Funding Risky Research

Chiara Franzoni°, Paula Stephan°° and Reinhilde Veugelers°°°

°School of Management, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy. Email: chiara.franzoni@polimi.it

°°Andrew Young School of Policy Studies, Georgia State University, Atlanta, GA 30302, USA & National Bureau of Economic Research. Cambridge, MA 02138, USA NBER. E-mail: pstephan@gsu.edu

°°° Department of Management, Strategy & Innovation, KU Leuven, Bruegel & Peterson Institute for International Economics. Email: reinhilde.veugelers@kuleuven.be

Preliminary Draft

April 12, 2021

Do not quote without authors' permission

The authors wish to thank Katalin Karikó for her readiness to respond to our emails.

1. Introduction

The Covid-19 pandemic has underlined how much society depends on the pace of scientific research and how effective science can be. The speed with which Covid-19 vaccines were developed and their high performance surpassed even the most optimistic expectations. This is especially the case for those based on the new designer mRNA technology which enabled the identification of a vaccine with high efficacy in less than three months after sequencing of the virus and holds huge promise for the future of vaccines and medicine more broadly.[1]

We draw on this exceptional moment for science to reflect on an important and pressing theme. Is the government funding system sufficiently supportive of the science needed for key breakthroughs, such as mRNA -based drugs? If the science needed for such breakthroughs requires risk-taking, particularly in its early phases, does our system of science funding encourage sufficient risk taking to induce scientists to explore transformative research paths?

In this contribution, we discuss risk-taking and the funding of risky science. We start in Section 2 by describing problems faced by Katalin Karikó, a scientist who did pioneering-research related to mRNA-based drugs. In section 3 we briefly describe measures developed to distinguish risky from non-risky research and the extent to which the citation footprint of risky research differs from that of non-risky research. We then review empirical work concerning the funding of risky research which suggests that funding is biased against risky research. Section 4 provides a framework for thinking about why funding agencies may eschew funding risky research. We focus first on how factors within the research system, such as pressure to show results in a short time period

[1] https://www.nature.com/articles/d41586-021-00019-w. Accessed March 4, 2021.

and the widespread use of bibliometrics, contribute to risk aversion. We then focus on three key players affecting research funding and the role the three play in determining the amount of risky research that is undertaken: (1) funding agencies, (2) panelists and (3) principal investigators. Section 5 closes with a discussion of interventions that government agencies and universities could do if they wish to increase the amount of risk that they fund.

2. mRNA-design: difficulties encountered advancing a risky agenda

We begin our journey by describing the development of designer messenger RNA (mRNA), the breakthrough technology used by Pfizer-BioNTech and Moderna to develop the first two vaccines against Covid 19 to obtain FDA approval in the US and EMA approval in the EU.

The mRNA is a protein-coding-single-stranded molecule, produced by cells during the transcription process, when the genes encoded in DNA are copied into the molecule of RNA. The discovery of mRNA was reported in Nature in May of 1961, a result of scientists' search concerning the synthesis of proteins coded in DNA. The path to synthesize mRNA in a test tube was made possible in 1984 when Paul Krieg and Doug Melton, scientists at Harvard, identified that by using SP6 RNA polymerase functional mRNA can be produced in vitro. By 1990, a group of scientists demonstrated that injection of synthetic, in vitro transcribed mRNA into animals led to expression of the encoded protein (Wolff et al. 1990). Soon, the scientific world realized that this system could potentially be used to turn human bodies into medicine-making factories and treat a variety of diseases, ranging from infection, to cancer to rare diseases and possibly mend such things as damaged heart tissue (Sahin, Karikó, and Türeci 2014). But, at the time, mRNA was not the only conceivable way to introduce protein expression into cells: other nucleic acid-based technologies were under investigation. Moreover, there remained two critical problems that needed to be addressed. In vitro-transcribed mRNA, when delivered to animals, could either be destroyed by the body as the body fielded an immune response before reaching its target, or worse yet, cause serious side effects (Sahin et al. 2014). No one knew how to make mRNA effective in humans despite years of interest on the part of scientists.

Katalin Karikó was determined, by all accounts, on finding a way to make synthetic mRNA non-inflammatory applicable to treat human diseases. Born and educated in Hungary, she came to the US in 1985, first as a postdoctoral fellow to Temple University, then to USUHS. In 1989, she moved to a faculty position at the Medical School of the University of Pennsylvania.[2] She submitted more than 20 grants, initially for smaller sums, to the University of Pennsylvania and the American Heart Association, then for larger sums to NIH. [3] As hard as she tried, she repeatedly failed to get funding for her research. "Every night I was working: grant, grant, grant," recounts Karikó. "And it came back always no, no, no."[4] Her inability to support her research on grants eventually resulted in her being taken off her faculty position by the university. In 1995 she accepted a non-faculty position, that she describes as "more like a post-doc position" at the University of Pennsylvania without any prospect of advancing.[5]

Two years later, Drew Weissman, an MD PhD immunologist, moved from NIH (where he had worked with Anthony Fauci) to the University of Pennsylvania. The same year, Karikó and Weissman met at the school's photo copying machine. While chatting informally, they recognized that they shared an interest in developing a synthetic mRNA vaccine against HIV. They realized the potential of combining their biochemistry, molecular biology and immunology expertise and decided to begin working together. At the time, Karikó was focused on mRNA-based therapy for treating cerebral diseases and strokes. With Weissman, Karikó switched focus to mRNA-based vaccines. Weissman supported the early-stage work partly on one of his existing NIH grants, which had no

---

[2] For a summary of Karikó's early career see https://www.wired.co.uk/article/mrna-coronavirus-vaccine-pfizer-biontech
[3] Emails from Kariko to coauthors, March 17, 2021 and March 18, 2021.
[4] https://www.statnews.com/2020/11/10/the-story-of-mrna-how-a-once-dismissed-idea-became-a-leading-technology-in-the-covid-vaccine-race/. Accessed March 4, 2021
[5] Correspondence with Kariko and https://www.ae-info.org/ae/Member/Karik%C3%B3_Katalin for Karikó's CV.

direct connection to mRNA research.[6] Their breakthrough occurred when they recognized that uridine was the nucleoside in the mRNA that provoked the human immune system. They discovered that, when replacing uridine with pseudouridine, another naturally occurring nucleoside in the mRNA, it could enter into cells without alerting the RNA sensors. Their research was eventually published in *Immunity* in 2005, after being rejected by several leading journals.[7] Karikó was the first author, Weissman the senior author. It eventually became a highly cited paper, receiving to date more than 1000 Google-Scholar citations, although it took until 2015 to reach its first 500 citations. As a result of their joint work, the two filed and obtained patents that were assigned to the University of Pennsylvania. These patents, in line with the Bayh-Dole act, acknowledge NIH grants, including those started before the foundational mRNA patent applications were filed. The patents were licensed exclusively to CellScript by the University of Pennsylvania. CellScript sublicensed the University's patent to the German-based firm BioNTech, incorporated in 2008, and US-based Moderna, incorporated in 2010. The subsequent development of the mRNA-based drugs was conducted by these companies with equity investments, but also with a large involvement of public money.[8]

Despite the 2005 discovery, the two found funding for mRNA research difficult to obtain. According to Weissman "We both started writing grants. We didn't get most of them. People were not interested in mRNA. The people who reviewed the grants said mRNA will not be a good therapeutic, so don't bother."[9] Weismann, however, continued to receive funding from NIH, some, but not all of it, for mRNA research; Karikó continued to have difficulty getting funding.[10] A 2007 R01 application that Karikó submitted to NINDS

---

[6] The grant was for the Role of gp-340 in HIV Infection and Transmissions.
[7] Gina Kolata, "Kati Kariko Helped Shield the World from the Coronavirus," New York Times, April 8, 2021. https://www.nytimes.com/2021/04/08/health/coronavirus-mrna-kariko.html.
[8] E.g. in 2013 DARPA awarded Moderna a grant for up to $25M for developing mRNA-based therapeutics https://investors.modernatx.com/news-releases/news-release-details/darpa-awards-moderna-therapeutics-grant-25-million-develop. Accessed March 4, 2021.
[9] Gina Kolata, "Kati Kariko Helped Shield the World from the Coronavirus," New York Times, April 8, 2021. https://www.nytimes.com/2021/04/08/health/coronavirus-mrna-kariko.html.
[10] Drew Weissman appears as the principal investigator on a total of 10 projects funded by the National Institutes of Health (NIH) between 1998 and 2021. https://reporter.nih.gov/search/OHy6bdrppUaRJWqIwDoUfQ/projects.

at NIH, for example, was not discussed at the study section meeting, having been judged by reviewers to be in the lower half of the applications. The proposal focused on the anti-inflammatory effects of neurotropics in ischemia stroke. Two reviewers described the proposed work as "novel," the third described the proposal as suffering from a "relative lack of novelty." Other comments from reviewers included statements such as "Preliminary data should be provided to support that the proposed experiments can be carried out," "insufficient preliminary data" and work related to one of the aims "seems very preliminary and, there is high likelihood, that these experiments, especially in vivo, will not work."[11] A 2012 application, joint with Drew Weismann, with the goal of developing "a new therapeutic approach to treat ischemic brain injury by delivering specific mRNAs" was scored but neither it nor the resubmission received a sufficiently strong score to be funded. Concerns included "preliminary data presented are insufficient to suggest that this approach is worthy of in-depth evaluation in a stroke model" and that the first aim of the study was "largely descriptive." [12]

In 2006 Karikó and Weissman founded the company RNARx, with the intention of using mRNA to treat anemia. Karikó was the CEO of the company from 2006 to 2013. In her role, she applied-for and received one STTR grant from NIH.[13] In 2013, Karikó became the senior vice president at BioNTech.

We have no way of knowing what would have played out in terms of research outcomes if Karikó's early applications for funding had not been turned down or if she had gotten research support from the University of Pennsylvanian in the early period. Perhaps mRNA-based vaccines would have been available for Swine Flu in 2009. But without the casual meeting with Weissman at the photo copy machine, she could also have given-up

---

[11]Reviews provided to the authors by Karikó March 18, 2021/
[12] The application was for the continuation of an R01 that Karikó "inherited" from Frank Welsh when he retired. Concerns included "preliminary data presented are insufficient to suggest that this approach is worthy of in-depth evaluation in a stroke model" and that the first aim of the study was "largely descriptive."
[13] Katalin Karikó was the principal investigator of an STTR award project funded by the NIH between 2007 and 2011: https://www.sbir.gov/sbirsearch/detail/294077; https://grantome.com/grant/NIH/R42-HL087688-02;
.

researching a way to make designer mRNA technology effective for drug development in humans. What we do know is that her early proposals were not funded and that the University of Pennsylvania moved her out of her soft-money faculty position. This could reflect a failure to address the problem of the immune system response, which was facilitated later on by her collaboration with Weissman. It could also reflect risk aversion on the part of review panels, that considered the area too risky to be "fundable" at the time, especially since Karikò had, at that time, few publications and citations, few preliminary results and no prior record of funding. More generally, the example tells us that the early-funding of designer mRNA research, now considered a promise of future medicine, was difficult.

Karikó is not the only scientist to hear "no, no, no." Similar anecdotal evidence is not difficult to find. A researcher at a top research institution in the US, in speaking of NASA and NSF, said their "programs are not very adventurous." And "what I experienced was that I couldn't get any new idea or anything I was really excited about funded by NSF. It never worked…the feedback is 'well this is too new: we don't know whether it's going to work." (Franzoni and Stephan 2021). James Rothman, the day after he shared the Nobel Prize in Medicine or Physiology in 2013 told an interviewer that "he was grateful he started work in the early 1970s when the federal government was willing to take much bigger risks in handing out funding to young scientists." Rothman went on to say "I had five years of failure, really, before I had the first initial sign of success. And I'd like to think that that kind of support existed today, but I think there's less of it. And it's actually becoming a pressing national issue, if not an international issue." (Harris 2013).

3. Risk aversion in science funding: A review of empirical evidence

Concerns that the selection of grant proposals is overly conservative has been growing in recent years. Commentators on science policy have long lamented that science funders are too conservative and risk averse and skimp on supporting breakthrough research (e.g. Laudel 2017; Mazzacato 2015; Viner, Powell, and Green 2004). Funding agencies are accused of placing too much emphasis on the downside of avoiding failure and too little emphasis on the upside potential of supporting truly courageous ideas (Azoulay, Graff Zivin, and Manso 2012; Nicholson and Ioannidis 2012). But do we have more than anecdotal evidence on risk bias by science funding agencies? Do we have any science of science funding insights on this?

Although research in the area is limited, a handful of recent empirical works have begun to address the topic. The results support the view of risk aversion in funding. Before we review this evidence, it is important to note that risk remains an ill-defined concept (Althaus 2005; Aven 2011; Hansson 2018). Moreover, it is difficult to measure. Here we follow Franzoni and Stephan (2021) and use the term *risk* in its *speculative* meaning, in the sense that risk refers to uncertainty concerning the outcomes of research, where the outcomes vary predominantly in the spectrum of gains and potentially lead to exceptional results, but also to no results.[14] We preface the literature review with ways to measuring risky research.

Measures of risky research

Empirically identifying risky research is challenging. Most researchers who study risk depend on partial measures that look at the degree to which research results deviate from past results and/or the building blocks upon which the research is based. Foster and colleagues (2015) adopt the first approach and distinguish between three types of papers based on the chemical relationships described in the work. Research that makes a *jump* explores previously unexplored chemical relationships -jumping beyond current knowledge. Such research arguably is more likely to fail but, if the research succeeds, is more likely to make a breakthrough. Research that explores relationships between previously studied entities, is subdivided into research that tests a *new* relationship, not published before, or research that *repeats* an analysis of a previously studied relationship. Foster and colleagues (2015) find that "jump" papers reporting highly innovative chemical combinations receive 52% more citations on average than "repeat" papers reporting known combinations, while "new" papers reporting moderately-innovative combinations

[14] We do not use the term *risk* in the *preventive* meaning, to mean the possibility of a negative event (e.g., a loss or harm), a use that is common in the Risk Analysis literature.

enjoy 30% more citations than those reporting known combinations. Their findings suggest that taking the risk associated with "jump" and "new" research makes it more likely to achieve high impact. But they also find this research is more likely to "fail." The authors thus find that the citation distribution associated with *jump* papers and *new* papers has a higher mean value and a higher variance than that of *repeat* papers, characteristics that we expect in risky research, suggesting their measure correlates with risk. The additional rewards associated with jump papers are, however, relatively small and may not compensate sufficiently for the possibility of failing, suggesting higher expected returns of a safer research path. (Foster et al. 2015: 886).[15]

Wang et al (2017) view scientific research as a combinatorial process and measure novelty in science by examining whether a published paper makes first time ever combinations of scientific knowledge components as proxied by referenced journals, accounting for the difficulty of making such combinations. Almost all new combinations made by novel papers cross subject categories. While recognizing that novelty is but one dimension of "risk," they show that novel papers have patterns consistent with risky research of a higher mean and higher variance in citation performance. They have a higher probability of becoming a highly cited paper, but at the same time also a higher probability to be a no/low cited paper. Wang et al (2017) also find strong evidence that novel research takes more time to be top-cited (Figure 1) and that it is published in journals having a lower impact as measured by the Journal Impact Factor. These findings suggest that bibliometric indicators based on short-term citation counts and Journal Impact Factors, may be biased against risky, novel research. Another finding is that the high impact enjoyed by highly novel papers comes from citations outside the field in which the article is published (Figure 2). They show that citations to novel papers are more likely to come from a broader set of disciplines and from disciplines that are more distant from their "home" field, suggestive that novel research has a tendency to both be best appreciated and to spark applications well beyond the disciplinary boundaries.

[15] ". They also find that papers based on *repeat* strategies were six times more likely to be published than those that used *new* or *jump* strategies during the period 1983-2008.

Figure 1 Delayed Recognition: Novel papers take more time to be cited

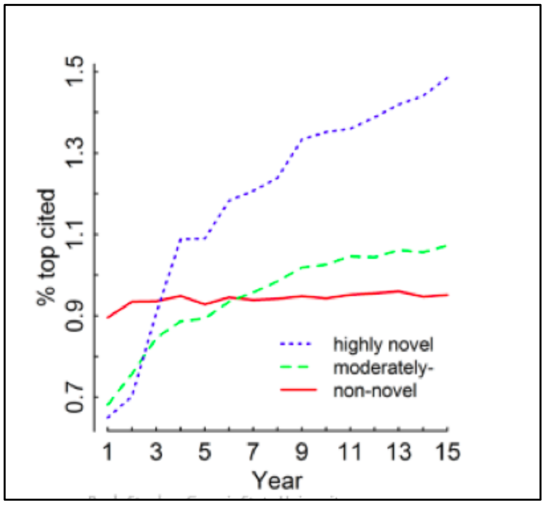
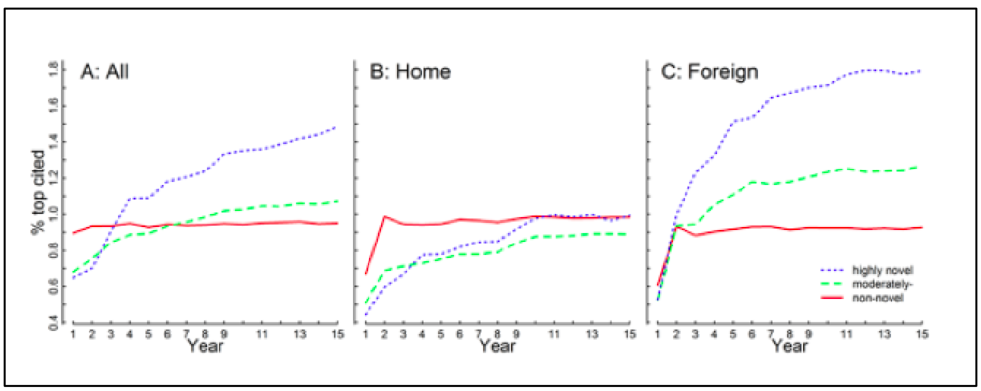


Figure 2 Trans-disciplinary impact over time



Source: Wang et al (2017)

Assessing bias in funding risky research

Using their novelty measure, Veugelers, Stephan and Wang (2021) examine whether the ERC, the most important funding agency of the EU, set up in 2007 with the explicit aim to fund "high gain/high risk" research, is biased against novelty.   They find that applicants to the ERC Starting Grant program with a history of highly novel publications are significantly less likely to receive funding than those without such a history. The major penalty for novelty comes during stage one, when panel members screen a large number of applications based on a short summary of the proposed research and a CV listing the candidate's main publications. The finding is consistent with the use of bibliometric indicators in grant selection as a source of bias against risky research.

In an experiment conducted at the Harvard Medical School, Boudreau and coauthors (2014) find that more novel research proposals, as measured by the percent of keywords not previously used, receive more negative evaluations during peer-review. The result is driven by proposals with particularly high levels of novelty. Their preferred explanation for this finding rests on bounded rationality of reviewers. To quote the authors: "experts extrapolating beyond the knowledge frontier to comprehend novel proposals are prone to systematic errors, misconstruing novel work. This implies that, rather than receiving unbiased assessments (with zero mean errors), novel proposals are discounted relative to their true merit, quality and potential." (Boudreau et al. 2014: 2779). Lanöe (2019), using a measure of novelty, finds evidence that funding decisions made by French National Research Agency are biased against risk-taking. Wagner and Alexander (2013) evaluate the SGER NSF program designed to support high risk, high reward research that ran from 1990 to 2006. Funding decisions were made entirely by program officers with no external review. The authors find that program officers routinely used but a small percent of available funds. The authors interpret the findings as suggesting that either officers were averse to funding risky research, despite the number of funded proposals that had transformative results or, that risk taking was not rewarded within NSF. Packalen and Bhattacharya (2018), using the vintage of ideas embodied in a paper as a proxy for novelty, find that NIH's propensity to fund projects that build on the most recent vintage of advances has declined over the last several decades.

Although the evidence discussed above is preliminary, it suggests that risky research is disfavored in the competition for funding. This seems the case not only when the funding is directed to 'standard' science, but even when a deliberate goal of the funding agency is to support high-risk, high gain, as in the case of the ERC.

Assuming that the preliminary evidence is correct, why do funding agencies eschew supporting risky research and thus possibly miss the opportunities of funding breakthroughs? Is this a conscious or unconscious choice, implanted in their modes of operating? And what can be done to encourage risk taking among funders? These are the questions we explore further in the next sections.

4. Why would funding agencies eschew supporting risky research?

To the best of our knowledge there is no research that directly addresses why the portfolio of funding agencies may be light on risky research. Moreover, there may be multiple factors that concur to play a role. In lack of research that informs on clear causes, we can only formulate a set of hypotheses. In this section we provide an overview of hypotheses that deserve more scrutiny in the future. We start with hypotheses that arise from outside the funding agencies and relate to the broader research system. We then discuss how these translate into a set of incentives and opportunities that could induce funding agencies to eschew supporting risky research and instead fund "safe" research at different levels of analysis. We consider three such levels: i) The funding agencies, ii) the panels in charge of selecting research (composed by panelists and research officers), and iii) the principle investigators who provide the supply of research by submitting grant proposals. Figure 3 provides a summary of the hypotheses.

Bias against risky research in the broader research system

Calls for more accountability when using public funds and the trend towards more and regular evaluations of policy programs put increasing pressure on publicly funded science institutions to show results, especially to show results which are aligned with political cycles. Shorter windows for results already bias against basic research programs in general; witness the heated discussions on the share of overall public R&D budgets for bottom-up basic research programs like ERC, NSF and NIH, compared to more applied and closer to market programs. But even within basic research programs, the pressure to show results quickly may discourage publicly funded agencies from funding more risky research. A major factor is the length of time it takes risky research to have an impact: novel breakthroughs typically take a long time to materialize. As shown in Figure 1, Wang et al (2017) find that the big impact for novel research requires a longer time window than for non-novel research. For the first three years after publication, the probability that a highly novel paper is among the top 1% of cited papers is below that of non-novel papers; but beyond these three years, highly novel papers start accumulating a lead and 15 years after publication, novel papers are nearly twice as likely to be in in the top 1% of highly cited papers. This longer time window for big impact means that it takes the community a longer time to learn about new approaches and switch from its established research paths to adopt approaches. Funding agencies, in an effort to maintain or expand funding, may feel that they cannot afford to wait for risky research to show its impact and opt instead to fund safer research that has measurable effects in the near term.

An important factor for discouraging risky research may be the general lack of tolerance for failure and the meager rewards for those that take uncertain paths within the science system. As hiring, promotion and funding decisions are important conditions to engage in research and as reputation is a key reward to doing science, a lower inclination of researchers to engage in risky research can be traced back to biases against risk in the science system in general.

Universities routinely make crucial career decisions, such as hiring, mid-year review, tenure and promotion. When these career evaluations are made using bibliometric indicators with relatively short-time windows (like journal impact factors or short windows for calculating citations), to measure research "quality", they can discourage risky research, as these measures appear to be biased against risk taking (Stephan, Veugelers, and Wang 2017).

Career status and progression is not only an important reward for scientists themselves, it is also an element that goes into the track record and reputation that funding agencies and their panels consider as part of the applicant's profile. Any bias against risk in career decisions may thus have indirect effects on funding decisions, which may in turn affect career progression negatively when a candidate's productivity in acquiring external funding is a crucial factor in determining career progression, as the story of Karikò illustrated.

The large number of researchers in the US on "soft money" positions, i.e. on positions where salary is funded from grants that the researcher is responsible for obtaining, also encourages the submission of proposals with little risk. If their research is not deemed fundable or comes up empty handed, the university can cut its losses and hire another individual into the position. It is notable that soft money positions have been on the rise in recent years. In the US, for example, the majority of basic medical faculty are hired in soft money positions and are responsible for bringing in most of their own salary (Stephan 2012). Soft-money positions also are common outside of medical institutions. Stephan documents that during the years when the NIH budget doubled, the majority of new hires were made into soft money positions (Stephan 2007). Soft money positions not only transfer risk to the faculty; they also discourage risk taking on the part of the faculty given the importance of continued funding.

Beyond affecting career progression and funding decisions, track record matters more generally as it affects a scientist's reputation and recognition. Although by no means the only reward to doing science, peer recognition and reputation are key drivers of scientists' choices (Merton 1957; Stephan and Levin 1992).[16] With peer recognition and reputation biased against risky research, scientists may be less prone to choose risky research paths. This is however not obvious; examples of scientists who have taken a risky course receiving a Nobel Prize, for example, are readily available [17] and there is research suggesting that prestigious prizes can encourage risk taking (Rzhetsky et al. 2015)

4.1 Principal investigators

The "lack of risky proposals" hypothesis

When Story Landis was Director of NINDS at NIH, she noticed that the amount of support the institute provided for what it classified as "basic-basic" research was declining, compared to what it was spending on "basic-disease" and applied research. Finding this

---

[16] Rewards also include the satisfaction derived from puzzle solving, and financial gain that often accompanies a successful research career (Stephan 2012; Stephan and Levin 1992). Cohen, Sauermann and Stephan (2020) also show that scientists are strongly motivated by an interest in contributing to society.

[17] Jim Allison, who shared the Nobel Prize for immunotherapy for cancer in 2018 is but one case in point.

disconcerting, the Institute set out to investigate why. Somewhat to their surprise, they found that the amount of dollars researchers requested to do "basic-basic" research had declined by 21%. Landis, when asked why she thought the decline was occurring replied: "My concern is that the decrease in the number of basic/basic applications reflects the perception that NINDS is only interested in disease-oriented research."[18] Basic research is not, of course, the same thing as risky research but the two are arguably close cousins. The example may suggest that researchers, anticipating that risky proposals have a difficult time at review, may simply refrain from conceiving and submitting risky research proposals. This is difficult to test, given a lack of data on proposals. But it is a plausible hypothesis. Some evidence consistent with a lack in supply of risky proposals is reported in Veugelers, Stephan & Wang (2021). They find that non-funded junior ERC applicants who fail in the second stage have significantly lower likelihood of producing novel papers after being rejected, compared to the successful ones. The evidence is consistent with rejected applicants learning that risk is not rewarded. Faced with the pressures to (re-) apply for funding, they adjust their research portfolio accordingly away from risky research, something which the successful applicants are "freed" from doing.

Overall, the reward-premium awarded by the science system for doing risky-research, compared to that of doing not-so-risky research, appears insufficient to encourage risk taking. The findings from Foster and colleagues (2015) reported supra, suggest that taking the risk associated with "jump" and "new" research makes it more likely to achieve high impact. But the additional rewards in terms of extra citations they find are relatively small and may not compensate sufficiently for the possibility of failing in terms of not getting published and its negative impact on the researcher's careers. Their results thus suggest that returns may be higher for following a safer research path. Stephan (2019) has called this the *"quad" effect*, referring to the fact that competitive female figure skaters attempt fewer quadruple jumps, arguably because the incremental score they can earn for completing a quad, compared to successfully completing a triple jump, is insufficient to compensate for the risk of failing to complete the quad jump. For male figure

[18] https://blog.ninds.nih.gov/2014/03/27/back-to-basics/. Accessed March 30, 2021.

skaters, scoring is different: the incremental score is larger and provides sufficient incentive to attempt the quad. The work of Uzzi et al. (2013) is consistent with the findings of Foster et al. (2015), and shows that "[t]he highest-impact science is primarily grounded in exceptional conventional combinations of prior work yet simultaneously features an intrusion of unusual combinations", suggesting that a risky approach when embedded in a more standard conventional approach can escape better the citation premium bias. Stated differently, a little bit of risk adds spice to the research; but conventionality is the dominant characteristic of highly cited papers.[19]

The "Loss aversion by principal investigators" hypothesis

The preferences of scientists for the level of risk involved in the projects they wish to pursue may not only reflect biases against risk in the reward structure of science, as discussed supra, but also loss aversion on the part of scientists. Behavioral psychology has shown that humans are generally loss-averse. They over-estimate the magnitude of perspective losses and under-estimate the magnitude of perspective gains (Kahneman and Tversky 1979; Tversky and Kahneman 1991). It may not be implausible to expect that scientists are no exception to the rule.

4.2 Research agencies

Lack of a portfolio approach hypothesis

A common approach taken by investors to stabilize the volatility of outcomes is to include in the same portfolio stocks that have uncorrelated outcomes, or that have outcomes that are negatively correlated, i.e. when one loses, the other gains. In finance, where the investors normally want to maximize the overall return of the portfolio and are risk-averse, a portfolio approach enables purchasing more risky stocks than when the investor buys stocks "one by one" without reference to what is in her portfolio. The "one-by-one" practice is frowned upon in the investment literature, given that the choice of a

[19] Papers characterized as having high medium conventionality coupled with a high tail "novelty" have a hit rate in the top 5 percent 9.2 times out of 100.

stock whose outcomes are highly correlated to those already in the portfolio may expose the investor to extreme gains, but also extreme losses, foregoing any advantages from using the portfolio to diversify away the risk.

The same logic holds to some extent for funding agencies. Agencies generally review proposals one by one, rank them in descending order of overall aggregated score, and then distribute funds according to the score until the budget is exhausted.[20] This "one by one" approach may arguably restrict the level of risk agencies take. To the extent that they are risk averse, the "one by one" approach only aggravates the risk-taking problem.

The "interdisciplinary bias" hypothesis

Review panels are often designed by funding agencies to be discipline-based. This, for example, is generally the case at NSF, and ERC. The latter for instance operates with 25-panels which are mostly discipline focused. It follows that investigators who want to propose research involving multiple disciplines often must make hard choices concerning which is the most appropriate panel to consider their proposal. It also means that their proposal may face obstacles at review that disciplinary-focused proposals do not face. Bromham, Dinnage and Hua (2016) studied more than 18 thousand proposals submitted to the Australian Research Council Discovery Program. They found that the probability of receiving funding decreased as the degree of interdisciplinarity of the proposal increased.[21] Banal-Estanol and colleagues (2019) studied the success rate of teams of co-investigators that sought funding at the UK Engineering and Physical Sciences Research Council. They showed that team-members with interdisciplinary backgrounds (i.e. who had balanced shares of publications in different fields) were penalized, even if those who were eventually funded were more successful ex-post.

---

[20] At some agencies, such as NSF, program officers have some leeway in making decisions, but this is not common.

[21] The study uses the Interdisciplinary Distance (IDD), a measure that takes into account the fields indicated as pertinent to the proposal by the principle investigator and the distance between the fields, based on the relative frequency with which the fields co-occur throughout the entire sample.

A penalty directed at interdisciplinary research may work against funding risky science, because, as noted supra, papers of high novelty are often interdisciplinary. Moreover, Wang et al (2016) also find that novel work that is highly cited is more likely to garner citations from outside, not from within its own field, suggesting that the research is appreciated more by others than by colleagues.

Peer review protocols conceal uncertainty hypothesis

Peer review opinions, especially for risky proposals, involve forecasting research outcomes in conditions of uncertainty (Knight 1921; Nelson 1959). However, protocols commonly used to elicit experts' opinions arguably provide little room for uncertainty, usually requiring reviewers to provide a single score on a numeric ordinal scale to represent a criterion. For example, the ERC requires a single score to rate the "ground-breaking nature and potential impact of the research project". Given the uncertainty of future outcomes, the request of a single-point estimate score can conceptually be thought of as the *median* value of the possible outcome distribution envisaged by the reviewer. Whereas in peer review of 'standard' science, the provision of a single point-estimate may be seen as a necessary time-saving compromise, in evaluations of risky research, the outcomes of interest can be expected to be in the tails and a single-point estimate may have little meaning. Furthermore, uncertainty regarding the outcomes is *the* key piece of information in this case (Morgan 2014; Morgan and Henrion 1990). It seems plausible that similar practices that demand experts to express a score that conceals, rather than represents uncertainty, may induce poor judgments.[22]

Practices that stress reviewers' agreement may disfavor risky science hypothesis

[22] The scholars of expert elicitation have elaborated and tested a number of techniques which are common used in drug-approval, risk analysis, climate-change forecasting, and other areas where uncertainty is key and expert opinions are the only way to collect information (Morgan and Henrion 1990).

It is customary in grant peer review to collect several expert opinions about each proposal before taking a decision. The underling idea is that *aggregation*[23] of a larger number of opinions improves accuracy (Kaplan, Lacetera, and Kaplan 2008; Snell 2015), because random errors likely cancel each-other out when averaging results (Larrick and Soll 2006). This would especially be important for risky proposals, which are more difficult to evaluate, hence more exposed to imprecisions and misjudgments.

The efficacy of this approach relies on two assumptions. First, that a large number of independent reviewers are available. Second, that the mechanisms for aggregating multiple views are unbiased towards risk. In practice, however, both assumptions are problematic. The costs of the review process (Ismail, Farrands, and Wooding 2008) and the unwillingness of reviewers constrain the number of opinions that can be collected (e.g., the NIH advices 4 reviews for each proposal. ERC panels solicit between 3 and 10 external reviews, but only for proposals that are short-listed to go to the second-stage of evaluation). Moreover, reviewers may not be independent and instead have correlated errors (biases), because they share the same background knowledge or beliefs (Clemen and Winkler 1999; Ottaviani and Sorensen 2015).

Another critical point when moving from multiple opinions to a single aggregated opinion and to a final *deliberation* (e.g., binary choice to fund or not)[24] relates to whether methods and rules used in the decision are unbiased towards risk. Prior studies of peer review have evidenced low levels of agreement among reviewers even in the evaluation of 'standard' proposals (Pier et al. 2018). Risky proposals probably spark even greater disagreement, given the larger uncertainty involved. Current practices at NIH and ERC use *behavioral* aggregation, i.e. *consensus meetings*, during which multiple views are confronted and disagreement resolved with discussion (Lamont 2009). However, behavioral

[23] The term *aggregation* (Bolger and Rowe 2015; List 2012) means the combination of multiple opinions. Aggregation can be computed with rules or algorithms (e.g., average, quantile average, ..) or can be done behaviorally, with a *consensus meeting* (Hora et al. 2013; Martini and Sprenger 2018).

[24] *Deliberation*, i.e. the binary choice to fund or not, can be directly dependent on the aggregation method or involve additional rules (e.g., aggregation with arithmetic average and deliberation in descending order of aggregated score until budget saturation.)

aggregation is exposed to groupthink (Cooke 1991; Lamont 2009) and may lead people to herd away from the truth, following influential opinions (Banerjee 1992; Mengel 2019). Consistent with the findings from Della Vigna and Pope (2020) that academics overestimate the accuracy of beliefs of highly-cited scholars, this may lead to a herding on their beliefs. Furthermore, the requirement of consensus may arguably induce a bias against risky research. Assuming that risky proposals lead to outcomes in the 'tails' of the distribution, i.e., either "hits" of "flops" (Azoulay, Graff Zivin, and Manso 2011), it is plausible that the related opinions would also be polarized. If this is the case, methods of aggregation and deliberation that do imply consensus may be systematically biased against risk-taking (Linton 2016). Alternative methods that do not imply consensus exist or are conceivable, such as gold cards, or lotteries among those above a given threshold (Fang and Casadevall 2016; Gross and Bergstrom 2019; Roumbanis 2019), but their limited use has not to date enabled analyses.

4.3 Panelists and research officers

The "insurance agent" hypothesis

Many agencies and panels are acutely aware that the future of their program depends upon supporting researchers who do not come up "empty-handed." They may look at the opportunity cost of funding more risky research and compare it with benefits from funding safer research. These concerns may be magnified by the size of the grant. It is one thing to place $200,000 on a project that may come up empty handed. It is entirely another to place $2M.

Such concerns can lead panels to place considerable emphasis on "what can go wrong" rather than "what can go right" during the review process. One of the "what can go wrong" concerns is that the proposed research cannot be accomplished. This concern undoubtedly fuels the heavy emphasis at many funding agencies on strong "preliminary findings" or, at some agencies, contingency plans, as part of the proposal. In so doing, the implicit requirement is that research be de-risked before it is funded. In this way, the panel supports research with little down-side risk, funding sure bets rather than research that is not a sure-bet but may have considerable up-side potential.

The "bibliometric screening and workload" hypothesis

Scientists and agencies collectively invest a huge amount of time in peer review. For example, the NIH evaluates approximately 80,000 applications annually, engaging over 2,000 reviewers per years and has more than 150 standing Study Sections.[25] The ERC averages 15 members on each of its 25 separate panels. The average panel member for the Starting Grants looks at 137 proposals per call; for Advanced Grants, 83 proposals.[26]

Given the heavy workload, it is not surprising that reviewers and panel members may seek ways to rapidly screen proposals, especially on a first pass. One of the easiest way to do so is to focus on the publishing record of the scientist proposing the research by examining readily available measures of citations to papers and the Impact Factor in which the paper is published on platforms such as Google Scholar and Scopus. Such was not always the case: as late as the early 1990s the only way to count citations was to laboriously look in the volumes published by the Institute of Scientific Information, which were usually only available in the panel member's institutional library.

Does a heavy focus on bibliometrics affect the panel's decision when it comes to supporting risky research? The work by Wang, Veugelers and Stephan (2017) suggests that the answer could be yes: they find that novel research is systematically less likely to be published in high Impact Factor journals.[27] Such a bibliometric bias against novel research can lead panels to select against individuals with a history of novel (risky) research, especially when the applicant is young and has a short history of citations—which, as noted

---

[25] https://grants.nih.gov/grants/peerreview22713webv2.pdf. Accessed January 2, 2021For Study Sections see https://public.csr.nih.gov/StudySections/StandingStudySections. Accessed January 2, 2021.
[26] Average for 2008-2013.
[27] The fact that these impact factors are calculated using relatively short citation windows, coupled with the supra mentioned finding of Wang et al (2017) that it takes a longer time window for novel research to become highly cited may explain why journal editors, striving for good scores on their impact factor, may be biased against novel research.

above, are biased against novelty. More generally, a focus on bibliometrics shifts the basis of decisions away from the substance of the proposal to an easily accessible metric.

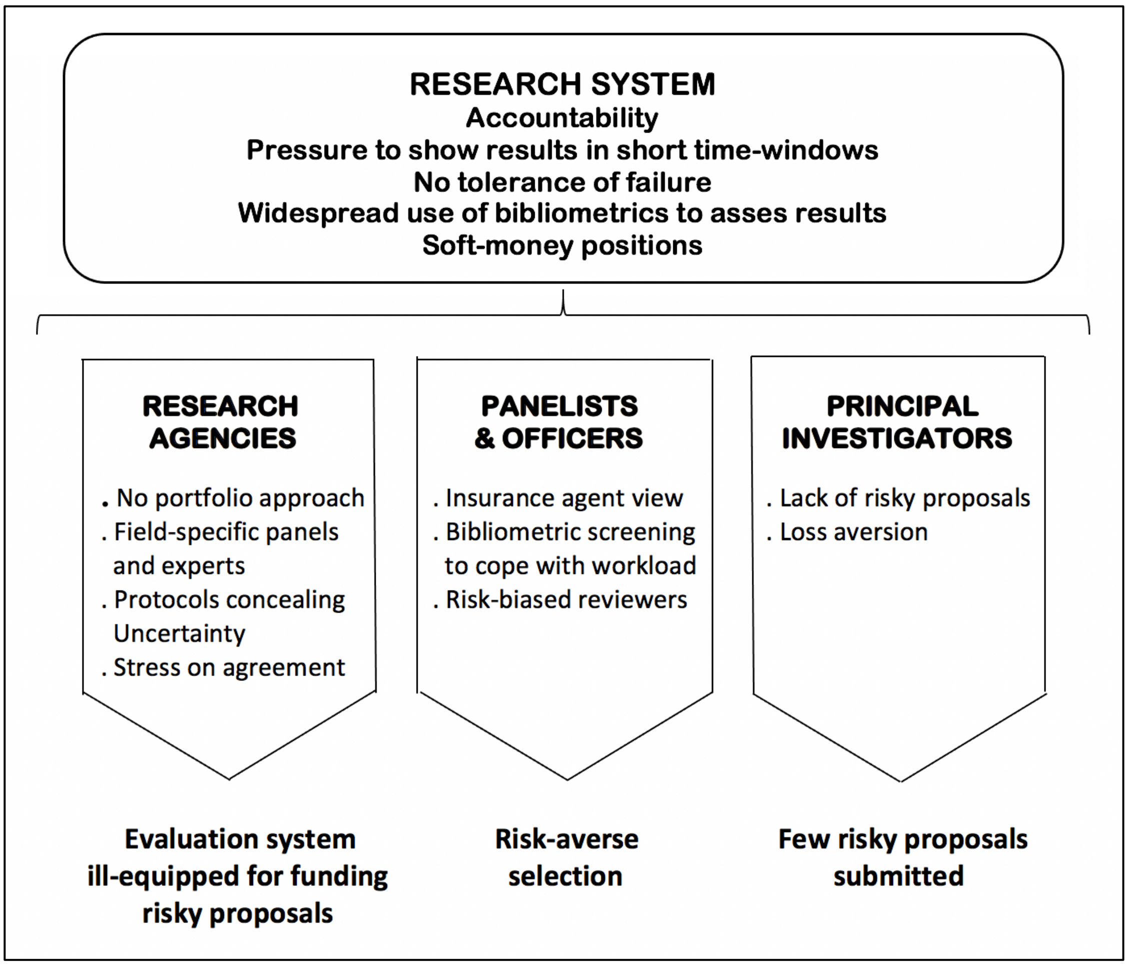
How workload affects the selection of novel R&D projects has not been studied for public funding agencies; it has, however, been studied in R&D departments of for-profit firms. The authors of one study find that a high panel workload reduces the panel's preference for novel research (Criscuolo et al. 2017).

"Risk-biased panel members" hypothesis

Reviewers are often selected based on the excellent expertise of their research profile. Selection of top/senior experts specialized in the exact same niche area of the proposal are typically assumed to be best choice for reviewers. But are top experts also the best reviewers for assessing risky research proposals? Which kind of reviewers are needed for an unbiased assessment of risky research? Does it require experience with risky research to be more willing to take risks in evaluating the proposals of others?

Unfortunately, we lack specific knowledge illuminating which kind of reviewer's expertise is best suited at assessing risky research, both in terms of willingness to take risks and in terms of capacity to accurately assess risk. Only a handful of studies have looked at reviewers' characteristics and related preferences for funding research proposals. Boudreau and colleagues (2014), look at the intellectual distance between the background of the reviewers and the research being proposed. They find that novel proposals were evaluated less favorably than non-novel ones, but this was not explained by the degree of intellectual proximity of the evaluator. Li (2017) studies NIH evaluations and finds that proposals which were related to the research of the panelist were scored more favorably than those unrelated. DellaVigna and Pope (2018) study the accuracy of reviewers when predicting the outcomes of experiments in behavioral social sciences. They find that scholars with expertise in an area were not more accurate than scholars who were not experts. More senior scholars, and academics with more citations were also no more accurate. In conclusion, the limited evidence available suggests that reviewers who are expert in the niche area are not more accurate in assessing research proposals. They may be more prone to fund research in their area, but they may not be more inclined to fund novel research in the area. This is a particularly worrisome result, given the preference for reviewers who are seen as top specialist.

Figure 3 Incentives and opportunities of risky research: A summary.



5. Possible ways to encourage risk taking to advance breakthrough research in science: Suggestions for Science Funders

Society needs scientific breakthroughs to tackle many of the challenges it faces. Because many of the paths to such breakthroughs are risky, its science system and particularly its public science funding system, need to ensure that risk taking is encouraged or, at a minimum, that the system is not biased against risky research. The previous sections have made clear that we cannot take for granted that this is the case. The findings and hypotheses that we have explored also suggest possible ways for moving forward.

The Karikó -Weissman mRNA case already provides some initial thoughts regarding ways to promote risky research. Their early joint work appears to have been partially supported on a grant Weissman obtained that did not directly relate to designer mRNA. Such a "backburner" strategy is not uncommon in science. Scientists regularly employ funding obtained to support another research objective, particularly in the very early exploratory stages of the other research, when it is still highly risky and without enough preliminary findings to apply for dedicated funding.

The mRNA example also shows the importance luck plays in pathbreaking research: if they had not met at the copy machine, Karikó and Weissman might never have formed the collaboration that led to mRNA vaccines. The probability that such lucky encounters occur, can be at least partly "engineered". A work environment enabling more open serendipitous encounters has the potential of leading to more risky research built on new unprecedented connections of knowledge pieces. The example also underlines the importance of taking an interdisciplinary approach: Karikó was trained as a biochemist, Weissman as an immunologist, a powerful combination to address the critical bottleneck for mRNA to be effective and safe in humans. An open environment enabling cross-disciplinary connections could thus already take away an important impediment for risky research. Moving the example from the mRNA novel scientific insights supported in part by NIH funding, into the development of mRNA-based vaccines for the market, supported in part by SBIR funding, DARPA, and eventually BARDA and Warp Speed, shows the importance of staging the funding of the new approaches.

We conclude with suggestions of ways to encourage risk taking (or at a minimum avoid a bias against risk) in science, combining insights from the Karikó-Weissman mRNA research with insights from the admittedly limited evidence and research on risk avoidance as it relates to science funding reviewed in the previous sections. As discussed supra, discouraging risk needs to be addressed within the entire science system: it cannot be solved by an individual program or funding agency. It requires a holistic perspective on the science enterprise, activating not only funders and their reviewers, but also universities and research centers, journal editors and their reviewers and, last but not least, researchers themselves. Without ignoring the holistic perspective, we nevertheless focus the discussion on suggestions for science funders if they would like to augment their support for risky research. Where appropriate, we suggest further research and experiments designed with the goal of advancing our knowledge of ways to promote risk taking in science.

Deemphasize bibliometrics

Funding agencies, in order to advance innovation, could insist on multiple ways to assess applicants' research record, avoiding an overly focused use of bibliometrics. They could refrain from asking or inferring that grant applicants provide short-term citation counts, and indicators based on short term windows, such as JIFs and Top field cited articles. They could instruct panels to abstain from discussing such indicators in reviews or at a very minimum instruct panels of the potential biases which using such indicators entails.

Diversity in panel composition

Funders could balance panel composition with a sufficiently large number of panel members holding diverse and independent perspectives. They could avoid a panel design which is narrowly disciplinary focused, as thus runs the danger of underappreciating the out-of-field broader impact from risky research. This is a more complicated task than selecting on the basis of top expertise in the field as a main panel selection criterium, as is commonly done and which runs the danger of wrongly relying on the assumed superior assessment of experts regarding possible outcomes. More importantly, much could be learned concerning the causal impact of panel composition on risky research selection by using random control design to run experiments on panel composition.

Allow for disagreement

Alternatives to the commonly used consensus or average procedures should be considered. Aggregation/deliberation rules could be adapted to the nature of the science that the grant aims at sponsoring. Because more risky research is more prone to extreme outcomes, it matters not only to have reviewers willing to take risk, but also an accurate assessment of these extreme outcomes and their probability of occurrence. This requires a large enough number of sufficiently uncorrelated risk-unbiased opinions. In addition, the evidence that risky research may lead to more polarized views, warns against aggregation methods that rely on consensus (e.g., behavioral aggregation), or that assumes distributions of opinions according to a bell-shape (e.g., arithmetic average). To learn more about alternative methods that do not imply consensus, experiments with such alternative procedures could be conducted, using random trials so that we can properly evaluate their impact on the selection of risky research.

Portfolio approach

The 'one by one' approach typically used in panels works against selecting risky proposals. At a minimum, panels need to think about correlation among the proposals they are funding. One sure sign of high correlation in terms of low risk is requiring that all successful proposals have convincing preliminary findings. More generally, a portfolio approach to address risk aversion could require panels to put in different baskets highly risky and moderately risky proposals and provide a way to choose proposals from each. In practice such a portfolio approach could be quite challenging to implement for research projects. First, portfolio theory requires that the research paths be sufficiently uncorrelated. This may not hold within panels that are focused on specific subdisciplines that share risk factors. Correlation between research paths, in and of itself, can be hard to determine, particularly when covering vastly different goals across different fields and with different research approaches. Second, there is the question of fairness: in building a portfolio approach some proposals may have to be eliminated in an effort to balance or de-risk the portfolio.

Staging

An approach to de-risking, commonly used for funding entrepreneurial projects in the Venture Capital industry, where it is referred to as the 'spray and pray' strategy (Lerner and Nanda 2020), is to fund in stages, where increasingly larger amounts of funding are allocated, depending on whether interim milestones are being met. Funding in stages can be combined to include a portfolio approach[28].

Although this strategy is used by DARPA,[29] the SBIR program and was recently introduced into similar EIC program in the EU, it is used less frequently among most science funding agencies. Can such a staging approach also be used by science funding agencies allowing them to take more risk? Interim evaluation is especially useful when initial estimates are unreliable, but can be quickly updated and when investments can start at small scale but can be quickly scaled up (Vilkkumaa et al. 2015). It is thus especially suitable for research that can make substantial steps forwards in a relatively short period of time and does not require large fixed-costs to be started (Ewens, Nanda, and Rhodes-Kropf 2018).[30] Some research fields meet these conditions, but the conditions are more the exception than the norm in the natural sciences, where the share of research that requires expensive equipment and substantial effort is large (Stephan 2010).

Loose-play early stage ideas

In the mRNA example, it was crucial that Weissman could use some of the funding he had already obtained to support early stage risky joint research with Karikó. Other

[28] Veugelers and Zachmann (2020), for example, proposed a combination of a staging and portfolio approach to fund vaccines projects and calculated what such an approach would cost to society to obtain a desired number of vaccines at the end.
[29] https://fas.org/sgp/crs/natsec/R45088.pdf Accessed August 30, 2020.
[30] In the VC industry, this has largely favored IT and digital companies.

researchers often do the same. But a necessary condition to doing so requires that the PI have existing funding that can be redirected. If these are funds obtained from regular science funding programs, this raises the question of selection bias in terms of who obtains such funding. Researchers with a track record in novel research may be biased against. Also, early career researchers do not have access to such funds; others, such as Karikó, have tried to get funding but have not succeeded. And even for those who have such funding, the rules of engagement may not allow use funds for other than the research described in the proposal.

Going beyond the "backburner" option, could *dedicated* loose-play funding for early stage risky explorations be operationalized? One approach is to have early stage funding readily and quickly available to researchers at their home institution. The California Institute of Technology, by way of example, had such a program whereby faculty could submit a short proposal to the Vice Provost for Research and get a decision in a matter of days. Funds ranged from $25,000 to $250,000 a year for a period of two years. The idea was to give faculty the wherewithal to engage in early-stage risky research that, given apparent risk aversion of granting agencies, was deemed not yet ready for submission. If the initial findings looked promising, and produced enough preliminary data, the faculty could then submit a full grant proposal. Other institutions, like ETH Zurich, provide generous base research funding to all their chairs which they can deploy for research at their own initiative.

Such dedicated programs for early stage risky research require resources, which institutions often do not have or they do not want to redirect away from their normal programs. One way to provide institutions with the financial means for such funding schemes would be for federal funders to shift a trench of resources to local institutions with the goal of encouraging risk taking. Dedicated loose-play programs have a number of pluses: it decreases the amount of effort that goes into preparation of fully-fledged proposals, it is reasonably easy to administer, and it has the potential of "de-risking" a research agenda before it goes up for formal evaluation at a granting agency. It also has downsides. It could, for example, promote favoritism at the local level. It requires willingness and capacity to support risky projects at the local level, or at least no bias against risky projects, something which may not be present, as argued supra. It also involves a willingness to provide salary support to applicants, especially applicants in soft money positions. A primary reason Karikó was turned down for the $10,000 Merck starter grant, administered by the University of Pennsylvania, that she applied for in 1991, was a request for salary support. The rejection letter singled this out, citing one reviewer who said that "the most substantial weakness is the use of the entire award for faculty salary support."[31] Politically, shifting funds from federal agencies to universities for such programs also involves granting agencies ceding some control to local institutions.[32]

Funding researchers rather than projects and for longer periods of time

Programs that fund researchers rather than projects for longer periods of time allow researchers to engage in more risky research. It gives the scope and time to researchers to redirect their research in case of failure. The example that readily comes to mind is The Howard Hughes Medical Institute (HHMI) that funds successful applicants for seven years, rather than for three to five years, as is common for most other funding organizations, where selection is based more on the applicant and his longer-term research strategy rather than a specific research project. HHMI, moreover, does not demand early results nor does it penalize researchers for early failure. Azoulay et al (2011) compare the research output of HHMI investigators to a group of similarly accomplished NIH investigators using propensity scoring. They find that HHMI investigators use more novel keywords and produce more hits and more flops, compared to the NIH investigators. Although it is not clear whether these results depend upon the longer duration of grants and the practice of HHMI to not demand early results nor penalize researchers for early failure or other variables, the results suggest that these practices encourage risk taking.

---

[31] Letter to Karikó dated April 29, 1991.

[32] NIH already does this by awarding training grants to institutions to administer. In the early years of NIH, candidates for training awards were selected at NIH.

Targeting science funding to risky breakthrough missions

In the Karikó -Weissman case, the vexing problem of immune response was a well-known scientific challenge, impeding the promising mRNA technology to be used as a modality for treating humans. As such, it could have been turned into a "mission" with dedicated funding. While the science funding system was either unable or unwilling to identify such a "mission" in the early research stages of mRNA research, a more "mission" oriented approach was followed in later stages, first when DARPA awarded Moderna up to $24.6 million in 2013 "to research and develop its messenger RNA therapeutics™" after having already awarded the company a "seedling" grant of .7 million to begin work on the project.[33] Subsequently, when the search for corona-virus vaccines became more pressing and BARDA and WarpSpeed entered the picture, a more targeted approach for funding their development was used.

DARPA is frequently heralded for its successes in funding mission-oriented high-risk, high-reward research. Azoulay, Fuchs et al (2019) identify the organizational flexibility and autonomy given to program directors as key success elements of the DARPA model. Individual discretion by the program officer, designing and selecting projects from across the distribution of reviewer scores, is seen as an antidote to the risk bias DARPA's peer reviewers hold. Autonomy of the program officers goes together with accountability with clear targets. Key is attracting program staff who are more like risk-taking, idea-driven entrepreneurs than administrators.

Can the DARPA model be replicated for avoiding the risk bias in funding basic research? Azoulay, Fuchs et al (2019) identify as *DARPAble* domains mission motivated research on nascent technologies within an inefficient innovation system. Given that the missions must be clearly identifiable, associated with quantifiable goals and trackable progress metrics, the authors deem the DARPA model not appropriate for funding basic research. The focus of basic research on improving our understanding is not a clearly

---

[33] https://investors.modernatx.com/news-releases/news-release-details/darpa-awards-moderna-therapeutics-grant-25-million-develop

defined mission. This is not to say that DARPA does not fund early stages of research, as specific scientific challenges may need to be met before the technology can progress. But because fundamental understanding is not in itself a goal of DARPA, it is not designed for basic research per se. While DARPA may not be a general model for funding basic research, it may nevertheless be inspirational for specific scientific challenges for which goals can be clearly defined.

Prizes

An alternative to a grants approach is to create prizes to encourage path-breaking research (e.g. Williams 2010, 2012)[34]. Although such prizes for risky breakthroughs can incentivize research, they shift the risk onto the shoulders of the researchers, given that prizes are only awarded conditional upon success. It may thus only incentivize researchers who are risk lovers, who have high (perhaps too high) estimates of their probability of success and have access to resources.

To conclude, science funding agencies should be encouraged to pave the way for promoting risk taking in scientific research, given that breakthrough research is often perceived as risky. The way forward is neither safe, nor is it clearly defined. It is a "risky" road, as it is not clear which paths will work and which will not. Perhaps the most important contribution funding agencies can make would be to support research which builds knowledge on the design of funding programs and reviewing practices related to risky proposals that have the potential of delivering breakthroughs. This support could entail not only financing such research, but also granting access to data and being the gardens where experiments can be conducted.

[34] https://www.nsf.gov/sbe/sosp/tech/williams.pdf
https://drive.google.com/file/d/17bthRCWoDcW4LfzukdhjFKINr1vkLWaz/view

6. References

Althaus, Catherine E. 2005. “A Disciplinary Perspective on the Epistemological Status of Risk.” *Risk Analysis* 25(3):567–88.

Aven, Terje. 2011. “On Some Recent Definitions and Analysis Frameworks for Risk, Vulnerability, and Resilience.” *Risk Analysis* 31(4):515–22.

Azoulay, Pierre, J. Graff Zivin, and Gustavo Manso. 2011. “Incentives and Creativity: Evidence from the Howard Hughes Medical Investigator Program.” *The Rand Journal of Economics* 42(3):527–54.

Azoulay, Pierre, Joshua Graff Zivin, and Gustavo Manso. 2012. *NIH Peer Review: Challenges and Avenues for Reform*. 18116. Cambridge, MA.Azoulay, Pierre, Erica Fuchs, Anna Goldstein, Michael Kearney.2019. *Funding Breakthrough Research: Promises and Challenges of the "ARPA Model"*, NBER.

Azoulay, Pierre and Danielle Li. 2020. *Scientific Grant Funding*. Vol. 26889.

Banal-Estañol, Albert, Inés Macho-Stadler, and David Pérez-Castrillo. 2019. “Evaluation in Research Funding Agencies : Are Structurally Diverse Teams Biased Against ?” *Research Policy* 48:1823–40.

Banerjee, A. V. 1992. “A Simple Model of Herd Behavior.” *Quarterly Journal of Economics* 107(3):797–817.

Bolger, Fergus and Gene Rowe. 2015. “The Aggregation of Expert Judgment: Do Good Things Come to Those Who Weight?” *Risk Analysis* 35(1):5–11.

Boudreau, Kevin, Eva Guinan, Karim R. Lakhani, and Christoph Riedl. 2014. “Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance 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):684–87.

Clemen, Robert T. and Robert L. Winkler. 1999. “Combining Probability Distributions From Experts in Risk Analysis.” *Risk Analysis* 19(2):187–203.

Cohen, Wesley M., Henry Sauermann, and Paula Stephan. 2020. “Not in the Job Description: The Commercial Activities of Academic Scientists and Engineers.” *Management Science* 66(9):4108–4117.

Cooke, Roger M. 1991. “Experts in Uncertainty. Opinions and Subjective Probabilities in Science.” New York Oxford: Oxford University Press.

Criscuolo, Paola, Linus Dahlander, Thorsten Grohsjean, and Ammon Salter. 2017. “Evaluating Novelty: The Role of Panels in the Selection of R&D Projects.” *Academy of Management Journal* 60(2):433–60.

DellaVigna, Stefano and Devin Pope. 2018. “Predicting Experimental Results: Who Knows What?” *Journal of Political Economy* 126(6):2410–56.

Ewens, Michael, Ramana Nanda, and Matthew Rhodes-Kropf. 2018. “Cost of Experimentation and the Evolution of Venture Capital.” *Journal of Financial Economics* 128(3):422–42.

Fang, Ferric C. and Arturo Casadevall. 2016. “Research Funding: The Case for a Modified Lottery.” *MBio* 7(2).

Fortunato, Santo, Carl T. Bergstrom, Katy Börner, James A. Evans, Dirk Helbing, Staša Milojević, Alexander M. Petersen, Filippo Radicchi, Roberta Sinatra, Brian Uzzi, Alessandro Vespignani, Ludo Waltman, Dashun Wang, and Albert László Barabási. 2018. “Science of Science.” *Science* 359(6379).

Foster, Jacob G., Andrey Rzhetsky, and James A. Evans. 2015. “Tradition and Innovation in Scientists’ Research Strategies.” *American Sociological Review* 80(5):875–908.

Franzoni, Chiara and Paula Stephan. 2021. *Uncertainty and Risk-Taking in Science: Meaning, Measurement and Management*.

Gross, Kevin and Carl T. Bergstrom. 2019. “Contest Models Highlight Inherent Inefficiencies of Scientific Funding Competitions." *PLoS Biology* 17(1).

Hansson, Sven Ove. 2018. "Risk." in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta. Stanford University.

Hora, Stephen C., Benjamin R. Fransen, Natasha Hawkins, and Irving Susel. 2013. "Median Aggregation of Distribution Functions." *Decision Analysis* 10(4):279–291.

Ismail, Sharif, Alice Farrands, and Steven Wooding. 2008. *Evaluating Grant Peer Review in the Health Sciences: A Review of the Literature*. Santa Monica.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47(2):263–92.

Kaplan, David, Nicola Lacetera, and Celia Kaplan. 2008. "Sample Size and Precision in NIH Peer Review." *PLoS ONE* 3(7).

Knight, Frank. 1921. *Risk, Uncertainty and Profit*. (1964 repr. New York, NY: Sentry Press.

Lamont, Michele. 2009. *How Professors Think. Inside the Curious World of Academic Judgement*. Cambridge, MA: Harvard University Press.

Lanoë, Marianne. 2019. "The Evaluation of Competitive Research Funding: An Application to French Programs." L'Universite de Bordeaux.

Larrick, Richard P. and Jack B. Soll. 2006. "Intuitions About Combining Opinions : Misappreciation of the Averaging Principle." *Management Science* 52(1):111–27.

Laudel, Grit. 2017. "How Do National Career Systems Promote or Hinder the Emergence of New Research Lines?" *Minerva* 55(3):341–69.

Lerner, Josh and Ramana Nanda. 2020. "Venture Capital's Role in Financing Innovation: What We Know and How Much We Still Need to Learn." *Journal of Economic Perspectives* 34(3):237–61.

Li, Danielle. 2017. "Expertise versus Bias in Evaluation: Evidence from the NIH." *American Economic Journal: Applied Economics* 9(2):60–92.

Linton, Jonathan D. 2016. “Improving the Peer Review Process: Capturing More Information and Enabling High-Risk/High-Return Research.” *Research Policy* 45(9):1936–38.

List, Christian. 2012. “The Theory of Judgment Aggregation: An Introductory Review.” *Synthese* 187:179–207.

Martini, Carlo and Jan Sprenger. 2018. “Opinion Aggregation and Individual Expertise.” Pp. 180–201 in *Scientific Collaboration and Collective Knowledge*, edited by T. Boyer-Kassem, C. Mayo-Wilson, and M. Weisberg. Oxford University Press.

Mazzucato, Mariana. 2015. *The Entrepreneurial State. Debunking Public vs. Private Sector Myths*. New York: Anthem Press.

Mengel, Friederike. 2019. *Gender Biases in Opinion Aggregation*.

Merton, Robert K. 1957. “Priorities in Scientific Discovery: A Chapter in the Sociology of Science.” *American Sociological Review* 22(6):635–59.

Morgan, M. Granger. 2014. “Use (and Abuse) of Expert Elicitation in Support of Decision Making for Public Policy.” *Proceedings of the National Academy of Sciences* 111(20):7176–84.

Morgan, Millett Grager and Max Henrion. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis 6th Edition*. Cambridge University Press.

Nelson, Richard R. 1959. “The Simple Economics of Basic Scientific Research.” *Journal of Political Economy* 67(3):297–306.

Nicholson, Joshua M. and John P. A. Ioannidis. 2012. “Research Grants: Conform and Be Funded.” *Nature* 492(7427):34–36.

Ottaviani, Marco and Peter Norman Sorensen. 2015. *Combining Reviews in Expert Panels*.

Packalen, Mikko and Jay Bhattacharya. 2018. *Does the NIH Fund Edge Science?* 24860.

Cambridge, MA.

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 of the United States of America* 115(12):2952–57.

Rosenberg, Nathan. 1992. “Economic Experiments.” *Industrial and Corporate Change* 1(1):181–203.

Roumbanis, Lambros. 2019. “Peer Review or Lottery? A Critical Analysis of Two Different Forms of Decision-Making Mechanisms for Allocation of Research Grants.” *Science, Technology, & Human Values* 44(6):994–1019.

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.

Sahin, Ugur, Katalin Karikó, and Özlem Türeci. 2014. “MRNA-Based Therapeutics-Developing a New Class of Drugs.” *Nature Reviews Drug Discovery* 13(10):759–80.

Snell, Richard R. 2015. “Menage a Quoi? Optimal Number of Peer Reviewers.” *PLoS ONE* 10(4).

Stephan, Paula. 2007. “Early Careers for Biomedical Scientists: Doubling (and Troubling) Outcomes Presentation Harvard University.” in *Presentation Harvard University*. Boston, MA.

Stephan, Paula. 2012. *How Economics Shapes Science*. Cambridge, MA: Harvard University Press.

Stephan, Paula. 2019. “Practices and Attitudes Regarding Risky Research.” Stanford University: Metascience Symposium.

Stephan, Paula E. 2010. “The Economics of Science.” Pp. 217–73 in *Handbook of Economics of Innovation*, edited by B. H. . Hall and N. Rosenberg. Amsterdam, NL: North Holland.

Stephan, Paula and Sharon Levin. 1992. *Striking the Mother Lode in Science: The Importance of Age, Place, and Time*. Oxford, UK: Oxford University Press.

Stephan, Paula, Reinhilde Veugelers, and Jian Wang. 2017. "Blinkered by Bibliometrics." *Nature* 544(7651):411–12.

Tversky, Amos and Daniel Kahneman. 1991. "Loss Aversion in Riskless Choice: A Reference-Dependent Model." *Quarterly Journal of Economics* 106(4):1039–61.

Uzzi, Brian, Satyam Mukherjee, Michael Stringer, and Ben Jones. 2013. "Atypical Combinations and Scientific Impact." *Science* 342(6157):468–72.

Veugelers, Reinhilde, Paula Stephan, and Jian Wang. 2021. "Excess Risk-Aversion at ERC." unpublished manuscript.

Veugelers, Reinhilde and Georg. Zachmann. 2020. "Racing against COVID-19: A Vaccines Strategy for Europe." *Policy Contribution* (April).

Vilkkumaa, Eeva, Ahti Salo, Juuso Liesiö, and Afzal Siddiqui. 2015. "Fostering Breakthrough Technologies - How Do Optimal Funding Decisions Depend on Evaluation Accuracy?" *Technological Forecasting and Social Change* 96:173–90.

Viner, Neil, Philip Powell, and Rod Green. 2004. "Institutionalized Biases in the Award of Research Grants: A Preliminary Analysis Revisiting the Principle of Accumulative Advantage." *Research Policy* 33(3):443–54.

Wagner, Caroline and Jeffrey M. Alexander. 2013. "Evaluating Transformative Research Programmes: A Case Study of the NSF Small Grants for Exploratory Research Programme." *Research Evaluation* 22(3):187–97.

Wang, Jian, Reinhilde Veugelers, and Paula Stephan. 2017. "Bias against Novelty in Science: A Cautionary Tale for Users of Bibliometric Indicators." *Research Policy* 46(8):1416–36.

Wolff, JA, RW Malone, P. Williams, W. Chong, G. Acsadi, A. Jani, and PL. Felgner. 1990. "Direct Gene Transfer into Mouse Muscle in Vivo." *Science* 23(247):1465–68.