NBER WORKING PAPER SERIES

INCENTIVES IN BASIC RESEARCH

Edward P. Lazear

Working Paper 5444

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1996

This work was supported in part by the National Science Foundation. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper 5444
January 1996

INCENTIVES IN BASIC RESEARCH

ABSTRACT

Individuals involved in basic research, like other workers, respond to incentives. Funding agencies provide implicit incentives when they specify the rules by which awards are made. The following analysis is an exercise in understanding incentives at an applied level. Specific rules are examined and analyzed to determine their incentive effects. For example, what is the effect of rewarding past effort? What happens when a few large awards are replaced by many small awards? How does the timing of an award affect effort? How does an agency choose which topics to fund? After having mapped out the responses of researchers to rules, socially optimal rules are derived. Research incentive issues have private business analogues, and the extension to the operation of the firm is discussed briefly.

Edward P. Lazear
Graduate School of Business
Stanford University
Stanford, CA 94305-5015
and NBER

There are many public and private agencies in the United States that provide financial support for basic and applied research. The most compelling reason for subsidizing research is the traditional argument that there is a conflict between private and social incentives in the case of investment in technological change. Even research with direct marketability will not be undertaken at the appropriate rate unless the inventor is entitled to the full rents from the resulting advance. But capturing most of the rent generally implies granting a monopoly to the inventor, which then results in inefficiently low quantities of the product being produced.[1]

With basic research, the problem is even more acute. It is virtually impossible for a scholar who does basic research to use the market as a vehicle for capturing the value of his innovation. Neither Plato, Isaac Newton, Adam Smith, Pascal nor Einstein received the full worth of their discoveries.[2]

Funding agencies, both private and governmental, have recognized the need to subsidize basic research.[3] The product of the research is then in the public domain and its use is unrestricted. The major problem with this approach is that replacement of market forces with committee decision making means that the normal guidance that market forces provide, however imperfect in this context, is unavailable. Substitute criteria must be established by which grants can be made.

---

[1]See Stigler (1968).

[2]Hirshleifer (1961) argues that information about the way that an innovation will affect prices can be capitalized and provides an incentive to innovate (actually to change). But it is difficult to believe that those forces can be an important motivator for basic research.

[3]Free rider problems are significant. Not only might one donor prefer to rely on the contributions of another, but an entire country might opt to allow another nation to advance basic knowledge.

While these are all important issues, the focus of this work is somewhat narrower, examining the incentive effects of various award schemes and their economic efficiency. The questions relate most closely to two lines of past research: The first is the tournament literature[4] which discusses the effect of awarding prizes on worker incentives. The second is the literature on methods of compensation,[5] which takes issues of risk aversion and capital constraints into account in determining the effect of various compensation methods on worker actions.

Although the basic approach is related to this earlier work, the thrust is new. The incentives literature has not been applied to the somewhat specialized issue of research funding. Further, there are some questions of general interest for incentives, which have not been addressed before. For example, how large a weight should be attached to past performance versus the quality of the current idea? At what stage in a project should an agency evaluate research and similarly, at what stage should a worker's performance be evaluated?

Some specific questions that will be answered are:

1. How much weight should be placed on past performance as opposed to the quality of the proposal?

2. Which areas of research should be funded? One advantage of a market system is that it rewards correct moves and punishes wrong ones. This disciplines individuals and forces them to choose projects that have private, and generally social value. A difficulty that government granting agencies face is that they must assess social value absent the market or at least before

---

[4]See Lazear and Rosen (1981), Nalebuff and Stiglitz (1983), Reinganum (1981), Loury (1979), Carmichael (1983), Green and Stokey (1983), and O'Keeffe, Viscusi, and Zeckhauser (1984) to name a few.

[5]See Lazear (1986), Brown (1990), Abowd (1990) and Leonard (1990), Fama (1991).

market information is conveyed.

3. Related, what topics should be funded? How does an agency decide whether to fund theory or empirical research and whether labor economics or industrial organization should receive preferential treatment?

4. Should the criteria be experience-adjusted so that it is more difficult for more senior researchers to obtain funding?

5. Should many small awards be made or is it better to give a few larger ones?

6. Should award amounts be fixed in advance or should they be continuous functions of the quality of the proposal?

The answers to many of these questions may appear to be common sense, but as the next few pages will argue, the issues are far more complicated than they first appear and rigorous analysis is required to provide reasoned answers.

The Model

The model is of the overlapping generations variety. Each individual lives two periods and there is a continuum of types; some researchers are inherently more productive than others. For now, assume that individuals cannot self-finance so that the decision to refuse an award implies that a project is not undertaken by that individual. Initially, the analysis is static so that the basic structure can be introduced.

The analysis of one of basic research. Since there is no market test for the value of the output, some other criterion for making awards must be used. This may involve peer reveiw. Alternatively, the granting of awards may be determined by an administrator. Both possibilities are nested in this model.

Receiving an award to support research involves elements of chance. But it also involves talent and hard work. One way to model the competition is as a lottery, where one's chances of winning depend on the proportion of tickets that one has in the urn. The number of tickets held by any individual, T_i, depends on the skill and effort of the entrant. Thus,

$$T_i = f(\mu_i, \delta_i), \qquad (1)$$

where μ_i is individual i's effort and δ_i represents his ability.

The total number of tickets in the urn is T, which is merely the sum of the tickets of all individual contestants. A key feature of the award structure is that a contestant can win at most one award. Thus, having a ticket drawn twice is assumed to give the contestant no more payoff than having only one ticket drawn. This is not a necessary or even an accurate assumption. Since the size of the award can vary somewhat, we could permit more tickets drawn to imply a larger award. The assumption that at most one prize is awarded actually makes the modeling somewhat more complicated, but reflects the idea that there are diminishing returns to grant getting effort.

Indeed, the problem would be much easier if awards were continuous, say depending on the measured output of the work, or on its effort.[6] A continuous structure is not adopted because the goal is to model the way that agencies actually behave, not the way they should behave. Later, when efficiency is considered, optimal (as opposed to existing) contracts will be analyzed.

---

[6]It would then resemble piece rates pay, the properties of which are well known.

But for now, the somewhat discrete structure is used because it is closer to the current agency award distribution pattern.

The mapping of μ and δ into a prize is stochastic because it depends on having one's ticket drawn. The non-deterministic aspect captures two phenomena: First, effort and ability do not exactly determine output; luck may be important. Second, output may not be observed perfectly by the granting agency; evaluation is not costless.[7]

Output of the process depends not only on application effort and ability, but also on completion effort, θ. Production, distinct from winning an award, is a mapping Q: $R^N \rightarrow R^1$, where each individual's output, q, depends on μ, δ, and θ, and where N is the number of contestants.

A contestant has T_i/T of the tickets in the urn. Suppose that the granting agency chooses to draw A tickets from the urn with replacement[8]. This is not quite the same as giving out A awards since some individuals may have more than one ticket drawn and will receive only one award. But A can be thought of as a parameter chosen by the agency that reflects the number of awards to be given, where the agency recognizes that for any given A, the number of awards

---

[7] The tournament framework emphasizes differences between rank among contestants. This lottery approach emphasizes winning versus losing, but not the rank among the winners.

This structure is neither more nor less general than a tournament model where winning depends on the rank of output. Tournaments postulate a specific functional form that transforms effort into an observation on rank. A specific production function is chosen and then an error distribution is specified which implies a distribution of ranks for given effort. In the lottery approach, both stages are compressed into one. The probability of having one's ticket drawn increases in effort and ability in a particular way.

[8]The formulas for statistics with replacement are much easier than those for non-replacement. Below, in the case of some specific examples, probabilities are computed both with and without replacement. At least in those contexts, the assumption is completely innocuous.

given is a random variable. The assumption merely simplifies the notation somewhat and has no substantive implications.[9]

Define A* as the expected number of awards given. In general, $A^* < A$ because some individuals probably will have their names called at least twice. The relation of A to A^* is somewhat complicated and depends on the distribution of tickets. But it is clear that A^* is non-decreasing function of A. Adding an additional draw in no way reduces the number of possible award sequences, but may (and usually does) add some. Thus, A*(A) > A*(A-1) for all A.

The probability that any contestant will receive an award is

$$P_i = 1 - (1 - Ti/T)^A,$$

or, defining π_i as the probability of failing to have a ticket drawn on any given try,

$$P_i = 1 - (\pi_i)^A . \tag{2}$$

The expected net return for a contestant is

$$\begin{aligned} \text{Expected net return} &= WP_i - k(\mu) - c(\theta) \\ &= W(1 - (\pi_i)^A) - k(\mu) - c(\theta) \\ &= W[\,1 - (1 - Ti/T)^A\,] - k(\mu) - c(\theta), \end{aligned} \tag{3}$$

[9]In real lotteries, the number of awards given is a random variable as well. Numbers are drawn and if no one's ticket matches the numbers, then no awards are given. If multiple tickets match, multiple awards are given, but the total amount of award money is split among winners. In the NSF context, the agency can choose to award less than the full amount of money, saving some for the next round. This happens when there are not enough good proposals, i.e., too many of the T_i's are low.

where W is the amount given on each award and $k(\mu)$ is the cost of application effort function, which reflects the disutility associated with putting forth application effort.[10] As is standard, k' and k'' are assumed to be positive. Similarly, $c(\theta)$ is the cost of completion effort function with $c', c'' > 0$.[11]

Application effort is distinguished from completion effort only by coming before rather than after the evaluation period. Specifically, application effort may be just as productive as completion effort. Many researchers do much of the work before they apply for the grant. Indeed, one of the questions to be answered is whether completion effort is needed at all. One possibility is to have all work done before the grant is approved and then to think of the grant as a prize for work well done. Structuring awards in this way may provide good incentives because it ensures that the work gets done before any money is doled out. This question is answered in a later section, but it is important to bear in mind that there need be nothing wasteful about application effort.

The individual must choose application effort μ so as to maximize expected net return in (3). Assume that the individual lives two periods. Assume further, relaxed below, that the two periods are independent, in terms of costs of effort, efficiency of production, and most important, in terms of probability of receiving an award. Then the problem in each period is to solve

[10]The prize, W, could be allowed to vary with measured output or input. Doing so would make the structure a piece rate and could guarantee efficient behavior. While analyzed below, the notion of continuous prize does not capture the essence of the award structure of most agencies, and certainly not NSF.

[11]The structure is quite general because the cost of effort function and production function offer sufficient freedom to capture the spirit of most alternative specifications.

(4) $$\max_{\mu} \; W(P_i) - k(\mu) - c(\theta)$$

Since there is no reward to completing the assignment, $\theta = 0$ for now and is dropped from discussion until later. The first-order condition is then

(5) $$W \frac{\partial P_i}{\partial \mu} - k'(\mu) = 0.$$

The individual takes the total number of tickets, T, as given. This is the standard assumption that individuals in the market behave as competitors, taking prices as given, assuming that their own actions have only negligible effects on the market.

The Less Able Try Harder:

How do more able individuals behave as compared with less able individuals? The answer is obtained by looking at $\partial\mu/\partial\delta$ along (5). From (5), we obtain

$$\frac{\partial\mu}{\partial\delta}\Big|_{(5)} = \frac{-\partial/\partial\delta}{\partial/\partial\mu}$$

or

(6) $$\frac{\partial\mu}{\partial\delta}\Big|_{(5)} = \frac{-\partial^2 P_i/\partial\mu\partial\delta}{\partial^2 P_i/\partial\mu^2 - k''/W}$$

The sign of (6) depends on the relation of P_i to μ and δ. Two general forms are presented. They describe reasonably the range of possibilities.

First, suppose that the reading of T_i by the evaluation panel depends linearly on ability and effort so that

$$T_i = \mu_i + \delta_i. \tag{7}$$

While this form may seem somewhat specific, in fact it is quite general, applying to any additively separable specification. Since no assumption about the distribution of δ has been made, it is possible to simply define δ as some function of the true distribution of ability. If $\delta = D(\delta^*)$ where δ^* is true ability, we can simply replace δ^* with δ and calculate the distribution of δ, given the underlying distribution of δ^*. Then (7) allows any function of the form

$$T_i = \mu_i + D(\delta^*_i).$$

Indeed, the same can be done for μ_i, but there are some restrictions. Since effort is produced at cost $k(\mu)$, it is necessary that the transformation of μ leave k', $k'' > 0$. Subject to that restriction, (7) can be reinterpreted as

$$T_i = M(\mu_i) + D(\delta^*_i),$$

where M and D are monotonically increasing functions. This specification encompasses a large class of additively separable functions.

Under the specification in (7), (6) becomes

$$(8) \quad \frac{\partial\mu}{\partial\delta} = -\frac{1}{1 + \dfrac{k'' T^2}{WA(A-1)\pi^{A-2}}}$$

Since $k'' > 0$, $\partial\mu/\partial\delta$ lies between -1 and 0.

Less able individuals put forth more effort, but not enough to eliminate the difference between the more able and less able. Thus, more able individuals are more likely to be given an award than less able ones, but their chances do not increase in proportion to their ability. The less able compensate for some of their disadvantage by working harder on their proposals.

Diminishing returns lies at the heart of this result. Even though the cost of effort, as modeled, is no higher to more able individuals they put forth less of it because the marginal return to effort decreases with T_i. The intuition is clear if we take it to the extreme. Suppose that 5 tickets were to be drawn and that an individual had 99 out of 100 of the outstanding tickets. That person would be almost certain to win one and only one award. (Since drawings are with replacement, the probability of winning an award is $1-.01^5 = .999999999$.) There would be no incentive for him to buy additional tickets by putting forth additional effort. The same is not true for the other individual who holds only one of the 100 tickets. He can greatly increase his chances of winning an award by putting forth additional effort. Since T_i is the sum of effort, μ, and ability, δ, more able people have less to gain from effort.

The diminishing returns argument is weakened, but generally not reversed by allowing ability and effort to interact. In the additively separable case, higher ability makes it more likely that an individual will win an award. But with additive separability, effort by more able individuals does not improve the probability that an award is received by any more than does effort by less able individuals. To remedy what may be an implausible assumption, consider a multiplicative specification

$$(9) \qquad T_i = \mu_i \delta_i.$$

As before, the specification is more general because μ and δ can be interpreted as functions of the true underlying effort and ability levels according to $M(\mu)$ and $D(\delta)$. Under this specification, (6) becomes

$$(10) \qquad \frac{\partial \mu}{\partial \delta} = \frac{1 - \frac{\delta(A-1)\mu}{T\pi}}{\frac{(A-1)\delta^2}{T\pi} + \frac{k''T}{WA\pi^{A-1}}}$$

The sign of (10) is ambiguous because there are two effects. The first is the diminishing returns factor which dominates in the additively separable specification. The second term in the numerator, which is subtracted, reflects this force. But it is preceded by a 1, which tends to make the derivative positive. This is because higher ability individuals receive a higher marginal return to effort. Since effort makes the number of tickets go up more for more able individuals, the most able are motivated to put forth more effort on this score. The more able get more observed output out of an hour spent on a proposal than the less able. This force is powerful and dominates unless the number of prizes are high. Substituting in the definition of π and rewriting, the necessary and sufficient condition for (10) to be negative is

$$\frac{T}{\mu\delta} < A - 1$$

When all participants are identical, the l.h.s. of the inequality is simply the number of participants. The r.h.s. is the number of draws, minus 1. As the number of draws gets large, therefore, the diminishing returns effect dominates the effect of ability in enhancing effort. Thus, an additional implication of this section is that allowing the number of awards to become large ensures that the less able try harder, even when effort and ability are complementary. If an agency would like to encourage the more able to work increase their effort relative to the less able, it is most likely to do so by keeping the number of awards few.

The intuition follows. If there are a large number of awards, the able gain little by putting forth much effort. Since they are at an advantage, and many prizes are to be given, an able individual is likely to receive one anyway, which discourages effort.

It is possible, of course, that the prize could fall enough to induce some contestants to drop out altogether. In a technical sense, this is already handled by the previous analysis. As long as k(0)=0 and k'(0) = 0, then it will always pay to put forth some level of effort. This may mean no more than reading the agency's announcement, or than thinking about a topic on which an application may be based. Individuals only drop out completely if the costs of exerting even the most trivial amount of effort is sufficiently high to induce individuals to avoid those costs.

When the agency raises the number of awards, it is implicitly raising the prize money awarded because the prize per award, *W*, is not reduced correspondingly. Next, the effects of increasing the number of awards, but lowering prizes per award will be considered.

Fewer Awards, Bigger Prizes:

Suppose that an agency has an expected budget G. If there are A^* expected awards to be given, then the expected budget is

$$G = WA^*. \tag{11}$$

What happens if A^* is increased, while W is decreased so as to keep G constant? The answer is that more awards are given, but the effort put into each proposal falls. Thus, increasing the number of awards while decreasing the amount per award so as to hold the total expenditure constant is output decreasing. The analysis follows.

To increase the number of awards, A^*, the number of draws, A, must increase. The approach is to examine how effort varies as A changes, taking into account the effect of changes in A on A^*.

Using the multiplicative specification for T_i in (9) and substituting in the budget constraint (11), the first-order condition becomes

$$\frac{GA}{TA^*}\pi^{A-1}\delta - k'(\mu) \tag{12}$$

Differentiating along (12), one obtains[12]

[12]To differentiate, it must be assumed that A*, A, and μ are continuous. However, A is an integer. Thus, the function μ(A) is defined only when A takes on integer values. But

$$(13)\quad \frac{\partial \mu}{\partial A}\Big|_{(12)} = \frac{G\pi^{A-1}\delta(1 + A\ln\pi - \frac{A}{A^*}\frac{\partial A^*}{\partial A})}{(\frac{GA(A-1)\pi^{A-2}\delta^2}{T^2 A^*} + k'')TA^*}$$

In the linear case, where (7) holds, the expression is the same except that the δ is replaced by 1.

The denominator is positive so the sign of the expression depends on the sign of the numerator. Key is the sign of

$$(14)\quad A\ \ln(\pi) + 1 - \frac{A}{A^*}\frac{\partial A^*}{\partial A}$$

The relation of A* to A depends on the distribution of T_i across contestants. In the special case where there are N identical contestants so that $T_i = T/N$, it can then be shown that

---

if $\partial\mu/\partial A < 0$ for all real values of A, then $\Delta\mu/\Delta A$ must be negative over integer values as well.

This calculation takes T as given. But for discrete changes in A, there will be an effect on T. If effort decreases, the total number of ``tickets" declines, i.e., it is easier to win an award with a given amount of effort and ability. The effect of the decline in T is to decrease effort further since $\partial\mu/\partial T$ is positive. A decrease in T is analogous to an increase in δ since both increase the probability of receiving an award.

(14) is negative for all positive values of A and N.[13] With identical contestants, an increase in A coupled with a corresponding decrease in W so as to keep expected prize money constant has the effect of reducing effort.

Note also that whether effort increases or decreases with the number of awards is independent of ability level of the research pool as long as all researchers are identical. The same logic implies that the decision to apply in the first place does not depend on ability. One can think of everyone as applying at some level of effort. Individuals who simply read the NSF announcement are putting forth a very low level of application effort. But the general specification would give some trivial number of tickets to them.

Crucial to this result is that ability is the same across contestants. If researchers can be sorted into different groups on the basis of ability, then effort is increasing among all researchers as the number of prizes falls. But in most cases, individuals with varying ability will be competing in the same contest; the contests will not be sorted perferctly by ability. Then the

---

[13] PN is a first-order approximation to A* so

$$\frac{\partial A*}{\partial A} = -N\pi^{A}\ln\pi \quad .$$

The sign of the numerator of (14) depends on the sign of

$$1 + A\ln\pi + \frac{A}{A^{*}}N\pi^{A}\ln\pi \ .$$

which is negative for $N \geq 2$, $A > 1$ and $0 < \pi \leq 1$.

The distribution of tickets among contestants is important in determining the relation of A to A^*. The more widely are tickets distributed, the closer is A to A^*. For example, suppose that there were 100 tickets, each held by a different individual. Then no one could have his named called twice, thus maximizing the number of awards for a given number of draws. At the other extreme, suppose that all tickets are held by one individual. Then the number of awards, A^*, equals 1, independent of the number of draws, A, since all additional draws are duplicates.

results become ambiguous. Intuition suggests that cutting the number of awards and raising the prize associated with each would favor the most able. An example supports the intuition.

Suppose that tickets are given by the multiplicative form so that the specification in (9) holds. Suppose further that costs of effort are given identically across contestants by

$$k(\mu) = \mu^2 / 2 \ .$$

There are two players. The able individual has $\delta_i = 5$, whereas the less able has $\delta_j=1$. The game now involves two players, each of whom has first order conditions given by (5). Each takes T as given. The equilibrium to the game solves the first order conditions for each player given the T that results in equilibrium.[14] The prize is derived in accordance with (11), such that the expected award amount equals a constant, in this case, 100.

The results are given in table 1.

[14]When there are only two players, the assumption of sampling with replacement is more problematic. The probability computations were done both ways, and the results were differed only in the fourth decimal place.

Table 1

A=	1	2	100
Variable			
μ_i	9.81	4.66	1.5 E-11
μ_j	1.96	4.66	5.8 E-11
T_i	49.0	23.3	7.2 E-11
T_j	1.96	4.66	5.8 E-11
p_i	.96	.97	.9999
p_j	.04	.30	.9999
W	100	78.2	50

Table 1 is intuitive. As A rises, W falls to offset the greater number of expected awards. Further, the more able individual's effort falls as A rises whereas the less able individuals effort rises as A initially rises. As A gets very large, there is virtually certainty that both individuals will receive one and only one award so the prize falls to 50. Note that effort for both individuals becomes very small. Thus, while effort rises for low ability individuals at low levels of A, as A continues to rise, effort falls even for low ability individuals. This is as it should be: As A rises, the probability that all individuals will receive an award rises. As it approaches 1, even low ability individuals gain little by putting forth additional effort.

Increasing the number of awards reduces the value of being a high ticket participant in the contest. For example, if the number of awards increased to a number greater than the number of applicants, then there would be no reason to expend any effort at all. Since researchers get at most one award, each applicant would be guaranteed an award. There would be no reason to put forth any effort because it would not alter the expected returns. As A increases, the probability of receiving an award becomes less and less sensitive to effort so the marginal return to effort declines.

The prediction is that in fields where few high level awards are given, each applicant would put more effort into the proposal than in fields were a large number of smaller awards are given. This does not imply that giving a few large awards is necessarily better than giving a large amount of small awards. The optimal level of effort has not yet been derived. The level could fall short or exceed that generated by some selected W, A pair.[15] What will be shown, however, is that any attempt to balance the budget will result in too little effort, and granting a few large awards is likely to come closer to the social optimum than granting many small awards.

Should Awards Be Accomplishment Based?

There are two primary effects of basing awards on past performance. First, completion effort increases among young researchers. Second, the distribution of awards shifts toward older individuals and older individuals put forth less effort. These are considered in turn.

---

[15]Another factor is ignored. If diversification is valuable and if grants affect the amount of completion effort put forth (see below), then the agency may prefer a larger number of awards to increase the chances that at least one researcher hits.

The effect on completion effort:

To see what happens to completion effort, the multiperiod aspect of the model must be fleshed out. Individuals are permitted to live for two periods, designated y for young and m for mature. When individuals are mature, the effort exerted when young is credited in deciding whether to make an award. Thus, young and mature individuals have tickets in the lottery as follows:

(15) a. $T_y = \mu_y \delta$

b. $T_m = (\mu_m + b\theta_y)\delta$

where θ_y is effort devoted to completion of the project when young.[16] The mature individual is given credit depending on the size of b. If $b = 0$, no credit is given. But there is no upper limit on the size of b, which is a parameter to be chosen by the granting agency.

The functional form in (15b) is chosen not only because it allows ability to affect the marginal value of effort, but also because it ensures that those who did not complete an award when young have some chance of receiving an award when mature. If the function were completely multiplicative, anyone with $\theta = 0$ would have $T_m = 0$.

The mature researcher's maximization problem is identical to the one period problem solved above, except that T_m includes credit for experience. The problem is

[16]It would be possible to allow θ_y to depend on receipt of an award. At least in some fields, capital constraints are binding so an individual denied an award may not have as large a value of θ_y. Conceptually, it is easy to handle this by allowing the cost of completing θ_y to depend on whether or not an award is given. In this section, these considerations are ignored to save on notation.

(16)
$$\operatorname*{Max}_{\mu_m} \; W(1-\pi_m^A) - k(\mu_m),$$

where

$$\pi_m \equiv 1 - \frac{(\mu_m + \theta_y b)\delta}{T}.$$

The first-order condition is

$$\frac{\delta\, WA\pi_m^{A-1}}{T} - k'(\mu_m) = 0,$$

as before. Note that completion is not an issue for mature researchers since there is no reward to putting forth completion effort. If awards are given to mature workers, they do not encourage completion among the mature. Instead, they affect the incentives of young researchers who want to be able to receive awards when mature.

The young researcher's problem is somewhat different since the young researcher must take into account that effort expended when young now affects the probability of receiving an award when mature and that the mature researcher will choose μ_m according to (16). The problem is then

(17)
$$\operatorname*{Max}_{\mu_y,\theta_y} \; W(1-\pi_y^A) - k(\mu_y) - c(\theta_y) + \frac{W(1-\pi_m^A) - k(\mu_m)}{1+r},$$

where $c(\theta)$ is the cost of completion effort and r is the interest rate. The first-order conditions are

(18a)
$$\frac{\partial}{\partial \mu_y} = \frac{WA\pi_y^{A-1}\delta}{T} - k'(\mu_y) = 0,$$

(18b)
$$\frac{\partial}{\partial \theta_y} = \frac{WA\pi_m^{A-1}b\delta}{T(1+r)} - c'(\theta) = 0.$$

To see that increasing the weight of past completion performance has a positive effect on completion effort, differentiate θ_y with respect to b, given (18b):

$$\text{(19)} \qquad \frac{\partial \theta_y}{\partial b}\Big|_{(18b)} = \frac{1 - \dfrac{b\delta\theta(A-1)}{T\pi}}{\dfrac{b^2\delta(A-1)}{T\pi} + \dfrac{c''T(1+r)}{WA\delta\pi^{A-1}}}$$

The denominator is positive so the sign of (19) depends on the sign of the numerator, i.e., on the sign of

$$1 - \frac{b\delta\theta(A-1)}{T\pi}$$

which is positive when b=0.

The conclusion is that completion effort, θ, rises as the weight, b, given to completion effort goes from zero to positive. Since the likelihood of receiving an award when mature depends on completion effort when young, young individuals work harder to complete the project.

Effort of mature workers:

When b increases, the initial impact is to raise the number of tickets that mature researchers have in the urn. But μ_m will adjust as well. In fact, for a given T, μ_m necessarily falls according to (8) which says that an increase in δ implies a decrease in μ_m. Adding $b\theta\delta$ is equivalent to increasing δ in (8). If b increases from zero to a positive number, then μ_m will be reduced according to (8), but not by as much as $b\theta$.

When awards are accomplishment based, mature researchers reduce effort as diminishing returns set in. The more certain a researcher is that he will get an award, the less likely he is to put forth effort. If one ticket is drawn, the value of having a second one drawn is zero. Experience based credit gives more tickets to older researchers and thereby decreases incentives. Thus, total tickets rise for the mature, but by less than the full amount of the experience credit.

Ability-Based Awards and Learning:

The previous results related to completion effort. A mature individual is given credit for effort put forth to complete the project when young. Now suppose that δ is not known when the individual is young, but must be learned. Suppose further that after one period, δ is learned perfectly. The agency might want to take δ into account in making awards. Recall that higher ability individuals put less effort into applying because they are more likely to get an award (see (8) or (10)). The granting agency can offset this effect by changing the standard on which individuals are given awards.

The agency can ``disqualify" some of an individual's tickets before the drawing is held. Since the more able have more tickets, disqualifying tickets among the most able of the mature applicants tends to offset the effect of diminished effort. While this sounds a bit strange, it really is not too far-fetched. The agency realizes that highly able applicants think they have it made and so do not put forth sufficient effort. To offset this, the agency simply makes it tougher on them so that their effort will increase. Of course, nothing has been said about the optimal amount of effort, but merely that effort can be increased by raising the standards for the most able.

The Optimum Amount of Research

The focus to this point has been on the researcher's supply problem. Now that the conditions that affect the researchers have been clearly laid out, it is necessary to determine what the agency should do. Thus, it is necessary to solve the social planner's problem.

First output must be defined. A simplifying assumption is that output is one-to-one with tickets. Then

$$q_{ji} = T_{ji}, \qquad j = y, m, \tag{20}$$

where q is output of young or mature individual i. This assumption is justified on two grounds. First, when both application effort and completion effort are productive, this is the natural specification. Recall that application effort is defined specifically as effort that occurs before the award is made and completion effort is defined as effort that occurs is after the award is made. Neither is necessarily wastesful effort. Thus, it is reasonable that society would care about the sum of the effort levels. Second, the assumption can be replaced by any other one without much difficulty. But if any assumption is likely to lead to efficiency, one would expect this one to do so since the social planner is focused on the same target, namely number of tickets, as are the researchers.

The social planner's problem is then

(21) $$\underset{\mu_{m_i},\mu_{y_i},\theta}{\text{Max}}\; q_{y_i} + q_{m_i} - [k(\mu_y) + c(\theta_y) + k(\mu_m)]$$

for each contestant i.[17] The first-order conditions are

(22)

a. $(\partial q_y / \partial \mu_y) - k'(\mu_y) = 0$

b. $(\partial q_y / \partial \theta_y) - c'(\theta_y) = 0$

c. $(\partial q_m / \partial \mu_m) - k'(\mu_m) = 0$

It is obvious from (22a,b,c) that the social planner would like to induce individuals to set their marginal costs of effort equal to its marginal social value, the standard result in any incentive problem. The funding agency has three instruments at its disposal, W, A, and b. With these, it must induce researchers to behave efficiently in accordance with (22). This is impossible in general.

A trivial counterexample proves the case. If output is additive in δ, μ, and θ, then $\partial q_m/\partial\mu_m$, $\partial q_y/\partial\mu_y$, and $\partial q_y/\partial\theta$ all equal 1. This implies that the optimal levels of application and completion effort are independent of ability. But it has already been shown that effort levels depend on ability in (8). Specifically, for a given set of awards, high ability individuals put forth less effort because of diminishing returns. The social planner would have them put forth the same level as their lower ability counterparts.

Even if there were only one level of ability, it would still be impossible as a general matter to obtain efficiency if the agency is restricted to using only three instruments, W, A, and b. This is seen as follows:

[17]As before, the contestant pool is taken as given. This assumption is relaxed below.

Consider the additive case. The first-order conditions in (22) imply that all marginal costs must be equal to 1. Substitute (22a,b,c) into the worker's first-order conditions in (16) and the additive case analogue (18a,b) to obtain

(23) a. $$\frac{WA\pi_m^{A-1}}{T} = 1$$

b. $$\frac{WA\pi_y^{A-1}}{T} = 1$$

c. $$\frac{WA\pi_y^{A-1}b}{T(1+r)} = 1$$

Eqq. (23b) and (23c) imply that

(24) $$b = (1 + r),$$

which is intuitive. In order to induce an individual to exert completion effort efficiently, it is necessary to pay him the value of that effort. Since payment is not received until the next period, it must be increased by the rate of interest. Also, since

$$\frac{WA\pi_m^{A-1}}{T} = 1$$

from (23a), the marginal present value of completion effort is necessarily 1.

While it is possible to set b correctly, it is impossible in general to attain the efficient levels of both μ_y and μ_m simultaneously. Eqq. (23a) and (23b) imply that

$$\pi_y^{A-1} = \pi_m^{A-1}.$$

The condition cannot hold as a general proposition. Suppose, for example, that there is learning by doing so that the marginal cost function for mature researchers is everywhere below the cost for young researchers. Then μ_m* is greater than μ_y* which implies that

$$\pi_y < \pi_m,$$

so the condition will not hold.

A solution to the problem is to make the award age contingent. Instead of having just one *W*, allow *W* to depend on age. Then both (23a) and (23b) can be satisfied for general cost conditions.

One way to do this is to tie the prize to the individual's salary, as NSF does. This helps not only on the age dimension, but also assists with the ability problem. Since high ability researchers do not put forth as much effort as low ability ones for a given *W*, tying *W* to salary helps to undo the inefficiency. Whether salary is the appropriate metric on which to base awards is, of course, another matter.

Discrete Awards or Continuous Payments:

If output is given by (20), then one must ask whether the agency is using the correct reward structure in the first place. Since output is continuous in effort, why not have the granting agency simply pay on the basis of effort? One difference between a piece rate scheme and one that pays *W* with some probability less than one is that the prize scheme changes the allocation and perhaps the total amount of risk.

That total risk can be decreased by using a prize structure is easily demonstrated. Suppose that output is given deterministically by (20), but observed output is equal to actual output plus measurement error. In the simplest world, where everyone had equal ability, output would be a constant. Paying continuously as a function of measured output adds noise to the amount paid and amount received, even though output does not vary. If measurement error is sufficient, a prize can do better, simply by increasing A until the probability that each person receives an award is arbitrarily close to 1. Effort can be altered by varying the amount per award.

It is also possible that the prize structure adds noise. Change the previous assumption and suppose that output is perfectly observable and identical across individuals. Payment of a straight piece rate guarantees that researchers' earnings and agency ``profits'' are constant. Payment of prizes adds noise. Some researchers may receive the prize, while others do not, even though each individual's output is identical. Further, because of the distinction between A and A^*, total prize money is stochastic.[18]

An intuitive reason for using fixed prizes over a continuous award structure has to do with fixed costs of production. In order to justify a minimum award size, it necessary that there be a fixed cost of production so that grants below some minimum size are not observed. Ignore completion effort for the moment and think of the award as a prize for past performance. If

$$\lim_{\mu \to 0} k(\mu) = k,$$

then no awards less than k would be observed. Fixed costs justify a minimum award size, but not a maximum. Agencies do impose rules that at least resemble caps. For example, NSF awards in economics are limited by placing a maximum on the amount of summer salary that can be reimbursed. Is there an argument for imposing the maxima?

[18]Green and Stokey (1983) have demonstrated that a nonlinear piece rate always dominates a tournament if there is no common noise and output is observable.

If there is, it is not based on the efficiency of fixed prizes over continuous awards. In fact, in the simplest case the argument goes the other way. Fixed prizes, where losers are constrained to receive zero, do not bring about first best and clear the market, whereas continuous piece rates do.[19] The proof follows:

First, piece rates clearly bring about efficiency and clear the market:

Efficiency requires that $k'(\mu) = 1$ from (22) (we ignore completion effort here). A linear piece rate has the form that compensation equals $S + R\mu$, where R is payment per unit of effort (or output). The worker solves $\max_{\mu} \; S + R\mu - k(\mu)$ with first-order condition $R = k'(\mu)$. If R is set to 1, then efficiency holds for any k, with $\mu = \mu^*$ defined as first-best effort. In fact, it does not matter that different types with different cost of effort functions are paid according to the same schedule. Efficiency is guaranteed for all. Furthermore, by choosing S such that $S + R\mu^* = k(\mu^*)$, all workers' rents are exhausted. Thus, there are no transfers that must be made from the public at large to researchers.

The same is not true for a prize structure. Efficiency is not obtainable in general at levels of income that clear the market. To see this, consider a quadratic cost function: $k(\mu) = k\mu^2$ where k is a parameter. Further, let the multiplicative specification hold with $\delta = 1$ (equivalent to additive with $\delta=0$). The researcher's maximization problem is

[19] That losers receive zero is important. If losers can be given some consolation prize (or can be forced to pay some entry fee), then first best can be achieved. This is the standard result of tournament theory with risk neutral agents.

(25) $$\max_{\mu} \; W\left(1 - (1 - \frac{\mu}{T})^A\right) - k\mu^2,$$

with first-order condition

$$\frac{WA}{T}\left(1 - \frac{\mu}{T}\right)^{A-1} = 2k\mu.$$

For efficiency, $2k\mu = 1$ from (22) so $\mu = 1/(2k)$. Substitution into the f.o.c. yields that the condition,

$$W = \frac{T}{A\left(1 - \frac{1}{2kT}\right)^{A-1}}$$

is required for efficiency. Substituting this condition and that $\mu = 1/(2k)$ into (25) gives

(26) $$\frac{T\left(1 - \left(1 - \frac{1}{2kT}\right)^{A}\right)}{A\left(1 - \frac{1}{2kT}\right)^{A-1}} - \frac{1}{4k}.$$

Recall that T is the total number of tickets, which is the product of tickets per contestant times the number of contestants. The number of contestants is exogenous and the number of tickets per contestant has already been shown to be $\mu = 1/(2k)$. Thus, T rises in N. It can be shown that the limit of (26) as N -> ∞ is 1/(4k), i.e.,

$$\lim_{N \to \infty} \frac{T(1-(1-\pi^{A}))}{A\pi^{A-1}} - \frac{1}{4k} = \frac{1}{4k}$$

which exceeds zero.

Recall that (26) is the expected value of competing. As a result, the award structure that induces researchers to put forth the efficient level of effort also provides for rents. Society may be unwilling to tolerate a situation where researchers earn rents. This also implies that individuals prefer to go into this occupation over others. Since rents are available, there is an excess supply of applicants.

If the agency insists on clearing the market, then it must choose too low an award to induce the first best level of effort. An alternative is to charge an application fee. The problem is that the concept of an application fee is somewhat antithetical to ex ante funding of research. Since one reason for funding the research before it is done is to provide funds for those who may not have access to private capital markets, charging an application fee would undermine the purpose of ex ante funding.

Can there be any reason then for placing a cap on the award level, say, as NSF does in economics? One possibility is that caps prevent bureaucrats from being gamed by rent-seeking researchers.[20]

Suppose that the agency can observe

$$\hat{q} = q + \lambda$$

where λ is effort that has no social value, but makes a project appear to be a good one. There is a cost of producing λ, given by $L(\lambda)$. If researchers are paid by the piece and $R = 1$ to obtain efficiency, then it is trivial to show that researchers will choose the efficient amount of effort, μ, but will augment it with socially unproductive λ which solves

$$L'(\lambda) = 1.$$

The agency will then spend $R(q+\lambda)$ per unit of output rather than Rq and its total expenditures will be correspondingly higher.

Alternatively, the agency can commit to offering an expected A^* awards, (i.e., drawing A tickets) each of which pays W. The problem that the researcher solves is then

$$\underset{\mu,\lambda}{\text{Max}} \; W\left(1 - \left(1 - \frac{\mu+\lambda}{T}\right)^{A}\right) - k(\mu) - L(\lambda).$$

There are two differences between giving prizes and paying a piece rate. First, when fixed prizes are given, the agency's expenditure level is not affected by the gaming by researcher. Since A and W are fixed in advance, changes in λ do not affect total expenditure.[21] Of course,

[20]This is related to the point of Milgrom (1988) and Baker (1992).

[21]This is not exactly correct. Since the relation of A^* to A depends on the distribution of tickets among contestants, there could be some minor effect on total expected expenditure. The direction of the effect is ambiguous, however.

there is no reason why a constraint on the total budget could not be imposed, either in the case of prizes or piece rate payment.

Second, when fixed prizes are given, the existence of positive values of λ reduces the amount of real effort, μ, that is provided. This follows directly from (8) and (10) since δ can be replaced by λ and the derivation is identical. To the extent that prizes induce too little effort, this is a drawback. But the interaction between μ and λ may mean that λ is not as high as it would be were an efficient piece rate paid. If a prize is interpreted as a capped payment, then the conclusion is that there is no unambiguous case to be made for prizes over piece rates in compensating a researcher.

When to make the award:

It is often suggested that NSF awards in economics are more like prizes for past work than they are support for future work. It is alleged that proposals that are not almost completed papers have little chance of success. True or not, the question is does it matter? What are the effects of paying prizes for past work instead of funding future work? There are a number of possible differences which relate to risk, capital markets, and information issues.

At one extreme, researchers could be given grants on the basis of their age-adjusted reputations, independent of any project. At the other, researchers could undertake the entire project on their own and then submit it to the agency upon completion. In the first case, researchers are essentially government employees who are paid a salary and are not residual claimants. Any output is the property of the government. In the second case, the researcher is the residual claimant. After producing the product, he may succeed in ``selling" it to the government, or he may fail. If the agency opts not to buy, i.e., turns down the ``proposal," the researcher may attempt to sell the work elsewhere.

Risk considerations push toward early awards. It is likely that researchers are more risk averse than agencies. As such, the natural outcome is for researchers to be ``hired" by the agency so that most of the residual goes directly to the agency.

Capital considerations lead to a similar conclusion. An inability to borrow is like risk aversion in that it changes the shape of the individual's cost function. High levels of expenditure are extremely costly when financed by individuals. Like risk aversion, capital market constraints further convexify the $k(\mu)$ function, making additional levels of effort extremely costly.

This consideration is likely to be very field specific. While the ability to borrow may not vary much by field, the size of a viable project does. In economics, a theorist may self-finance a project simply by devoting his own time to it. An anthropologist who must fund a field study simply cannot do the work without physical capital for travel, supplies and support, in addition to his own time.

Awarding the grant early reduces the cost of completing the work because it extends funds to researchers who could not otherwise obtain the funds.[22] But holding the award until work is completed provides better incentives to finish the work. In the two period context, this is modeled as follows:

If the award is given early, i.e., after only μ effort has been expended, then the cost of completion effort θ is, as before, given by $c(\theta)$. But if the award is given after completion then

[22]Of course, it is necessary to argue that the inability to obtain funds is due to a market failure so that completing the capital market by award funds early improves efficiency.

the cost of completion effort is $\tilde{c}(\theta)$ where $\tilde{c}'(\theta) > c'(\theta)\ \forall\ \theta$.

When the award is given early, the maximization problem is exactly the one in (17) with first-order conditions (18a) and (18b). The incentive to complete the project comes only through b, the experience rating effect, and works only for young researchers. Mature researchers solve (16) and set $\theta_m = 0$.

When the award is given late, after all work is completed, there are a number of changes. First, mature researchers maximize (16′) rather than (16):

(16′)
$$\max_{\mu_m, \theta_m} W(1 - \tilde{\pi}_m^A) - k(\mu_m) - \tilde{c}(\theta_m),$$

where $\tilde{\pi}_m \equiv 1 - \dfrac{\delta(\mu_m + b\theta_y + \theta_m)}{\tilde{T}}$ and $\tilde{T} = \sum_i \tilde{T}_i$. The first-order conditions are

(16′a)
$$\frac{WA\tilde{\pi}_m^{A-1}\delta}{\tilde{T}} - k'(\mu_m) = 0$$

(16′b)
$$\frac{WA\tilde{\pi}_m^{A-1}\delta}{\tilde{T}} - \tilde{c}'(\theta_m) = 0$$

Eq. (16′b) implies that $\theta_m > 0$ if $\tilde{c}'(0) < \dfrac{WA\tilde{\pi}_m^{A+1}}{\tilde{T}}$. Thus one benefit to paying after completion is that mature researchers have incentives to finish the project.

The problem for young researchers is

(27) $$\max_{\mu_m, \theta_m} \; W(1 - \tilde{\pi}_y^A) - k(\mu_y) - c(\theta_y) + [W(1 - \tilde{\pi}_m^A) - k(\mu_m) - \tilde{c}(\theta_m)]/(1 + r),$$

where θ_m, μ_m come from (16′a,b) above. The first-order conditions are

(28a) $$\frac{WA\tilde{\pi}_y^{A-1}\delta}{\tilde{T}} - k'(\mu_y) = 0$$

(28b) $$\frac{\delta\, WA\tilde{\pi}_y^{A-1}}{T} + \frac{\delta\, WA\tilde{\pi}_m^{A-1} b}{T(1 + r)} - \tilde{c}'(\theta_y) = 0.$$

Eq. (28b) differs from (18b) in that the current-period award affects completion effort when the award is given at the end of the project. But since $\tilde{c}'(\theta) > c'(\theta) \; \forall \; \theta$, effort is lower for any given award when the award is made ex post. A better solution is to make the award early and increase b to create incentives to complete the project. Then all that is lost is completion effort among mature researchers.

Information issues aside, the only reason to give the award after completion is to provide incentives to mature researchers who do not replay the game. Over an extended research career, this problem may not be so great.

The most important implication of this section is that fields in which capital considerations are most important would be expected to weight heavily past experience. If capital intensive fields use experience weighting, the adverse consequences on completion effort of not paying for completed work would be offset by reputation effects.

Which Fields Should be Funded?

One problem that an agency faces is deciding on the allocation of funds across fields. For example, the entire NSF Economics Budget is rounding error on the typical budget in the physical sciences. Is it appropriate to place so much emphasis on physical over social science?

Obviously, this is not the kind of question that economics is well suited to answer. But some points can be made. Most important is that choosing to fund a field depends not only on the expected return, but also on the variance in return. Funding a project is likef purchasing an option. It is a standard theorem in finance that the value of an option increases in variance. Applied here, this implies that risky projects should be favored out of proportion to their expected value. The analysis follows:

Consider two fields, A and B. Output in A is less risky, for convenience assume completely deterministic. Output in B is riskier. There are τ time periods. Suppose that information about the value of a project is obtained (perfectly) after observing the output for one period. Abstracting from effort considerations then, output is given by

$$q^A = \mu_A \qquad (29)$$

$$q^B = \mu_B + \epsilon.$$

where ϵ reflects an ex ante unobservable term that is specific to field B. Let $\epsilon \sim f(\epsilon)$.

The problem for the agency is to decide whether to fund field A or field B. The agency is attempting to maximize expected social output so the rule is to choose field B iff

(30) $$F(\mu_A - \mu_B)(\mu_B + \underline{\epsilon} + (\tau - 1)\mu_A) + (1 - F(\mu_A - \mu_B))\tau(\mu_B + \bar{\epsilon}) - \tau\mu_A > 0$$

where $\underline{\epsilon} \equiv E(\epsilon \mid \epsilon < \mu_A - \mu_B)$ and $\bar{\epsilon} \equiv E(\epsilon \mid \epsilon > \mu_A - \mu_B)$.

To see how this equation was derived, consider figure 1:

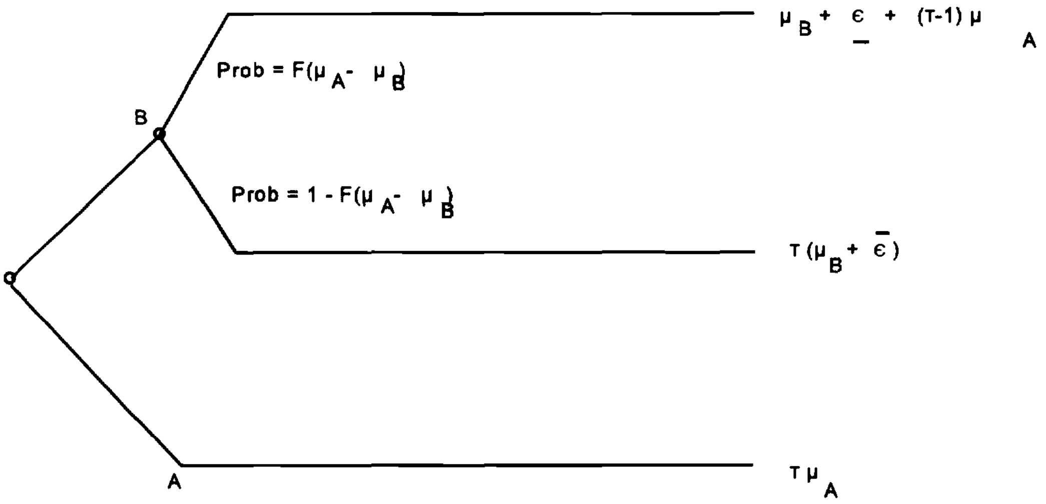

Figure 1

If field A is funded, output over the τ periods is $\tau\mu_A$ because nothing is ever learned to change the agency's decision about funding A. This is shown as the bottom branch in figure 1. If field B is funded, two things can happen. After one period the agency may learn that B was the right choice and it continues to fund B. Under these circumstances, expected output is $\tau(\mu_B + \bar{\epsilon})$. This is shown as the middle branch in figure 1. Alternatively, the agency may discover that it made a mistake and switch to funding A. Then expected output is

$$\mu_B + \epsilon + (\tau\text{-}1)\mu A \quad ,$$

shown on the top branch of figure 1. A switch back to *A* is worthwhile whenever $\mu_B + \epsilon < \mu_A$ or when $\epsilon < \mu_A - \mu_B$, which happens with probability $F(\mu_A - \mu_B)$. Conversely, $1 - F(\mu_A - \mu_B)$ of the time, the agency continues to fund *B*.

Even if $\mu_A > \mu_B$, it may well be optimal to fund *B* because doing so yields τ-1 periods after the information is obtained to enjoy potentially high output in *B*. Mistakes can be reversed without much cost, but unless *B* is tried, very high levels of output may be lost. For example, if $\epsilon \sim U[-3,3]$, $\mu_A = 5$ and $\mu_B = 3$, (30) is negative for $\tau < 4$, but positive for $\tau > 4$. Even though $\mu_A > \mu_B$, the possibility of high output in *B* makes it worthwhile to fund that field when there is a long enough period over which to recoup.

Some comparative statics are informative. Differentiating the l.h.s. of (30),

(31) a. $$\frac{\partial}{\partial \tau} = (1 - F)(\mu_B + \bar{\epsilon} - \mu_A) > 0,$$

since $(\mu_B + \bar{\epsilon} - \mu_A) > 0$ is required in order for *B* ever to be worthwhile.

b. $$\frac{\partial}{\partial \mu_A} = -F - (1 - F)\tau + f(\ )[\mu_B + \underline{\epsilon} + (\tau - 1)\mu_A - \tau(\mu_B + \bar{\epsilon})] < 0,$$

since $\tau(\mu_B + \bar{\epsilon}) > \mu_B + \epsilon + (\tau\text{-}1)\mu_A$ is required in order for *B* ever to be worthwhile.

c. $$\frac{\partial}{\partial \mu_B} = F + (1 - F)\tau - f(\ )[\mu_B + \underline{\epsilon} + (\tau - 1)\mu_A - \tau(\mu_B + \bar{\epsilon})] > 0$$

d. $$\frac{\partial}{\partial \underline{\epsilon}} = F > 0$$

e. $$\frac{\partial}{\partial \bar{\epsilon}} = (1 - F) > 0$$

The results in (31) make intuitive sense. First, the value of choosing field *B* over *A* increases with τ. The longer is the period over which returns are to be received relative to the learning period, the higher is the value of the high variance option. Fields where fixed effects have longevity should be favored over those where innovations are transitory. Eq. (31) also provides the major rationale for funding young researchers: To the extent that research output is correlated over an individual's life, it is better to discover a young star than an old star. The option value of betting on younger researchers is higher than betting on older (undiscovered) ones.

Second, (31b,c) imply that raising the expected value of *A* relative to *B* reduces the option value of *B*. This is obvious. Even though there is some value to taking risks, that value must offset the cost of forgoing *A*. The opportunity cost increases in the difference between μ_A and μ_B. There is no point in funding an obvious loser, even if the risk is quite high.

Third, an increase in either ϵ or $\bar{\epsilon}$ increases the value of funding field *B*.[23] When ϵ rises, the cost of rejecting *B* after having observed the low output decreases. When $\bar{\epsilon}$ rises, the value of accepting *B* after having observed the high output increases. Thus, if bad outcomes are not likely to be too bad, or more important, if good outcomes are very good, *B* is favored relative to *A*.

This section provides a justification for funding economics. While the expected payoff may or may not be as high as other fields, the option value is likely to be very large. For example, discovering a way to increase economic growth rates by 1/10 of 1% per year would have enormous effects on long term output. It also provides a justification for an emphasis on

[23]It is quite possible to increase $\bar{\epsilon}$ or $\underline{\epsilon}$ without changing the probability *F*. E.g., a distribution with a shorter but higher lower tail can result in a higher value of $\underline{\epsilon}$ than a long left-tailed distribution, even though the area to the left of $\mu_A - \mu_B$ is the same.

basic over applied research and perhaps for theory over empirical research. Basic theoretical research is a high variance activity. Most of the time, the value is small, but occasionally a fundamental breakthrough occurs. When one does, the value may be extremely high and may lead to a whole new area of research. If the theoretical research does not lead anywhere, it can be cut off rather quickly without too much social cost.

Other Issues:

A number of topics have been addressed in this essay, but others remain. Some of the more important questions for future research are listed here:

1. Should cost affect the probability of receiving an award? W has been fixed throughout, but it is possible to interpret a fixed W as fixed for a given category of proposal. Different categories could have different prizes or prizes could be made continuous as a function of cost.

2. How long a period should be funded without review? The timing of an award has been considered, but only to the point of asking whether it should be given ex ante or ex post. Another question relates to the number of years to be funded. We abstracted from this by awarding W for the project and did not ask whether one project was submitted per year, per month or per decade. Should the duration of the grant depend on past performance and/or on the quality of the proposal?

3. Which criteria should an agency use to judge proposals? Implicit in the lottery analogy was a rule for converting effort and ability into the probability of receiving an award. But there was no specific consideration of the appropriate mapping.

4. Who should judge the proposal? Presumably peers have the best information about the quality of the work, but their incentives may not be as pure as those of a third party, especially one who does not interact on a continuing basis with the researchers.

Extensions:

The analysis may have broader implications. Most directly, managers in firms must assign projects to particular workers, much like an agency's obligation to fund specific proposals. How does a firm decide whether to base assignments on the quality of a worker's idea or on the worker's past performance and level? This relates to the shape of the firm's pyramid. When individual-specific fixed effects are important, the firm may be more hierarchical, with assignments going to senior people who can then delegate. When fixed effects are less important, a rectangular structure may be more appropriate.[24] Additionally, when experience rating must be used as an incentive, the firm tends to be more hierarchical, with senior workers delegating assignments to junior ones. The question of how long a period should be funded is similar to asking how much free reign should a worker be given.

In politics, it is well known that incumbency is a major advantage in any campaign. But if fixed effects are important, or if job-specific human capital matters, it is reasonable to bet on incumbents rather than on campaign promises and ``new ideas." The issue of how often voters return incumbents to office is closely related to the question of how to weight past performance relative to the merits of the current proposal in awarding research grants.

---

[24]See Sah and Stiglitz (1986) for a discussion of organizational architecture.

Finally, journal editors receive complaints by younger authors that referees are too heavily influenced by the identity of the author. Blind refereeing is similar to awarding a grant on the basis of the current proposal rather than past accomplishments because both attempt to judge the work independently of the author's reputation.

Summary and Conclusions:

There are many arguments for subsidizing research. But the specific nature of the subsidy can have significant effects on the performance of the researchers. The results of this work generalize to any worker incentive problem, but the focus here is on the details of the research subsidization business. The most significant results are:

1. Under a fixed award structure, less able researchers try harder than more able ones. They put more effort into their proposals, but the additional effort does not fully compensate for their lower ability. As a result, the more able are still more likely to win an award. Toughening the standards for more able researchers would offset the diminished effort that more able researchers put forth.

2. Increasing the number of awards while decreasing the amount per award so as to leave expected prize constant reduces effort for the able and in the limit, even for the least able. Even though the expected award is the same, the marginal value of effort in obtaining the award declines so effort falls. This, coupled with the result that prizes tend to induce too little effort suggests that awarding many small prizes creates incentives for less than the efficient amount of effort.

3. Young researchers are induced to complete their projects when older researchers are

given credit for past jobs well done. The effect is to increase effort among the young and to reduce effort among older researchers.

4. Efficiency can be improved by making awards age contingent. Mature researchers can be given higher awards to offset their tendency to reduce effort. If salary and ability are correlated, tying awards to salary as NSF does, is efficiency enhancing.

5. If losers are constrained to receive zero, then there is no award structure that induces efficiency and also clears the market. Charging an application fee is an alternative, but may defeat the purpose of issuing grants in the first place.

6. Holding awards until after the project is completed will induce individuals to complete the project. But late awards may severely limit the kinds of projects that can be undertaken to those that can raise their own capital. An alternative is to award early, and to motivate completion effort by making future awards contingent on past performance.

7. Since high risk fields have high option value, moving in the direction of high risk funding increases the expected payoff to research.

References

Abowd, John "Does Performance-Based Managerial Compensation Affect Corporate Performance?" *Industrial and Labor Relations Review*, Vol 43, 3, Feb. 1990 Special Issue, 52s-73s.

Baker, George. ``Incentive Contracts and Performance Measurement," *Journal of Political Economy* (June 1992).

Brown, Charles. ``Firms' Choice of Method of Pay," *Industrial and Labor Relations Review* 43 (February 1990): 165-S-182-S.

Carmichael, H. Lorne. ``The Agent-Agents Problem: Payment by Relative Output," *Journal of Labor Economics* **1** (January 1983): 50-65.

Fama, Eugene F. ``Time, Salary, and Incentive Payoffs in Labor Contracts," *Journal of Labor Economics* 9 (January 1991): 25-44.

Green, Jerry R., and Stokey, Nancy L. ``A Comparison of Tournaments and Contracts," *Journal of Political Economy* **91** (June 1983): 349-64.

Hirshleifer, J. ``The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review* **51** (September 1961): 561-74.

Lazear, Edward P. ``Salaries and Piece Rates," *Journal of Business* **59** (July 1986): 405-31.

Lazear, Edward P., and Rosen, Sherwin. ``Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy* **89** (October 1981): 841-64.

Leonard, Jonathan "Executive Pay and Firm Performance," *Industrial and Labor Relations Review*, Vol 43, 3, Feb. 1990 Special Issue, pp 13s-29s.

Loury, Glenn C. "Market Structure and Innovation," *Quarterly Journal of Economics* (August 1979):395-410.

Milgrom, Paul R. ``Employment Contracts, Influence Activities, and Efficient Organization Design," *Journal of Political Economy* **96** (February 1988): 42-60.

Nalebuff, Barry J., and Stiglitz, Joseph E. ``Prizes and Incentives: Toward a General Theory of Compensation and Competition," *Bell Journal of Economics* **14** (Spring 1983): 21-43.

O'Keeffe, Mary; Viscusi, W. Kip; and Zeckhauser, Richard J. ``Economic Contests: Comparative Reward Schemes," *Journal of Labor Economics* **2** (January 1984): 27-56.

Reinganum, Jennifer F. ``Dynamic Games of Innovation," *Journal of Economic Theory* 25 (1981): 21-41.

Sah, Raaj Kumar, and Stiglitz, Joseph E. ``The Architecture of Economic Systems: Hierarchies and Polyarchies," *American Economic Review* 76 (September 1986): 716-27.

Stigler, George J. ``A Note on Patents," in *The Organization of Industry,* 123-25. Homewood, IL: Richard D. Irwin, 1968.