Searching where Light Shines: The Availability of Information on Past Cumulative Output of Scientists and its Influence on Funding and the Consequent Performance of the Scientists

Anand Nandkumar
Indian School of Business
AC8, Level 1, Office # 8126
Gachibowli,
Hyderabad - 500 032
91-40-2318-7163
91-40-2318-7032 (fax)
email: Anand_Nandkumar@isb.edu

Reddi Kotha
Lee Kong Chian School of Business
Singapore Management University
50 Stamford Road
178899, Singapore
65 6828 0401
65 6828 0777 (fax)
email: reddikotha@smu.edu.sg

Harsha Tadikonda
Indian School of Business
Gachibowli,
Hyderabad - 500 032
email: Harsha_Tadikonda@isb.edu

First draft dated 30 October 2013

Abstract

We argue that difficulties associated with evaluating new projects of academic scientists' result in evaluators relying on a simple rule to allocate research funding-past cumulative output of a scientist. We identify and hone in the impact past cumulative output has on funding. We suggest that variation in knowledge that underlies a scientist's portfolio attenuates or exacerbates the evaluation uncertainty thus influencing funding. We then focus on the consequences of such an allocation. There are three broad possible outcomes: that cumulative output based funding leads to a victorious cycle of improvement by those with high cumulative output and receive more funding; cumulative output based allocations are neither efficient nor inefficient; cumulative based allocations lead to a mismatch between funding and productivity. We test these predictions using a sample of 29,859 academics who had and had not received funding by federal government at a large mid-western university from 1970 to 2005. Our results proffer evidence that is consistent with evaluation uncertainty influencing funding. Furthermore, the results suggest that there is a mismatch between amount of funding based on cumulative output and productivity of scientists. We suggest an alternative decision strategy that could have been used: past productivity and discuss the improvement that could have resulted. We conclude with a discussion of the findings of the paper to the literatures on resource allocation under uncertainty, social ranking based theories, and design of evaluation systems in organizations and society.

There is resurgence in the behavioural literature on how individuals use simple rules under uncertainty to make decisions. Benartzi & Thaler (2001) show with laboratory and field data that individuals allocate their savings equally between options offered in retirement plans. Bardolet, Lovallo & Fox (2011) suggest that CEOs of multi-division firms make equal capital budget allocations to their divisions, due to the cognitive burden involved in processing information. Another simple rule: whether a customer purchased in the past nine months was used by retail marketing managers to predict repurchases by customers performed better than other sophisticated rules using more variables (Wubben & Wangenheim, 2008). Simple rules have been argued to be in use when there is time pressure and uncertainty. While past studies focus mostly on a single individual making inferences about probabilities, there are many organizational situations wherein groups allocate resources to candidates. Consider for instance, senior managers rather than just the CEO allocating capital to divisions, in large multi-business firms, the treasury making allocations to government departments, or committee of peers making funding grants to academic scientists. There are two aspects that set these organizational situations from those explained by prior work using simple rules. First, the starting point is there is a group rather than a single individual decision maker allocating resources. Groups make systematically different than individuals (Stasser and Titus, 1985). Second, these allocations are to candidates or departments, rather than allocation of saving to "face-less" asset classes. Hence it is a possibility social influences may play a role in the allocation by made by groups to their peers. Therefore, in this paper we propose to study a situation wherein evaluators make recommendations on allocation of resources to projects of candidates.

We suggest that uncertainty associated with new projects leads decision makers relying on easily available and seemingly pertinent information. We draw on the sociology of science and status literatures to motivate why evaluators pay attention to past cumulative output to allocate resources (Merton, 1968; Allison & Stewart, 1974;Allison, Long, & Krauze, 1982). Status is a social ordering of actors based on perceptions of deservingness in the eyes of the evaluators (Podolny, 1994).

We argue that cumulative output which is used to infer social ordering may contain heterogeneous components. As the components measure uncertainty of the domain for instance: newness of science (Azoulay etal, 2011; 2013) then the reliance on social ordering should increase as the project level uncertainty increases (Podolny, 2001). However, if the components of past output are diffused and hence imply unclear aggregation of cumulative output (Zuckerman, 1999), then reliance on social ordering may become muted. Thus we explore the conditions under which the positive relationship between reliance on social ordering information: cumulative output and funding is amplified or muted by the type of components involved in the past output.

After building predictions on why resource allocation occurs in the pattern it does we then turn to the consequences. It is an open question whether allocation of resources, when based on simple rules, are efficient or a societal drag. An analytical strategy pioneered by Becker (1993) examines the *ex post* performance of preferred and non-preferred groups. In our setting we would compare the performance of those scientists who had high or low cumulative output and when they received low or high funding. There are three broad possible outcomes: i) that cumulative output based funding leads to a victorious cycle of high productivity of those with high output and receive more funding; ii) cumulative output based allocations are neither efficient nor inefficient; iii) cumulative output based allocations lead to a mismatch between funding and productivity. The first outcome would suggest that it is functionally good for a society to allocate more resources to those with high cumulative output. The second result while theoretically interesting would suggest that society's scarce resources are not being whittled away. Whereas the third result implies that resources are sub-optimally allocated.

We test these predictions using data on funding of academic scientists by U.S. federal government from 1970 to 2005 at a large mid-western university. The sample consists of 29,859 unique individuals associated with the university in that period. We predict the allocation of research grants to scientists and their consequent output. Our results are consistent with the view that the past cumulative output predicts resource allocation. The relationship between cumulative output and resource allocation varies by the components of the scientists past output. Furthermore, the performance of scientists' reveals that there is mismatch. That is those scientists who had higher cumulative output and had received higher funding and have lower productivity.

This study makes three contributions to extant literatures. First, the literature on behavioural resource allocation has shown that simple rules are used by individuals to explain a vast majority of resource allocation. In this study we show that simple and easily available information influences allocation consistent with the behavioural view. There is a debate in the behavioural literature on consequences of using simple rules. Some have argued that these lead to sub-optimal allocations (Kahneman, 2011). Others have suggested these simple rules are just as efficient as the next best alternative available to the decision makers (Gigerenzer, 1991). Our results support the former view rather than the later. In *ex post* analysis we propose an alternative decision rule: allocation based on past productivity of a scientist. Information on past productivity could have been made available to decision makers as past cumulative output. We then compare the performance of allocation based on past cumulative output and past productivity. While we recommend caution against extrapolating from the current environment in the study to other situations, the results of the *ex post* analysis suggest that the alternative rule we recommend performs better. Second, we draw extensively on the literature on status to motivate the causal mechanisms for why past cumulative output is relived upon by decision makers. To this literature by showing that reliance on past cumulative output varies even at the same level of cumulative out when the heterogeneity in the cumulative output is considered. When past projects consist of wide breath of knowledge components that do not allow for meaningful aggregation of past output, hence reliance on cumulative output to infer social ordering diminishes in such situations.

Third, this study contributes to the literature in sociology of science on inequality. In this literature the persistence of inequality in output of scientists is well documented. Since it was highlighted by Merton (1968) there have been vigorous debate on the actual mechanisms that lead to such an inequality. Consistent with Merton's argument a view suggests that strong adherence to universalistic evaluation actually accelerates cumulative advantage by concentrating resources among those best equipped to use them (Zuckerman, 1977). Another view suggests that the innate talent of scientists is unlikely to be as inequitably distributed as the output of scientists is. Therefore inequality must be bad for society (Turner and Chubin, 1979). However, Denrell (2003) cautions against observing outcomes of survivors and making inferences, without examining inputs. Furthermore, evidence for the cumulative advantage is tenuous and equivocal (Allison, Long, & Krauze, 1982). It is tenuous since only the elements of the causal chain have been tested in anyone study due to lack of adequate data on resource allocation. In a survey, Fox (1983) remarks that "..*since investigators have lacked these data on resources, their findings support the cumulative advantage hypothesis, only indirectly".* Similar comments have been echoed by Azoulay, Stuart, Wang (2013) who have called for studies focused on resource allocation to tease apart the causal mechanisms through which a cumulative advantage may occur to scientists.

Scholars in economics of science have also stressed the need for studying resource allocation. For instance Stephan (pp1124, 1996) in a review of economics of science remarks that rather than focus on life-cycle effects on output of scientists, researchers should focus on the role of resource allocation: *"this leads one to wonder if we should not use our talents as economists to develop a different approach to the study of scientists that stresses the importance of resources in the process of discovery rather than the importance of the finiteness of life".* Arora and Gambardella (2005) take up this challenge. In a carefully done study they examine funding by National Science Foundation (NSF) to economists. They find that while past track record was important predictor who got funded; past track record did not influence the amount that was funded. They suggest that since economists probably need a few months of salary support, graduate assistance and a computer, it was unlikely that there would be much variation in the grant proposals. Presumably funding amounts vary more greatly in science and engineering disciplines. Hence, in this study by starting with resource allocation and then examining the long term performance consequence we are able to provide to the literatures in sociology of science and economics of science a more complete causal analysis. Our results show that greater resources indeed flow to those with high cumulative output but their productivity does not keep improving (in fact it is negative).

We explore an alternative decision rule that could have been used by decision makers: information on past productivity of a scientist. By showing evidence of mismatch and by providing a solution that improves the relative use of societal resources this paper has important implications for policy. It is important that we stress the concept of relative use of society's resources. It does not make sense to speak in the absolute. The growth in publication output is constrained by number of journals. Even if we were to double the funding in a year it would make no difference to number of papers published if there was no comparable growth in the number of journals (Arora and Gambardella, 2005). Therefore, we are careful to make the case of relative productivity of scientists contingent upon a track record and a degree of funding.

The rest of the paper is organized as follows. In the next section we review the literature and build arguments for our hypotheses. The section after that is the methods section followed by the results section. We conclude with a discussion of the findings for extant literatures on behavioural resource allocation, organization status and policy implications.

THEORY

The theory development is structured as follows. First we review the literature on behavioural decision making and summarize the central insights on the use of simple rules to make decisions. Then we review the literatures on sociology of science and organization status to suggest why peer-evaluators may rely on past cumulative output to allocate resources. Then we follow the literature on status to suggest two moderators of the positive relationship between cumulative output and funding. Finally, we articulate competing hypotheses on the relationship between funding based on cumulative output and productivity of scientists. These competing hypotheses reflect the debate in literature in sociology of science.

Behavioural Approach

Assume for now a simple closed-system. Wherein there are producers who interact with the same evaluators. Positive evaluations lead to more resources or access to consumers. Consumers rely on evaluators as they may lack expertise to do the evaluation themselves or do not have the time to perform the evaluation themselves. Several market interactions share this fundamental structure: listed companies and stock market analysts; movie producers and movie critics. If the producers and evaluators are a stable set and interact regularly then it is possible that the relative deservingness of a producer may be well understood by the evaluators in the long run (Lant, 1992). For an exception of this general assumption see

Denrell (2003). Furthermore, if the evaluators themselves are evaluated and rewarded based on how well their *ex ante* predictions correspond to subsequent performance of the producers then there would be all the more reason for the consumers to rely on the evaluations of the evaluators.

Now let us relax these assumptions in a closed-system. What if producers entered and exited the market for evaluation irregularly? Furthermore, what if the evaluators themselves entered and exited the market? In addition, what if the evaluators had competing demands on their time and were not paid based on the reliability of their evaluations? All of this would suggest that such an open-system would have higher margin for error than a canonical closed-system. However the mean value of evaluations in such an open system may still be unbiased even if the variation in the evaluations is large.

Under such an open system let us introduce a behavioural perspective. The behavioural view suggests that decision makers do not have the limitless processing capacity to process all information pertinent to a decision. That is they are boundedly rational (Simon, 1957). Furthermore, the behavioural view assumes that decision makers do not optimize but satisfice. That is decision makers do not examine all solutions to a problem to find the best solution but stop when they have a good enough solution. This essentially presupposes that decision makers have a stopping rule, i.e., a level of aspired performance. Once a solution meets the level of aspired performance, then search stops. There is a large body of organization literature that has examined how aspirations are set and how firms and managers react to performance above or below aspirations (for a review see: Greve, 2003).

More pertinent for the purpose of this paper is how decision makers evaluate alternatives. The behavioural literature suggests that decision makers may seek to conserve cognitive energy by examining fewer cues, reducing the effort of retrieving cue values, simplifying the weighting of cues, integrating less information, and examining fewer alternatives (Shah and Oppenheimer, 2008). This then leads to questions like what cues do the decision makers pay attention in a context and what are the performance consequences of such cognitively miserly behaviour. Below we motivate why peer-evaluators pay attention to cumulative output of the scientists and why the cumulative output of a scientist influences the grant amount she receives. Before we do so, we will briefly describe the constraints and motivations of the peer-evaluators in our context.

Constraints and Motivations of Peers Evaluators

The organizational structure of most federal research grant agencies is as follows. There is a program director in charge of a field. Chemistry for instance may contain up to 8 or more program directors at National Science Foundation (NSF). Once a proposal is received the program director sends it to reviewers, who are experts in the topic. The reviewers are asked to rate a proposal as excellent, very good, good, fair or poor. They are also asked to support their rating with a written evaluation. The program director then decides on whether to fund a project and how much to fund a project. The evaluations of the reviewers are not binding on the program director. Peer evaluators typically do not get paid for their reports. It is a voluntary service. In some agencies evaluations are made by a panel of experts. The panels may meet three times a year to decide on proposals. The panels typically provide a ranking of projects to the program director. The program director then decides on funding. Again panel members are not permanent and not usually paid. Thus, it is clear that peer evaluators have no feedback system on the ultimate performance of the projects they had evaluated.

Contrast such a system with investment analysts covering listed company stocks. An analyst is typically an expert in a domain and assigned to few stocks to cover. Analysts periodically evaluate the same set of companies. Analysts make forecasts about earnings and share price. There is regular feedback loop to the analysts when earnings data are released and by the reaction of the stock market. Thus in such a system there is opportunity to learn from feedback at regular intervals on the gap between evaluation and actual outcome; whereas in a system like the funding of science, evaluators typically lack a feedback loop. Therefore peer evaluations by scientists may be noisier than evaluations by analysts covering public companies[1].

Evaluation Difficulty in Funding Scientific Projects: Past Cumulative Output

It is a relatively easy starting point for us to motivate why peer-evaluators pay attention to the past output of the scientists. In academia promotion, tenure and social prestige are based on publications of a scientist[2]. Hence publications are the primary currency for a scientist. Peer evaluators who are themselves mostly academics may focus on publications of the applicant. NSF for instance recommends that evaluators focus on: the significance of the investigation, ability of the applicant, capacity of the institution to support the research. Cole, Rubin and Cole (1977) in a review of NSF peer reviewed proposals suggest that heavy emphasis is placed on first two factors.

New science projects are uncertain. Even for most experts it is hard to predict the outcomes of a new scientific investigation. Hence evaluators, under a cloud of uncertainty about a project, may be forced to pay more attention to more readily available information about the producers. In other words they may weigh the cue on past performance more than they weight the cue on quality of the current project since processing information on quality of a project is harder. This would lead to those scientists with a moderate quality project but a

[1] It is important that we state that we do not think scientists who serve as evaluators shirk intentionally. Our explanation focuses more on the organization structure of the open system that constraints evaluators' ability to learn and expend effort due competing demands on evaluators' time. Stephen (1996) summarily puts it that scientists do not shirk.

[2] The highest status accrues to those with elite prizes the like Nobel Prize. These prizes are typically awarded to those who are first to solve an important problem.

high past research performance to receive a higher score than those scientists with presumably higher quality project but a much lower past research performance. At the limit of this weighting function in favour of past research performance we should see that those with higher past performance get more funding than those with lower past performance, regardless of the quality of the proposed projects.

Now it is a matter of fleshing out what dimension of past performance that most reviewers focus. Since the modal applicant is unlikely to have won elite prizes which are by definition sparingly bestowed (Zuckerman, 1977), evaluators have to rely on more widely observed dimension of past performance[3]. Since almost all grant giving agencies require applicants to provide their resume, one piece of information that evaluators have easy access is to is the information on the publications of the applicants. Information that is readily available and comes to attention easily is more likely to be relied upon in making decision (Tversky and Kahneman, 1973). Furthermore, evaluators are more favourably disposed to a producer's current project if they had past encounters with the producer's prior work. This may lead to the weighting cue on project quality to be influenced upward. This is more likely to be so in case of those scientists with higher past publications. Evaluators are more likely to have read prior work of a scientist with a greater number of publications than a scientist with fewer publications[4]. It has also been suggested the recognition itself is a heuristic and may guide decision making when choosing between alternatives (Goldstein and Gigerenzer, 1999). An alternative that is more readily recognized is weighted more positively. Extending this insight to case by case evaluations it is may be plausible that scientists with higher cumulative publications are more likely to be recognized by evaluators and hence their

[3] This is not to say for winners of elite prize such information becomes even more salient than the project quality or other dimensions of past performance. Also the construct "elite prizes" fits our theory of simple and easy to observe rule. We are constrained by pragmatic reasons not to make this part of our predictions. We explain the enormous empirical challenges with collecting this data in the methods section.

[4] Cumulative citations should also work in the same direction. We do not have hypothesis on cumulative citations as this may not be a simple and as readily available to evaluators. Hence it is a control variable.

projects viewed more positively. Hence an evaluator may judge the quality of a producer's project to be higher if the producer has higher number of past publications. Therefore we predict that:

> *H1: Ceteris paribus the count of cumulative publication is positively related to the amount of funding.*

Varying Types of Uncertainty: New Science

In the hypothesis above we had focused on two salient cues: project quality and past performance of a scientist to make our arguments. Now we turn to situation wherein just the project quality is much harder to evaluate. Some scientific projects may be at the frontiers of science and may use new tools and techniques which even specialists in their own domain may not fully understand. This would imply that peer evaluators are more likely to be uncertain about the quality of a project should it use newer scientific tools and techniques in its proposal. This view is reinforced by Azoulay, Graff Zivin, and Manso (pp:531;2011) who suggest that *"... peer-review panels in charge of allocating awards, are notoriously risk averse and often insist on <u>great deal of preliminary evidence</u> before deciding to fund a project. This often leads researchers to resubmit their application several times and to multiply the number of applications, taking time away from productive research activities. It is often-hear complaint among academic biomedical researchers that study sections' prickliness encourages them to <u>pursue relatively safe avenues that build directly on prior results</u>, at the expense of truly exploratory research (emphasis added)".* Thus to the extent new science forms a part of a proposal it may be evaluated to be of lower quality.

Since academic scientists research agenda are stable and path dependent (Dosi, 1982). That is scientists work in a domain and don't typically change their research program frequently, as it takes many years to learn the tools of the trade in a domain. Thus, scientists working with new science in past are more likely to pursue with their new science agenda. Hence the extent of new science used in the past output is likely to be positively related to content of new science in the current project proposals. Given that new science is harder to evaluate it follows that projects of those scientists with more new science content in their past output will be judged to be of lower quality. Therefore we predict that:

H2a: The use of new science by a scientist in her past output is negatively related to grant amount.

Now consider the joint variation in project level uncertainty and past cumulative output of scientists. When uncertainty is high evaluators turn towards social cues to make decisions (Podolny, 2001). Therefore when a science projects contain higher new content evaluators may rely to a greater extent on social cues. Past publications of a scientist are again the most widely available information to the evaluators. To the extent past publications also convey a social order then it follows that evaluators should place a greater emphasis on past publications under uncertainty. Recall that promotions and tenure in academia are based on past publications. It has been extensively argued in sociology of science that two paper of equal quality one by a more well-known scientist and another by not as well-known scientist would result in the paper by well-known scientists to be rated higher for quality (Merton, 1968), cited more often, cited at a faster rate. Therefore to the extent a scientist with higher cumulative publications is more well-known, her work of equal quality should be rated higher than a scientist with lower cumulative publications.

Assuming for now that the quality of the projects are similar. Take the case of a two scientists: one scientist with high cumulative publication and the other with low cumulative publications. Assume further that both have projects that have similar high level of newness of science component. Then the evaluators should prefer the project of the scientists with higher cumulative publications. Not only this evaluators weighting of cumulative publications cue should increase as faster rate since uncertainty about the project (new science) makes them to rely even more on cumulative publications (Podolny, 2001). This would imply that under condition of uncertainty preference of those with higher cumulative publications would be much stronger than when evaluators were more certain.

> *H2b: The new science* <u>*positively*</u> *moderates the positive relationship between cumulative publications and grants. Such that relative increase in grants when compared to those with low cumulative publications to those with high cumulative publications is much greater when the science used is new.*

Varying Types of Uncertainty: Diffused Knowledge base

Producers may vary in the extent their output spreads over several domains. Zuckerman (1999) argues that companies that are atypically in their diversification face an illegitimacy discount. This occurs due to analysts in investment banks being less likely to cover such companies. Analysts are less likely to cover such companies because they may lack the expertise to evaluate operations of a company that is spread across wide variety of sectors, as a typically evaluator is specialized. Similarly in our setting a scientist whose projects combine a wide variety of knowledge domains may be at a disadvantage when compared to a scientist whose output is more concentrated. This is due to the fact that the funding institution can easily match an applicant who is specialized to an evaluator with expertise in the domain. Since projects that combine wide spectrum of knowledge are less likely to find evaluators who share similar expertise or enough common ground (Kotha, George, Srikanth, 2013), lacking wide expertise evaluators may rate such projects of lower quality. We predict that those scientists with a wider disbursed knowledge base in their past output would receive lower funding as their projects may be rated as being lower in quality.

H3a: The width of science in a scientist's past output is negatively related to grant amount

Now consider what role social cues would play in such a situation with a producer whose past output is diffused across a wide spectrum. Typically under uncertainty social cues should play a greater role (Podolny, 2001). However if the past output of a scientist is diffused across a wide spectrum of domains an evaluator examining the past output would more strongly discount the portion of the output not in her domain of expertise. Hence, scientists with diffused past output would only get a fraction of credit based on their social standing they would have had otherwise gotten if only their past output was more concentrated. Conversely scientists whose past output is concentrated would be matched easily with a specialist with similar knowledge. Hence the past output of such scientists would be rated higher. Whereas the past output is more diffused and as cumulative output increases the perceptions regarding quality increase at a much lower than the increase in quality perceptions had the output been more concentrated. Therefore we suggest that:

H3b: The width of science by a scientist negatively moderates the positive relationship between cumulative publications and grants. Such that relative increase in grants when compared to those with low cumulative publications to those with high cumulative publications is much greater when the width is low.

Performance Consequence:

What are the performance consequences of an allocation based on past cumulative output of a scientist? There is a fierce debate within the sociology of science on the societal optima of social ranking based funding. One set of scholars have argued for a victorious cycle between social order based allocation and return to society. These scholars suggest that scientific talent is unequally distributed. Some scientists have higher innate ability: "scared spark". Scientists with higher talent are more likely to have a higher output. Furthermore, if these scientists with higher innate talent receive more resources they are best positioned to take advantage of the larger funding. Since their superior innate ability allows them to solve scientific problems faster and better than those without such innate ability.

Furthermore, colleagues may shower those with perceived higher ability with positive externalities. These scientists would get more feedback on their projects (Zuckerman and Merton, 1972). The expectation of their colleagues may also motivate those with higher past output to work harder. In contrast to the spiralling success explanation for those with high past output, scientists with little past output face disheartening hurdles. They lack resources. They are unlikely to be unilaterally approached by colleagues offering advice and encouragement. The long, lonely and arduous process of research without intermediate positive feedback loop from colleagues may discourage them to such an extent that they stop believing in themselves and do not extend much effort. Thus this vicious cycle would result in vast difference between the productivity of those with high cumulative output when compared to those with low cumulative output.

Therefor it may be functionally better for society if more resources were provided to those with the higher cumulative output. This view would predict there would be a positive relationship between the joint effect of increasing cumulative output and increasing funding on future output of scientists.

> *H4a: The amount of funding received positively moderates the relationship between past cumulative output and future output.*

A contrasting view suggests that social perceptions of quality and actual quality may diverge. When there is a divergence between social perception and actual quality and when increasing allocations are made to those with higher social standing, then, even if resource allocation lead to improvements in the output of those who received more resources, it may still be the case that the output of such scientists may not be commensurately high enough to justify the additional funding. That is the productivity of scientists with high cumulative output and high levels of funding may be lower. One reason for this could be the flight of funding to those with past track record. Since new projects are uncertain, capital may be attracted to those with cumulative output at an increasing rate than they could utilize and or improve their ability.

There may be cognitive limits on the number of research assistants and collaborators that a researcher can manage[5]. Increasing grants may lead to a scientist reaching this threshold and grants beyond this level would lead to negative productivity. Thus this view would suggest there would be a mismatch between grants based on social ordering and the consequent performance of the scientist. Therefore those scientists who have high cumulative publications and receive large grant amounts may have a lower output.

H4b: The amount of funding received negatively moderates the relationship between past cumulative output and future output.

Note we do not make an explicit prediction when the joint effect of high past cumulative output and funding has no effect on future output. But, it is worthwhile mentioning that should the joint effect be not significantly positive or negative it would imply that the society is no worse off. It is still open to speculation if the lack of a significant result is due appropriate funding based on true quality or based on the learning and improvement made by those with high cumulative output when they got higher funding.

[5] Furthermore this may be due to the fact that improvement in grant writing which is a specialized function. Some studies have shown that it takes up nearly 30% of a scientist's time. Grant writing may be crowding out the time needed for her to do research. Or experience with grant writing may lead to better ability to structure the grant proposal such that experienced writers get more grants than the quality of their proposal merits. We control for the past count of grants received.

METHOD

Research Site

We test the predictions in a sample drawn from the population of individuals at a large mid-western research university in the U.S. An individual enters our sample when it is the first time she has one of the following: received a grant or published a paper, and the university was listed as the institution of affiliation. We then track that individual from the time of entry till end of 2010. We gathered the grant data from university dean's office, which collects data on all grants received by the employees and students of the university. In addition using the Scopus database on scientific publications we collected the number of scientific publications that relates to a scientist in a year by matching on the first and last names of the scientists. Using this procedure we acquired all the scientific publications from 1970 through 2010. Those that are not in our sample and at the university in that period are employees and students who never published a paper or never got a grant in the sample period. There are 29,859 unique individuals in our sample.

Variables

Dependent variables

Federal funding. The first dependent variable is the total grants made by federal government agencies to a scientist in a three year forward window. We use three year window since federal grants are typically granted for 3 to 5 years. The modal researcher is likely to have one or less federal grants at any given point of time. Extending and reducing the grant window does not influence our results. The average grant size per scientist conditional on getting a federal grant was $605,773

Publications. The second dependent variable is the total count of publications that a scientists had in a three year forward window. Publications are the most widely used measure of output for tenure and promotion in academic. Publications are also the most commonly used measure of output by researchers in sociology of science (Allison and Stewart, 1974; Allison et al., 1982) and economics of science (Stephan, 1996) literatures.

Weighted Citations. The third dependent variable is the weighted count of citations received. One practical hurdle of using citations is truncation: older articles have had more time to be cited, and hence are more likely to reach the tail of the citation distribution. Moreover, disciplines could also vary in their propensity to cite prior art. To overcome these issues, we calculate this variable as follows. From Scopus database we first acquired the number of forward citations to every article written by a focal scientist. Given that citations vary by discipline and by publishing year cohort, we weighted the citations by the average for that discipline-year cohort. For every scientist's article, we calculated the weight by dividing the number of citations for that focal specific article by an average that represents the average number of forward citations for the relevant discipline-year pair. We then multiplied the actual number of citations for the focal article by 1+ weight to get the number of weighted citations for a particular article. For a scientist year pair, the dependent variable represents the sum of the weighted citations across all articles written by the focal scientist until the year of observation. For instance suppose a scientist had 10 and 5 forward citations for his article published in 2004 and 2005 respectively in the areas of cellular biology. Suppose the average number of forward citations for any article published in that area for those respective years were 2 and 4. The weights for years 2004 and 2005 will be 10/2=5 and 5/4=1.25 respectively. The weighted citations for this scientist for the year 2005 will be 10*(1+5)=60 and for the year 2006, will be 5*(1+1.25)=11.25. Since we cumulative weighted citations, the values for this focal scientist would be 60 for the year 2005 and 71.25 for the year 2006.

Explanatory variables.

Cumulative output. It is the total count of past publications of a scientist calculated until the focal year for that scientist. This variable we have argued is the simple, easily available and seemingly pertinent information that evaluators pay attention. This measure may capture the time varying quality differences between scientists and the social standing of the scientists. Hence as a robustness, following the insight of work that has tried to measure past perform and then strip the overlap between performance and status from a measure of status (Castellucci and Ertug, 2010), we explored the robustness of our principal results using an alternative measure of social ordering. This measure is constructed as follows. We follow a simple orthogonal transformation of past cumulative output from past productivity. Past productivity may not be as readily observed by decision makers as past cumulative output was. Hence past productivity is unlikely to influence social ordering. But past productivity is related to future productivity (Becker, 1993). Therefore we strip from the past cumulative output the common correlation with past productivity to arrive a orthogonal measure of social ordering. We label this variable orthogonal cumulative output. We report results using the orthogonal measure in the robustness section.

Newness of knowledge. We follow Azoulay, et al. (2011) and use the age of a Scopus keyword as our measure for newness of knowledge. A keyword is said to be born in the first year it appears in any article indexed by Scopus database. In essence, this measure captures the extent to which a scientist's research is novel relative to the world's research frontier. For every article published by a scientist, we first calculated age for each article. This is calculated as the difference between the year in which the keyword pertaining to the focal scientist's article was born minus the focal year. Thus smaller values of this year imply the relative newness of an article. When there were multiple keywords for an article we took the average age of all keywords referenced by that article. The variable that we use in our empirical analysis represents the average age of all articles authored by the focal scientist until the focal year. In regressions, for ease of interpretation we reverse code the variable to measure increasing values of newness of science, we transform this variable as standard deviations.

Specialist. We measure the width of knowledge by computing the degree of overlap of Scopus keywords corresponding to the focal scientist's articles The variable that we use in our empirical analysis represents the average overlap of all articles authored by the focal scientist in the immediate 3 years including the focal year In essence, an overlap of 1 indicates the perfect specialization of work around a single area and when overlap is close to zero, it reflects the spread-out nature of a scientists body of work. . Once again for purposes of easy interpretation, we transform this variable as standard deviations

Control variables.

Supply of federal grants: This variable represents the total grants made by the US federal government made to a specific discipline that a scientist is affiliated with in a year. Thus variation in this variable reflects the amount of total federal grants available to scientists belonging to a discipline across different U.S. universities. We identified the discipline(s) that a focal scientist relates to by using the department she was affiliated with, in a year. Interdisciplinary departments were mapped to multiple disciplines. For example, department of computer and electrical engineering was mapped to both computer science and electrical engineering.

Other grant measures: We control for grants measured in 1985 dollars made by other entities by including three more grant variables that reflect the grant made in the immediately preceding three year window by the University (University grants), for profit institutions (For profit grants) and other institutions (Other grants) that include grants made by a variety of heterogeneous grant institutions that are neither university, federal or for profit institutions.

Department grant measures: In addition we also control for the total grants over the immediately preceding three year period to the department that a focal scientist is affiliated to. As with grants made to individual scientists, department grants take four forms: grants made by a federal agency to departments (Department federal grants) grants made the University itself to departments (Department university grants), grants made by for profit institutions to departments (Department for profit grants) and grants made by other institutions to departments (Department other grants).

Cumulative total grants: In robustness analysis we control for the cumulative total grant dollars raised by the focal scientist until the focal year.

Time dummies: In all our empirical specifications we control for macroeconomic unobserved time effects using 35 time dummies. The left out year in all our specifications is a dummy variable that represents the year 1970.

Empirical Strategy

Time invariant quality and discipline. What are some issues that may confound us from observing or lead us to spuriously observe results that are consistent with test of our predictions? The first issue is scientists vary in their ability to raise grants which may also vary by the discipline that they work on -- disciplines may vary in the amount of funding to produce research output. To account for these sources of heterogeneity we use scientists fixed effects in all our estimations. Given that scientists typically belong only to one department the scientist fixed effects also accounts for heterogeneity in disciplines. Inclusion of department fixed effects in addition to inventor fixed effects do not alter our results. Inclusion of fixed effects also estimates within scientist effect over time. Given that it is also plausible that there might be unobserved period effects, we also include time dummies in all our estimations.

Time varying quality. The more complex problems are regarding unobserved quality, especially that of time varying quality. Since we control for unobserved differences in scientist ability using fixed effects in all our estimations what we are left to worry about is the unobserved time varying quality. To the extent that time varying quality is related to professional life cycle tenure of the scientists in the profession may account for some time varying quality. Studies on life cycles effect of scientists suggest that initially scientists have increasing productivity in early in their careers. This increase levels off and then drops towards the end of their career. All else equal, tenure should control for some time varying ability differences between scientists.

Self-selection to funding. In addition it is plausible that selection issues plague the identification of the effect of federal grants on performance. For instance, scientists with inferior status may not apply for federal grants As robustness, we account for this possibility using variation in the availability of federal government funding which varies by discipline and year This is based on the assumption that while an increase in overall federal funding available will likely influence the likelihood that a scientist applies for federal funding, it is likely to be orthogonal to output or performance. Hence as robustness we estimate a selection equation in which we use the total amount of federal grants for a discipline as an exogenous source of variation.

RESULTS

A brief description of the variables and summary statistics are reported in Table 1. Table 2 is the table with correlations between the variables used in this study.

-INSERT TABLES 1 &2 ABOUT HERE-

We start with a simple descriptive statistics to motivate our empirical analysis. In Table 3, we first explore how cumulative publications of a scientist is related to the amount of federal funding raised by the scientist during a three year period. To this end, we compare the amount of federal grants raised by scientists with "high" (higher than median amount of cumulative publications) and "low" (higher than median amount of cumulative publications) cumulative publications. Table 3 shows that a scientist with "high" cumulative publications on average raises about 0.22 million whereas as scientist with "low" cumulative publications on average raises only 0.01 million. Thus as hypothesized in H1 scientists with "high" cumulative publications raise more federal grants. Next, we compare the amount of federal grant raised by a scientist that pursues "new" science. To this end, we compare federal grant raised by a scientist with higher than median newness of knowledge ("New" category) with that of a scientists with lower than median newness of knowledge ("Old" category). Table 3 also shows that scientists that pursue "old" science categories on average raise more grant than those that pursue "new" science. Finally, those with diffused knowledge on average also raise less in grants than specialists.

-INSERT TABLE 3 ABOUT HERE-

While these estimates are indicative of what follows, they do not control for a variety of factors. Accordingly we now implement regressions.

Funding estimations. We test H1-H3b by estimating fixed effects regression in which we predict the total grants raised by a scientist in a 3 year period forward window that includes the current year and the two immediate following years. Model 1 of Table 4, contains all the control variables and it the baseline model. We start by testing how cumulative publications influence grants in Model 2 of Table 4. To this end we include cumulative publication. The coefficient of cumulative funding is positive and significant ($b=.257$; $p<.01$). This suggests that a standard deviation increase in cumulative publications (about 15.97 publications) increases the federal grant raised by a scientist by $ 220,711 dollars over a 3 year period. This result supports H1. In Model 3, we test H2a and H3a. H2a suggests that scientists that pursue new science should raise less federal grants while H3a suggests that specialists should attract more federal grants than those with wider knowledge base. To this end, we include keyword age and keyword overlap as additional covariates. Results of Model 3 support hypothesis 2a, the coefficient of new science is negative and significant (b=-.085; p<.01). This result suggests that a standard deviation increase in the newness of science decreases the amount of federal grant raised by a scientist by $72,998. Results of Model 3 support hypothesis 3a, the coefficient of specialists is positive and significant (b=.086; p<.01). Note that the hypothesis 3a predicted diffused which is inverse of specialist would be negative to funding. For ease of interpretation we use specialists to test the hypothesis. A standard deviation increase in the concentration increases the amount of federal grant raised by a scientist by about $73,856. These results support H2a and H3a which state that scientists that pursue old science and specialists should raise more federal grant than those that pursue new science and those with diffused knowledge.

-INSERT TABLE 4 ABOUT HERE-

Performance estimations. We now test our hypotheses on performance. We start with testing the consequence of funding decisions based on cumulative output of a scientist on the number of publications produced by a scientist in three year period that includes the current year and two years following the current year. We once again implement a fixed effects regression with time dummies. In addition we include a variety of controls that reflect the types of grants raised by the focal scientist – grants raised from UARF, for profit institutions and from a variety of other grant sources. We also control for the different types of department level funding using the four grant type dollars raised by departments. Finally we also control for the tenure of a focal scientist using tenure and tenure square. In Model 2 of Table 5 the joint effect of federal funding and cumulative publications is negative and significant (b=-.014; p<.01). This suggests that those who get high federal grants and have high cumulative publications are not as productive per dollar of funding received. This result suggests that a standard deviation increase in federal grant increases along with one standard deviation increase in cumulative publications decreases the publications produced by a scientist by 0.06 publications over a three year period. Thus this result supports H4b. In specification 4 of Table 5, we replicate specification 2, with forward citations as a dependent variable and get similar results. In essence, these results of funding and performance estimations indicate that although cumulative publications of a scientist enables her to raise more federal grants, but a federal dollar yields less in output for a scientist with higher cumulative publications.

-INSERT TABLE 5 ABOUT HERE-

Robustness Analysis

What other information could the decision makers have used? Gigernzer (1991) argues that the efficiency of heuristics: simple rules should be compared with an alternative decision strategy that could have been used by decision makers. Furthermore, this alternative decision strategy should have reasonable expectations of the decision makers' ability to process information. Since we cannot observe project quality what else can we recommend to decision makers in our context. Perhaps past productivity: output per dollar received is a good predictor of future productivity (Becker, 2009). Information on total grants raised by a scientist was collected by us from the university dean's office. It is possible that grant agencies could have asked this information of each applicant. Armed with information on total grants and with information on publications from the CV of a scientist a decision maker could calculate the past productivity: grant amount per for paper published. Larger values indicate relative inefficiency.

We test our insight in estimations reported in Table 6. For easy of interpreting the grant amount per publication as an efficiency measure rather than an inefficiency measure we reverse code the variable. Thus high value of this variable means that a scientist is more productive per dollar of grant money received. In Model 1 & 2 we estimate funding amount granted based on the past efficiency of the scientists using grant per publication-efficiency variable. This variable is not significant (b=.047; p<.36; in Model 2). Thus decision makers appear to pay no attention to past productivity in making grants.

In Model 3 & 4 we estimate the publications and weighted citations respective. We use the interaction of federal grant amount and grant per publication-efficiency interaction term to see if those with past productivity and receive high amount of grants are more productivity. We find that the coefficients of this interaction term is positive and significant for publications (b=.282; p<.01) and for citations (b=.322; p<.01). Thus this indicates if those with one standard deviation higher productivity efficiency along with one standard deviation higher federal grants would have had published 1.03 more papers and would have had about 5.37 more citations to their new work. Therefore we infer that by paying attention to easily available information: past output decision makers appear to ignore information that could have been accessed and proved to be a much better predictor of relative productivity

-INSERT TABLE 6 ABOUT HERE-

DISCUSSION

Resource Allocation: Formal Models

Formal models of decision making make a distinction between *ex ante* optimal allocation and *ex post* optimal allocation. The point being what is consider today as an optimal allocation given the information that available today on consumers taste and preferences does not become sub-optimal at later when consumer preferences have changed. However in our setting unlike tastes and fads of consumers we are evaluating the deservingness of scientific projects. Presumably change in consumer tastes have would have no significance to the causal relationships to scientific domain. Hence the *ex ante* allocation should not be systematically biased in ways which we can anticipate.

We however find evidence that is consistent with the idea that relative more resources are allocated to those with more output and this increases with uncertainty at the project level. This leads to a mismatch between funding and productivity. Had the decision makers relied on information on past productivity, which would have been collected by the program directors, then this mismatch would be ameliorated. In post analysis we find that indeed the current system does not pay attention to past productivity when making funding decisions. In Table 6 Model 2 past productivity has no effect on funding (b=.047;s.e=.049). Implying that decision makers ignored information on productivity: either because they did not possess this information or had the information but deliberately ignored it. In Model 3, Table 6 the joint effect of past productivity and funding is positively related publication performance (b=.282;s.e=.048). Furthermore, Model 4, Table 6 the joint effect of past productivity and funding is positively related citations received as well (b=.322;s.e=.048) This would be consistent with formal decision making models on making the best use of information that could have been collected *ex ante* with very little extra cost. The distributed structure of decision making: relying on community of peer evaluators who may not have complete information on past funding and who do not get feedback on the ultimate performance of the projects so that they can learn means that this system even in long run shows evidence of the mismatch. The evidence we found is also consistent with Arora and Gambardella (2005) who find that NSF does not fund early career economists to the extent their subsequent productivity would suggest. This is because early career scientists lack a detailed past record.

Micro Behavioural Decision Making Models

We had motivated the paper with the decision makers' reliance of simple, easily available and seemingly pertinent information. The literature in micro tradition of behavioural decision making makes a distinction between availability of information and the recognition heuristic. Availability is when information about a choice comes more readily to mind then the choice is judged to be more frequent in a population. Whereas recognition heuristic suggests if one of two objects is recognized and the other is not, then decision makers infer that the recognized object has higher value with respect to the criterion on which evaluation is based on (Goldstein & Gigerenzer, 2002). The distinction between the two may seem not so obvious to those not entrenched in this literature. But, this has led to heated exchange over the past decade (Marewski, Pohl, Vitouch, 2011).

What can our data say to this stream of literature? First, information about those who have higher cumulative publication is more available. When examining multiple proposals like the panels at NSF do it is possible that those proposals by those who have higher publications are more easily recognized than the unknown scientist with low cumulative publications. At this level we cannot tell apart recognition from availability. The proponents of recognition suggest that it should lead to efficient or as good outcomes as rules that used more complex information processing (Gigerenzer and Gaissmaier, 2011). Here we are able to say to this stream of literature whenever organizational pathologies lead for divergence between evaluators' perception of quality and innate quality then perhaps recognition heuristic would not work as well as more involved strategy.

We propose one such strategy that could have been easily employed the program directors at the funding agencies: use of information on past productivity. We find that past productivity involves one more calculation. Information on past productivity could have been easily gathered if the funding agencies included just a simple request for two units of information: total publications and total grants of an applicant with every application. Hence we suggest that in an open system when evaluators lack information on past productivity and decision makers do not use information on past productivity to make allocations such systems even in the long run may lead to sub optimal allocations.

Organization Status

We draw heavily from the literature on organizational to make predictions that argue that social ordering influence should be greater under uncertainty. We find evidence that is consistent with this prediction. By arguing that diffused output of a producer diminishes the influence of the social ordering in allocation we highlight a condition wherein social ordering effect may be muted. Again this is not something that is new to the literature. The joint effect of diffused output and social ordering being a negative moderator of the relationship between social ordering and allocation may have been anticipated by the literature. But to our knowledge we are the first to show such an effect. Thus we find evidence that is consistent with the prior arguments made in the organization status literature.

What is perhaps of more interest to this literature is the extent and conditions under which social ordering based allocations are beneficial to those with high and low status. Our results suggest that improvements made by those with higher social ordering and who receive larger funding is not sufficiently high enough to be as productive as those with lower grants and same social standing or lower social standing and smaller grants. Hence from a societal perspective allocation for scientific projects based on social ordering appear to be sub optimally allocated.

It is an open question what else would decision makers rely on if they cannot rely on social ordering as the market is rife with uncertainty. Some have argued that such markets fail (Akerlof, 1970). This may be the reason why society and not private firms fund research (Partha and David, 1994). What could be improved upon is the design of the system of funding. Some speculation could be that part of the evaluators group are blind to information on ordering of producers. Blind review has been extensively suggested as potential solution in the past only to be rejected. The reason offered that identify of scientists becomes apparent anyway to experts in the field. It is an open question if getting one of the several peer evaluators on a project to generate rating that is blind to information of the producer background would improve allocation efficiency.

Uncertainty, Social Ordering, and Performance

We had not made specific predictions on the performance of how the interaction effect of cumulative publications with funding on performance would change under sub samples by type of past output: high and low science; diffused and concentrated. The reason being we have no detailed ex ante guidance from the theory of simple rule based decision making or the social ordering literature on how the science opportunities would vary within new science and old science or work by concentrated and diffused scientists. There is wider body of literature on innovation that has studied this (Katila & Ahuja, 2002; Fleming &.Sorenson, 2004). For a Review of this literature see Klevorich et al.,(1995).

The most important evidence we have regardless of which sub-sample we focus on the relationship between joint effect of cumulative publications and funding is negative to performance. Thus within departments and within a time period and by scientist who do similar projects as measured by new science, old science, diffused and concentrated there is a negative relationship between the interaction term of cumulative publications and funding when predicting output. With that established here are results of unreported robustness estimations on the interaction of cumulative publications and funding in the sub-samples estimating performance. The coefficient of the interaction is negative and significant in high science age ($b=-.01$; $p<.01$); low science age ($b=-.04$; $p<.01$); diffused ($b=-.06$; $p<.01$); concentrated ($b=-.02$; $p<.01$) predicting performance. We get similar results in estimations predicting citations. Therefore the most basic evidence we can provide from this paper is that there is mismatch in all types of project settings, within a discipline, within a year, and with a host of other controls.

Policy Implications

There has been a fierce debate in the past on fairness of federal agencies granting research funding. One vicious attack has been that there is an “old boys” network where established scientists get more funding. This is alleged with support of the program directors who send proposals from established scientists to their friends. This has been firmly denied in a NSF initiated research of by Cole, Rubin, and Cole (1977). Our explanation and results suggest a much more benign cause for the mismatch than the “old boys” hypothesis. Our theory predicts that organization structure of evaluation in an open system, with open entry and exit of producers and evaluators, delayed and plausible lack of feedback loop about subsequent performance, the availability of social ordering information under conditions of uncertainty, and crucially the non-availability or ignoring past productivity information lead evaluators to be influenced systematically by social deservingness of the applicant. The influence of social ordering is magnified under conditions of uncertainty.

In our sample had the funding agencies collected and provided the evaluators data on past productivity of scientists this would have improved the allocation of funding. Past productivity is highly related to future productivity. But past output data is highly visible to evaluators in CVs of the scientists. Perhaps cumulative grants are not necessary disclosed. Applicants could have been asked to provide historic information on their total grants raised. Armed with this single piece of data and information on publications the evaluators have at their disposal a more effective indicator of scientist productivity (see Table 6: Models 3 & 4 wherein the interaction of productivity and funding is positive and significant).

While this advice works in our sample we are hesitant to propose this as a universal solution without new tests of the solution. Consider the case of initial tournament of cooperation between software programs in which “tit-for-tat” won. After the tournament, analysis of the data revealed that a different strategy than “tit-for-tat” would have performed much better. But when new competition was run with “tit-for-tat” , the new strategy, and other strategies; “tit-for-tat” still performed the best. This is due to the fact that “tit-for-tat” was the best strategy for a variety of environments. Over fitting a strategy to current environment may be counter productivity if future environments are different. We may be open to the same fallacy. Future studies should confirm if the rule we propose is indeed better in variety of new environments.

Limitations

Better measure of status. We had checked the robustness of our results with a productivity based measure of super stars (Zucker and Darby, 1996). Furthermore, we had tried to separate out signal of past productivity from social ordering by orthogonal variable transformation as used in a strand of status literature. We had conducted robustness analysis using the two new variables of past productivity and orthogonal measure of social ordering. Our results are largely similar for the theory variable.

A better measure would be awards of prestigious prizes in the discipline (Merton, 1968, Azoualy, Stuart and Wang, 2013). There were nearly 30,000 prizes in science in early 1990 and the number was growing (Zuckerman, 1992). Collecting information on prizes across all domains is almost an impossible data exercise. If information on prizes was hard to collect then an affiliation based measure based on the network of co-authors would also be preferable to an output based measure (Podolny, 1994). Again collecting data on the universe of scholars in each domain and calculating the affiliation preference of each scientist in our sample is also an almost impossible empirical exercise. But consider how the bias would have influenced our results. If a better measure of status then those with "truly high status" would have garnered more resources. They may also be just as productive or more productive than others with similar funding. Again this would lead us not to find effect for the mismatch. Thus having found the result it would suggest had we measured status better and indeed took out the more productive "truly high status" scientists then the mismatch would be even greater.

Project level data. We have no access to individual project level applications. We follow others who have collected aggregate measure at the level of the scientists and argued that lacking the former data this is an acceptable solution. The reason being scientist have research agendas than span multiple years. These research trajectories of scientists do not suddenly change (Dosi, 1982). We are cautions to use producer level data on the inputs and estimate producer level performance. This removes some concerns surfaced in the recent work that effect of social ordering may be overstated when social ordering is measured at the producer level and used to estimate project level performance.

Conclusion

This study adds to the literatures on formal and behavioural model of decision making and organization design under uncertainty. We study federal funding of 28,859 scientists from 1970 to 2005. We find support for our prediction that evaluation uncertainty leads to reliance on easily available and seemingly pertinent information: past cumulative publications. To the extent past cumulative publications capture social ordering this reliance we argue should increase with uncertainty at the project level. We were open to performance consequences of such social ordering influencing funding: it could be societally beneficial or sub optimal. Our results suggest the latter to be the case. We then explore the possibility or an alternate decision making strategy, which we suggest could be relatively easily implemented: past productivity based allocation. Our analysis confirms that past productivity is better ex *ante measure* of future productivity. Hence our study has implications for design of evaluation systems at organizational and societal level.

REFERENCES

Akerlof, G. A. (1970). The market for" lemons": Quality uncertainty and the market mechanism. *The quarterly journal of economics*, 488-500.

Allison, P. D., Long, J. S., & Krauze, T. K. (1982). Cumulative advantage and inequality in science. *American Sociological Review*, 615-625.

Allison, P. D., & Stewart, J. A. (1974). Productivity differences among scientists: Evidence for accumulative advantage. *American sociological review*, 596-606.

Arora, A., & Gambardella, A. (2005). The impact of NSF support for basic research in economics. *Annales d'Economie et de Statistique*, 91-117.

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.

Bardolet, D., Fox, C. R., & Lovallo, D. (2011). Corporate capital allocation: a behavioral perspective. *Strategic Management Journal*, *32*(13), 1465-1483.

Becker, G. S. (1993). Nobel lecture: The economic way of looking at behavior.*Journal of political economy*, 385-409.

Becker, G. S. (2009). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago Press.

Benartzi, S., & Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American economic review*, 79-98.

Castellucci, F., & Ertug, G. (2010). What's in it for them? Advantages of higher-status partners in exchange relationships. *Academy of Management Journal*,*53*(1), 149-166.

Cole, S., Cole, J. R., & Rubin, L. (1977). *Peer review and the support of science*. WH Freeman.

Denrell, J. (2003). Vicarious learning, undersampling of failure, and the myths of management. *Organization Science*, *14*(3), 227-243.

Dosi, G. (1982). Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change.*Research policy*, *11*(3), 147-162.

Etzkowitz, H. (2003). Research groups as 'quasi-firms': the invention of the entrepreneurial university. *Research policy*, *32*(1), 109-121.

Fleming, L., & Sorenson, O. (2004). Science as a map in technological search.*Strategic Management Journal*, *25*(8-9), 909-928.

Fox, M. F. (1983). Publication productivity among scientists: A critical review.*Social Studies of Science*, *13*(2), 285-305.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". *European review of social psychology*, *2*(1), 83-115.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, *62*, 451-482.

Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: the recognition heuristic. *Psychological review*, *109*(1), 75.

Greve, H. R. (2003). A behavioral theory of R&D expenditures and innovations: Evidence from shipbuilding. *Academy of Management Journal*, *46*(6), 685-702.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Katila, R., & Ahuja, G. (2002). Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of management journal*, *45*(6), 1183-1194.

Klevorick, A. K., Levin, R. C., Nelson, R. R., & Winter, S. G. (1995). On the sources and significance of interindustry differences in technological opportunities. *Research policy*, *24*(2), 185-205.

Kotha, R., George, G., & Srikanth, K. (2013). Bridging the mutual knowledge gap: coordination and the commercialization of university science. *Academy of Management Journal*, *56*(2), 498-524.

Lant, T. K. (1992). Aspiration level adaptation: An empirical exploration.*Management science*, *38*(5), 623-644.

Marewski, J. N., Pohl, R. F., & Vitouch, O. (2011). Recognition-based judgments and decisions: What we have learned (so far). *Judgment and Decision Making*, *6*(5), 359-380.

Merton, R. K. (1968). The Matthew effect in science. *Science*, *159*(3810), 56-63.

Partha, D., & David, P. A. (1994). Toward a new economics of science.*Research policy*, *23*(5), 487-521.

Podolny, J. M. (1994). Market uncertainty and the social character of economic exchange. *Administrative Science Quarterly*, 458-483.

Podolny, J. M. (2001). Networks as the Pipes and Prisms of the Market1.*American Journal of Sociology*, *107*(1), 33-60.

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: an effort-reduction framework. *Psychological bulletin*, *134*(2), 207.

Simon, H. (1957). A bounded-rationality model of rational choice. *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*.

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of personality and social psychology*, *48*(6), 1467.

Stephan, P. E. (1996). The economics of science. *Journal of Economic literature*, *34*(3), 1199-1235.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, *5*(2), 207-232.

Turner, S. P., & Chubin, D. E. (1979). Chance and eminence in science: Ecclesiastes II. *Social Science Information*, *18*(3), 437-449.

Wübben, M., & Wangenheim, F. V. (2008). Instant customer base analysis: Managerial heuristics often “get it right”. *Journal of Marketing*, *72*(3), 82-93.

Zucker, L. G., & Darby, M. R. (1996). Star scientists and institutional transformation: Patterns of invention and innovation in the formation of the biotechnology industry. *Proceedings of the National Academy of Sciences*,*93*(23), 12709-12716.

Zuckerman, H. (1977). *Scientific elite: Nobel laureates in the United States*. Transaction Books.

Zuckerman, H. (1992). The proliferation of prizes: Nobel complements and Nobel surrogates in the reward system of science. *Theoretical Medicine*, *13*(2), 217-231.

Zuckerman, E. W. (1999). The categorical imperative: Securities analysts and the illegitimacy discount. *American journal of sociology*, *104*(5), 1398-1438.

Zuckerman, H., & Merton, R. K. (1968). Age, aging, and age structure in science. *Aging and society/by Matilda White Riley and Anne Foner, in association with Mary E. Moore, Beth Hess, and Barbara K. Roth*.

Table 1: Description of the Variables

Variable	Description	Source of variation	N	Mean	Std. deviation
Federal funding 3 year forward	Cumulative 1985 dollars in million funded by a federal agency in the current year and two year after the current year, in standard deviations	Scientist, year	739,038	0.02	1.00
Publications 3 year forward	Cumulative publications in the current year and two year after the current year, in standard deviations	Scientist, year	739,038	0.01	1.00
Weighted forward citations 3 year forward	Cumulative forward citations, in standard deviations in the current year and two year after the current year, weighted by the average for the department. Calculated as forward citations for a scientists divided by the department mean number of forward citations for that year cohort.	Scientist, year	739,038	0.01	1.00
Federal funding 3 year	3 year cumulative 1985 dollars in million funded by a federal agency in the immediately preceding three year period in standard deviations	Scientist, year	613,577[a]	0.02	1.00
Cumulative publications	Cumulative publication until the current year, in standard deviations	Scientist, year	739,038	0.03	1.00
New science	Calculated as current year minus the year in which the keyword first appeared in our data, in standard deviations	Scientist, year	415,248[b]	-0.15	1.01
Specialist	Overlap in keywords between previous and current year. Overlap calculated as total overlapping keywords between the previous and current year divided by total keywords related to the focal scientist	Scientist, year	411,764[c]	-0.04	0.98
3 year university funding	3 year cumulative 1985 dollars in million funded by the university in the immediately preceding three year period.	Scientist, year	613,577[a]	0.002	0.044
3 year for profit	3 year cumulative 1985 dollars in million funded by for profit institutions in the immediately preceding three year period	Scientist, year	613,577[a]	0.003	0.07
3 year other	3 year cumulative 1985 dollars in million funded by other institutions in the immediately preceding three year period	Scientist, year	613,577[a]	0.007	0.10
Department university funding	3 year department level cumulative 1985 dollars in million funded by the university in the immediately preceding three year period	Scientist, year	736,204[d]	0.007	0.08
Department for profit grants	3 year department level cumulative 1985 dollars in million funded by for profit institutions in the immediately preceding three year period	Scientist, year	736,204[d]	0.002	0.10
Department grants from other	3 year department level cumulative 1985 dollars in million funded by other institutions in the immediately preceding three year period	Scientist, year	736,204[d]	0.35	1.36
Department federal grants	3 year department level cumulative 1985 dollars in million funded by federal agencies in the immediately preceding three year period	Scientist, year	736,204[d]	0.002	0.009
Total cumulative grants raised	Total cumulative grants in million raised by the focal scientist until the current year	Scientist, year	613,577[a]	0.46	3.92
Tenure	Number of elapsed calendar year from year of year of joining (2 years from year of first publication or grant)	Scientist, year	739,038	13.09	7.89
Tenure squared	Square of tenure	Scientist, year	739,038	233.71	280.33
Grant per publication	Grant per published paper in millions of dollars, in standard deviations	Scientist, year	613,577[a]	-0.04	1.00
Year dummies	35 year dummies one for each year for years 1970 through 2005	Year	-		
Scientists fixed effects	29,859 scientist fixed effects for the funding and 2582 fixed effects for performance regressions	Scientist	-		

Notes: [a] We started with a panel of all scientists that worked for the university. For about 125,461 scientist years we were not able to match publications with grant data.
[b] For about 323790 scientist years we were not able to acquire Scopus keywords.
[c] For about 3484 scientist years we were not able to acquire Scopus keywords for two consecutive calendar years.
[d] We could not conclusively map about 185 scientists comprising of 2834 scientist years to a department.

Table 2: Correlation matrix

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	Federal grants 3 year forward	1																
2	Cumulative publications	.18	1															
3	Specialist	.07	.35	1														
4	New science	-.03	-.14	-.22	1													
5	Tenure	.04	.25	.09	.09	1												
6	Tenure squared	.04	.27	.08	.08	.94	1											
7	3 year university	.12	.21	.06	-.03	.04	.05	1										
8	3 year for profit	.16	.19	.06	-.03	.03	.04	.03	1									
9	3 year other	.11	.14	.06	-.03	.04	.04	.11	.14	1								
10	Department university funding	.01	.02	.00	.00	.02	.02	.02	.00	.01	1							
11	Department for profit grants	.01	.00	.00	-.01	.01	.01	.00	.03	.01	.00	1						
12	Department grants from other	.00	.00	.02	-.09	.02	.02	-.01	.01	.00	-.02	.00	1					
13	Department federal grants	.00	.00	-.01	.00	.00	.00	.00	.00	.00	.00	.01	.11	1				
14	Total cumulative grants raised	.39	.41	.14	-.06	.14	.15	.18	.22	.23	.02	.01	.00	.00	1			
15	3 year federal grants	.28	.21	.08	-.04	.05	.05	.13	.15	.16	.02	.01	.00	.00	.70	1		
16	weighted citations	.19	.39	.17	-.08	.05	.05	.19	.15	.16	.01	.00	.00	.00	.24	.18	1	
17	Publications	.20	.57	.21	-.08	.04	.05	.19	.18	.18	.02	.00	-.01	.00	.27	.20	.63	1
18	Grants per publication	.02	.00	.00	.01	.01	.01	.01	.01	.04	.01	.00	.00	.00	.06	.03	.01	.01

Table 3: Federal funding, publications and weighted citations

	Federal funding in millions of $		Difference (High - low)
Cumulative pubs High versus low	0.107	0.003	0.104***
	(0.00)	(0.001)	(0.002)
New versus old science	0.028	0.077	-0.049***
	(0.002)	(0.002)	(0.002)
Specialist versus diffused knowledge	0.175	0.019	0.157***
	(0.002)	(0.002)	(0.003)

Notes: ***Significant at 1 percent. ** Significant at 5 percent. * Significant at 10 percent All specifications use 1040 observations. High cumulative publications represents scientist years with higher than sample median number of cumulative publications. New science represents scientist years with higher than sample median value of the variable New science. Specialist represents scientist years with higher than sample median value of the variable Specialist

Table 4: Fixed effects regressions of funding amounts, dependent variable three year forward federal grants raised in standard deviations

	Model 1	Model 2	Model 3	Model 4
Cumulative publications (SD)		**0.257*** **		**0.308*** **
		(0.004)		**(0.007)**
Specialist (SD)			**0.086*** **	**0.023*** **
			(0.004)	**(0.005)**
Cumulative pubs X specialist				**0.015*** **
				(0.002)
New science (SD)			**-0.085*** **	**-0.016*** **
			(0.005)	**(0.006)**
Cumulative pubs X New science				**0.106*** **
				(0.006)
Control variables				
Tenure	0.007***	0.003	0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Tenure squared	-0.000	-0.000***	-0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
3 year university funding	1.645***	1.410***	1.640***	1.416***
	(0.043)	(0.043)	(0.043)	(0.043)
3 year for profit	2.681***	2.497***	2.669***	2.474***
	(0.032)	(0.032)	(0.032)	(0.032)
3 year other	0.055**	0.071***	0.051**	0.062***
	(0.022)	(0.021)	(0.022)	(0.022)
Department university funding	-0.037*	-0.025	-0.031	-0.025
	(0.021)	(0.021)	(0.021)	(0.021)
Department for profit grants	0.022	0.025	0.022	0.026
	(0.019)	(0.019)	(0.019)	(0.019)
Department grants from other	-0.004**	-0.004**	-0.004**	-0.004**
	(0.002)	(0.002)	(0.002)	(0.002)
Department federal grants	-1.056	-0.755	-0.889	-0.862
	(1.882)	(1.871)	(1.886)	(1.876)
Total cumulative grants raised	-0.084***	-0.105***	-0.087***	-0.105***
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	-0.101**	0.041	0.041	0.104**
	(0.041)	(0.041)	(0.042)	(0.042)
Observations	739,038	739,038	411,764	411,764
Number of scientists	39,967	39,967	29,662	29,662
Within R-Squared	0.05	0.06	0.05	0.06

Notes: ***Significant at 1 percent. ** Significant at 5 percent. * Significant at 10 percent. All specifications include 35 time dummies

Table 5: Fixed effects regressions of performance, dependent variables three year forward publications and weighted forward citation in standard deviations

	Model 1	**Model 2**	**Model 3**	**Model 4**
	Publications	*Publications*	*Weighted citations*	*Weighted citations*
3 year federal funding (SD)	0.026***	0.065***	0.035***	0.066***
	(0.004)	(0.006)	(0.005)	(0.006)
Cumulative publications (SD)		-0.093***		0.058***
		(0.010)		(0.010)
Fed grant X cumulative pubs		**-0.014*****		**-0.013*****
		(0.002)		**(0.002)**
Control variables				
3 year university funding	0.018***	0.022***	0.024***	0.023***
	(0.002)	(0.002)	(0.002)	(0.002)
3 year profit funding	0.665***	0.807***	0.784***	0.801***
	(0.067)	(0.068)	(0.068)	(0.069)
3 year other funding	0.332***	0.323***	0.427***	0.425***
	(0.045)	(0.045)	(0.045)	(0.045)
tenure	0.023***	0.037***	0.028***	0.018***
	(0.005)	(0.005)	(0.005)	(0.006)
tenure square	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Department grants from other	-0.003	-0.003	0.029**	0.027*
	(0.014)	(0.014)	(0.014)	(0.014)
Department for profit grants	0.030	0.022	0.073	0.069
	(0.092)	(0.092)	(0.093)	(0.093)
Department university funding	-0.270***	-0.285***	-0.167*	-0.164*
	(0.092)	(0.091)	(0.093)	(0.092)
Department federal grants	29.898	29.960	-16.939	-14.357
	(37.259)	(37.128)	(37.704)	(37.631)
Constant	2.297***	2.364***	0.587***	0.569***
	(0.166)	(0.166)	(0.168)	(0.168)
Observations	21,705	21,705	21,705	21,705
Number of scientists	2582	2582	2582	2582
R-squared	0.04	0.05	0.03	0.04

Notes: ***Significant at 1 percent. ** Significant at 5 percent. * Significant at 10 percent. All specifications include 35 time dummies

Table 6: Fixed effects regressions with grants per publication as independent variable, dependent variables three year forward federal funding, three year forward publications and weighted forward citation in standard deviations

	Model 1 *Fed funding*	**Model 2** *Fed funding*	**Model 3** *Publications*	**Model 4** *Weighted citations*
3 year federal funding (SD)	-	-	-0.025***	-0.024**
			(0.010)	(0.010)
Grant per publication-efficiency (SD)	**0.048**	**0.047**	0.143**	0.036
	(0.037)	**(0.049)**	(0.057)	(0.057)
Fed X Grant per publication-efficiency (SD)			**0.282*** **	**0.322*** **
			(0.048)	**(0.048)**
Control variables				
Specialist (SD)		0.049**		
		(0.021)		
Grant per pub-efficiency X specialist		-0.031		
		(0.075)		
New science (SD)		-0.017**		
		(0.008)		
Grant per pub-efficiency X New science		0.021		
		(0.024)		
Tenure	0.009***	0.007***	0.021***	0.027***
	(0.002)	(0.002)	(0.005)	(0.005)
Tenure squared	-0.000***	-0.000***	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
3 year university funding	1.158***	1.150***	0.018***	0.023***
	(0.043)	(0.044)	(0.002)	(0.002)
3 year for profit	2.150***	2.136***	0.636***	0.751***
	(0.032)	(0.032)	(0.067)	(0.068)
3 year other	-0.204**	-0.210**	0.334***	0.430***
	(0.022)	(0.022)	(0.045)	(0.045)
Department university funding	-0.030	-0.030	-0.270***	-0.170*
	(0.022)	(0.022)	(0.091)	(0.093)
Department for profit grants	0.019	0.019	0.023	0.066
	(0.019)	(0.019)	(0.092)	(0.093)
Department grants from other	-0.004***	-0.004***	-0.004	0.028*
	(0.001)	(0.001)	(0.014)	(0.014)
Department federal grants	-1.646	-1.598	29.970	-16.838
	(1.913)	(1.919)	(37.219)	(37.664)
Constant	-0.098**	-0.054	2.254***	0.562***
	(0.043)	(0.044)	(0.166)	(0.168)
Observations	414,479	411,764	21,705	21,705
Number of scientists	29,856	29,662	2582	2582
Within R-Squared	0.02	0.03	0.04	0.03

Notes: ***Significant at 1 percent. ** Significant at 5 percent. * Significant at 10 percent. All specifications include 35 time dummies