Paying for science: a theory of UKRI grant funding

Thomas Fink
London Institute for Mathematical Sciences, Royal Institution, 21 Albemarle St, London W1S 4BS, UK
(Dated: November 14, 2024)

How to allocate research funding is an important and contested problem. We analyse the methodology used by UK Research and Innovation to set grant budgets and derive an equation for the unrestricted portion. We find the unrestricted fraction can vary widely, and give a strategy for maximising it. This is important, because the viability of research institutions depends on being able to cover the cost of their research.

Introduction

Academic research in Britain is funded by UK Research and Innovation. The government agency allocates £8bn to research each year, much of it through competitive schemes for research grants. The way in which it does so has several strengths: flexibility in the subject, size and duration of the grant; a simple 12-page application; and moderate success rates of 22–30%, depending on the council [1].

UKRI uses a particular methodology for determining grant budgets which, while formulaic, is not intuitive. We analyse this methodology and derive an equation for the portion that is unrestricted. We find that the unrestricted fraction can vary widely, and give a strategy for maximising it. This is important, because the viability of universities and especially research institutions depends on being able to cover the cost of the research they do.

In this paper, we define the restricted portion of a grant to be the directly incurred costs: money that we would certainly spend were we to get the grant, such as paying postdoc salary. But we don't include directly allocated costs: money we could optionally spend, such as additional office space or a new member of staff. The unrestricted portion is the total value of the grant minus this.

Results

Research grants differentiate between established scientists, which we call principal investigators or PIs, and less established scientists, which we call postdocs. PIs tend to be already employed by the research organization over the period of the requested grant, so the cost of their salary is a sunk cost. But postdocs are usually brought in specially to work on the research proposed in the grant.

Throughout we will use the following notation:

- a amount of postdoc time per year on a grant, in years,
- b amount of PI time per year on a grant, in years,
- p postdoc annual salary,
- q PI annual salary,
- s support costs rate per man-year, set by UKRI.

Note that the amount of postdoc time per year need not be an integer. For example, if in a three-year grant there are two two-year postdocs, $a = 4/3$. A single grant can have multiple PIs as well as multiple postdocs. The support costs rate s is set by UKRI for universities and research institutions based on their historical accounts. The 2024 lower and upper quartiles for s are £62,809 and £80,431 for non-laboratory research, and £69,815 and £89,872 for laboratory research. In this paper we don't consider the cost of equipment [2].

Restricted and unrestricted income

A typical UKRI grant is between two and four years, but other periods are possible, such as 18 months. However, let's just consider a single year of a grant, which we can linearly scale up to multiple years as needed.

As a simplfying assumption, we take the on-costs of a postdoc and PI to be $1/4$ of the salary. These include the UK National Insurance contribution and any statutory pension, which we take to be $1/8$ of the salary; and travel, computing and publication fees, which we also take to be $1/8$ of the salary. For example, a postdoc with a salary of £40,000 costs $5/4 \cdot £40,000 = £50,000$.

What UKRI calls direct costs are the salary and on-costs of the postdocs and PIs: $5/4\, ap + 5/4\, bq$. What UKRI calls indirect costs is the total number of man-years times s: $(a + b)s$.

UKRI notoriously only funds 80% of the full cost of a research grant. The remaining 20% is meant to be obtained from other grants, student fees, the Research Excellence Framework or other sources. So the total amount that is actually applied for—and, if successful, received—is $4/5$ of the sum of the direct and indirect costs, namely,

$$T = (4/5\, s + p)a + (4/5\, s + q)b. \tag{1}$$

The unrestricted portion U of this grant is T minus the postdoc costs of $5/4\, ap$:

$$U = (4/5\, s - p/4)a + (4/5\, s + q)b. \tag{2}$$

If, for example, we have a year of postdoc time ($a = 1$) and three months of PI time ($b = 1/4$), the equations have a simple form: $T = s + q/4 + p$ and $U = s + (q - p)/4$.

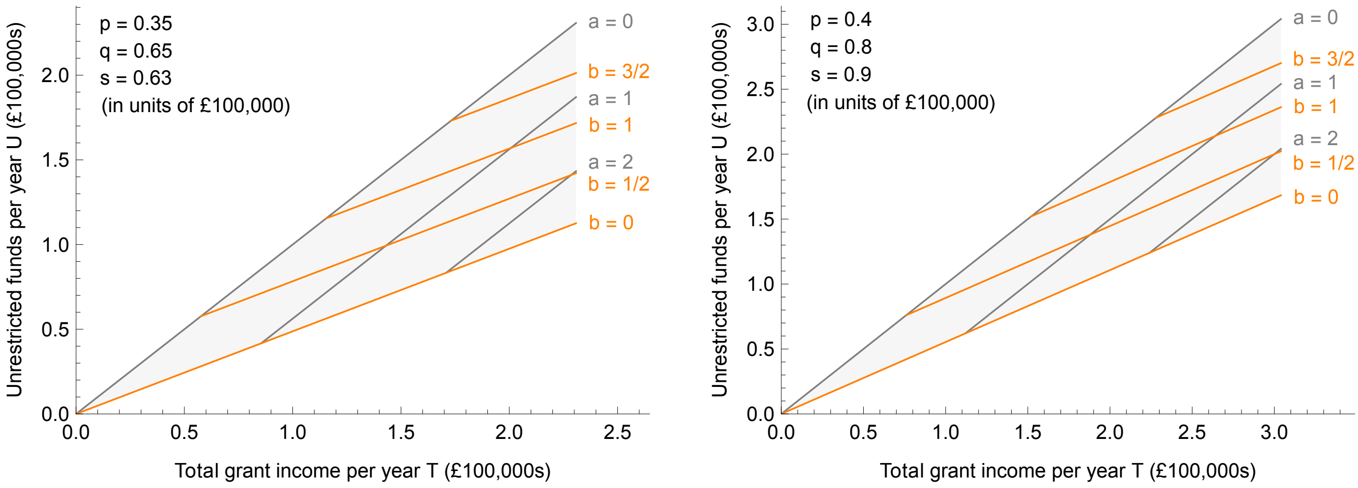

FIG. 1: **Unrestricted funds as a function of total grants funds.** The unrestricted funds U must lie within the grey cone. The structure of the cone is determined by the postdoc salary p, PI salary q and the UKRI support costs rate s. The location of U within the cone is determined by a and b. In the left plot we set (in units of £100,000) $p = 0.35$, PI salary $q = 0.65$ and $s = 0.63$, and on the right we set $p = 0.4$ and $s = 0.9$. The orange lines are given by eq. (3) for various values of b. The grey lines lines are given by eq. (4) for various values of a.

Returning to eq. (2), and using the typical values $p = 0.4$, $q = 0.4$ and $s = 0.8$ (all in units of £100,000), we have for the grant total and the unrestricted portion of the total

$$T = 1.04a + 1.44b,$$
$$U = 0.54a + 1.44b.$$

Here we see the explicit dependence of T and U on the two key quantities a and b that are most under our control. For example, for $a = 1$ and $b = 1/4$, $T = 1.4$ and $U = 0.9$. Over a period of four years, this amounts to a grant of $4T =$ £560,000 with the unrestricted portion being $4U =$ £360,000.

Maximising unrestricted income

At fixed grant total T, we can solve eq. (1) for the amount of postdoc time a. Inserting the result into eq. (2) eliminates a in favour of T:

$$U = \frac{(4/5\, s - p/4)T + (s + 5/4\, q)pb}{4/5\, s + p}. \quad (3)$$

Plotting this as a function of T gives the orange lines in Fig. 1 for $b = 0$, $1/2$, 1 and $3/2$, with p, q and s fixed.

In a similar way, we can solve eq. (1) for the amount of PI time b. Inserting the result into eq. (2) eliminates b in favour of T:

$$U = T - 5/4\, ap. \quad (4)$$

This form of U we already knew, and indeed used it to motivate eq. (2). Plotting this as a function of T gives the grey lines in Fig. 1 for $a = 0$, 1 and 2, with p fixed.

The reason that the orange lines cannot pass through the top grey line, and the grey lines cannot pass through the bottom orange line, is because a and b must be non-negative. The touching points are determined by $T \geq (4/5\, s + p)a$ and $T \geq (4/5\, s + q)b$.

Since $\frac{\partial U}{\partial b}$ in eq. (3) is always positive, to maximise U we want b to be as big as possible. Likewise, since $\frac{\partial U}{\partial a}$ in eq. (4) is always negative, to maximise U we want a to be as small as possible.

Discussion

The main take-home message of this paper is that the amount of unrestricted income U can vary a lot, depending on the choices of a and b. It can lie anywhere in the grey cones in Fig. 1, the structure of which are set by p, q and s. For a given grant total T, U increases with b and decreases with a.

Fraction of unrestricted funds

To see how the size of U can vary for grants of the same size T, here is an example. Consider two versions of a three-year grant, both of size $T =$ £528,000. In both versions we set the postdoc salary to be £40,000, the PI salary to be £80,000, and the support costs rate to be £80,000. In units of £100,000, $p = 0.4$, $q = 0.8$ and $s = 0.8$. In the first version of the grant, we choose $a = 41/26 = 1.58$ (so there is an average of 1.58 postdocs per year) and $b = 1/12$ (so a PI spends a month per year on the grant). This gives an unrestricted portion of $U =$ £291,000, which is 55% of the total. In the second version, we choose $a = 1$ and $b = 1/2$. This gives an unrestricted portion of $U =$ £378,000, which is 72% of the total.

The fractions above are just two of the values that U/T can take. What is the range of values as we vary a and b? The minimum occurs for $b = 0$ and the maximum for $a = 0$, giving the bounds

$$\frac{s - {}^{5}\!/_{16}\, p}{s + {}^{5}\!/_{4}\, p} \leq \frac{U}{T} \leq 1. \qquad (5)$$

These bounds are the slope of U in eq. (3) and eq. (4). For the typical values $p = s/2$, eq. (5) becomes ${}^{27}\!/_{52} \leq {}^{U}\!/_{T} \leq 1$. Multiplying eq. (5) by T, the left and right sides are the bottom orange lines and top grey lines in Fig. 1.

Principal investigators versus postdocs

How much should an established scientist invest in doing research, versus training and managing junior scientists to do it? In other words, how should we balance PI time b and postdoc time a? The answer depends on the type of science being done, the function of the organization doing the research, and scientific culture itself.

There are three broad types of science: experimental, computational and theoretical. Experimental science is more labor intensive than computational science, which in turn is more labor intensive than theoretical science. Experimental and computational science require a hierarchy of skills, and some of the less-skilled labor falls to postdocs. So as we move from experiment to computation to theory, we would expect postdoc time a to decrease and PI time b to increase.

The structure of the research organization also plays a role in balancing a and b because it determines how much time PIs can dedicate to research. Universities exist to pass on knowledge and create knowledge, that is, to teach and do research. At universities, in general half or more of a PI's time is taken up by teaching and administrative duties. On the other hand, at dedicated research institutes, such as the Institute for Advanced Study and the London Institute for Mathematical Sciences, PIs tend to spend most of their time on research. So they are inclined to choose higher values of b.

The balance between PIs and postdocs also depends on whether we seek an exponential growth in the number of active scientists—a movement that began after WWII—or rather seek to keep the number steady. If one in a handful of postdocs goes on to an academic career, then at steady state a PI should have just a handful of postdocs over the course of his or her entire career. In other words, values of a should be lower than values of b, which is counter-cultural in most universities today. The current culture of exponential growth is unsustainable, particularly in basic science, because there are not enough long-term job opportunities compared to the number of postdocs looking for jobs.

Universities versus institutes in Britain

The three factors that influence how to balance a and b above apply to all research organizations. However, in Britain, there is another aspect, which we discuss here. The London Institute for Mathematical Sciences is Britain's only Independent Research Organisation in the physical sciences. In its grant applications is uses a postdoc salary of £45,000 ($p = 0.45$) and its UKRI support costs rate is £91,065 ($s = 0.91$). Inserting these values into eqs. (1) and (2) gives for each year of the grant

$$T = 1.18a + (q + 0.73)b,$$
$$U = 0.62a + (q + 0.73)b,$$

where q is the salary of the PI.

[1] UK research council success rates fall as budgets tighten, *Times Higher Education*, 18 Sep 2023.
[2] https://www.ukri.org/who-we-are/policies-standards-and-data/funding-assurance-programme/.