Off the beaten path: what drives scientists' entry into new fields?

Peer-reviewed author version

KELCHTERMANS, Stijn; Neicu, Daniel & Veugelers, Reinhilde (2022) Off the beaten path: what drives scientists' entry into new fields?. In: INDUSTRIAL AND CORPORATE CHANGE, 31 (3) , p. 654-680.

DOI: 10.1093/icc/dtab054
Handle: http://hdl.handle.net/1942/37461

Off the Beaten Path:

What Drives Scientists' Entry into New Fields?

Stijn Kelchtermans (KU Leuven, Hasselt University)[1]

stijn.kelchtermans@kuleuven.be

Daniel Neicu (European Commission)[2]

daniel.neicu@ec.europa.eu

Reinhilde Veugelers (KU Leuven, Bruegel, CEPR)

reinhilde.veugelers@kuleuven.be

August 2021

Abstract

Given that venturing into unknown territory carries substantial risk, scientists don't take the decision to enter a new field lightly. This paper analyses a broad set of factors associated with the risks and rewards from entry into new-to-the-researcher scientific fields, including individual capacities and preferences as well as incentives stemming from career progression and access to funding. Using a panel of researchers in biomedical sciences and science & engineering from a large European research university, we find that productivity affects new field entry as such, but is not associated with entry into fields that are very distant to one's current expertise. Scientists in more senior ranks, with larger co-author networks and collaborating with PhD students are more likely to enter new fields but these factors don't represent an additional push to enter very remote fields. Such 'long jumps' are more likely to be made by above-average talented, rather than merely productive researchers. Finally, accounting for its endogeneity, we find that funding does not make new field entry more likely.

Keywords: economics of science; new field entry; scientific funding

Funding: This work was supported by KU Leuven (GOA/12/003, BOF HUB-KUB 3H110407), FWO (G.085816.N) and NBB (NBB/15/012).

Acknowledgements: The authors thank ECOOM for allowing access to the publication data. Publication data are sourced from the Clarivate Web of Science Core Collection.

"In this age of specialization, men who thoroughly know one field

are often incompetent to discuss another."

Richard Feynman, Caltech lunch forum, 1956

1 Introduction

Science is known to be a risky endeavour, where payoffs in terms of scientific success can be highly uncertain (Arrow, 1962). Given that 'wrong moves' in one research trajectory may have long-lasting consequences, scientists arguably don't take the decision to enter a new field lightly. These choices define whether they develop a more generalist profile that spans disciplinary boundaries, rather than being a specialist in a narrower set of fields, and have ramifications for the production of scientific knowledge at an aggregate level.

Prior work studying what drives scientists to diversify their research agendas and enter research fields which are new to them has zoomed in on specific fields, like biomedical sciences (Azoulay *et al.,* 2011) and/or specific factors - like career incentives, teams, or funding - as potential explanations for why scientists expand the scope of their research. For example, Franzoni & Rossi-Lamastra (2017) find that scientists who have obtained tenure are more likely to diversify their research portfolio and theorize that the associated job security mitigates the risks that come with diversifying into new fields. Azoulay *et al.* (2019) find that, following the unexpected deaths of preeminent life scientists, individuals from outside fields are more likely to enter and succeed. Azoulay *et al.* (2011) showed that the HHMI funding program stimulates biomedical scientists to explore new lines of investigation within biomedical research. Myers (2020) finds that it requires a substantial amount of (expected) funding to induce scientists to make even small changes in the direction of their research.

Our analysis contributes to the small but growing empirical literature studying scientists' decisions to redirect their research orientation, by looking at the decision to enter new fields. We take a comprehensive approach, covering a wide set of influencing factors for researchers, across different scientific fields. We build on the economics of science literature to develop a framework that outlines the risks and rewards associated with new field entry, incorporating individual capacities and preferences as well as incentives stemming from career progression and the competitive funding system. Since not all new field entries are alike in terms of their distance to the scientist's existing expertise and since distance will affect the risk-reward trade-off, we not only analyse them as binary outcomes but also consider the distance of new field entries and whether they are particularly remote from the scientist's current field(s) of research. In our empirical analysis, we can use a rich dataset from a leading European research university (KU Leuven, Belgium). The set of determinants we consider ranges from talent, past track record, experience with new field entry and co-authorship teams, over tenure and rank, to research funding, while also taking into account career age, gender and scientific field differences. For funding, we assess whether research grants affect new field entry and its distance, taking into account its endogeneity. More specifically, we disentangle whether drivers such as talent and career progression affect new field entry directly as well as indirectly through their effect on funding.

Our unbalanced panel (1992-2001) of 734 KU Leuven researchers in biomedical and exact sciences shows that the effects of the drivers of new field entry tend to differ depending on whether entry is distant or not. We find in our econometric analysis that high-quality scientists are not more likely to enter new fields per se. However, conditional on new field entry, they are more likely to make 'long jumps'. In contrast, scientists' past productivity record is associated with new field entry, but it is not associated with (long) distance, conditional on entry. Researchers' past new field entries have little bearing on whether they are more likely to repeat that behaviour, suggesting no evidence of experience nor increasing returns to new field entry. Career incentives matter: each successive step in the academic hierarchy provides an additional push to enter new fields. Yet, they do not matter for more distant entries, suggesting they induce a balanced risk-reward trade-off. For research funding, we find in our sample no relation with new field entry as such. But, conditional on entry, tentative results indicate that academics are more inclined to venture into more distant research fields if they obtained a grant.

The remainder of the paper is organized as follows. The next section outlines our framework for assessing the risk-reward trade-off that scientists face when they consider entering new fields. Section 3 discusses the data and operationalizes the variables used in the analysis. Sections 4 and 5 present the results and robustness checks. Section 6 concludes and reflects on policy implications.

2 Framework for Assessing New Field Entry

Implicitly or deliberately, scientists take into account a multitude of factors when deciding whether to embark on research that branches out into scientific fields they are less familiar with. To analyse these decisions, we build on the economics of science literature to develop a framework that outlines the risks and rewards associated with new field entry, incorporating individual capacities and preferences, incentives stemming from career progression, and the competitive funding system.

2.1. The Risk-Reward Trade-Off

The scientific reward system is a critical factor that shapes scientists' decisions on which research trajectories to follow. More specifically, with rewards going disproportionally to 'big hits', there's a ruthless competition to win the race for being the first to come up with big scientific discoveries (Dasgupta & David, 1994). This system may push scientists to take more risk by going into new territory in order to increase their odds to come up with a big win. At the same time, the 'winner takes all'-logic may hold back a scientist to enter a new field unless she has strong beliefs that she has a comparative advantage to win the race.

Given that the nature of scientific competition is focused on big wins, how to trade off the rewards and risks when deciding on entering new fields? Similar to financial portfolio investment, one could consider new field entry as a diversification strategy to lower the overall risk of failure. At the same time, diversification into new areas may avoid the diminishing returns from continuing in existing areas and reaching dead ends (Hackett, 2005). Diversification constitutes an even more valuable coping mechanism in more mature fields, where the risk looms larger that a scientist finds herself in an exhausted area of research (Siow, 1998).

Besides limiting the downside risk, another line of reasoning in favour of new field entry is that it increases the 'upside potential' of one's research. A central tenet of the literature on creative thinking is that creativity requires breaking away from prior experience (Simonton, 1999), allowing for new perspectives that arguably characterize major breakthroughs in science. Wang *et al.* (2017) show how novel science, characterized as making new, first time ever, combinations of existing knowledge components, although entailing a higher risk of failure, is more likely to become a big hit, although this may take some time. From this perspective, choosing a more novel research trajectory by venturing into new areas may offer the researcher the prospects of a higher return but comes with a higher risk.

Manso (2011) describes this risk-return trade-off as a choice between *exploration* of new untested areas and the *exploitation* of well-known areas. Exploration reveals information about potentially superior actions, but is also likely to waste time when pursuing inferior ideas. Exploitation of well-known trajectories ensures reasonable payoffs, but may prevent the discovery of superior paths to discovery. Although entering new areas offers the scope for superior outcomes, researchers have to consider that entering an unfamiliar field comes with an opportunity cost in the form of foregone further specialisation in existing areas of expertise. That cost is arguably non-negligible given the ever-accumulating knowledge that one needs to assimilate in order to stay abreast of the latest findings in one's domain, and which is considered an important reason for the observation of increased specialization in science (Jones, 2009). Persistently focussing on a narrow set of topics may not only be needed to acquire the necessary competences, but also to convert accumulated expertise in published results (Hackett, 2005) as well as to establish and defend one's reputation in the field (Cole & Cole, 1967).

Influencing the risk-reward trade-off when venturing into new areas, is how distant the new areas are from the areas which the researcher is familiar with. The more different and uncorrelated the new fields are, the higher the advantage from a risk diversification perspective. Within the exploration versus exploitation perspective, going into new but closely related territories may reduce the risk of failure, but it may also lower the potential rewards - particularly the potential for big rewards - as these may require more, and highly different, areas to be combined (Wang *et al.*, 2017). Hence, going into very distant new areas may offer even higher rewards, but with commensurate higher risk.

2.2. Factors Shaping the Risk-Reward Trade-Off from Venturing into New Areas

Which researchers are more likely to venture into new areas? Which factors shape their decisions to enter new fields within a risk-reward trade-off, balancing between *exploring* new untested avenues versus *exploiting* well known ones?

As also pointed out by Borjas & Doran (2015a), underlying scientists' decisions of research direction is a complex trade-off of factors. It is worth emphasizing that most existing work on research orientation has looked at a particular driver, often within a single discipline (typically the life sciences). With our framework and empirical analysis we capture some of that complexity better than before by providing a broader view of the drivers of new field entry as well as their interactions. In what follows, we look at a comprehensive set of factors related to both scientists' capacity and preferences for reward and risk taking (section 2.2.1) as well as incentives for reward & risk taking generated by the science system environment in which the individual researcher operates (section 2.2.2). For science system incentives, we look at (i) scientific career schemes and (ii) access to funding. We also discuss the impact of specific demand and supply shocks in scientific disciplines as drivers of researchers' reorientation decisions (section 2.2.3).

2.2.1 Competences and Preferences for Exploration

Talent and Productivity Record in Existing Areas of Expertise

The capacity to be creative in science and to deal with its risk-reward trade-off is linked to talent, sometimes labelled the "sacred spark"-hypothesis. More talented scientists are better able to generate more novel ideas and be more successful in bringing these novel approaches to a success (Simonton, 2003), particularly when entering more distant new areas.

Scientists weigh their talent – and the associated capacity to capitalize on new field entry - and their preference for risk against the cost of new field entry. Although highly talented researchers may be apt at avoiding the pitfalls associated with entering a new field, they are not certain that their existing expertise will be transferable to the new field or that productive cross-fertilization between fields will take place. When their talent is specific rather than "general", talented researchers face a higher opportunity cost for venturing into new territories. The uncertainty on the transferability of their talent implies a higher risk when entering new fields, compared to remaining on familiar turf where their talent has proven to be valuable. Past work looking at inventors has shown that successful people tend to favour exploitation over exploration, generating increasingly incremental ideas (Audia & Goncalo, 2007).

In sum, while a scientist's talent and productivity record in his existing areas of expertise arguably play an important role when deciding to enter new fields, it is ex ante not clear whether one's talent and productivity record relate positively or negatively to new field entry. Also, the relation with making long jumps is unclear. Talent may allow making bigger jumps, but at the same time a particular concern for 'long jumps' to remote fields is that one's existing expertise will not be as useful in the new field, with the more seasoned talents and those with longer productivity records facing a higher opportunity cost.

Teams

Science is increasingly being done in (larger) teams, which has been ascribed to the increasing 'burden of knowledge' and easier long-distance communication (Wuchty *et al.*, 2007; Jones *et al.*, 2008; Jones, 2009; Agrawal *et al.*, 2013). While the scientist's decision to enter a new field is ultimately her own, the team(s) she is part of shape the social context within which she does her work and the access she has to complementary knowledge and expertise from team members. A researcher's team may thus be an important factor shaping the scientist's new field entry decision. More specifically, co-authors may serve as a steppingstone to enter new fields of research if they act as a conduit to different bodies of knowledge (Lee *et al.*, 2015). While they did not consider entry into new fields as such, Ayoubi *et al.* (2017) show the role of teams for scientists' tendency to tap into knowledge that is new to them but known to their team members. Being part of a team, particularly of a team crossing diverse competences and fields, may thus provide access to knowledge allowing to better value the benefits and risks from exploring other fields.

Preference for and Experience with New Field Entry

Scientists may have a preference for exploration, enjoying the challenges from pursuing risky avenues: the "puzzle joy" referred to by Stephan (1992). They may also have an innate talent or ability for exploration, being particularly apt at bridging fields and making new combinations. Intrinsic preferences for exploration, attitude towards risk and ability to bridge fields will result in persistency of new field entries, i.e. a scientist's propensity to enter new fields positively depends on her track record of past new field entries. In addition, familiarity with a bigger and broader range of scientific fields implies more opportunities to exploit synergies when considering to venture into yet another field, making it increasingly more attractive to enter new fields.

However, since there is a limited number of scientific fields someone can reasonably enter, in particular fields with sufficient scope for successful synergies, there may be decreasing marginal returns of new field entry for scientists with a larger stock of past field entries.

Career Age

Life cycle effects have been demonstrated to hold with respect to scientific productivity, with performance decreasing as one ages (Levin & Stephan, 1991). They may also affect decisions to enter new fields, independent from their effects on productivity. A shorter remaining time horizon may reduce the appetite for taking a more risky exploration approach and making the associated investments to master new fields, given that the returns may take a long time to materialize.

2.2.2 Science System Incentives for Exploration versus Exploitation

Scientists do not work in a vacuum but are part of an institutional system that provides various (dis)incentives for entering new fields. We look at two important science system components: career progression and research funding. Within the risk-reward framework, how the system rewards big successes and how it rewards, allows or even punishes failures, will be critical for the trade-off scientists face when contemplating whether to enter new fields. Holmstrom (1989) and Manso (2011) have argued a.o. that incentive schemes that motivate innovation must exhibit tolerance for failures. Institutions making promotion and funding decisions not necessarily honor that principle.

Career Incentives

An important and highly visible feature of the academic organizational context is the tenure and hierarchical rank system, in which scientists move up from assistant professor in a tenure track position to full professor. Career concerns constitute an important driver of scientists' behaviour (Dewatripont *et al.*, 1999). As argued above, while it may hold less promise for novel contributions, exploitation avoids the uncertainty that comes with new field entry. Hence, it reduces the risk of lack of output. When tenure and promotion decisions are heavily dependent on a well-filled track record, they bias against risk taking and failure, reducing the incentives for new field exploration. Junior scientists may suppress their inclination to explore other scientific fields before they achieve tenure and have progressed along the promotion ladder. They may especially avoid the riskier longer jumps. After securing tenure, and particularly once the full professor rank has been attained, promotion-related concerns no longer are at play, reducing the need to play-it-safe. Full professors can afford easily to venture into novel research directions, including more risky distant directions, despite the higher risks. Franzoni & Rossi-Lamastra (2017) indeed find evidence that tenured scientists are more likely to diversify their research.

Funding

As access to research funds is an important condition to be able to do research in most fields, successfully obtaining funding becomes a prominent incentive for researchers. How research funders value rewards and risks will thus affect research orientation decisions. More specifically, when research grants are awarded on the basis of performance without due recognition of risk taking, i.e. tolerating failure, they bias against exploration in favour of exploitation. Geuna (2001) already observed that many European universities have adopted a contractual-oriented approach to funding with an emphasis on research output, which has now become the dominant paradigm for allocating research money (Jonkers & Zacharewicz, 2016). Funding agencies are also being accused of becoming increasingly less tolerant of risk (Hicks, 2012; Alberts *et al.*, 2014). Boudreau *et al.* (2016) indeed found that grant proposals that deviate strongly from evaluators' expertise have lower chances of being positively evaluated. This trend of decreasing risk tolerance in research funding is associated with an increasing use of bibliometric measures. A metrics-based approach of research evaluation and funding is heavily debated, claiming that it encourages "me-too science" (e.g. Alberts, 2013). Consistent with this view, Wang *et al.* (2017) show that traditional bibliometric indicators such as journal impact factors and short-term citation counts bias against novel research, which requires a longer time window to realize its full impact than the one typically used to calculate citation windows. Reliance on this type of indicators may thus fuel risk aversion by funding agencies, selecting relatively safe projects that exploit existing knowledge at the expense of those that explore untested approaches (Walsh, 2013; Stephan *et al.*, 2017). In sum, the scientific funding system discourages scientists' entry into new fields, the more so when it relies on assessing past performance with standard bibliometric indicators

Conversely, there are also reasons to expect a positive relation between research funding and new field entry. Funding provides researchers with the resources to pursue curiosity-driven research avenues, creating the conditions for taking more risk and branching into new, more distant, research directions (Hollingsworth, 2004; Heinze *et al.*, 2009). This is the case particularly if the funding is substantial and covers a period of time long enough to provide the researcher with the protective space to tackle ambitious questions and if it allows sufficient degrees of freedom to reorient the research when failures occur.

Empirical evidence on the impact of funding on research direction is scarce. Most prior studies on funding are concerned with its effect on productivity and collaboration (Bozeman & Gaughan, 2007; Goldfarb, 2008; Defazio, Lockett, & Wright, 2009; Auranen & Nieminen, 2010; Jacob & Lefgren, 2011; Grimpe, 2012; Hottenrott & Lawson, 2013; Kelchtermans & Veugelers, 2011, 2013; Whalley & Hicks, 2014; Gush *et al.*, 2015). In one of the few studies on changes in research direction following the acquisition of grants, Azoulay *et al.* (2011) showed that the HHMI program stimulates biomedical scientists to explore new lines of investigation within biomedical research, an effect which can be attributed to the characteristics of the HHMI funding scheme, notably its policy of supporting young, promising talents with longer term and substantial funding, providing them with the protective space to develop their innovative ideas.

2.2.3. Discipline-specific demand and supply shocks

The existing empirical literature studying researchers' field reorientation decisions focuses mostly on well-defined shocks within specific fields as drivers for such decisions. For the field of mathematics, Borjas & Doran (2015a) study how the Russian immigration after the collapse of the USSR affected US researchers to reorient their research. They found that this shock caused US mathematicians to move away from fields that received large numbers of Soviet émigrés. Diminishing returns due to increased competition, rather than beneficial human capital spillovers, seem to have dominated their reorientation decisions. In Borjas & Doran (2015b) they look at how awards may affect a scientist's tendency to enter a new field, by studying the case of the Fields award in mathematics. Resembling the response to the receipt of sizeable research funding (cf supra), a highly regarded award "buys the scientist freedom" i.e. it shifts the risk balance as the scientist can afford to be less productive in post-award excursions. Comparing

Fields medalists with similarly brilliant contenders, they find that the winners' productivity declined after the award, because the medalists were found to begin studying unfamiliar topics at the expense of decreased output. In another study of an exogenous supply shock driving researchers' reorientation decision, Furman & Teodoridis (2020) study the impact of the arrival of a new research technology. They look at the unexpected arrival of a new automated motion-sensing research technology, as a consequence of the hacking of the Microsoft Kinect system. Their analysis shows how this shock induced researchers to pursue more diverse ideas, distant from their original research trajectories. Interestingly, they found that the effect did not only hold for researchers active within the motion-sensing research field, but also - and even more so - for researchers in other areas.

Another example of a supply-side shock influencing scientists' research orientation decisions is provided in Murray *et al.* (2016). They look at a shock in the cost of doing research, more particularly at the case of a lowering of the access cost to use genetically engineered mice, finding that it attracts new researchers to the community investigating mouse genetics.

Next to supply shocks, also sudden changes in the demand for research in certain fields can be drivers of researchers' reorientation. Bhattacharyaa & Packalen (2011) investigate whether researchers in life sciences respond to 'demand', and find that switches in scientists' orientations are being driven by changes in disease prevalence, controlling for research opportunities in fields.

3 Data & Methodology

In our empirical analysis, we can use a rich panel dataset from a leading European, comprehensive research university (KU Leuven, Belgium). KU Leuven is a research-intensive comprehensive university, which is consistently ranked among the top 100 universities in the world in university rankings (ARWU/Shanghai and THE).[3] As a research-intensive university, its recruitment, promotion and research funds allocation policy follows international practices geared towards research excellence. The panel (1992-2001) has 734 professors in biomedical and exact sciences for whom we have information on their individual characteristics, scientific discipline, publications, career progression and funding.[4] This dataset allows us to look at a broad set of factors influencing new field entry decisions of researchers across different scientific disciplines.

3.1. Dependent Variable: New Field Entry

To study the decisions of researchers to enter into new scientific fields, we construct an indicator with the value 1 for every year a scientist publishes in a field in which she hasn't published before, provided she publishes at least once more in that field in subsequent years.[5] The latter condition implies that we exclude one-time only excursions into other fields i.e. we require a minimum level of 'stickiness' for new field entries.[6] Our measure relies on evidence of new field entry through a publication, which excludes new field entry that does not result in a publication record.[7]

To determine new field entries, we use publication data from the Web of Science (WoS) dating back to 1971, which allows tracing the fields in which a scientist has published over the years. The taxonomy of scientific fields we use is a somewhat simplified version of the 255 "subject categories" used for classifying each WoS publication.[8] Our classification of 194 fields covers all domains of science, including social sciences and arts & humanities. The subset of 129 scientific fields covering the domains of science and engineering account for the bulk of publication activity in our sample. The yearly average number of scientific fields a researcher publishes in in our sample is 4.4, which is very similar to numbers reported in other studies. For example, Franzoni & Rossi-Lamastra (2017), relying on semantic clustering of article titles and abstracts, find that scientists publish on average in 4.1 different "research themes" per year.

In order to test for any differences depending on how distant the newly entered field is (conditional on entry), we consider two additional dependent variables. First, we consider the *distance* of a scientist's main field (i.e. in which she has the most publications to date) to the newly entered field(s). To construct this variable, we first calculate the yearly cosine similarity matrix between fields based on cross-citations between journals, using their respective WoS subject categories (Zhang *et al.*, 2010; Stephan *et al.*, 2017).[9] The distance between each pair of fields is then defined as 1 minus the similarity score. Closer fields will have a smaller distance score, fields that don't cite each other have a distance equal to one. The dependent variable *distance* is calculated as the distance between the researcher's main field and the newly entered field, again conditional on the entry being 'sticky'.[10] We further leverage the distance measure by analysing the probability of 'long jumps' relative to one's main field, defined as those field entries that are more than 1 standard deviation away from one's main field, taking into account all new field entries in the sample in the same year. Similarly, we identify 'short jumps' as those field entries that are less than 1 standard deviation away from one's main field. We estimate the distance and short/long jump models conditional on entry, in order to separate — in line with our theoretical framework — *(i)* the decision whether or not to enter a new field from *(ii)* the distance of new field entry, as the latter measures how much risk one takes when entering a new field.[11]

With our detailed scientific field classification, it is quite common in our sample for researchers to enter a new field: 486 out of the 734 researchers in our sample, i.e. 66%, enter a new scientific field at least once in the considered period (1992-2001) and at the scientist*year level about 22% of observations correspond to new field entries (see the 'Entry in new field' column, bottom row, in Table 1). Conditional on entering a new field, scientists enter on average 1.54 new fields per year, Most of these new field entries don't seem to be drastic reorientations: only a small share of the scientists in the sample (9.2%) switches main field (i.e. the one accounting for the majority of all their publications) during the observation window.

Conditional on entering a new field, the mean distance of these new field entries is 0.92.[12] On average about 13% of new entries are 'large jumps', but there is a large standard deviation of 0.34 (Table 2). Conditional on making at least one new field entry, a researcher makes on average 0.28 long jumps, and 21.0% of scientists who enter a new field make a long jump. 16% of new field entries are nearby entries or 'short jumps'. A researcher makes on average 0.34 short jumps, conditional on making at least one new field entry and 27.2% of scientists who enter a new field make a short jump. When looking at the within-descriptives for these groups, we find that the average distance of a long jump is 0.99 (i.e. close to the maximum of 1) with a standard deviation of only 0.009 while for a short jump it is 0.75 (s.d.=0.063) and for the remainder category ('average jumps') it is 0.92 (s.d.=0.051). Once the threshold for a long jump has been crossed, there isn't much heterogeneity i.e. long jumps are a quite distinct category.

Table 1 further shows that new field entry patterns differ substantially across disciplines. Scientists in *Biology* and *General & Internal Medicine* are most likely to enter a new field, while *Mathematics* displays the lowest probability of new field entry but – conditional on entry – the latter shows a higher prevalence of long jumps.[13]

[Table 1 about here]

As the number of new field entries for a researcher in a given year is smaller than two, we don't estimate a count model but rather focus on the probability of new field entry in a given year, and conditional on entry, we estimate the distance from one's main field and the probability of making a long (or short) jump.

The estimation sample for the probability of new field entries has 734 researchers observed yearly from 1992 till 2001, leaving a sample of 4,706 observations. The estimation sample for distance and long jumps includes only those observations in which a new field entry is observed, leaving 1,022 observations (N=486).

3.2. Covariates

In this section we operationalize the framework discussed in section 2. We report descriptive statistics for the different covariates used in our empirical analysis in Table 2.

[Table 2 about here]

Talent and Productivity

As a proxy for talent, we include a pre-sample measure of a scientist's research quality by looking at his average number of citations per publication before 1992 relative to the average number of citations per publication in the scientist's main discipline in the entire Web of Science in 1991. The average value is 0.97, which indicates the sample is on average comparable to the population of researchers. Talent heterogeneity is substantial, with a citation rate almost double the world average (1.88) for the scientists above the world average, and less than half (0.43) for those below. Researchers whose pre-sample publication quality is above, respectively below, the world average do not seem to differ materially in terms of their new field entry decisions (see the right panel of Table 2).

In addition, we compare the quantity of each scientist's publication output in every year to her colleagues in the university who are active in the same main discipline.[14] Using k-means clustering, we classify researchers each year into one of three mutually exclusive clusters (low, medium and high).[15] About 13% of our sample observations are in the high productivity cluster. Membership of a high productivity cluster is associated with a higher likelihood of new field entry.

Experience with New Field Entry

To capture a scientist's experience with new field entry, we include the stock of unique scientific fields in which the researcher has been active up to that point. Besides a scientist's inclination to diversify into new areas, this variable also captures the potential for synergies of a new field entry with the scientist's existing scope of expertise or rather the risk for diminishing returns to new field entry. Sample researchers are on average active in 9 fields.

Table 2 shows no apparent link between the size of a researcher's stock of fields relative to peers in her discipline and the propensity to enter additional fields, although scientists with an above-median stock of fields seem more likely to make long jumps. To further characterize the breadth and spread of experience across fields, we calculate the Herfindahl concentration index, with 1 denoting full specialization in a single field. Splitting the sample by the median shows no major differences in field entry behaviour, apart from a slightly higher occurrence of long jumps for the less specialized scientists.

Career Age

The average seniority (i.e. years since starting employment at the university), is about 10 years. Consistent with their smaller publication track record, junior faculty are more prone to enter new fields, but when they do so, they are more likely to make short jumps while long jumps are more prevalent among the more senior scientists.

Teams

We measure the size of a scientist's team by counting all unique co-authors of a researcher on all his publications in a given year, both including internal and external (i.e. non-KU Leuven) co-authors. The average annual team size is 13.9 with a large standard deviation (20.0). These relatively high numbers are driven by lab-based disciplines like *Biosciences* (26.6) and *Clinical and Experimental Medicine* (22.9).[16] To assess whether especially co-authors from other disciplines may promote new field entry, we include the percentage of co-authors that share their main discipline with the focal researcher. On average, 41% of co-authors are active in the same main field as the focal researcher. We also include the number of PhD students that graduate in a given year and who were supervised by the focal scientist. Doctoral students may act as 'scouts' for the supervisor in terms of exploring novel ideas that break into other fields.

Table 2 shows that researchers with above-median sized teams and numbers of PhD students are more likely to enter new fields. Supervision of more PhD students is positively associated with making long jumps as well as, somewhat surprisingly, smaller teams and teams with higher shares of co-authors in the same main discipline.

Career Incentives

Incentives and risks stemming from promotion perspectives are captured by a scientist's rank, for which we distinguish between assistant professor, associate professor, professor and full professor. Note that for the university that we study, tenure coincides with the promotion to associate professor, so we do not include a separate control for tenure. Tenure and promotion decisions are based on an assessment by faculty committees of the performance of the candidate on research excellence, teaching and internal and external services.

Slightly over one quarter of researcher*year observations (28%) fall into the assistant professor group. The probability of new field entry is fairly constant across ranks. Conditional on new field entry, the appetite for making long jumps increases with rank.

Funding

The funding scheme we observe is the Research Fund of the KU Leuven (known as "BOF"), which represents a quarter of all research funding available to researchers at the university.[17] The fund is financed by the Flemish government and its overall size is determined by the share of the KU Leuven in the output of scientific publications and citations in the Flemish region. The university manages the fund autonomously through a competitive process involving international peer reviewers. In the observation period, the funding scheme contained two major funding programs, open to applicants across the whole university and accepting proposals that are entirely driven by the curiosity of the scientist. In other words, the studied funding programs are 'neutral' with respect to new field entry as they do not a priori steer researchers in a certain direction. Of the pre-screened applications that are allowed to submit a full proposal less than 50% obtains funding. Grants vary between 475,000 Euro for individual researchers or small teams up till 1,625,000 Euro for larger teams, with funding periods from four to five years. As such, this funding scheme, with its bottom-up nature, large budgets and long time windows should allow for a protective space to enter new research areas.

In our sample, 42% of researchers receive a BOF grant at least once in 1992-2001 (Table 2). Funding is positively correlated with entry into new fields, with 1 out 4 researchers who have received a BOF grant entering new fields, compared to 1 out of 5 for non-funded researchers. Conditional on entering new fields, funded researchers are also more likely to make long jumps.

Table 3 compares funded and non-funded researchers and shows that they differ systematically in terms of their characteristics. Funded scholars have a significantly higher pre-1992 quality of research, with a normalized citation rate of 1.23 versus 0.86. They are also twice as likely to be in the high productivity cluster (22% versus only 11% of unfunded researchers). The annual co-author network of funded researchers is on average more than twice the size compared to colleagues without BOF grants, and with a higher share of co-authors from the same main discipline. Funded researchers are more often found among the higher (tenured) ranks and more senior scientists. Receiving a BOF grant also varies substantially by discipline (not shown in the table). For example, bioscientists have the highest funding rate at 54% i.e. more than half is funded at least once in the sample period, compared to only 16% of mathematicians.

[Table 3 about here]

As funded scientists differ systematically from unfunded scientists, not only in terms of new field entry, but also in terms of the covariates, we need to take selection effects into account. Funding might go primarily to those scientists who are more inclined to enter new fields anyway. Furthermore, factors like talent, track record and career stage are not only direct drivers of new field entry (cf supra), but also increase the likelihood of receiving funding. We estimate a 2-stage linear probability model in which we account for the probability to receive funding in the first stage, using an instrumental variable approach to better identify the true effect of funding on new field entry.[18] As an individual-specific instrument, we use the number of university colleagues who share the same discipline with the focal scientist and who obtained funding in the same scheme (BOF) in the corresponding year. The rationale of the instrument, which is in the spirit of the literature employing instruments based on 'local conditions' (e.g. Berger *et al.*, 2005), is that funding of colleagues relates to funding of the focal scientist, either positively because a given discipline may be particularly resource-demanding or is seen as a strategic priority by funding panels and therefore attract more funding, or negatively because grant budgets are fixed implying a zero-sum game. At the same time, grants given to colleagues in the same discipline are not expected to affect the research direction of the focal scientist. The validity of the instrument would be hindered if funded scientists used their funds in such a way that it would change the research agenda of their colleagues in the same discipline. This risk is mitigated as the grants in this scheme are mostly used to hire PhD students and postdocs, who are less likely to affect the research orientation of the focal scientist. Furthermore, the instrument includes all funded colleagues in the same main discipline, which is much broader than direct colleagues in the same department, mitigating concerns that the instrument captures funding streams that also affect the focal scientist's propensity (not) to enter a new field.[19] We report formal validity tests of the instrument as well as a robustness check in Appendix A1.

We lag the covariates in both the funding and the field entry equation by one year to mitigate endogeneity arising from simultaneity. For the funding variables, we use a two-year lag to account for the time it takes before a grant translates into published results. We have tested the robustness of our results using different lags of funding and obtain very similar results for longer lags.[20]

3.3. Control variables

Given that some fields may serve more easily as a steppingstone to enter additional fields (see Table 1), we control for a scientist's past experience by including dummy variables for each **scientific field** in which the scientist published in the past. The highly granular classification of 193 fields provides an individualized 'map' for each scientist. In a robustness check, we control for a scientist's main discipline (using the more aggregate classification discussed in section 3.1), defined as the one in which she has the majority of her publications, based on her lifetime publication record. This more general classification saves on degrees of freedom at the cost of not fully capturing all field specificities. It also restricts the analysis of distance, as it ignores the distance between subfields within the aggregate discipline, which is where most of the new field entry takes place. The findings for this alternative specification are robust to the main results using the more granular field fixed effects, as we report in section 4 below.

To capture field-specific demand and supply shocks requires combining field and year fixed effects, which would consume considerable degrees of freedom. In the robustness analysis, we will report on (more aggregate) discipline-specific time effects.

In addition, we control for a number of individual characteristics. Although male and female scientists seem equally likely to enter new fields (Table 2), we control for **gender** in the multivariate analysis. We also include an indicator for whether a professor was part of the **post-1992 entry cohort** to capture a spike in hiring at the university in that year. This variable also corrects for the fact that we have fewer observations to calculate stocks for the most recent entrants in the sample. We control for whether a scientist is working **full-time** at the university and his **teaching load** (measured as the average number of teaching hours per week). A correlation table including all covariates is included in Table 6 in Appendix A3.

4 Results

Table 4 shows our econometric results for the 3 dependent variables i.e. new field entry (0/1) and, conditional on entry, the distance to the newly entered field and the occurrence of 'long jumps'. In general, significance levels for results on distance are weak, hampered by a lack of sufficient heterogeneity, as the low standard deviation reported in Table 1 shows. For results on distance, we primarily rely on the results for the 'long jumps': although they are less frequent than field entries as such, there is more heterogeneity to be exploited (see Table 1). Results for the 'short jumps' are mentioned in the robustness analysis. We first discuss the results in the single equation entry models before we turn to the discussion of the 2-stage linear probability models, in which we account for the probability to receive funding in the first stage.

[Table 4 about here]

Talent and productivity. While talent, as measured by the ratio of one's citations per publication relative to the world average, does not seem to be significantly more associated with new field entry (see model 1 in Table 4), one's past productivity record, as measured by belonging to the medium or high productivity cohort in the previous year, shows a significant positive relation with the likelihood of new field entry.[21] The role of quality and productivity differs once we consider how far the newly entered field is from the researcher's current expertise, conditional on new field entry (see models 2 & 3). The researcher's quality is positively associated with distance, in particular with the probability of making long jumps (0.027, significant at the 5% level). For the researchers who make new field entries, the more productive ones don't appear more likely to enter more distant fields.

Teams. Researchers with a higher number of co-authors are more likely to enter new fields and the same holds for researchers who supervise more doctoral students (model 1). However, the number and relatedness of one's co-authors do not have an effect beyond the entry decision as such. Conversely, the effect of PhD students persists: supervising more doctoral students is not only associated with new field entry but also with entering more distant fields (model 2). This suggests that PhD students perform a 'scouting role' for the scientist, exploring avenues that are more remote from her main research. This distance effect is however limited: in the long jump model, PhD students no longer show a significant positive effect.

Experience with new field entry. A researcher's stock of new field entries is negatively associated with additional entry (model 1), supporting diminishing returns to new field entry and diversification. Perhaps even more surprising, is that, conditional on new field entry, experience with new field entry does not help to make 'long jumps', on the contrary: the larger the stock of new field entries in the past, the less likely 'long jumps' are being made (model 3), suggesting a diminishing tendency to new field entry. However, there is no significant evidence that, conditional on a researcher's stock of fields, more specialization - as measured by higher values for the Herfindahl index - explains whether she would enter additional fields.[22] Also in the distance and 'long jump' models we do not find a significant effect for specialization within a given scope of fields covered.

Career age. Career age (seniority) is negatively associated with new field entry: older researchers are significantly less likely to enter new fields. Conditional on entry, career age is not associated with the distance of entries nor the probability to make 'long jumps' i.e. the appetite for risk of senior researchers seems confined to new field entry as such without providing a further nudge to venture into more distant fields (models 2 & 3).

Career incentives. All else equal (including seniority), higher ranks are increasingly associated with new field entry compared to assistant professors (model 1). So incentives for entering new fields operate throughout the career hierarchy, with not only tenure (which coincides with the associate professor rank in the data) but also having reached higher steps on the career ladder providing an additional push to take on the risk of entering a new field. Full professors are the most likely to enter new fields.[23] However, conditional on entry, rank is no longer associated with the distance of entries or the probability to make 'long jumps' (models 2 & 3), a result similar to what was found for seniority.

Funding. Looking first at the single equation results (models 1-3), shows that funding positively affects the likelihood of new field entry, but not significantly. Conditional on new field entry, although affecting distance negatively (but not significantly), funding does show a significant positive association with making long jumps. Accounting for its endogeneity, with the instrument being highly significant (see models 4a-5a-6a)[24], funding shows a negative but again insignificant effect on new field entry (model 4b).

Conditional on entry, funding shows a negative association with the distance of newly entered fields (β=-0.116, significant at the 10% level). Interestingly, funding is positively associated with long jumps with a large and significant coefficient (β=0.499, significant at the 5% level). In sum, the pattern that seems to emerge is that funding does not make scientists significantly more adventurous: if anything, the evidence, although inconclusive, seems to go more in the direction of making them less likely to go into new fields or enter more distant fields. But the results suggest a dual role in the sense that funding is positively associated with entering very distant fields.

The 2-stage estimation procedure allows comparing the effects of covariates, disentangling direct effects on entry from effects running indirectly through funding, taking into account their effect on likelihood to get funded. The 1^{st} stage results in models 4a-5a-6a show which covariates are significant for selection into funding (model 4a for the total sample, relevant for the 2^{nd} stage results on new field entry in model 4b; models 5a and 6a for the subsample of new field entry observations, relevant for the 2^{nd} stage results on distance and long jumps in models 5b and 6b). These 1^{st} stage results show, not unexpectedly, the importance of talent and productivity for the selection into funding (although only significantly in the full sample), but also a positive association with the career stage and team size variables and, interestingly, a negative association with the stock of new field entries (especially for new field entry as such, model 4a). All else equal, these results seem to suggest that funding selects on past track record but biases against scientists who have a larger stock of entered scientific fields.

In general, taking into account selection into funding, and with funding not significantly affecting new field entry (if properly instrumented), the 2^{nd} stage results for the covariates do not change sign or significance for new field entry, suggesting that the total effects of the covariates are mostly direct. But the positive career rank effects are much stronger in the 2nd stage results than in the single equation model, indicating that for these covariates the direct positive association with new field entry is dampened by their indirect effect through funding.

Taking into account the selection into funding is more important for the covariates' effects on 'long jumps'. Comparing the single equation and 2nd stage 2SLS results shows that the overall negative significant stock of fields effects from the single-equation analysis disappear in the 2nd stage results, indicating it is entirely due to the indirect effect running through funding. In other words, the overall lower inclination of profiles with a larger stock of fields to enter into new fields is an indirect effect from their lower likelihood to receive funding, all else equal.

Controls. The field fixed effects are important controls, stressing that the scientific context matters. Also, their granularity matters: in particular the result for stock of fields is sensitive to using fixed effects at the level of the fine-grained set of scientific fields rather compared to the 12 broader disciplines.[25] The robustness section includes a more elaborate discussion of field-specific shocks.

Other controls, like gender or teaching load, are not significant. A noteworthy observation is that the cohort that entered the university in or after 1992 on average stays closer to fields they are familiar with (as opposed to the older cohort) but they are more likely to make long jumps than their colleagues who have been at the university longer. Finally, when adding individual fixed effects as controls, many control variables become insignificant, as there is no or too little variance across time per researcher.

5 Robustness Checks

As highlighted in our theoretical framework, time trends and supply or demand **shocks** may shift a scientist's risk-reward trade-off. Therefore, we performed supplementary analyses to assess the robustness of our findings with respect to field and time trends. First, while the models in Table 4 already account for scientists' seniority and a 1992 dummy to capture a spike in hiring at KU Leuven, additionally controlling for *general time trends* through year fixed effects in both stages of the 2SLS models[26] does not materially affect the results (available from the authors). We limit the discussion here to the most noteworthy differences to the results in Table 4. The positive relation of researcher quality with long jumps is not confirmed. Also the incentive effect of research funding for making long jumps no longer is statistically significant.

While adding year fixed effects helps to control for common trends across disciplines, it does not account for *field specific supply or demand shocks* . Estimating models with fixed effects for the combination of 194 fields and 10 years is not feasible with the available data. As an alternative, we estimate a specification with fixed effects at the intersection of years and scientists' main disciplines (see Table 1 for a list of disciplines), i.e., at a more aggregate level than scientific fields. Although this approach can only partially capture field-specific events like breakthrough discoveries or a sudden inflow of other scientists (e.g. Borjas & Doran, 2015a), the results are informative (Table 5 in Appendix A2). Focusing on the 2SLS models and a comparison with the main findings in Table 4, we find that the positive and statistically significant association between quality and long jumps remains intact. The increasing propensity of higher ranks to enter new fields is no longer statistically significant. Again, the evidence for a positive relation of funding with long jumps disappears, suggesting that this relationship is primarily driven by idiosyncratic shocks such as the emergence of new scientific opportunities or that such positive shocks attract funding rather than long jumps being linked to funding per se.[27]

We also investigate alternative specifications for distance. In case a scientist has several new field entries in a single year[28], in the main analysis we measured the distance using the field farthest away. Results are robust when we use the average or even minimum distance as an alternative outcome measure in such cases. We also looked for which factors would be more associated with 'short jumps' compared to the rest. The (2nd stage) **regression analysis for short jumps** shows few statistically significant relations with our covariates. Conditional on entry into new fields, talent & productivity, career age, teams, career incentives and funding do not significantly affect the likelihood of short jumps. Only the stock of fields is positively related to short jumps while the Herfindahl shows a negative relation. Although both these effects are only weakly statistically significant, they suggest that scientists who entered more fields in the past but who remain more specialized rather than diversified, are more likely - when they enter new fields - to enter close-by fields.

Additional robustness checks (available from the authors) confirming the main results include the following. We controlled for productivity using lagged publication counts rather than the discrete categories. Omitting field fixed effects altogether has little effect on the results for quality, productivity, rank and funding. Finally, we did not find evidence of a non-linear relationship between (the distance of) new field entries and a scientist's stock of fields.

6 Conclusions

Our analysis contributes to the small but growing empirical literature studying scientists' decisions to enter research fields which are new to them. Exploring new research fields may be a more risky avenue compared to further exploiting familiar fields, but it holds the potential of bigger breakthroughs, which often require novel approaches. When deciding whether to enter a new field, researchers thus face a difficult risk-reward trade-off, the more so, the more distant the new fields are. We build on the economics of science literature to develop a framework that covers a wide set of factors influencing the risk-reward trade-off when deciding on new field entry, incorporating individual capacities and preferences as well as incentives stemming from career prospects and the competitive funding system.

In our empirical analysis, we use a rich dataset from a leading European research university (KU Leuven, Belgium). The set of determinants we consider ranges from talent, past track record, experience with new field entry and co-authorship and PhD teams, over tenure and rank, to research funding, while also taking into account career age, gender and scientific field differences.

The results show that entry into new fields is quite common. About two thirds of researchers in our sample go into a new field at least once in a 10-year period. Conditional on making at least one new field entry, a researcher makes on average 0.28 long jumps and 0.34 short jumps. Both the likelihood of entering a new field and, conditional on new field entry, the likelihood of making a 'long jump' vary substantially within our sample.

The econometric analysis first of all shows that individual competences and preferences matter, as well as the incentives associated with career progress and the competitive funding system. But a different set of factors matters for new field entry in general versus how far the newly entered field is from the researcher's main field. This contrast suggests the importance of considering distance when looking at the risk-reward trade-off faced by researchers when deciding to enter new fields.

While for new field entry a scientist's past productivity record plays a role rather than research quality, it is quality and not productivity that makes long jumps significantly more likely. Researchers with more experience in new field entry are significantly less likely to make additional new field entries and are especially less likely to make long jumps. This is consistent with diminishing returns to new field entry, rather than experience effects. Researchers with larger co-author and PhD student networks are more likely to enter new fields, but it does not help to make long jumps. The inclination to enter new fields diminishes with career age, but conditional on entering new fields, does not affect the distance of the jump.

Second, the analysis shows that context matters. The career stage of a scientist has implications for how risk and rewards are traded off, with full professors most likely to enter new fields, although not when it comes to making 'long jumps'. Finally, although the descriptive statistics showed a positive correlation between funding and new field entry, the econometric analysis correcting for other drivers which are associated with higher probability of funding (like productivity), and which accounts for selection into funding, no longer shows that funding serves as a mechanism to venture into new territory. However, conditional on entry, funded researchers seem more likely to make long jumps. Our results also show that the negative selection into funding of researchers with experience in new fields, fully accounts for their lower likelihood to make long jumps. Once correcting for this selection, profiles with larger stocks of new field entry are no longer significantly less likely to make 'long jumps'. In sum, pivotal changes in working context and when they occur in scientists' careers have an effect on how their research agendas develop and whether a given scientist, over time, spans many disciplinary boundaries or rather specializes deeply in a few fields.

At the level of the aggregate science system, individual decisions accumulate and result in a particular composition of specialists and generalists, and varying degrees thereof in various fields, with ramifications for the productivity of the science system.

Many of the contextual factors we identified can be influenced through policy. Perhaps the primary policy lever for influencing research orientation that emerges from our analysis is funding. To the extent that society wants scientists to come up with novel breakthrough findings and that these are primarily found by spanning fields, our results suggest that the funding system in our sample seems to perform poorly, on average, in terms of selecting researchers with a track record of entering new fields. It also performs poorly in terms of inducing researchers to enter new fields, although there is some tentative evidence that it fosters long jumps. This result ties into the debate on science funding, with prominent scientists calling into question the design of funding mechanisms that reflect funders' risk-aversion and bias against more risky, novel scientific endeavours (Lane, 2009; Stephan *et al.*, 2021). To the extent that policy makers aim to increase diversity of research avenues pursued, or at least not bias against this, our findings show that grant schemes may not have the desired effect and may fail to provide scientists with the protective space to take more risk and venture into new territory.

Our analysis is limited by the available data, which suggests several avenues for further research. One obvious restriction is limited heterogeneity since we analyse one university and a limited set of research grants. Hence our results cannot be readily generalized across institutions and different funding schemes. As the work of Azoulay *et al.* (2011) suggests, the precise characteristics of the funding program matter in terms of shaping the precise incentives, with sizeable, long-term, people-focused funding creating a context that is more favourable to developing novel research questions.[29] Also the source of funding plays a role, with prior work showing that industry funding may shift academic research agendas from basic towards more applied science (Florida & Cohen, 1999; Gulbrandsen & Smeby, 2005; Glenna *et al.*, 2011; Banal-Estañol *et al.*, 2015). Therefore, an extension of our analysis could consider a wider variety of funding schemes in order to better understand the effects of design choices. Our results on 'long jumps' suggest that funding as such may not be a decisive lever to induce risk-taking by scientists, but further research is needed to better understand this relation. Considering the project level may be useful here: if a scientist manages a portfolio of grants or holds a big, long grant, then she may deliberately design some of the projects within her portfolio as 'high-risk, high-reward' (Azoulay *et al.*, 2011). Franzoni & Stephan (2021) also highlight various 'hedging strategies' that scientists may use to avoid that an entire grant is lost by taking on risky research that doesn't pan out. Another explanation for the 'long jump' effect of funding might be that certain scientists use funding differently than others, i.e. it may depend on past productivity or talent, such that scientists who have built a 'reputational buffer' can better afford duds, in line with Borjas & Duran's (2015b) work on the effects of scientific awards.

To assess the impact of individual entry decisions, follow-up work could consider more explicitly the relation between new field entry and 'long jumps' in terms of scientific performance, i.e. how productive scientists are after a new field entry, and how likely larger jumps are to produce publications that are highly novel ("new to the world") and outliers in terms of impact, or whether such entries result in failures. This would allow to further assess the risk reward trade-off when entering new fields and the factors shaping this trade-off.

References

Agrawal, A., McHale, J., & Oettl, A. (2013). *Collaboration, Stars, and the Changing Organization of Science: Evidence from Evolutionary Biology* (Working Paper No. 19653). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w19653

Alberts, B. (2013). Impact Factor Distortions. *Science*, *340*(6134), 787–787. http://doi.org/10.1126/science.1240319

Alberts, B., Kirschner, M. W., Tilghman, S., & Varmus, H. (2014). Rescuing US biomedical research from its systemic flaws. *Proceedings of the National Academy of Sciences*, *111*(16), 5773–5777.

Andrews, I., Stock, J. H., & Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, 11, 727-753.

Angrist, J., and S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In *The Rate and Direction of Inventive Activity* (R. Nelson, pp. 165–180). Princeton, NJ: Princeton University Press.

Audia, P. G., & Goncalo, J. A. (2007). Past success and creativity over time: A study of inventors in the hard disk drive industry. *Management Science*, *53*(1), 1–15.

Auranen, O., & Nieminen, M. (2010). University research funding and publication performance—An international comparison. *Research Policy*, *39*(6), 822–834. http://doi.org/10.1016/j.respol.2010.03.003

Ayoubi, C., Pezzoni, M., & Visentin, F. (2017). At the origins of learning: Absorbing knowledge flows from within the team. *Journal of Economic Behavior & Organization*, 134, 374-387.

Azoulay, P., Graff Zivin, J. S., & Manso, G. (2011). Incentives and creativity: evidence from the academic life sciences. *The RAND Journal of Economics*, *42*(3), 527–554.

Azoulay, P., Fons-Rosen, C., & Graff Zivin, J. S. (2019). Does science advance one funeral at a time?. *American Economic Review*, 109(8), 2889-2920.

Banal-Estañol, A., Jofre-Bonet, M., & Lawson, C. (2015). The double-edged sword of industry collaboration: Evidence from engineering academics in the UK. *Research Policy*, *44*(6), 1160–1175.

Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G., & Stein, J. C. (2005). Does function follow organizational form? Evidence from the lending practices of large and small banks. *Journal of Financial Economics*, 76(2), 237-269.

Bhattacharya, J., & Packalen, M. (2011). Opportunities and benefits as determinants of the direction of scientific research. *Journal of health economics*, 30(4), 603-615.

Borjas, G. J., & Doran, K. B. (2015a). Cognitive mobility: Labor market responses to supply shocks in the space of ideas. *Journal of Labor Economics*, 33(S1), S109-S145.

Borjas, G. J., & Doran, K. B. (2015b). Prizes and productivity how winning the fields medal affects scientific output. *Journal of human resources*, 50(3), 728-758.

Boudreau, K. J., Guinan, E. C., Lakhani, K. R., & Riedl, C. (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science*, 62(10), 2765-2783.

Bozeman, B., & Gaughan, M. (2007). Impacts of grants and contracts on academic researchers' interactions with industry. *Research Policy*, *36*(5), 694–707. http://doi.org/10.1016/j.respol.2007.01.007

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: methods and applications.* Cambridge university press.

Cole, S., & Cole, J. R. (1967). Scientific output and recognition: A study in the operation of the reward system in science. *American sociological review*, 377-390.

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

Defazio, D., Lockett, A., & Wright, M. (2009). Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU framework program. *Research Policy*, *38*(2), 293–305. http://doi.org/10.1016/j.respol.2008.11.008

Dewatripont, M., Jewitt, I., & Tirole, J. (1999). The economics of career concerns, part I: Comparing information structures. *The Review of Economic Studies*, 66(1), 183-198.

Florida, R., & Cohen, W. M. (1999). Engine or Infrastructure? The University Role in Economic Development. In Branscomb, L. M., Kodama, F., & Florida, R. (Eds.), *Industrializing Knowledge: Universityv Industry Linkages in Japan and the United States*. MIT Press, London.

Franzoni, C., & Rossi-Lamastra, C. (2017). Academic tenure, risk-taking and the diversification of scientific research. *Industry and Innovation*, 24(7), 691-712.

Franzoni, C., & Stephan, P. (2021). Uncertainty and Risk-Taking in Science: Meaning, Measurement and Management (No. w28562). National Bureau of Economic Research.

Furman, J. L., & Teodoridis, F. (2020). Automation, research technology, and researchers' trajectories: Evidence from computer science and electrical engineering. *Organization Science*, 31(2), 330-354.

Geuna, A. (2001). The changing rationale for European university research funding: are there negative unintended consequences? *Journal of Economic Issues*, *35*(3), 607–632.

Glenna, L. L., Welsh, R., Ervin, D., Lacy, W. B., & Biscotti, D. (2011). Commercial science, scientists' values, and university biotechnology research agendas. *Research Policy*, *40*(7), 957–968.

Goldfarb, B. (2008). The effect of government contracting on academic research: Does the source of funding affect scientific output? *Research Policy*, *37*(1), 41–58. http://doi.org/10.1016/j.respol.2007.07.011

Grimpe, C. (2012). Extramural research grants and scientists' funding strategies: Beggars cannot be choosers? *Research Policy*, *41*(8), 1448–1460.

Gulbrandsen, M., & Smeby, J.-C. (2005). Industry funding and university professors' research performance. *Research Policy*, *34*(6), 932–950. http://doi.org/10.1016/j.respol.2005.05.004

Gush, J., Jaffe, A. B., Larsen, V., & Laws, A. (2015). The Effect of Public Funding on Research Output: The New Zealand Marsden Fund. *NBER Working Paper*, (w21652).

Hackett, E. J. (2005). Essential tensions: Identity, control, and risk in research. *Social studies of science*, 35(5), 787-826.

Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, *41*(2), 251–261. http://doi.org/10.1016/j.respol.2011.09.007

Heinze, T., Shapira, P., Rogers, J. D., & Senker, J. M. (2009). Organizational and institutional influences on creativity in scientific research. *Research Policy, 38*(4), 610-623. doi:10.1016/j.respol.2009.01.014

Higham, K. W., Governale, M., Jaffe, A. B., & Zülicke, U. (2017). Unraveling the dynamics of growth, aging and inflation for citations to scientific articles from specific research fields. *Journal of Informetrics*, 11(4), 1190-1200.

Hollingsworth, R. (2004). Institutionalizing excellence in biomedical research: the case of Rockefeller University. In D. H. Stapleton (Ed.), *Creating a Tradition of Biomedical Research*. New York: Rockefeller University Press.

Holmstrom, B. (1989). Agency costs and innovation. *Journal of Economic Behavior & Organization*, 12(3), 305-327.

Hottenrott, H., & Lawson, C. (2013). Fishing for complementarities: Competitive research funding and research productivity. KU Leuven - Faculty of Economics and Business. Retrieved from https://lirias.KU Leuven.be/bitstream/123456789/429644/1/MSI_1315.pdf

Jacob, B. A., & Lefgren, L. (2011). The impact of research grant funding on scientific productivity. *Journal of Public Economics*, *95*(9), 1168–1177.

Jones, B. F. (2009). The burden of knowledge and the 'death of the renaissance man': Is innovation getting harder? *The Review of Economic Studies*, *76*(1), 283–317.

Jones, B. F., Wuchty, S., & Uzzi, B. (2008). Multi-university research teams: Shifting impact, geography, and stratification in science. *Science*, *322*(5905), 1259–1262.

Jonkers, K., & Zacharewicz, T. (2016). *Research Performance Based Funding Systems: a Comparative Assessment*. Seville, Spain: Institute for Prospective Technological Studies, Joint Research Centre. Luxembourg: Publications Office of the European Union. EUR, 27837.

Kelchtermans, S., & Veugelers, R. (2011). The great divide in scientific productivity: why the average scientist does not exist. *Industrial and Corporate Change*, *20*(1), 295–336. http://doi.org/10.1093/icc/dtq074

Kelchtermans, S., & Veugelers, R. (2013). Top research productivity and its persistence: Gender as a double-edged sword. *Review of Economics and Statistics*, *95*(1), 273–285.

Lane, J. (2009). Assessing the impact of science funding. *Science*, *324*(5932), 1273–1275.

Lee, Y.-N., Walsh, J. P., & Wang, J. (2015). Creativity in scientific teams: Unpacking novelty and impact. *Research Policy*, *44*(3), 684–697.

Levin, S. G., & Stephan, P. E. (1991). Research productivity over the life cycle: Evidence for academic scientists. *The American Economic Review*, 114–132.

Lewbel, A., Dong, Y., & Yang, T. T. (2012). Comparing features of convenient estimators for binary choice models with endogenous regressors. *Canadian Journal of Economics/Revue Canadienne D'économique*, *45*(3), 809–829.

Manso, G. (2011). Motivating innovation. *The Journal of Finance*, 66(5), 1823-1860.

Murray, F., Aghion, P., Dewatripont, M., Kolev, J., & Stern, S. (2016). Of mice and academics: Examining the effect of openness on innovation. *American Economic Journal: Economic Policy*, 8(1), 212-52.

Myers, K. (2020). The elasticity of science. *American Economic Journal: Applied Economics*, 12(4), 103-34.

Olea, J. L. M., & Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3), 358-369.

Simonton, D. K. (1999). *Origins of genius: Darwinian perspectives on creativity*. Oxford University Press.

Simonton, D. K. (2003). Scientific creativity as constrained stochastic behavior: the integration of product, person, and process perspectives. *Psychological Bulletin*, *129*(4), 475.

Siow, A. (1998). Tenure and other unusual personnel practices in academia. *Journal of Law, Economics, & Organization*, 152-173.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 557-586.

Stephan, P. E., & Levin, S. G. (1992). *Striking the mother lode in science: The importance of age, place, and time.* Oxford University Press, USA.

Stephan, P., Veugelers, R. & J. Wang, (2017). Blinkered by bibliometrics, *Nature*, 544, 411-412.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In *Identification and inference for econometric models: Essays in honor of Thomas*

Rothenberg, ed. D.W.K. Andrews and J.H. Stock, 80-108. Cambridge: Cambridge University Press.

Walsh, D. (2013). Not Safe for Funding: The N.S.F. and the Economics of Science. *The New Yorker.* Retrieved from http://www.newyorker.com/tech/elements/not-safe-for-funding-the-n-s-f-and-the-economics-of-science

Wang, J., Stephan, P. E. & Veugelers, R. (2017). Bias Against Novelty in Science: A Cautionary Tale for Users of Bibliometric Indicators. *Research Policy, 46, 1416-1436.*

Whalley, A., & Hicks, J. (2014). Spending wisely? How resources affect knowledge production in universities. *Economic Inquiry*, *52*(1), 35–55.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036–1039.

Zhang, L., Liu, X., Janssens, F., Liang, L., & Glänzel, W. (2010). Subject clustering analysis based on ISI category classification. *Journal of Informetrics*, 4(2), 185–193.

7 Tables

Table 1: New field entry (averages across researchers and years, by main discipline)

	Entry in new field (0/1)		*Number of new fields* $^{\Delta}$		*Distance to new field* $^{\dagger\Delta}$		*Long jumps (0/1)* $^{\Delta}$	
	mean	*s.d.*	*mean*	*s.d.*	*mean*	*s.d.*	*mean*	*s.d.*
Agriculture	0.25	0.43	1.65	0.99	0.94	0.07	0.26	0.44
Biosciences	0.23	0.42	1.52	0.80	0.89	0.09	0.02	0.15
Chemistry	0.16	0.37	1.46	0.80	0.95	0.08	0.25	0.43
Engineering	0.18	0.39	1.54	0.97	0.97	0.06	0.36	0.48
Geosciences	0.21	0.41	1.55	0.74	0.95	0.08	0.23	0.43
Mathematics	0.09	0.29	1.65	0.70	0.99	0.01	0.47	0.51
General & internal medicine	0.29	0.46	1.67	1.05	0.89	0.08	0.03	0.18
Non-internal medicine††	0.23	0.42	1.51	0.94	0.93	0.07	0.09	0.28
Neuroscience	0.11	0.32	1.10	0.32	0.92	0.07	0.00	0.00
Physics	0.17	0.37	1.21	0.60	0.96	0.05	0.32	0.47
Biomedical	0.18	0.39	1.66	0.97	0.89	0.07	0.03	0.19
Biology	0.29	0.45	1.53	0.81	0.89	0.08	0.07	0.26
All disciplines	0.22	0.41	1.54	0.94	0.92	0.08	0.13	0.34

† Distance between researcher's main field prior to year t and the newly entered field. Maximum distance in case multiple fields were entered in a single year.

†† E.g. dermatology, ophtalmology, otolaryngology, dentistry…

$^{\Delta}$ Conditional on new field entry

Table 2: Descriptive statistics (yearly averages across researchers, N=734)

			Entry (0/1)	*Nr of fields*	*Distance*	*Long jump (0/1)*	*Short jump (0/1)*
Variable	*mean*	*s.d.*	*mean*	*mean*	*mean*	*mean*	*mean*
Talent and productivity							
Quality$_{pre-1992}$ (citation rate vs world avg)	0.97	1.06					
≥ 1 (above world average)	1.88	1.23	0.23	1.52	0.92	0.14	0.13
< 1 (below world average)	0.43	0.33	0.21	1.55	0.92	0.13	0.18
Productivity cluster							
Low	0.52	0.38	0.16	1.49	0.92	0.13	0.18
Medium	0.34	0.31	0.26	1.55	0.92	0.13	0.15
High	0.13	0.26	0.31	1.60	0.92	0.14	0.15
Teams							
Co-authors (#)	13.92	20.00					
≥ median†	25.40	34.19	0.27	1.55	0.92	0.12	0.15
< median†	2.74	3.52	0.17	1.51	0.92	0.15	0.18
Co-authors in same discipline (%)	0.41	0.31					
≥ median†	0.56	0.41	0.22	1.52	0.92	0.15	0.16
< median†	0.13	0.18	0.22	1.58	0.92	0.10	0.17
PhD graduates (#)	0.36	0.53					
≥ median†	1.53	0.93	0.27	1.55	0.94	0.17	0.11
< median†	0.00	0.00	0.20	1.52	0.91	0.09	0.18
Past new field entry							
Stock of fields (#)	8.97	6.21					
≥ median†	12.60	6.61	0.22	1.53	0.93	0.15	0.11
< median†	4.99	3.26	0.21	1.55	0.91	0.10	0.23
Herfindahl index of stock of fields	0.33	0.23					
≥ median†	0.46	0.24	0.22	1.51	0.92	0.11	0.18
< median†	0.19	0.09	0.21	1.55	0.93	0.15	0.15
Career age (years)	10.33	7.67					
≥ median†	15.41	6.85	0.19	1.53	0.94	0.17	0.10
< median†	4.16	2.00	0.25	1.54	0.91	0.10	0.21
Career incentives: rank							
Assistant professor	0.28	0.40	0.22	1.55	0.91	0.09	0.23
Associate professor	0.25	0.34	0.21	1.52	0.91	0.10	0.20
Professor	0.21	0.34	0.21	1.52	0.93	0.16	0.10
Full professor	0.25	0.41	0.23	1.55	0.94	0.17	0.12
Funding (0/1)							
Funded*	0.42	0.49	0.26	1.56	0.92	0.15	0.17
Unfunded	0.58	0.49	0.20	1.53	0.92	0.12	0.16
Controls							
Teaching load (hours/week)	3.99	3.78					
≥ median†	7.36	3.96	0.21	1.55	0.93	0.14	0.12
< median†	1.56	1.31	0.23	1.53	0.91	0.12	0.20
Male (1/0)	0.91	0.29	0.22	1.54	0.92	0.13	0.16
Female (1/0)	0.09	0.29	0.23	1.49	0.91	0.16	0.21
Post-1992 entry cohort (1/0)	0.45	0.50	0.25	1.60	0.90	0.12	0.23
Full-time employment (1/0)	0.91	0.27	0.22	1.53	0.92	0.13	0.16
Total	-	-	0.22	1.54	0.92	0.13	0.16

* *Funded* = 1 if the researcher ever had a BOF grant in 1992-2001.

† The median is relative to other researchers in the same main discipline.

Table 3: Comparison of funded and unfunded researchers

	Funded	Unfunded		
Variable	*mean*	*mean*	*t-statistic*	*p-value*
Talent and productivity				
Quality$_{pre\text{-}1992}$ (citation rate vs world avg)	1.23	0.86	11.08	0.00
Productivity cluster				
Low	0.32	0.60	-18.75	0.00
Medium	0.47	0.29	11.93	0.00
High	0.22	0.11	10.07	0.00
Teams				
Co-authors (#)	23.14	11.15	14.20	0.00
Co-authors in same discipline (%)	51.47	37.46	10.98	0.00
PhD graduates (#)	0.71	0.31	14.11	0.00
Past new field entry				
Stock of fields (#)	11.53	8.38	15.40	0.00
Herfindahl index of stock of fields	0.28	0.34	-8.39	0.00
Career age (years)	11.60	9.77	7.58	0.00
Career incentives: rank				
Assistant professor (%)	6.29	25.37	-15.56	0.00
Associate professor (%)	23.32	28.43	-3.63	0.00
Professor (%)	26.82	22.87	2.91	0.00
Full professor (%)	43.58	23.33	14.28	0.00
Controls				
Teaching load (hours/week)	4.66	4.39	2.05	0.04
Male (1/0)	0.93	0.90	2.73	0.01
Post-1992 entry cohort (1/0)	0.23	0.39	-10.32	0.00
Full-time employment (1/0)	0.98	0.91	9.21	0.00

Table 4: Regression results for new field entry

				Two-stage models				
	Single equation models			**1st stage (funding)**		**2nd stage (entry)**		
	New field entry model 1	Distance model 2	Long jump model 3	New field entry model 4a	Distance / Long jump model 5a	New field entry model 4b	Distance model 5b	Long jump model 6b
Talent & Productivity								
Quality$_{pre\text{-}1992}$	-0.007	0.003	0.034***	0.024***	0.019	-0.002	0.005*	0.027**
	(0.006)	(0.002)	(0.011)	(0.009)	(0.017)	(0.007)	(0.003)	(0.013)
Medium productivity group$_{t\text{-}1}$ †	0.070***	0.000	-0.037	0.050***	0.033	0.080***	0.003	-0.048*
	(0.016)	(0.006)	(0.024)	(0.016)	(0.030)	(0.018)	(0.007)	(0.026)
High productivity group$_{t\text{-}1}$ †	0.070***	0.013	-0.027	0.031	0.001	0.074***	0.011	-0.021
	(0.024)	(0.009)	(0.036)	(0.027)	(0.048)	(0.025)	(0.010)	(0.037)
Teams								
Co-authors	0.001***	0.000	0.000	0.001*	0.001	0.001***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Co-authors in same discipline	-0.018	-0.005	-0.003	0.052**	0.014	-0.006	-0.003	-0.009
	(0.017)	(0.008)	(0.028)	(0.021)	(0.041)	(0.019)	(0.008)	(0.030)
PhD graduates	0.019**	0.005*	-0.001	0.044***	0.033**	0.029**	0.008**	-0.014
	(0.009)	(0.003)	(0.017)	(0.010)	(0.016)	(0.011)	(0.003)	(0.019)
Past new field entry								
Stock of fields	-0.191**	-0.025	-0.315***	-0.107**	-0.135*	-0.071*	-0.016	0.083
	(0.089)	(0.015)	(0.114)	(0.052)	(0.073)	(0.038)	(0.018)	(0.073)
Herfindahl index of fields	0.035	0.009	0.059	-0.06	-0.007	0.021	0.008	0.06
	(0.036)	(0.018)	(0.078)	(0.046)	(0.093)	(0.038)	(0.018)	(0.085)
Career age (years)	-0.008***	0.000	0.003	-0.003	-0.005	-0.009***	0.000	0.004
	(0.001)	(0.001)	(0.002)	(0.002)	(0.004)	(0.001)	(0.001)	(0.003)
Career incentives: rank ††								
Associate professor	0.029	-0.006	0.028	0.087***	0.150***	0.055**	0.014	-0.052
	(0.022)	(0.009)	(0.032)	(0.027)	(0.045)	(0.028)	(0.015)	(0.052)
Professor	0.060**	-0.002	0.033	0.177***	0.163***	0.106***	0.020	-0.057
	(0.027)	(0.011)	(0.036)	(0.038)	(0.061)	(0.040)	(0.018)	(0.064)
Full professor	0.109***	-0.006	0.054	0.260***	0.218***	0.173***	0.021	-0.054
	(0.037)	(0.016)	(0.053)	(0.049)	(0.081)	(0.057)	(0.022)	(0.082)
Funding								
Peers who acquired funding$_{t\text{-}2}$				0.420***	0.395***			
				(0.072)	(0.123)			
Funding$_{t\text{-}2}$	0.019	-0.006	0.061**			-0.21	-0.116*	0.499**
	(0.019)	(0.007)	(0.028)			(0.148)	(0.068)	(0.234)
Controls								
Teaching load	-0.001	0.001	0.002	-0.010***	-0.008	-0.003	0.000	0.006
	(0.002)	(0.001)	(0.005)	(0.003)	(0.006)	(0.003)	(0.001)	(0.006)
Male	-0.013	-0.001	-0.047	-0.005	0.008	-0.017	-0.001	-0.046
	(0.025)	(0.011)	(0.039)	(0.032)	(0.051)	(0.026)	(0.012)	(0.042)
Post-1992 entry cohort	0.036	-0.022**	0.083***	-0.041	-0.116**	0.032	-0.030***	0.118***
	(0.022)	(0.008)	(0.030)	(0.033)	(0.054)	(0.023)	(0.010)	(0.043)
Full-time position at the university	-0.014	-0.004	-0.016	0.090***	0.117**	0.004	0.008	-0.065
	(0.029)	(0.011)	(0.048)	(0.031)	(0.057)	(0.031)	(0.014)	(0.059)
Field fixed effects	*included*	*included*	*included*	*included*	*included*	*included*	*included*	*included*
Observations	4,706	1,022	1,022	4,706	1,022	4,706	1,022	1,022
Researchers	734	486	486	734	486	734	486	486

***, **, and * indicate significance levels of 1%, 5% and 10%, respectively. Robust standard errors (clustered by researcher) in parentheses. All models include an intercept.

† The reference category is the low productivity group. ††The reference category is assistant professor.

Models 5b and 6b build on the same first stage (model 5a).

8 Appendix A1: Instrument Validity

As discussed in section 3.2, we use as an instrument for funding the number of university colleagues who share the same discipline with the focal scientist and who obtained funding in the same scheme (BOF) in the corresponding year. In this appendix, we report on formal validity checks and a robustness check using an alternative definition of the instrument.

Validity Tests

The instrument accounts for significant variation in the funding of the focal scientist as the funding variable and the IV correlate 0.27. As a formal test of whether the instrument is weak, we checked the F-statistic for the Wald test of the instrument in the structural model. In model 4a (new field entry, Table 4) the robust F-statistic equals 34.47, and for model 5a and 6a it equals 10.21, which are all above the commonly used rule of thumb of 10 (Staiger & Stock, 1997). We also verified whether the respective F-statistics exceed the critical value of 16.38 for the just-identified case when allowing for at most a 10% distortion in the Wald test (Stock & Yogo, 2005). The robust F-statistic for model 4a (34.47) is greater than this critical value. In model 5a and 6a, which are run on a smaller sample - since here we look at the distance of new fields conditional on entry - we cannot rule out that the instrument is weak but based on the result for model 4a, this is not an unreasonable assumption.

An important caveat is that these tests (Stock & Yogo, 2005) are based on the restrictive assumption that errors are homoscedastic. However, Andrews, Stock & Sun (2019) find that in settings with a single endogenous variable the test for weak instruments proposed by Montiel-Olea & Pflueger (2013), which is based on a non-homoscedasticity robust F-statistic, actually equals the robust F-statistic reported above and also equals the Kleibergen-Paap F-statistic, which implies that these tests can be used with the Stock & Yogo (2005) critical values, as we did above.

While these tests are reassuring with respect to the relevance of the instrument, as a further verification of the robustness of our findings we employed methods that are robust to weak instruments. We re-estimated the model using the method in Andrews, Moreira & Stock (2007) but, due to multicollinearity issues, we opted for a specification with dummies for scientists' main disciplines rather than - as in the results reported in the paper - past fields. Despite using these more aggregate controls, which prevents a head-on comparison of the results, the results are informative with respect to the instrument. The first-stage F-statistic for funding in the new field entry model equals 46.6 ($p=0.000$) with a 2^{nd} stage parameter estimate of -0.34 and a corrected p-value (based on the likelihood ratio test in Andrews *et al.*, 2007) of 0.029, making the negative funding effect marginally significant in this model (it was insignificant in the main model 4b in Table 4). For the distance and long jump models (conditional on entry), we find an F-statistic for the instrument of 4.12 ($p=0.043$) and a funding coefficient $\beta=-0.25$ ($p=0.007$) in the distance model and $\beta=-1.36$ ($p=0.0003$) in the long jump model.

By and large, these results are not indicative of a weak instrument problem and the signs and statistical significance of the funding coefficient confirm the pattern reported in section 5, i.e., that funding is not associated with new field entry (rather the opposite) and conditional on entry it is negatively related with the distance of the new field entry, except for long jumps.

Since these validity checks of the instrument still assume i.i.d. errors, we also estimated the models using the LIML estimator using a robust estimate of the variance-covariance matrix, which has been shown to have better finite-sample properties if instruments are not strong, but assumes joint normality of errors in the structural and first-stage equations. Using this approach, we find that the instrument is statistically significant at the 1% level in all models, with the estimates for the funding parameter in the 2nd stage again consistent with earlier results (all models including field fixed effects). In the entry model β=-0.21 (SE=0.133, p=0.116); in the distance model β=-0.12* (SE=0.064, p=0.072); and in the long jump model β=0.499** (SE=0.215, p=0.020), with the asterisks denoting the same statistical significance levels as in Table 4.

Robustness Check

We performed a robustness check using a sharpened instrument, which measures funded scientists in the same main discipline (e.g. biosciences, mathematics, etc.: see Table 1 for an overview of the included disciplines) but including only those who belong to different departments. By considering only funded colleagues in other departments, the likelihood that the focal scientist's research agenda is affected by funded colleagues in other departments is lower since their research agendas are on average more remote to the one of the focal scientist than those of departmental colleagues. So, for example, for a funded oncologist the instrument contains funded colleagues from all departments in the discipline of internal medicine (‘Neurosciences’, ‘Rehabilitation Sciences’, ‘Chronic Diseases and Metabolism’, etc.) but not those from the department of Oncology. When excluding same-department funded colleagues, the mean of the instrument drops slightly to 0.21, compared to 0.23 for the instrument as used in the paper (i.e. scientists in the same main discipline and including department colleagues). Re-estimating the models with the new instrument generally shows robust results (available from the authors), the only notable exception being the effect of funding in the 'long jump' model (β=0.24), which is still positive (as in Table 4) but loses statistical significance. The first-stage tests for the relevance of the instrument show that the robust F-statistic equals 24.84 for the new field entry model (34.47 for model 1 in Table 4), and 7.06 for the 1st stage of the distance and long jump models (10.21 in the corresponding models in Table 4). In other words, for the distance and long jump models we cannot rule out that the stricter instrument is weak and the result does not confirm the positive relation of funding with making long jumps.

9 Appendix A2: Robustness Check

*Table 5: Results with discipline*year fixed effects*

	Single equation models			Two-stage models: 1st stage (funding)		Two-stage models: 2nd stage (entry)		
	New field entry model 1'	Distance model 2'	Long jump model 3'	New field entry model 4a'	Distance / Long jump model 5a'	New field entry model 4b'	Distance model 5b'	Long jump model 6b'
Talent & Productivity								
Quality$_{pre\text{-}1992}$	-0.005	-0.001	0.022**	0.039***	0.038**	-0.013	0.000	0.051**
	(0.008)	(0.002)	(0.010)	(0.010)	(0.016)	(0.011)	(0.005)	(0.020)
Medium productivity group$_{t-1}$ †	0.065***	-0.004	-0.059**	0.064***	0.036	0.053**	-0.004	-0.033
	(0.016)	(0.006)	(0.024)	(0.018)	(0.034)	(0.021)	(0.007)	(0.038)
High productivity group$_{t-1}$ †	0.062**	-0.009	-0.082**	0.043	0.037	0.053**	-0.008	-0.053
	(0.025)	(0.009)	(0.040)	(0.031)	(0.055)	(0.027)	(0.010)	(0.057)
Teams								
Co-authors	0.001**	0.000	-0.000	0.001**	0.001	0.001	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Co-authors in same discipline	-0.038**	0.005	0.000	0.064***	0.014	-0.051**	0.005	0.011
	(0.017)	(0.008)	(0.030)	(0.022)	(0.042)	(0.021)	(0.007)	(0.042)
PhD graduates	0.021**	0.004	0.014	0.057***	0.051***	0.010	0.004	0.051*
	(0.008)	(0.003)	(0.016)	(0.011)	(0.015)	(0.014)	(0.006)	(0.028)
Past new field entry								
Stock of fields	0.004**	0.001*	0.007***	0.000	0.002	0.004**	0.001*	0.008**
	(0.002)	(0.001)	(0.002)	(0.003)	(0.005)	(0.002)	(0.001)	(0.004)
Herfindahl index of fields	0.051	-0.001	0.042	-0.092*	-0.050	0.068*	-0.002	0.002
	(0.034)	(0.016)	(0.063)	(0.048)	(0.101)	(0.038)	(0.016)	(0.094)
Career age (years)	-0.009***	0.001	-0.001	-0.003	-0.005	-0.008***	0.001	-0.005
	(0.001)	(0.001)	(0.002)	(0.002)	(0.004)	(0.002)	(0.001)	(0.005)
Career incentives: rank ††								
Associate professor	0.017	-0.005	0.002	0.103***	0.156***	-0.000	-0.003	0.108
	(0.022)	(0.010)	(0.030)	(0.029)	(0.046)	(0.029)	(0.019)	(0.077)
Professor	0.042	-0.005	0.002	0.159***	0.157**	0.014	-0.003	0.111
	(0.027)	(0.011)	(0.040)	(0.039)	(0.063)	(0.041)	(0.020)	(0.092)
Full professor	0.086**	-0.012	-0.005	0.259***	0.220**	0.037	-0.009	0.154
	(0.035)	(0.015)	(0.054)	(0.051)	(0.093)	(0.062)	(0.027)	(0.135)
Funding								
Peers who acquired funding$_{t-2}$				-0.671***	-0.431*			
				(0.149)	(0.233)			
Funding$_{t-2}$	0.015	-0.003	0.024			0.208	-0.020	-0.709*
	(0.018)	(0.006)	(0.028)			(0.202)	(0.111)	(0.418)
Controls								
Teaching load	0.001	0.001*	0.008*	-0.010***	-0.007	0.003	0.001	0.003
	(0.002)	(0.001)	(0.004)	(0.003)	(0.007)	(0.003)	(0.001)	(0.007)
Male	-0.003	-0.002	-0.040	-0.013	-0.033	-0.001	-0.003	-0.064
	(0.026)	(0.011)	(0.044)	(0.037)	(0.072)	(0.026)	(0.012)	(0.062)
Post-1992 entry cohort	0.038*	-0.004	0.037	-0.100***	-0.150**	0.058*	-0.007	-0.074
	(0.022)	(0.009)	(0.038)	(0.036)	(0.059)	(0.030)	(0.018)	(0.080)
Full-time position at the university	-0.045	-0.017	-0.090*	0.141***	0.140***	-0.072*	-0.015	0.011
	(0.029)	(0.012)	(0.053)	(0.029)	(0.052)	(0.041)	(0.019)	(0.080)
Discipline*year fixed effects	*included*	*included*	*included*	*included*	*included*	*included*	*included*	*included*
Observations	4,706	1,022	1,022	4,706	1,022	4,706	1,022	1,022
Researchers	734	486	486	734	486	734	486	486

***, **, and * indicate significance levels of 1%, 5% and 10%, respectively. Robust standard errors (clustered by researcher) in parentheses. All models include an intercept.

† The reference category is the low productivity group. ††The reference category is assistant professor.

Models 5b' and 6b' build on the same first stage (model 5a').

10 Appendix A3: Correlations

Table 6: Correlation table (N=4,706)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Entry	1	1.00																					
Nr of fields	2	0.83***	1.00																				
Distance	3	n/a[Δ]	0.84***	1.00																			
Long jump	4	n/a[Δ]	0.24***	0.36***	1.00																		
$\text{Quality}_{\text{pre-1992}}$	5	0.02	0.01	0.02	0.05***	1.00																	
Productivity cluster: high	6	0.06***	0.05***	0.07***	0.03*	0.09***	1.00																
Productivity cluster: medium	7	0.08***	0.07***	0.07***	0.00	0.11***	-0.28***	1.00															
Productivity cluster: low	8	-0.14***	-0.12***	-0.13***	-0.05**	-0.16***	-0.35***	-0.27***	1.00														
Stock of fields	9	0.11***	0.09***	0.11***	0.01	0.13***	0.27***	0.19***	-0.36***	1.00													
H-index of stock of fields	10	-0.04**	-0.03*	-0.04*	0.00	-0.02	-0.03*	-0.04**	0.07***	-0.48***	1.00												
Associate professor	11	-0.01	-0.01	-0.01	-0.02	0.03*	-0.12***	0.01	0.07***	-0.09***	0.00	1.00											
Professor	12	-0.01	-0.02	-0.01	0.01	0.02	-0.02	0.05**	-0.03*	0.04**	-0.05***	-0.32***	1.00										
Full professor	13	0.00	0.01	0.01	0.03	0.03	0.26***	0.01	-0.18***	0.23***	0.00	-0.38***	-0.33***	1.00									
Seniority	14	-0.10***	-0.09***	-0.10***	0.01	-0.00	0.13***	-0.02	-0.04**	0.20***	-0.02	-0.26***	0.10***	0.63***	1.00								
Co-authors	15	0.11***	0.10***	0.10***	-0.01	0.10***	0.35***	0.15***	-0.33***	0.43***	-0.14***	-0.07***	0.00	0.18***	0.09***	1.00							
Co-authors in same discipline	16	0.03*	0.03*	0.03	-0.02	0.12***	0.17***	0.26***	-0.27***	0.18***	-0.04**	-0.02	-0.04**	0.06***	-0.04**	0.19***	1.00						
PhD graduates	17	0.06***	0.05***	0.06***	0.06***	0.09***	0.26***	0.06***	-0.24***	0.17***	-0.05***	-0.13***	0.02	0.32***	0.20***	0.18***	0.11***	1.00					
Funded	18	0.04**	0.03*	0.04*	0.03*	0.14***	0.13***	0.14***	-0.22***	0.22***	-0.12***	-0.04**	0.05**	0.21***	0.16***	0.21***	0.14***	0.25***	1.00				
Teaching load	19	-0.05***	-0.05**	-0.04**	0.04**	-0.05***	0.06***	-0.07***	0.03*	-0.07***	0.06***	-0.21***	0.05**	0.53***	0.50***	-0.07***	-0.07***	0.21***	0.05***	1.00			
Male	20	-0.01	-0.00	-0.00	-0.02	0.06***	0.08***	0.04*	-0.09***	0.06***	0.01	0.04**	0.03	0.15***	0.17***	0.06***	0.01	0.09***	0.05***	0.12***	1.00		
Post-1992 entry cohort	21	0.06***	0.06***	0.05***	0.00	-0.07***	-0.11***	-0.00	0.05***	-0.11***	-0.00	0.04**	-0.25***	-0.42***	-0.59***	-0.06***	0.03*	-0.22***	-0.19***	-0.43***	-0.19***	1.00	
Full-time employment	22	0.00	-0.01	0.00	-0.02	0.05**	0.10***	0.08***	-0.14***	0.13***	-0.05***	-0.14***	0.01	0.17***	0.06***	0.08***	0.11***	0.07***	0.13***	0.19***	0.00	-0.02	1.00

* p<0.05, ** p<0.01, *** p<0.001

[Δ] Distance and long jumps are only defined conditional on entry=1.

[1] Corresponding author.

[2] The findings, interpretations and conclusions expressed in this paper are entirely those of the authors and should not be attributed to the European Commission. Possible errors and omissions are those of the authors and theirs only.

[3] https://www.kuleuven.be/english/research/

[4] The data comprises the population of professors in biomedical and exact sciences employed by KU Leuven in 1992-2001. We only use the 734 scientists who publish at least once in that period, resulting in 4,706 observations. For these active scientists we constructed full publication and citation records (including their publications before 1992) from the Web of Science (WoS, Clarivate Analytics). In order to retrieve each researcher's publication history, we performed searches on scientists' names (last name and first initial of the first name, as registered in personnel records) and, to avoid false matches due to homonyms, affiliation. For the latter, we used the *Organization-Enhanced* field provided by Clarivate Analytics, which links all possible variations of an organization's name found in the Web of Science. We also added the search term "University Hospital Leuven" to ensure all publications by clinical scientists affiliated to the university's medical departments were included in the data. The retrieved publications are restricted to those on which the scientists mention their KU Leuven affiliation. Although this may imply incomplete coverage of publication records, mobility of researchers is known to be very limited in our sample, especially in the time period we consider.

[5] Scientist's past field entries are coded as 0/1 variables as scientists typically don't enter with many publications at once into a new field i.e. it is rather the distinction between 'no entry' and 'entry' that matters.

[6] This restriction also partially addresses the concern that a field entry may be linked to co-authorships with colleagues whose main expertise is in fields that are new to the focal scientist. Also here, repeated publications in that field (even with the same team of co-authors) mitigates the concern that the focal scientist merely supplies own expertise without having much exposure to the new field.

[7] We consider the following types of publications: "Article", "Article; Book Chapter", "Article; Proceedings Paper", "Letter", "Meeting Abstract", "Note", "Proceedings Paper", "Review". Note that we observe all of a scientist's publications i.e. also those in lower-ranked journals in the Web of Science, which mitigates the concern that we miss any new field entries.

[8] The Web of Science classifies every journal and book that the *Core Collection* covers in "subject categories". We adjusted this classification so it displays a consistent granularity across the entire spectrum of scientific fields. In particular, some WoS subject categories consist of several sub-categories (e.g. *chemistry* is subdivided in *analytical*, *applied*, *inorganic* and *nuclear chemistry*), while others do not (e.g. *microbiology* is not further subdivided). We use the term "scientific field" throughout the paper for any of the 194 fields. We also use a more aggregate journal-based classification of 12 broad "disciplines", developed by the Centre for Research & Development Monitoring (ECOOM, KU Leuven), shown in Table 1.

[9] For example, the most similar scientific fields in 1992 were *'Remote Sensing'* and *'Imaging Science & Photo Technology'*, with a cosine similarity value of 0.692. The least similar in the same year (with a non-zero cosine value) are *'Astronomy & Astrophysics'* and *'Biochemistry & Molecular Biology'*. In 2002, *'Cell Biology'* and *'Biochemistry & 'Molecular Biology'* were the most similar, while *'Astronomy & Astrophysics'* and *'Health Policy & Services'* were the farthest apart of all fields.

[10] In case the researcher enters more than 1 field in a year, we take the maximum distance between the researcher's main field and all of the entered fields.

[11] While we control for a rich set of regressors, it is still conceivable that the 2 parts of the model are not independent, which could be addressed by estimating a selection model. Since we do not observe a variable that reasonably satisfies the exclusion restriction, we did not pursue this approach.

---

[12] The high mean value for distance (i.e. close to 1) reflects the on average low cross-citations between journals from the 194 different fields, resulting in an average pairwise cosine similarity of 0.02. Related fields do show higher cosine similarity values, e.g. *Cell & Tissue Engineering* and *Transplantation* (0.11) or *Genetics & Heredity* and *Biochemistry and Molecular Biology* (0.57).

[13] Note that some disciplines account for a fairly small number of scientists. The smallest disciplines in the sample are *Neuroscience* (12 scientists), *Geosciences* (18) and *Mathematics* (25). The average number of scientists per discipline is 61.

[14] For the list of main disciplines, see Table 1.

[15] Robustness checks using different numbers of clusters in prior work (Kelchtermans & Veugelers, 2013) showed that researchers cluster into 3 distinct productivity groups with the average numbers of publications of the 3 clusters being statistically significantly different. If the researcher published in more than 1 of the 12 disciplines in a year, we classify her based on the highest cluster performance.

[16] These splits by discipline are not reported in Table 2 but are available from the authors.

[17] We don't have information on funding received from other sources, but these other sources are typically highly correlated with the BOF funding we observe in the sample.

[18] To the best of our knowledge, there is no implementation of estimators that fully accommodates a dependent and endogenous variable that are both binary, with the exception of special regressor methods (Lewbel *et al.*, 2012). Because the requirements for the special regressor cannot be fulfilled by any variables in our data, we estimate a linear probability model (LPM) through 2SLS. It has been noted in the literature that LPM marginal effects are often very similar to average marginal effects that can be obtained from binary choice models (e.g. Cameron & Trivedi, 2005) and that they provide good approximations of true marginal effects, even if they do not fit choice probabilities perfectly (Angrist & Pischke, 2009). A LPM may underestimate standard errors, which is why the results should be interpreted with some caution. However, statistical significance levels of the covariates in the new field entry equation are consistent with

those of the estimates of a logit model (the latter without the endogeneity correction of funding). Moreover, the consistency of the results of using a binary DV and the distance to new fields provides additional reassurance of the robustness of the results.

[19] In other words, *disciplines* are defined very broadly and the action — in terms of scientific progress and associated funding — is rather going on at the more disaggregate *field* level. For example, Higham *et al.* (2017) found substantial heterogeneity in publication growth for specific fields within the discipline of physics. Furthermore, besides constructing the instrument at the aggregate level of 12 broad disciplines, in the robustness check in Appendix A1 we further restrict the instrument by only counting funded scientists outside the focal scientist's own department. For example, for an engineer, we measure how many peers in the discipline of engineering received BOF funding in the same year, excluding engineers from the same department (engineering departments are, for example, *Computer Science*, *Chemical Engineering*, *Architecture*, etc.). Hence, the instrument measures peer funding external to the scientist, and the link with the focal scientist is not likely to be based on sharing the same 'hot' discipline.

[20] The only noteworthy difference is when we use a shorter 1-year lag for the funding variable, in which case we no longer obtain a statistically significant estimate for the funding variable. We attribute this to the time it takes before acquired funding translates into publications that reflect potential new field entry.

[21] Given that quality is measured as the average citation rate of a scientist's pre-sample (=before 1992) papers, it is necessarily more noisy for the cohort that entered the university in or after 1992 since for assistant professors (which is the starting rank for the vast majority of those who enter the university) the measure will be based on a limited number of publications. A robustness check (not reported) splitting the cohort that entered before 1992 from the one that entered in or after 1992 shows that the results for quality (but also productivity) are indeed mainly driven by the pre-1992 entry cohort, with these characteristics positively and highly significantly associated with the entry into new fields.

[22] In a robustness check where we add researcher fixed effects, the stock of fields is negatively linked to the acquisition of funding but there is no longer a significant relation with new field entry. In the individual fixed effects model, we also find that the more concentrated one's portfolio of past fields (*Herfindahl index*

of fields), the more likely new field entry becomes. In other words, once we control for the inherent taste for new field entry, researchers with more specialized expertise seem more conducive to leverage it in new fields.

[23] Wald tests for pairwise comparisons of the coefficients for associate professor ($\beta = 0.055$), professor ($\beta = 0.106$) and full professor ($\beta = 0.173$) show they are significantly different at the 5% level.

[24] The positive sign of the instrument in the first stage indicates a positive correlation between a scientist's funding and her peers in the same discipline. As discussed in section 3.2, this suggests differences between disciplines in terms of how resource-demanding they are or their strategic priority for the university.

[25] A scientist's stock of fields is negatively and significantly associated with new field entry in the model with field fixed effects while it is positive and significant in the model with discipline fixed effects (for both the single equation model and the 2nd stage IV regression). This difference suggests that it is primarily the past entry in specific fields that explains additional new field entry.

[26] Note that with 273 covariates included, this model puts a lot of structure on the data, especially on the sample of 1,022 observations where entry is observed. In the new field entry model, most year dummies are statistically significant in the 1st stage (consistent with rising research budgets in the observation period) while none are significant in the 2nd stage. In the distance model, only one of the year dummies is statistically significant (in both stages) versus none in the long jumps model.

[27] Note how the first-stage results in Table 5 are informative about the mechanism behind the instrument. The zero-sum logic discussed in section 3.2 is apparent from the negative sign, which indicates that, for example, the more of an engineer's colleagues in the broad discipline of engineering are funded in particular year, the less likely it is she is awarded funding. However, once we take into account the specific fields that the engineer has been active in, as in Table 4, the parameter of the instrument takes on a positive sign, indicating that funding is not uniformly allocated across fields within a discipline. This makes intuitive sense: an obvious reason would be differences in their need for resources, or because funding panels consider certain fields strategically important. From Table 4 and 5 we gather that the zero-sum game logic

only holds in the aggregate, reflecting budget constraints, but once taking into account field specialization the instrument picks up a positive correlation.

[28] This is not common: on average, 1.24 publications (s.d.=0.66) are associated with a scientist's entry into new fields in a given year.

[29] The BOF funding system that we analyse is an open competition across fields with each proposal assessed by two external referees and the university research council. While we cannot rule out that the effects are contingent on specific features of this funding scheme, it does reflect the principles of other systems that competitively allocate research grants.