The Impact Of NSF Support For Basic Research In Economics

Ashish Arora
Heinz School of Public Policy
Carnegie Mellon University
Pittsburgh, PA 15213
ashish@andrew.cmu.edu

Alfonso Gambardella
Dept. of Economics
University of Urbino
Urbino, Italy

November 1996

The is paper draws upon a report submitted to the economics program at the NSF by Arora. We are indebted to Ruth Williams of the NSF for help in getting access to the data, and for patiently answering my questions and queries. Dan Newlon and Lynn Pollnow illuminated us about the many intricacies of the NSF grant procedures, and I thank them for their help and support. We very greatful to Paul David long and stimulating conversations, and to Seth Sanders and Dan Black for helpful suggestions and advice. Wei Kong provided enthusiastic and skillful research assistance. Data collection analysis was supported by a grant from the Heinz School.

ABSTRACT

This paper studies the relationship between NSF funding and the publications of US economists using data on 1473 applications to NSF during 1985-1990, 414 of which were awarded a research grant. We first outline a basic methodology for assessing the impact of the NSF support for basic research in Economics. In doing so, we shall also point to key conceptual and measurement problems. Second, we provide empirical evidence about the factors that influence the NSF allocation decision, the effects of this decision on the production of publications, and the extent to which these effects differ among researchers at different stages of their career.

Our main findings are as follows:

RESOURCE ALLOCATION: The results show that past performance affects expected budget received in two ways. The first effect is through the probability of being selected. Past performance affects the expected budget granted in an indirect way as well. The data indicate that PIs with better track records ask for larger budgets. Budget asked affects the amount received given selection, a PI with a better track record has a higher probability of getting selected, and given selection, gets a constant fraction of a larger requested budget.

IMPACT OF FUNDING: The effect of NSF funding seems to be more pronounced at earlier stages of the career of economists. The marginal product of NSF funding is greater for young PIs. The estimated elasticity is about 0.7. As an example, this means that for a selected young PI, an additional $10,000 grant would produce 12 more (quality-adjusted) publications units. In turn, this corresponds to one single-authored paper in a journal like the Review of Economics and Statistics. The marginal product of NSF funding (given selection) is practically zero for intermediate cohorts (between 5 and 15 years since PhD), and it is slightly higher for senior researchers.

1. INTRODUCTION

This paper studies the relationships between NSF funding and the publications of US economists. It investigates the factors that influence the NSF allocation decision, the effects of this decision on the production of publications, and the extent to which these effects differ among researchers with different characteristics. The paper focuses on questions like—What determines NSF selection of proposal? Given selection, what determines the amount of grant supplied? On average, do PIs who were selected by NSF produce more than other PIs with similar characteristics who were not selected? Given selection, what is the marginal product of NSF funds?

Today little is known today about the productivity of scientific research, the marginal product of public research funds, or the selection process in research grant applications. Therefore, evidence about these facts is of interest. However, in order to interpret these facts for prescribing policy, one needs more. One needs a structural model of NSF behavior, the decision of researchers to apply for grants, and the "production function" of publications, such as the one developed in Arora, David, and Gambardella (1996). However, such models are complex and require rich data. Although the data available on the NSF computer database are very detailed, our research convinced us that sensible structural modeling would have to await richer data.

This paper, therefore, does not present a formal structural model that one would need to answer some policy questions of interest. We do discuss the sorts of assumptions one would have to make in order to derive meaningful policy implications from the results presented here. One goal of the paper is to single out some of the relevant problems in this field, show how they would affect the interpretation of our results, and how further research might solve some of these problems.

The next section describes the data and discusses some important caveats. Section 3 discusses important conceptual issues and the limitations of the analysis. Section 4 looks at the selection of proposals by NSF, and the determinants of grant received. Section 5 focuses on the effects of NSF funding on the production of publications. Section 6 summarizes the main findings and concludes the paper.

2. DATA AND LIMITATIONS OF THE DATA

The paper uses data on 1473 applications to NSF during 1985-1990. Of these slightly less than one third (441) were selected for funding. These data can be classified in three categories:

i) *characteristics of the principal investigator (PI) or of the proposal*: PI's sex and institution (name of university or other organization); type of institution (PhD granting school, non-PhD school, organization, etc.); years since PhD; quality-adjusted number of publications of the PI in the five year window before application; budget requested; number of co-PIs in the proposal; quality-adjusted number of publications of each co-PI in the five year window before application; NSF reviewers' score for the proposal.

ii) *NSF decision variables*: index of selection; amount of NSF grant.

iii) *production of publications*: quality-adjusted number of the PI's publications during the five year window after application; quality-adjusted number of publications of each co-PI in the five year window after application.

All data, apart from publication output, were supplied by NSF. The data on publications were obtained from a search on Econlit. Each publication was weighted by the impact factor of the top 50 economic journals published in Liebowitz and Palmer (1984: table 2) to adjust for quality. The journals range from the Journal of Economic Literature (100) to the Journal of Development Economics (2.29). If a journal was not in the top 50, the publication was given an impact factor of 1. The number of publications was adjusted for co-authorship by dividing the impact factor of the publication by the number of co-authors—e.g. a paper co-authored with two other people published in JEL would count 33.3. The Appendix describes these data in greater detail. It also reports the list of the top 50 journals along with their impact factors. Table 1 defines all the variables that will be used in this analysis. Table 2 reports descriptive statistics.

Two important aspects of the empirical set up should be noted.

Unit of analysis:

The unit of analysis is a proposal (a grant application). A little more than 10% of the proposals had co-PIs listed, and the percentage showed a slight increase over time. In the results presented here, we ignored co-PIs. Specifically, we ignored the effects of co-PI characteristics on selection. We also ignored the publication output of co-PIs.[1] We did use the number of co-PIs as control variables. However, to the extent that some co-PIs are successful PIs in their own right in proximate years, this potentially creates measurement problems that we do not address.

Quality of Journal:

Weighting publications by citation is customary but also time consuming and expensive.[2] We choose to weight by journal quality. The measure journal quality used here is based on weighted citations, adjusted for the length of the journal. Citations are weighted by the quality of the journal where they occur. In other words, both the citation weights and the journal impact factor are determined jointly through an iterated procedure, described in Liebowitz and Palmer (1984).

3. CONCEPTUAL ISSUES, CAVEATS, AND QUALIFICATIONS

Social Impact:

[1]Estimates produced by using the total output of all the name researchers on a proposal did not give qualitatively different results for the production function of publications. Accordingly, as a first cut at the data, ignoring the publication output of co-PIs seemed reasonable.

[2]Our output measure refers to publications between 1985-90 for the earliest round of proposals, and between 1990-95 for the latest round.

Exercises such as ours can say very little about the social value, or even the social impact, of NSF-funding (or, for that matter, of any other public research agency). In using scientific publications to measure scientific output, one is faced with an obvious constraint: The aggregate output of publications is constrained by the growth rate of journals. Thus, at the aggregate level, one cannot say how different allocations of resources would increase the total scientific output.[3] To take a simple example, if there are only 5 economic journals in the US which published 20 papers per year, different allocations of resources will always produce 100 papers per year. The constraint in the number of journals makes the aggregate "social welfare" analysis of science meaningless, unless one comes up with (meaningful) measures of scientific output that are different from the number of publications.[4]

Although the question "By how much would aggregate output increase if one increased research funding by 10%?" cannot be meaningfully answered, the question "By how much would the output of Prof. X increase if NSF were to increase Prof X's grant by 10%?" From the viewpoint of the individual researcher, whose career prospects depend crucially on publication output, this is an interesting and important question. The question is also important for public funding agencies such as the NSF. *How much* money should be spent is not the only decision that NSF must make. Equally, if not more, important is the question of *how* a given amount of money should be spent. The impacts on individuals are an important component of the decision on how the NSF budget should be allocated.

Functional form:

Regression analysis also imposes functional form restrictions. Therefore, we show both non-parametric (differences of conditional means) as well as more conventional regression results. Regressions are necessary because even parsimonious representations of characteristics in the non-parametric analysis led to cells with few observations. We tried regressions with both log and level specifications. The results are robust to the choice of specification. We paper the log specification here.

Unobserved Heterogeneity:

A fundamental problem in studies such as the one presented here is that it is not easy to distinguish between the different effects that are being sought. For instance, in assessing the impact of factors that determine selection or the quantity of funds supplied, one needs to find variables that affect selection and not the amount of budget granted, and vice versa. This is the classical problem of identification. The problem is particularly difficult when we try to measure

---

[3]For an exercise based on the opposite point of view, see Adams and Grilliches (1996).

[4]This would be a problem even if publications were weighted by citations. The reference list of any economic article is of approximately constant length. Hence, the number of citations is also constrained by the number of journals (assuming fixed number of pages per journal). Although US based economists are not the only ones publishing in journals, they account for a large fraction of English language journals. See also Arora, David, and Gambardella (1996).

the impact of NSF support. Simply put, it is difficult to find variables that would affect selection (or the budget), but not publication output. The point is straight forward. Suppose that NSF selects more able researchers from amongst a set of otherwise similar researchers. (Indeed, the NSF is supposed to fund more promising projects, and these are likely to be proposed by more able researchers.) These researchers will, on average, be more productive. Only a (small) part of their greater productivity is properly attributable to NSF; the rest is due to their greater ability. If, as is likely, ability is imperfectly measured, regression estimates would tend to overstate the impact of NSF selection on output.[5] There are other aspects to the issue of unobserved heterogeneity, such as those relating to non NSF resources. These issues are further discussed in section 4 below.

[5]For a more extensive discussion of these problems in the economics of science, see Arora, David, and Gambardella (1996).

4. SELECTION AND GRANT RECEIVED

■ *Selection.* To examine selection, it is useful to start by comparing, in table 3, the sample means of all our variables conditional upon selection (AWARD=1) and non-selection (AWARD=0). These are suggestive of the correlations between selection and characteristics of the PIs.

As expected, there is a sizable difference between the sample means of PAFTER (total five year publication output after year of application) and PBEFORE (total five year publication output before year of application) conditional upon selection and non-selection. On average, PAFTER for a *selected* PI is 101.2, while for a *non-selected* PI is only 51. Similarly, the sample means of PBEFORE are respectively 108.6 and 51. *Selected* PIs are then more productive, both before and after selection.

As far as different age cohorts are concerned, the shares of ASSOC and PROF are higher in the *selected* sample than in the applicant sample as a whole, and that of ASSIST lower. Senior professors (SPROF) have the same share in *selected* sample as in the total sample of applicants. The percentage of MALE is not correlated with selection. (Female PIs account for only 8% of the total sample.)

The average SCORE of a *selected* proposal is 1.93 versus 2.78 for a *non-selected* proposal. Further analysis on the SCORE variable revealed that practically all the proposals with an average reviewer score greater than 3 (about 25% of total applications) were rejected. By contrast, all the proposals with an average reviewer score of 1 (i.e. all reviewers indicated that the proposal was outstanding) were selected. (These proposals however are only 1% of total applications.) Finally, more than 80% of the proposals with average score between 1 and 2 (about 22% of total) were selected, while only 20% of the proposals with average score between 2 and 3 (about 50% of the total sample) were selected. This points very clearly to the high correlation between SCORE and selection. It also suggests that for proposals with an average score around 2-2.5, the NSF panel has considerable discretion. Put differently, for a substantial fraction of proposals, PI's characteristics and other factors matter.

Table 4 shows the results of a probit regression with AWARD as the dependent variable. The table confirms the high correlation between SCORE and AWARD. It also suggests that, given SCORE, other factors influence selection. For instance, PBEFORE has a sizable and statistically significant impact for the three cohort groups—ASSOC, PROF, and SPROF. By contrast, it has a negligible impact for ASSIST. This is not unexpected, as past publications of younger PIs are a less informative signal of their ability. Other things being equal, PIs from ELITE institutions are more likely to be selected.[6] The interpretation of this could range from unobserved ability, greater contacts with the NSF, to simple reputation effects.

---

[6]This is a somewhat arbitrary list of 10 universities: Harvard, Stanford, MIT, Chicago, Princeton, Yale, Columbia, UC Berkeley, Wisconsin, and Northwestern.

■ *Grant Received.* Given selection, do quality and reputation, or the other characteristics of the PIs also affect the amount of grant received? To examine this question, the paper presents the results of a sample-selection, maximum likelihood estimation of log(REC_DOL) on various PI characteristics. In the analysis presented below, the specification of the probit-selection equation (AWARD) is the same as the one shown in the third column of table 4. The dependent variable of the regression for the selected sample is the natural logarithm of REC_DOL.

In thinking about this problem, there are two issues that have to be taken into account. First, as noted earlier, in order to identify the factors that determine selection separately from those that determine the amount of the grant, one is to find variables that affect the former and not the latter. We use SCORE as the variable that affects selection, and not the actual size of the grant. In essence, we assume that NSF uses the reviewer scores only in selecting proposals, but not in deciding what fraction of the requested budget to fund for the selected proposals.

The second issue is related. The decision about how much money to grant to each selected proposal is very likely to depend on the budget requested by the proposal. More generally, the budget requested provides a measure of the scale of the research that each PI is prepared to perform if selected. This would suggest that the expected grant received per proposal is influenced by: a) the factors that affect selection; b) the factors that affect how much money the applicants ask for in their proposals. In turn, the latter is a choice variable, and reflects how the NSF system is perceived to work.

In practice, factors that affect the budget requested are likely to be correlated with the same factors that influence selection. Upon regressing the log of RQS_DOL on the same variables that influence selection (not shown here), the former is correlated with PBEFORE, ELITE, and with senior PIs. In other words, senior PIs, or PIs from ELITE institutions, or with better track records, ask for larger grants. To examine these issues, table 5 presents two basic specifications of the sample selection estimations. Both use AWARD as the dependent variable of the probit equation and log(REC_DOL) as the dependent variable of the selected sample. The probit regressors are the same as those used in column 3 of table 4. (Since the results of the probit regression in the sample selection estimation are similar to those in table 4, they are not reported in table 5.)

In the first column of table 5, the regressors of the REC_DOL equation are the PI characteristics analyzed in table 4, column 3, except SCORE. In the second column, log(RQS_DOL) is added to the set of regressors of the *selected* equation, and in the third column log(RQS_DOL) is interacted with the cohort dummies—ASSIST, ASSOC, PROF, SPROF. In short, in the first column of table 5, the analysis does not attempt to distinguish between the factors that influence REC_DOL directly and those that influence REC_DOL through RQS_DOL. In the second and third columns, the analysis investigates whether the various PI characteristics such as PBEFORE, are correlated with REC_DOL even after one controls for RQS_DOL.

The first column of table 5 shows that PBEFORE affects the size of the grant received, but the effect of this variable does not change significantly with the cohort of the PI. Moreover, the effect of PBEFORE on the grant seem to be fairly small. This suggests that PBEFORE is relatively more important for selection than for the amount of grant given selection. Among the set of selected proposals, a PI with twice as many past publications as another PI, obtains a grant that is only 4-6% higher than the latter. The second column of table 5 shows that RQS_DOL has a positive and significant effect on the amount of the grant. Among the set of selected proposals, a PI with a requested budget twice as big as that of another selected PI receives a grant that is 35% higher. The third column of table 5 suggests that this effect differs across cohorts. A junior PI (ASSIST) with a requested budget twice as large as another's, obtained on average, a grant 80% higher. This percentage declines to 70%, 43%, and 23% for ASSOC, PROF, and SPROF respectively.

Interestingly enough, when using RQS_DOL among the regressors the correlation between the selection and the selected equation becomes statistically less significant. This suggests that RQS_DOL may take into account some of the unobserved heterogeneity present in the previous regressions. The effect of past publications also becomes insignificant. PBEFORE may then have an indirect rather than direct influence on the size of grants given selection. PIs with better track records ask for more money, and it is this, rather than a direct effect of past publications, that affects the size of grants. This also points to the fact that future studies should carefully model the determinants of dollar amount requested.[7]

5. PRODUCTION OF PUBLICATIONS

■ *Indirect effects of NSF funding and unobserved heterogeneity.* To estimate the production function of publications, and the effect of NSF funding, one has to deal with two issues.

The first one is that the resources of our PIs other than the NSF grant are not observed. For the unselected proposals, this means that the resources of their PIs are not observed at all. One can only measure the non-NSF resources using observed PI characteristics. Therefore, non-NSF resources are measured with error. The problem arises because this measurement error may be correlated with NSF funding. For instance, NSF grants could be a signal to the overall community about the ability of a given PI, and this may encourage additional funding from other sources (*"leveraging*"). Alternatively, NSF funding may "*crowd out*" other funding. Ultimately, this means that it is not possible to distinguish between the direct and indirect effects (leveraging; crowding out) of NSF funding on publication output.

The second problem, already discussed at some length, is the one about possible sources of "unobserved heterogeneity" in the production function of publications. If the NSF selection decision is based upon expected future publication output, then when regressing publications on PI characteristics and, say, an index for selection like AWARD, the estimated coefficient of the latter has an upward bias.

---

[7]The result that budget granted is influenced by past publications through budget asked emerged also from the study conducted by Arora, David, and Gambardella (1996), who examine a completely different sample of scientific research projects in molecular biology in Italy.

Both of these are serious problems, but ones that we can do little about. In the empirical analysis discussed below, the first problem is ignored. Instead of trying to distinguish between the direct and indirect effect of NSF funding, the analysis looks at the overall, net effect of NSF. The second problem is addressed largely, but not very satisfactorily, through the choice of functional form: We tried the Heckman selection correction which assumes linearity in the specification and normality of the errors.

■ *Non-parametric analysis.* To test the robustness of the linear specification we begin by comparing the performance of selected proposals with a "matched" sample of similar but unselected proposals. The characteristics along which proposals were matched are (i) age cohort of PI; (ii) SCORE; (iii) Institution (ELITE). To further reduced potential sources of heterogeneity, only male PIs were considered. This amounted to dividing 1330 PIs into one of 40 categories. So for example, E1ASSIST is the category of PIs from elite institutions, who obtained their PhD no more than 5 years before the grant, and whose proposal received a score of 1.

Within each category, we compared the selected and unselected proposals. However, many of the categories had insufficient number of observations for some cell. From the original 40 categories, only 16 of them, those in which both the awarded and non-awarded subcategory had at least 4 cases, were selected for further analysis. This amounted to selecting only categories with SCORE equal to 2 or 3, because practically all the PIs with SCORE = 1 were awarded, and practically all PIs with SCORE higher than 3 were rejected. To further reduce heterogeneity, the median value of PBEFORE was calculated for each category, and each category was further subdivided into 2, based on whether PBEFORE was above or below the median. This left us with 32 categories.

For each proposal in each of the 32 categories the difference between PAFTER and PBEFORE was calculated (GROWTH = PAFTER - PBEFORE), and then we calculated the average of GROWTH in each category for awarded and non-awarded proposals. We tested whether the differences are significant using *t* tests. The hypothesis that the variances of awarded group and the variances of unselected group are equal was also tested to determine the use of t-test with equal or unequal variances.

The detailed results of this analysis are reported in the Appendix, table A1. Here we just summarize these results. In 20 of the 32 categories, the average of GROWTH in the awarded group is higher than that in the non-awarded group. Of these, the difference is statistically significant at 0.1 level in just four categories—namely LOW E2PROF, LOW NE2SPROF, HIGH NE3ASSIST, and LOW NE3SPROF. In the remaining 12 categories, the average of GROWTH in the awarded group is lower than that in the non-awarded group. Only in one category, HIGH NE2PROF, the difference is significant at the 0.1 level.[8] In sum, the non-parametric analysis

[8]Similar analyses using different types of categories were also performed. For instance, the 1330 male PI proposals were divided into categories according to age cohort, and PBEFORE, but using SCORE classes 1-2, 2-2.5, 2.5-3, 3-4, greater than 4. Ultimately, 24 categories were obtained in which there were at least 4 cases of awarded and non-awarded proposals. Of these, 17

suggests that there is some evidence that the awarded PIs exhibit a higher growth in publications. However, each of the categories that were obtained through this procedure had a low number of observations, and this increased the standard errors.

Moreover, as one can see from table A1, publication output shows strong evidence of regression to the mean. Put differently, PIs with higher than average performance in the past are more likely to show lower than average performance in the future. This may be a life cycle effect. It may also be a timing issue. If NSF funding decisions are based in part on past performance then proposals getting funded may be from PIs who have published a lot, and hence, have less research in the pipeline. In either case, this is a serious issue because in virtually every category where non-selected proposals show higher than average growth, they also show lower than average past publications. In other words, one must control for the level of past performance more accurately than the simple non-parametrics will allow. Accordingly, we move to regression analysis.

■ *Regression analysis: OLSQ.*

Table 6 reports estimates from an OLSQ regression of log(PAFTER) using all 1473 observations in the sample. Apart from various PI characteristics (MALE, ELITE, PHD, NOPHD, age categories), the regressors included log(PBEFORE) interacted with cohort dummies, and (1-AWARD) also interacted with cohort dummies.[9] Table 6 suggests that NSF awards are significantly correlated with log(PAFTER). Moreover, the estimated coefficients of (1-AWARD) suggest that the effect of NSF grants is more pronounced for junior PIs, and it progressively declines with seniority.

With the caveat that we are ignoring the issue of unobserved heterogeneity, some simple experiments using the estimated coefficients of (1-AWARD) may be helpful. A typical proposal from a junior PI (ASSIST) would produce only 62.5% of the publications she would produce if awarded. Since the sample mean of PAFTER for awarded ASSIST PI is 122.3, this implies that her output would be lower by 46 units, which corresponds roughly to two publications in the Journal of Economic Theory. Similar experiments can be performed for the other cohorts. For instance, the sample means of awarded ASSOC and PROF are respectively 107.7 and 84.2. This implies that, if not awarded, an awarded ASSOC PI whose $PAFTER_O$ is equal to the sample mean would produce 31 fewer quality adjusted publication units—roughly equal to one American Economic Review paper. Similar calculations for PROF PI imply a reduction of 18.6 publication units, equivalent to one publication in the Journal of Econometrics. The typical senior PI would produce 5.5 fewer publication units if not selected, which correspond to a publication in Oxford Economic Papers or Southern Economic Journal.

---

categories showed an average GROWTH that was greater for awarded vis-à-vis non-awarded proposals.

[9] The strong collinearity between SCORE and AWARD implies that both cannot be used.

■ *Production function of awarded proposals.* Table 7 presents sample selection, maximum likelihood estimates of the production function of publications.[10] The sample selection probit equation is the usual one, with AWARD as the dependent variable. The specification is the same as the one in the third column of table 4, and the results are very similar to that regression. (Hence they are not reported in table 7.) The regression equation uses log(PAFTER) as the dependent variable, and this is regressed on PI characteristics, the log of PBEFORE interacted with the cohort dummies, and the log of REC_DOL also interacted with the cohort dummies. The second column uses log(RQS_DOL) interacted with cohort dummies.

The elasticities of log(REC_DOL) for the different cohorts denote the marginal effect of NSF-funding on the publication output of the PIs in that cohort. Note that this elasticity is sizable and significant for the young PIs (ASSIST); it is practically zero for the two intermediate categories (ASSOC and PROF), but is larger than zero for SPROF. Thus, for junior PIs not only is selection important, so also are additional resources. At the sample mean, REC_DOL for an ASSIST PI is 6.9 (69,000 dollars), and PAFTER is 122.3. Given that the estimated elasticity of NSF funding is 0.73, this means that a 10,000 dollars higher grant for a selected ASSIST PI would increase publication output by 13, which corresponds to one publication in the Review of Economics and Statistics. For SPROF, REC_DOL at the sample mean for awarded PI is 9.8, and the corresponding sample mean for PAFTER is 93.6. This means that at the margin, an addition of 10,000 to the grant would increase output by 2, which corresponds to a publication in the Journal of Development Economics. Finally, note that the elasticities of RQS_DOL follow a pattern that is similar to those of REC_DOL.

5. CONCLUSIONS

These results *suggest* that NSF has a small positive impact on average on scientific output. The average impact of NSF support appears to reflect a much larger one for PIs at an early stage in their careers, and a much smaller one for PIs in the middle stages. Such a conclusion would be consistent both with anecdotal evidence, as well as the more formal theories about path dependencies in career profiles (e.g. Dasgupta and David, 1994).

These results are merely suggestive, and should be interpreted with caution. Researchers differ in ability and rich as our data are, we cannot hope to measure all the important dimensions. If so, one may incorrectly ascribe to the NSF what is due to ability. Furthermore, the process we are trying to map is a quintessentially dynamic one. Researchers enter the profession with different reputations (perceived ability). Their credentials and luck affect their ability to obtain research support, and in turn, this affects their publication output. In the next stage, their publication output is an important part of their perceived ability, and is a key to obtaining further research support.

---

[10] In this analysis, OLSQ regressions of log(PAFTER) using only the AWARD = 1 observations were also performed. The results are very similar to the sample selection regressions presented in table 7.

What this means is that even the distribution of unmeasured (latent) ability may differ systematically across the age cohorts of applicants. Only the "stars" may apply to the NSF at early stages of the career. Correspondingly, lack of success may discourage researchers from applying to an agency such as the NSF, especially late in their career when they may have alternative sources of support, or non research interests. Such differences would affect the interpretation of the differences in the estimates for the various age cohorts. For instance, PBEFORE is likely to a better signal of ability for relatively senior PIs. In turn, this might explain a part of the bigger effect of AWARD for junior PIs.

This paper has raises more questions than it has answered. Many of the questions are not new. Others, especially those relating to the dynamic process by which public agencies allocate resources for scientific research, are novel. Sensible science policy requires a systematic examination of these questions.

APPENDIX 1

Data of Publication productivity scores:

The original database used to create the publication productivity scores is ECONLIT, which contains over 300,000 records on publications in journals, books, and PhD dissertations from 1969 to 1995. We only used publications in journals.

The first step in creating publication scores was to find the complete list of journal articles by those NSF applicants during the 1969-1995 period. This was accomplished through matching queries that match the names of the authors in the ECOLIT records with the names of NSF applicants. This resulted in about 30,000 articles. For each article the author was given a publication score equal to the impact factor of the journal divided by the number of authors of that article.

The impact factors were taken from Liebowitz and Palmer (1984: table 2, column 4). These impact factors are based on impact adjusted citations per article. -1980 citations to articles published during the period 1975-1979. The following is the list of the top 50 journals with their impact scores.

1 Journal of Economic Literature 100
2 Brookings Papers on Economic Activity 96.86
3 Journal of Financial Economics 62.15
4 Journal of Political Economy 59.12
5 Bell (Rand) Journal of Economics 39.45
6 American Economic Review 34.48
7 Journal of Monetary Economics 33
8 Economica 31.63
9 Econometrica 31.6
10 Review of Economic Studies 30.36
11 Journal of Mathematical Economics 24.73

12 Journal of Law and Economics 22.89
13 Journal of Economic Theory 22.28
14 Journal of Public Economics 19.65
15 International Economic Review 19.04
16 Journal of Econometrics 17.32
17 Journal of Industrial Economics 16.55
18 Quarterly Journal of Economics 16.17
19 Economic Journal 14.96
20 Journal of Finance 14.63
21 American Economic Review Papers and Publications 14.613
22 Journal of International Economics 14.12
23 Journal of Human Resources 13.63
24 Review of Economics and Statistics 12.4
25 Public Finance 11.92
26 National Tax Journal 9.9
27 Journal of Money, Credit, and Banking 9.88
28 Canadian Journal of Economics 9.43
29 Manchester School of Economic and Social Studies 9.38
30 Industrial and Labor Relations Review 8.95
31 Journal of Legal Studies 8.43
32 Journal of Business 8.29
33 Journal of Urban Economics 8.07
34 Economic Inquiry 7.88
35 Scandinavian Journal of Economics 7.11
36 Journal of Accounting Research 6.98
37 Environmental Economics Review 6.66
38 Public Finance Quarterly 5.52
39 Oxford Economic Papers 4.86
40 Southern Economic Journal 4.83
41 British Journal of Industrial Relations 4.75
42 Applied Economics 4.39
43 Kyklos 4.3
44 Journal of Environmental Economics and Management 4.16
45 Journal of Royal Statistical Society, Series A 4.14
46 Public Choice 4.09
47 Journal of Financial and Quantitative Analysis 3.44
48 Journal of the American Statistical Association 3.02
49 Inquiry 3.01
50 Journal of Development Economics 2.29

If the article is published by one of the 50 top journals, the publication score for this article, as to the NSF applicant, is the impact score of the publishing journal divided by the number of authors of this article. Any journal not among the top 50 journals is given the basic score of 1. Thus, the publication score of an article published in a non-top-50 journal is 1 divided by the number of authors in that article.

The publication scores by an NSF applicant was then grouped by the year of publication to have the annual publication scores during the period from 1976 to 1995 for each applicant. These yearly scores were used to create the variables of the sum of publication scores during the periods five years before and after the fiscal year in which an NSF applicant has at least one application for NSF funding.

Bibliography

Adams, J., and Griliches, Z, 1996, "Measuring Science: An exploration", NBER working paper, MA.

Arora, A., David, P, and Gambardella, 1996, "Reputation and competence in publicly funded science", working paper, Heinz School, CMU, Pittsburgh.

Dasgupta, P,.and David, P, 1994, "The new economics of science", Research Policy.

Liebowitz and Palmer , 1984, "The impacts of Economics Journals", Journal of Economic Literature, Vol. XXII , March 1984

TABLE 1: DEFINITION OF VARIABLES

AWARD = dummy equal to 1 if project was awarded NSF grant.

PAFTER = quality-adjusted number of publications in the 5 year window after the grant.

LAFT = Natural log of PAFTER
PBEFORE = quality-adjusted number of publications in the 5 year window before the grant.

LBEF = Natural log of PBEFORE
ASSIST = dummy equal to 1 for PI's who received their PhD less than 5 years before the grant.

ASSOC = dummy equal to 1 for PI's who received their PhD between 6 and 10 years before the grant.

PROF = dummy equal to 1 for PI's who received their PhD between 11 and 15 years before the grant.

SPROF = dummy equal to 1 for PI's who received their PhD 16 or more years before the grant.

PHD = dummy equal to 1 if PI belongs to a PhD granting school.

NPPHD = dummy equal to 1 if PI belongs to a school that does not offer a PhD degree.

OTHER = dummy equal to 1 if PI belongs to institutions other than PHD or NOPHD (e.g. organizations like NBER, foundations, etc.)

ELITE = dummy equal to 1 if PI belongs to the 10 leading economic departments (MIT, Harvard, Yale, Stanford, Berkeley, Chicago, Northwestern, San Diego, Wisconsin Madison, Columbia) or top organizations

DCOPI1 = dummy equal to 1 if project has 1 co-PI.

DCOPI3 = dummy equal to 1 if project has 2 or more co-PIs.

WMWNE = dummy equal to 1 if PI's institutions is in the West, Mid-West or North-East regions of the United States.

MALE = dummy equal to 1 if PI is male.

D8890 = dummy equal to 1 if application was in years 1988-1990.

SCORE = Average referee score on the project.

REC_OL = dollar received (in 10,000 dollar units)

RQS_DOL = dollar requested (in 10,000 dollar units)

TABLE 2: DESCRIPTIVE STATISTICS

NUMBER OF OBSERVATIONS: 1473

	Mean	Std Dev	Minimum	Maximum
PAFTER	65.08	85.68	0.00	1052.00
PBEFORE	67.21	89.61	0.00	929.00
ASSIST	0.26	0.44	0.00	1.00
ASSOC	0.22	0.41	0.00	1.00
PROF	0.19	0.39	0.00	1.00
SPROF	0.33	0.47	0.00	1.00
PHD	0.80	0.40	0.00	1.00
NOPHD	0.04	0.20	0.00	1.00
DCOPI2	0.23	0.42	0.00	1.00
DCOPI3	0.04	0.20	0.00	1.00
ELITE	0.35	0.48	0.00	1.00
WMWNE	0.83	0.37	0.00	1.00
MALE	0.93	0.26	0.00	1.00
SCORE	2.54	0.82	1.00	5.00
RCVD	2.55	4.93	0.00	37.19
RQST	10.83	8.44	0.00	85.31

TABLE 3: SAMPLE MEANS, AWARD = 1 and AWARD = 0

	AWARD=1 (N=414) MEAN	AWARD=0 (N=1059) MEAN
PAFTER	101.20	50.96
	(112.92)	(67.28)
PBEFORE	108.56	51.04
	(124.17)	(65.10)
ASSIST	0.21	0.28
	(0.41)	(0.45)
ASSOC	0.25	0.20
	(0.44)	(0.40)
PROF	0.21	0.18
	(0.41)	(0.38)
SPROF	0.33	0.33
	(0.47)	(0.47)
PHD	0.74	0.82
	(0.44)	(0.38)
NOPHD	0.03	0.05
	(0.16)	(0.22)
DCOPI2	0.20	0.24
	(0.40)	(0.43)
DCOPI3	0.03	0.05
	(0.18)	(0.21)
ELITE	0.51	0.28
	(0.50)	(0.45)
WMWNE	0.91	0.80
	(0.29)	(0.40)
MALE	0.92	0.93
	(0.27)	(0.26)
SCORE	1.93	2.78
	(0.53)	(0.78)
RCVD_DOL	9.08	0.00
	(5.23)	(0.00)
RQST_DOL	12.55	10.16
	(8.23)	(8.43)

Note: Standard Deviation in parenthese

TABLE 4: PROBIT EQUATION: Dependent Variable=AWARD

Parameter	Specification1 Estimate	Specification2 Estimate	Specification3 Estimate
D8890	-.11	-.10	-.11
	(.10)	(.10)	(.10)
ASSIST	2.09	1.79	1.99
	(.29)	(.31)	(.34)
ASSOC	2.43	2.04	1.98
	(.30)	(.32)	(.43)
PROF	2.29	1.94	1.98
	(.30)	(.31)	(.39)
SPROF	2.25	1.96	1.86
	(.29)	(.30)	(.32)
PHD	-.16	-.14	-.15
	(.12)	(.13)	(.13)
NOPHD	.09	.18	.18
	(.27)	(.27)	(.27)
DCOPI2	-.16	-.15	-.15
	(.10)	(.10)	(.10)
DCOPI3	-.21	-.19	-.17
	(.20)	(.20)	(.20)
ELITE	.29	.27	.27
	(.10)	(.10)	(.10)
WMWNE	.32	.32	.32
	(.13)	(.13)	(.13)
MALE	-.42	-.52	-.50
	(.15)	.16	(.16)
SCORE	-1.11	-1.08	-1.09
	(.07)	(.70)	(.07)
LBEF	...	.09	...
		(.03)	
LBEF1	...	...	.04
			(.05)
LBEF2	...	...	.11
			(.08)
LBEF3	...	...	.09
			(.07)
LBEF4	...	...	.13
			(.04)
N(all)	1473	1473	1473
N(positive)	414	414	414
Log Likelihood	-645.80	-639.95	-638.80

Note:
1. Standard Error in parentheses
2 LBEF1 is equal to LBERF * ASSIST. LBEF2, LBEF3 etc are similarly defined.

TABLE 5: AMOUNT OF GRANT (SAMPLE SELECTION (MAXIMUM LIKELIHOOD) ESTIMATION)

Probit dependent variable: award (total observation = 1473; positive observations = 414); regression dependent variable: log (rec_dol)

Parameter	Specification1 Estimate	Specification2 Estimate	Specification3 Estimate
D8890	.16 (.08)	.10 (.08)	.11 (.07)
ASSIST	10.17 (.26)	6.53 (1.75)	1.96 (.64)
ASSOC	10.43 (.29)	6.60 (1.83)	3.01 (1.56)
PROF	10.52 (.25)	6.60 (1.86)	5.96 (1.40)
SPROF	10.54 (.26)	6.71 (1.85)	8.32 (2.26)
PHD	-.07 (.07)	-.09 (.06)	-.05 (.08)
NOPHD	-.42 (.24)	-.40 (.20)	-.31 (.18)
DCOPI2	.23 (.08)	.15 (.07)	.10 (.08)
DCOPI3	.28 (.22)	.06 (.22)	.05 (.26)
ELITE	.02 (.08)	-.02 (.06)	-.03 (.12)
WMWNE	.39 (.12)	.29 (.12)	.20 (.15)
MALE	.35 (.12)	.26 (.11)	.22 (.10)
LBEF1	.07 (.04)	.03 (.03)	-.03 (.04)
LBEF2	.04 (.04)	.03 (.03)	-.01 (.05)
LBEF3	.05 (.04)	.04 (.04)	-.03 (.04)
LBEF4	.05 (.03)	.03 (.03)	.03 (.04)
LRQST	...	.36 (.17)	...
LRQST1	...	...	.80 (.06)
LRQST2	...	...	.70 (.10)
LRQST3			.43 (0.10)
LRQST4	...	...	.23 (.17)
SIGMA	.60 (.04)	.52 (.06)	.50 (.13)
RHO	-.38 (.17)	-.38 .14	-.54 (.71)
Log of LF	-994.44	-932.99	-908.49

Notes Heteroskedastic consistent standard errors (Eicker-White) in parentheses

TABLE 6: OLSQ PUBLICATION REGRESSION, FULL SAMPLE

Dependent variable: Log publications (after proposal date)
Number of observations: 1473

Variable	Estimated Coefficient	Standard Error
D8890	.06	.08
ASSIST	3.06	.24
ASSOC	1.20	.26
PROF	.75	.32
SPROF	.81	.24
PHD	−.28	.10
NOPHD	−.97	.18
DCOPI2	−.18	.07
DCOPI3	−.23	.13
ELITE	.14	.09
WMWNE	.20	.09
MALE	.13	.13
(1−AWARD)1	−.47	.14
(1−AWARD)2	−.34	.12
(1−AWARD)3	−.25	.14
(1−AWARD)4	−.06	.13
LBEF1	.32	.04
LBEF2	.64	.04
LBEF3	.72	.05
LBEF4	.69	.03

Log of likelihood function = -2304.54
R-squared = .51

Note:

1.Heteroskedastic consistent standard errors in parenthesis.
2. The variables (1−AWARD)1, (1−AWARD)2, (1−AWARD)3, (1−AWARD)4 and LBEF1, LBEF2, LBEF3, LBEF4, 1−4 correspond respectively to interaction with ASSIST, ASSOC, PROF, SPROF.

TABLE 7: SAMPLE SELECTION ESTIMATION (MAXIMUM LIKELIHOOD) -- PRODUCTION OF PUBLICATIONS

DEPENDENT VARIABLE: LAFT

Parameter	Estimate	Estimate
D8890	.16 (.13)	.22 (.14)
ASSIST	-4.11 (2.36)	-5.71 (2.42)
ASSOC	1.05 (2.02)	3.35 (1.82)
PROF	1.40 (2.43)	3.39 (2.99)
SPROF	-1.43 (1.64)	.65 (1.12)
PHD	-.28 (.13)	-.27 (.17)
NOPHD	-.87 (.46)	-.90 (.46)
DCOPI2	-.03 (.13)	-.03 (.13)
DCOPI3	-.11 (.32)	-.09 (.38)
ELITE	-.10 (.13)	.11 (.22)
WMWNE	.37 (.23)	.42 (.29)
MALE	-.48 (.21)	-.37 (.21)
LRCVD1	.73 (.22)	...
LRCVD2	.05 (.19)	...
LRCVD3	-.08 (.22)	...
LRCVD4	.19 (.15)	...
LRQST1	...	.86 (.23)
LRQST2	...	-.17 (.13)
LRQST3	...	-.25 (.22)
LRQST4	...	-.01 (.06)
LBEF1	.16 (.07)	.12 (.09)
LBEF2	.63 (.09)	.64 (.12)
LBEF3	.87 (.07)	.86 (.09)
LBEF4	.83 (.05)	.84) (.08)
SIGMA	1.02 (.05)	1.01 (.02)
RHO	.09 (.10)	.04 (.76)
Log of LF	-1234.20	-1231.62

Notes: Heteroskedastic consistent standard errors in parenthesis

APPENDIX 1

Data of Publication productivity scores:

The original database used to create the publication productivity scores is ECONLIT, which contains over 300,000 records on publications in journals, books, and PhD dissertations from 1969 to 1995. We only used publications in journals.

The first step in creating publication scores was to find the complete list of journal articles by those NSF applicants during the 1969-1995 period. This was accomplished through matching queries that match the names of the authors in the ECOLIT records with the names of NSF applicants. This resulted in about 30,000 articles. For each article the author was given a publication score equal to the impact factor of the journal divided by the number of authors of that article.

The impact factors were taken from Liebowitz and Palmer (1984: table 2, column 4). These impact factors are based on impact adjusted citations per article. -1980 citations to articles published during the period 1975-1979. The following is the list of the top 50 journals with their impact scores.

1 Journal of Economic Literature 100
2 Brookings Papers on Economic Activity 96.86
3 Journal of Financial Economics 62.15
4 Journal of Political Economy 59.12
5 Bell (Rand) Journal of Economics 39.45
6 American Economic Review 34.48
7 Journal of Monetary Economics 33
8 Economica 31.63
9 Econometrica 31.6
10 Review of Economic Studies 30.36
11 Journal of Mathematical Economics 24.73
12 Journal of Law and Economics 22.89
13 Journal of Economic Theory 22.28
14 Journal of Public Economics 19.65
15 International Economic Review 19.04
16 Journal of Econometrics 17.32
17 Journal of Industrial Economics 16.55
18 Quarterly Journal of Economics 16.17
19 Economic Journal 14.96
20 Journal of Finance 14.63
21 American Economic Review Papers and Publications 14.613
22 Journal of International Economics 14.12
23 Journal of Human Resources 13.63
24 Review of Economics and Statistics 12.4

25 Public Finance 11.92
26 National Tax Journal 9.9
27 Journal of Money, Credit, and Banking 9.88
28 Canadian Journal of Economics 9.43
29 Manchester School of Economic and Social Studies 9.38
30 Industrial and Labor Relations Review 8.95
31 Journal of Legal Studies 8.43
32 Journal of Business 8.29
33 Journal of Urban Economics 8.07
34 Economic Inquiry 7.88
35 Scandinavian Journal of Economics 7.11
36 Journal of Accounting Research 6.98
37 Environmental Economics Review 6.66
38 Public Finance Quarterly 5.52
39 Oxford Economic Papers 4.86
40 Southern Economic Journal 4.83
41 British Journal of Industrial Relations 4.75
42 Applied Economics 4.39
43 Kyklos 4.3
44 Journal of Environmental Economics and Management 4.16
45 Journal of Royal Statistical Society, Series A 4.14
46 Public Choice 4.09
47 Journal of Financial and Quantitative Analysis 3.44
48 Journal of the American Statistical Association 3.02
49 Inquiry 3.01
50 Journal of Development Economics 2.29

If the article is published by one of the 50 top journals, the publication score for this article, as to the NSF applicant, is the impact score of the publishing journal divided by the number of authors of this article. Any journal not among the top 50 journals is given the basic score of 1. Thus, the publication score of an article published in a non-top-50 journal is 1 divided by the number of authors in that article.

The publication scores by an NSF applicant was then grouped by the year of publication to have the annual publication scores during the period from 1976 to 1995 for each applicant. These yearly scores were used to create the variables of the sum of publication scores during the periods five years before and after the fiscal year in which an NSF applicant has at least one application for NSF funding.

Appendix 2

Comparison of GROWTH between Selected and Non-Selected Groups within Various PI's Categories

PI Category	Av. GROWTH AWARD=1	Diff. with Av. GROWTH for AWARD=0	N. obs.
Score class [2;2.5)			
ASSIST;LOW	65.7	-34.6	20
	(19.7)	(24.3)	
ASSIST;HIGH	63.2	-29.3	20
	(44.0)	(49.2)	
ASSOC;LOW	32.7	-60.1	19
	(27.9)	(29.9)	
ASSOC;HIGH	-16.0	-13.9	18
	(29.1)	(32.9)	
PROF;LOW	18.3	-10.0	17
	(11.4)	(16.4)	
PROF;HIGH	9.3	-15.5	20
	(17.8)	(25.3)	
SPROF;LOW	28.0	14.3	30
	(13.2)	(20.0)	
SPROF;HIGH	17.5	-12.0	29
	(17.0)	(23.2)	
Score class [1;2)			
ASSIST;LOW	15.4	-0.2	14
	(17.1)	(32.3)	
ASSIST;HIGH	-28.1	-5.7	13
	(33.1)	(47.3)	
ASSOC;LOW	-11.3	16.6	16
	(13.5)	(20.8)	
ASSOC;HIGH	23.1	-44.4	16
	(37.4)	(71.0)	
PROF;LOW	12.9	-13.9	14
	(18.6)	(36.6)	
PROF;HIGH	-45.8	125.8	14

	(9.0)	(9.0)	
SPROF;LOW	16.5	6.2	23
	(14.9)	(23.9)	
SPROF;HIGH	-8.5	-0.0	25
	(11.0)	(17.3)	
Score class [2.5;3)			
ASSIST;LOW	(*)	(*)	15
ASSIST;HIGH	70.5	7.1	15
	(24.5)	(44.3)	
ASSOC;LOW	17.5	-25.8	17
	(14.5)	(21.5)	
ASSOC;HIGH	-60.5	-1.5	16
	(3.5)	(12.8)	
PROF;LOW	(*)	(*)	15
PROF;HIGH	(*)	(*)	16
SPROF;LOW	56.0	-54.6	26
	(47.6)	(47.8)	
SPROF;HIGH	-6.5	7.7	25
	(12.6)	(19.8)	

Notes

1. (*) denotes that there are no observations for AWARD=1.
2. LOW;HIGH indicates “low” or “high” past publications as defined in text. ASSIST, ASSOC, PROF, SPROF are the age cohorts. Only male PIs are considered.