >95%ile		0.0022				0.0019		
		(0.0024)				(0.0024)		
Pivoting			-0.1841**				-0.2340**	
			(0.0099)				(0.0099)	
Pivoting × Novelty >95%ile			-0.1221**				-0.1241**	
			(0.0471)				(0.0461)	
Standing out				-0.1727**				-0.1346**
				(0.0176)				(0.0174)
Standing out × Novelty >95%ile				-0.2286**				-0.1837**
				(0.0712)				(0.0693)
Mean of dependent variable	0.464	0.455	0.455	0.419	0.464	0.455	0.455	0.419
Std. Dev. of dependent variable	0.499	0.498	0.498	0.493	0.499	0.498	0.498	0.493
Adjusted R^2	0.5045	0.4869	0.4864	0.4388	0.5182	0.5029	0.5059	0.4594
Nb. of investigators	33,348	36,799	36,538	31,867	33,348	36,799	36,538	31,867
Nb. of investigator-grant-cycles	89,960	102,035	101,336	73,218	89,960	102,035	101,336	73,218

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Columns with base controls are very similar to Table 2, column 2, while those with grant output controls are very similar to Table 2, column 3. Each estimate includes a risk-taking measure as a continuous variable: extreme tail outcomes scaled so a change of 1 in the regressor reflects a change of 10 in the long-run vintage-adjusted citation percentile spread (columns 1 and 5), disruption scaled so a change of 1 in the regressor reflects a change of 10 in the maximum disruption index percentile (columns 2 and 6), pivoting (columns 3 and 7), and standing out (columns 4 and 8). All estimates include an indicator variable for above the 95%ile in novelty as well as an interaction term between this indicator variable and the risk-taking measure. Robust standard errors were used, clustered at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

Table A2. Effect of career stage on the risk taking grant renewal penalty

A. New Investigator

	Base controls				Grant output controls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New investigator	0.0568**	0.0651**	0.0850**	0.1418**	0.0737**	0.0830**	0.0832**	0.1563**
	(0.0075)	(0.0095)	(0.0166)	(0.0259)	(0.0074)	(0.0092)	(0.0162)	(0.0253)
Extreme tail outcomes	-0.0119**				-0.0062**			
	(0.0006)				(0.0007)			
Extreme tail outcomes × New investigator	-0.0038**				-0.0038**			
	(0.0012)				(0.0012)			
Disruption		-0.0144**				-0.0074**		
		(0.0007)				(0.0007)		
Disruption × New investigator		-0.0022†				-0.0027*		
		(0.0013)				(0.0013)		
Pivoting			-0.1658**				-0.2249**	
			(0.0108)				(0.0108)	
Pivoting × New investigator			-0.0418*				-0.0202	
			(0.0209)				(0.0204)	
Standing out				-0.1393**				-0.1005**
				(0.0192)				(0.0190)
Standing out × New investigator				-0.1110**				-0.1114**
				(0.0315)				(0.0308)
Mean of dependent variable	0.463	0.454	0.455	0.419	0.463	0.454	0.455	0.419
Std. Dev. of dependent variable	0.499	0.498	0.498	0.493	0.499	0.498	0.498	0.493
Adjusted R^2	0.5046	0.4871	0.4861	0.4388	0.5187	0.5035	0.5064	0.4602
Nb. of investigators	33,403	36,876	36,538	31,867	33,403	36,876	36,538	31,867
Nb. of investigator-grant-cycles	90,075	102,250	101,336	73,218	90,075	102,250	101,336	73,218

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Columns with base controls are very similar to Table 2, column 2, while those with grant output controls are very similar to Table 2, column 3. Each estimate includes a risk-taking measure as a continuous variable: extreme tail outcomes scaled so a change of 1 in the regressor reflects a change of 10 in the long-run vintage-adjusted citation percentile spread (columns 1 and 5), disruption scaled so a change of 1 in the regressor reflects a change of 10 in the maximum disruption index percentile (columns 2 and 6), pivoting (columns 3 and 7), and standing out (columns 4 and 8). All estimates include an indicator variable for being a new investigator as well as an interaction term between this indicator variable and the risk-taking measure. Robust standard errors were used, clustered at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

B. Early stage investigator

	Base controls				Grant output controls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Early stage investigator	0.0472**	0.0515**	0.0328	0.1151**	0.0507**	0.0584**	0.0182	0.1268**
	(0.0119)	(0.0139)	(0.0243)	(0.0354)	(0.0116)	(0.0135)	(0.0235)	(0.0344)
Extreme tail outcomes	-0.0120**				-0.0064**			
	(0.0006)				(0.0007)			
Extreme tail outcomes × Early stage investigator	-0.0079**				-0.0072**			
	(0.0018)				(0.0018)			
Disruption		-0.0144**				-0.0075**		
		(0.0006)				(0.0007)		
Disruption × Early stage investigator		-0.0054**				-0.0053**		
		(0.0018)				(0.0018)		
Pivoting			-0.1705**				-0.2277**	
			(0.0101)				(0.0101)	
Pivoting × Early stage investigator			-0.0220				0.0101	
			(0.0292)				(0.0282)	
Standing out				-0.1514**				-0.1124**
				(0.0181)				(0.0179)
Standing out × Early stage investigator				-0.1321**				-0.1333**
				(0.0428)				(0.0417)
Mean of dependent variable	0.463	0.454	0.455	0.419	0.463	0.454	0.455	0.419
Std. Dev. of dependent variable	0.499	0.498	0.498	0.493	0.499	0.498	0.498	0.493
Adjusted R^2	0.5043	0.4865	0.4854	0.4381	0.5180	0.5024	0.5053	0.4591
Nb. of investigators	33,403	36,876	36,538	31,867	33,403	36,876	36,538	31,867
Nb. of investigator-grant-cycles	90,075	102,250	101,336	73,218	90,075	102,250	101,336	73,218

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Columns with base controls are very similar to Table 2, column 2, while those with grant output controls are very similar to Table 2, column 3. Each estimate includes a risk-taking measure as a continuous variable: extreme tail outcomes scaled so a change of 1 in the regressor reflects a change of 10 in the long-run vintage-adjusted citation percentile spread (columns 1 and 5), disruption scaled so a change of 1 in the regressor reflects a change of 10 in the maximum disruption index percentile (columns 2 and 6), pivoting (columns 3 and 7), and standing out (columns 4 and 8). All estimates include an indicator variable for being an early stage investigator as well as an interaction term between this indicator variable and the risk-taking measure. Robust standard errors were used, clustered at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

C. National Academy of Sciences membership

	Base controls				Grant output controls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NAS member	-0.0273	-0.0367	0.0487	-0.2406**	-0.0072	-0.0251	0.0657	-0.2415**
	(0.0177)	(0.0239)	(0.0409)	(0.0816)	(0.0177)	(0.0240)	(0.0412)	(0.0817)
Extreme tail outcomes	-0.0130**				-0.0073**			
	(0.0006)				(0.0007)			
Extreme tail outcomes × NAS member	0.0050				0.0042			
	(0.0032)				(0.0032)			
Disruption		-0.0152**				-0.0083**		
		(0.0006)				(0.0007)		
Disruption × NAS member		0.0059†				0.0064†		
		(0.0034)				(0.0034)		
Pivoting			-0.1713**				-0.2246**	
			(0.0098)				(0.0098)	
Pivoting × NAS member			-0.0594				-0.0686	
			(0.0572)				(0.0574)	
Standing out				-0.1780**				-0.1395**
				(0.0174)				(0.0172)
Standing out × NAS member				0.2926**				0.3131**
				(0.0976)				(0.0978)
Mean of dependent variable	0.463	0.454	0.455	0.419	0.463	0.454	0.455	0.419
Std. Dev. of dependent variable	0.499	0.498	0.498	0.493	0.499	0.498	0.498	0.493
Adjusted R^2	0.5042	0.4864	0.4854	0.4380	0.5179	0.5024	0.5052	0.4591
Nb. of investigators	33,403	36,876	36,538	31,867	33,403	36,876	36,538	31,867
Nb. of investigator-grant-cycles	90,075	102,250	101,336	73,218	90,075	102,250	101,336	73,218

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Columns with base controls are very similar to Table 2, column 2, while those with grant output controls are very similar to Table 2, column 3. Each estimate includes a risk-taking measure as a continuous variable: extreme tail outcomes scaled so a change of 1 in the regressor reflects a change of 10 in the long-run vintage-adjusted citation percentile spread (columns 1 and 5), disruption scaled so a change of 1 in the regressor reflects a change of 10 in the maximum disruption index percentile (columns 2 and 6), pivoting (columns 3 and 7), and standing out (columns 4 and 8). All estimates include an indicator variable for being a member of the National Academy of Sciences as well as an interaction term between this indicator variable and the risk-taking measure. Robust standard errors were used, clustered at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$. NAS—National Academy of Sciences.

Table A3. Descriptive statistics: Risk taking and scientific impact

	Renewed grant cycles					Not renewed grant cycles				
	Mean	Median	Std. Dev.	Min.	Max.	Mean	Median	Std. Dev.	Min.	Max.
FDA-approved drug patent citations	0.15	0	1.54	0	106	0.10	0	1.23	0	60
Subsequent grant cycle publications	9.95	8	9.27	1	184					
Subsequent grant cycle citations	698.16	377	1,143.23	1	67,451					

Note: Outcomes are measured at the investigator-grant-cycle level. Subsequent grant cycle publications and citations are only available for renewed grant cycles. FDA—Food and Drug Administration.

Table A4. Effect of risk taking on renewed grants' outcomes

A. FDA-approved drug patent citations

	Base controls				Grant output controls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Extreme tail outcomes	0.1044**				-0.0310			
	(0.0219)				(0.0283)			
Disruption		0.2317**				0.1147**		
		(0.0344)				(0.0381)		
Pivoting			1.2773**				1.7312**	
			(0.4946)				(0.6284)	
Standing out				0.6809				-0.0528
				(0.7323)				(0.8420)
Pseudo-R^2	0.2682	0.2767	0.2624	0.2390	0.3125	0.3164	0.3171	0.3003
Log pseudolikelihood	-19,402	-19,949	-20,272	-12,656	-18,227	-18,853	-18,767	-11,637
Nb. of investigators	12,288	13,299	13,215	11,764	12,288	13,299	13,215	11,764
Nb. of investigator-grant-cycles	28,389	31,377	31,147	21,655	28,389	31,377	31,147	21,655

Note: Estimates stem from a Poisson model where the unit of analysis is the investigator-grant-cycle, with the sample restricted to those investigator-grant-cycles which were renewed. The dependent variable is patent citations to grant cycle publications from patents associated with FDA approved drugs. Columns with base controls are very similar to Table 2, column 2, while those with grant output controls are very similar to Table 2, column 3. Each estimate includes a risk-taking measure as a continuous variable: extreme tail outcomes scaled so a change of 1 in the regressor reflects a change of 10 in the long-run vintage-adjusted citation percentile spread (columns 1 and 5), disruption scaled so a change of 1 in the regressor reflects a change of 10 in the maximum disruption index percentile (columns 2 and 6), pivoting (columns 3 and 7), and standing out (columns 4 and 8). Robust standard errors were used, clustered at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$. FDA—Food and Drug Administration.

B. Subsequent grant cycle publications and citations

	Publications				Citations			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Extreme tail outcomes	0.0806**				0.0370**			
	(0.0022)				(0.0042)			
Disruption		0.0827**				0.0891**		
		(0.0026)				(0.0049)		
Pivoting			-0.3738**				0.3035**	
			(0.0377)				(0.0700)	
Standing out				0.1682*				-0.1443
				(0.0723)				(0.1384)
Pseudo-R^2	0.1775	0.1699	0.1508	0.1356	0.3348	0.3428	0.3312	0.2914
Log pseudolikelihood	-146,470	-162,279	-164,875	-108,770	-10,805,529	-11,609,355	-11,746,335	-7,475,852
Nb. of investigators	14,941	16,222	16,110	14,159	14,941	16,222	16,110	14,159
Nb. of investigator-grant-cycles	34,920	38,819	38,523	25,190	34,920	38,819	38,523	25,190

Note: Estimates stem from a Poisson model where the unit of analysis is the investigator-grant-cycle, with the sample restricted to those investigator-grant-cycles which were renewed. The dependent variable is publications (columns 1-4) and citations (columns 5-8) in the next grant cycle after renewal. The controls used are very similar to Table 2, column 2. Each estimate includes a risk-taking measure as a continuous variable: extreme tail outcomes scaled so a change of 1 in the regressor reflects a change of 10 in the long-run vintage-adjusted citation percentile spread (columns 1 and 5), disruption scaled so a change of 1 in the regressor reflects a change of 10 in the maximum disruption index percentile (columns 2 and 6), pivoting (columns 3 and 7), and standing out (columns 4 and 8). Robust standard errors were used, clustered at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

Table A5. Descriptive statistics: Alternative measures of risk taking and novelty

	Renewed grant cycles					Not renewed grant cycles				
	Mean	Median	Std. Dev.	Min.	Max.	Mean	Median	Std. Dev.	Min.	Max.
Range, median-min citation percentile	32.79	31	18.58	0	97	31.98	30	18.29	0	96
Range, max-median citation percentile	17.19	15	11.68	0	84	19.76	18	12.62	0	92
Range, winsorized max-min citation percentile	39.33	38	18.92	1	99	41.65	41	18.66	1	99
Variance, citation percentile	377.65	311	334.37	0	4,579	424.31	354	369.50	0	4,950
Disruption index percentile, mean	39.71	39	15.43	0	100	43.96	44	16.10	0	100
Disruption index without Nk term percentile, max	60.28	63	22.18	0	94	62.55	66	22.57	0	94
Fraction new MeSH vs. investigator′s prior work	0.40	0	0.16	0	1	0.42	0	0.16	0	1
Cosine dissimilarity vs. investigator′s prior work	0.54	1	0.20	0	1	0.58	1	0.20	0	1
Fraction unique MeSH vs. other investigators	0.39	1	0.14	0	1	0.40	0	0.15	0	1
Cosine dissimilarity vs. other investigators	0.87	1	0.08	0	1	0.88	1	0.08	0	1
MeSH pair age, mean	25.58	25	7.48	0	63	28.30	28	8.18	1	67
Fraction new bigrams	0.04	0.04	0.03	0	1	0.03	0.03	0.03	0	0.5
Average new bigrams	1.14	1.00	0.85	0	13	0.92	0.78	0.79	0	13

Note: Outcomes are measured at the investigator-grant-cycle level. Measures are only defined for those cycles with associated publications (see footnote 10 for details).

Table A6. Robustness: Uniform sample

		Full sample		Uniform sample	
		Base controls	Grant output controls	Base controls	Grant output controls
Extreme tail outcomes					
	<10%ile	0.051**	0.022**	0.058**	0.027**
		(0.004)	(0.005)	(0.005)	(0.006)
	>90%ile	-0.044**	-0.021**	-0.040**	-0.017**
		(0.004)	(0.004)	(0.005)	(0.005)
Disruption					
	<10%ile	0.089**	0.054**	0.097**	0.072**
		(0.012)	(0.012)	(0.027)	(0.027)
	>90%ile	-0.022**	0.008†	-0.022**	0.006
		(0.004)	(0.004)	(0.005)	(0.005)
Pivoting					
	<10%ile	0.041**	0.045**	0.042**	0.051**
		(0.004)	(0.004)	(0.006)	(0.006)
	>90%ile	-0.028**	-0.047**	-0.033**	-0.047**
		(0.004)	(0.004)	(0.005)	(0.005)
Standing out					
	<10%ile	0.027**	0.011*	0.018**	0.007
		(0.005)	(0.005)	(0.006)	(0.006)
	>90%ile	-0.013*	-0.019**	-0.018**	-0.020**
		(0.005)	(0.005)	(0.005)	(0.005)

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Each column and independent variable measure reflect the results of a separate regression. Columns with base controls are very similar to Table 2, column 2, while those with grant output controls are very similar to Table 2, column 3. Columns 3 and 4 limit the sample to those investigator-grant-cycles for which all four measures of risk taking are defined (n = 63,899). The independent variables are indictor variables for <10%ile and >90%ile for each measure of risk taking, with 10-90%ile as the omitted category. Robust standard errors were used, clustered at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

Table A7. Heterogeneity by degree of pivoting in proposed grant abstract

	Full sample		Greater overlap between proposal and prior work		Less overlap between proposal and prior work	
	Base controls	Grant output controls	Base controls	Grant output controls	Base controls	Grant output controls
<10%ile, fraction new MeSH pairs	0.041**	0.045**	0.043**	0.045**	0.041**	0.045**
	(0.004)	(0.004)	(0.006)	(0.006)	(0.004)	(0.004)
>90%ile, fraction new MeSH pairs	-0.028**	-0.047**	-0.033**	-0.062**	-0.028**	-0.047**
	(0.004)	(0.004)	(0.010)	(0.009)	(0.004)	(0.004)
Mean of dependent variable	0.455	0.455	0.409	0.409	0.455	0.455
Std. Dev. of dependent variable	0.498	0.498	0.492	0.492	0.498	0.498
Effect bottom %ile group, in s.d. units	0.083	0.091	0.088	0.091	0.083	0.091
Effect top %ile group, in s.d. units	-0.056	-0.094	-0.067	-0.126	-0.056	-0.094
Adjusted R^2	0.4844	0.5035	0.4715	0.4915	0.4844	0.5035
Nb. of investigators	36,538	36,538	21,617	21,617	36,538	36,538
Nb. of investigator-grant-cycles	101,336	101,336	44,228	44,228	101,336	101,336

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Each column reflects the results of a separate regression. Columns with base controls are very similar to Table 2, column 2, while those with grant output controls are very similar to Table 2, column 3. Columns 3 and 4 limit the sample to those investigator-grant-cycles with below median fraction new MeSH terms between the grant proposal abstract and an investigator's publications in the preceding 5 years, while Columns 5 and 6 limit the sample to above the median fraction new MeSH. The independent variables are indictor variables for <10%ile and >90%ile for pivoting, with 10-90%ile as the omitted category. Robust standard errors were used, clustered at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

Table A8. Robustness: Investigator name frequency

		Full sample		Below median name frequency		Above median name frequency	
		Base controls	Grant output controls	Base controls	Grant output controls	Base controls	Grant output controls
Extreme tail outcomes							
	<10%ile	0.051**	0.022**	0.051**	0.019**	0.055**	0.025**
		(0.004)	(0.005)	(0.006)	(0.007)	(0.006)	(0.007)
	>90%ile	-0.044**	-0.021**	-0.044**	-0.018**	-0.045**	-0.022**
		(0.004)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)
Disruption							
	<10%ile	0.089**	0.054**	0.109**	0.072**	0.072**	0.037*
		(0.012)	(0.012)	(0.016)	(0.016)	(0.017)	(0.017)
	>90%ile	-0.022**	0.008†	-0.027**	0.004	-0.019**	0.011†
		(0.004)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)
Pivoting							
	<10%ile	0.041**	0.045**	0.033**	0.038**	0.050**	0.052**
		(0.004)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)
	>90%ile	-0.028**	-0.047**	-0.030**	-0.050**	-0.025**	-0.044**
		(0.004)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)
Standing out							
	<10%ile	0.027**	0.011*	0.026**	0.009	0.028**	0.012†
		(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
	>90%ile	-0.013*	-0.019**	-0.013†	-0.018*	-0.011	-0.017*
		(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Each column and independent variable measure reflect the results of a separate regression. Columns with base controls are very similar to Table 2, column 2, while those with grant output controls are very similar to Table 2, column 3. Columns 3 and 4 limit the sample to those investigator-grant-cycles with below median name frequency, calculated relative to the corpus of PubMed, while Columns 5 and 6 limit the sample to above median name frequency. The independent variables are indictor variables for <10%ile and >90%ile for each measure of risk taking, with 10-90%ile as the omitted category. Robust standard errors were used, clustered at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

Table A9. Robustness: Reliance on NIH funding

	Full sample		Below median fraction pubs. with NIH support		Above median fraction pubs. with NIH support	
	Base controls	Grant output controls	Base controls	Grant output controls	Base controls	Grant output controls
Extreme tail outcomes						
<10%ile	0.051**	0.022**	0.047**	0.021**	0.059**	0.024**
	(0.004)	(0.005)	(0.006)	(0.006)	(0.007)	(0.007)
>90%ile	-0.044**	-0.021**	-0.044**	-0.021**	-0.045**	-0.021**
	(0.004)	(0.004)	(0.006)	(0.006)	(0.005)	(0.006)
Disruption						
<10%ile	0.089**	0.054**	0.083**	0.053**	0.101**	0.050**
	(0.012)	(0.012)	(0.015)	(0.015)	(0.019)	(0.019)
>90%ile	-0.022**	0.008†	-0.019**	0.009	-0.025**	0.003
	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)
Pivoting						
<10%ile	0.041**	0.045**	0.034**	0.038**	0.051**	0.052**
	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)
>90%ile	-0.028**	-0.047**	-0.024**	-0.043**	-0.032**	-0.053**
	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)
Standing out						
<10%ile	0.027**	0.011*	0.030**	0.013†	0.024**	0.008
	(0.005)	(0.005)	(0.007)	(0.007)	(0.008)	(0.008)
>90%ile	-0.013*	-0.019**	-0.009	-0.014*	-0.015†	-0.023**
	(0.005)	(0.005)	(0.007)	(0.007)	(0.008)	(0.008)

Note: Estimates stem from a linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator variable for if the grant cycle was renewed. Each column and independent variable measure reflect the results of a separate regression. Columns with base controls are very similar to Table 2, column 2, while those with grant output controls are very similar to Table 2, column 3. Columns 3 and 4 limit the sample to those investigator-grant-cycles with below median fraction of career publications in PubMed acknowledging any NIH funding, while Columns 5 and 6 limit the sample to those with above median fraction career publications with any NIH funding. The independent variables are indictor variables for <10%ile and >90%ile for each measure of risk taking, with 10-90%ile as the omitted category. Robust standard errors were used, clustered at the investigator and grant. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

Table A10. Robustness: Oster's Delta

	Base controls		Grant output controls	
	Coefficient estimate	Oster's Delta	Coefficient estimate	Oster's Delta
Extreme tail outcomes	-0.0128**	-4.550	-0.0072**	-16.076
	(0.0006)		(0.0007)	
Disruption	-0.0151**	-48.926	-0.0081**	2.675
	(0.0006)		(0.0007)	
Pivoting	-0.1727**	-23.281	-0.2262**	-9.879
	(0.0097)		(0.0097)	
Standing out	-0.1717**	-2.264	-0.1328**	-2.073
	(0.0173)		(0.0171)	

Note: Each cell in columns 1 and 3 stems from a separate linear probability model where the unit of analysis is the investigator-grant-cycle and the dependent variable is an indicator for if the grant cycle was renewed. This is very similar to columns 2 and 3 of Table 2A, 2B, 2C and 2D for the extreme tail outcomes (row 1), disruption (row 2), pivoting (row 3) and standing out (row 4), respectively, except a single continuous regressor is used for each measure of risk taking. Robust standard errors were used, clustered at the investigator and grant. Columns 2 and 4 report the corresponding δ parameter from Oster (2019), which can be interpreted as the degree of selection on unobservables relative to observables necessary to explain away the effect of risk taking on grant renewal. We follow Oster's recommendation of setting $R^{max} = 1.3 \times R^2$. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.