Merit rating and formula-based resource allocation

Juha M. Alho

University of Joensuu, Joensuu, Finland

Mikko A. Salo

University of Joensuu, Joensuu, Finland

Keywords

Funding, Hierarchy, Universities, Model

Abstract

Formula-based allocation schemes are often proposed as a means of making the allocations of public funding more equitable, and more efficient. Unfortunately, these goals may be undermined by imperfect measures of merit. In particular, if the measures are subject to random variation or manipulation, the formula-based schemes may lead to disincentives and lack of efficiency. Develops a simplified model of a hierarchical allocation system and analyse stochastic models for the most important sources of variation in the system. Practical proposals for limiting the variability are studied and the framework is applied to the Finnish university system.

Introduction

The allocation of public funds to local governments, such as states or municipalities, or to local institutions, such as regional universities or their departments, is frequently suspected of being inequitable or inefficient unless explicit measures of merit or need are used as a basis for dividing the funds (e.g., Caudle *et al.*, 1987; Huffman and Just, 1994; Johnes, 1996; Mayston, 1998; Roy, 1985). A well known example is a series of legal proceedings in the 1970s in which a number of US courts found the system for funding the public schools based on property tax unconstitutional (e.g. Fisher, 1996, pp. 503-17). It was claimed that "right to an education in public schools is fundamental interest which cannot be conditioned on wealth" (Musgrave and Musgrave, 1989, pp. 29-31).

The public funding schemes that use merit or performance rating follow Aristotle's equity principle, which states that goods should be divided in proportion to each claimant's contribution to the common cause (Young, 1994). For example, Serban and Burke (1998) argued that the interest in the 1990s in the use of performance criteria in higher education increased accountability and improved institutional performance. The merit-based allocation systems also function by equalizing the capacity of public universities to provide degrees and research: each unit is given a fair chance to do its best.

As stated by Nagel (1988), in evaluating public policies, effectiveness and efficiency are as important as equity. While efficiency refers to keeping the costs down, effectiveness means achieving the goals, and any unanticipated side-effects.

In this study, we consider simplified formula-based allocation systems. We focus

on the special problems of small units as recipients. The primary issue we will address is the inherent uncertainty of information concerning small units. For example, in a small department within a university, such performance measures as degrees completed, papers published etc. can be highly variable. We will use statistical models to quantify such variability

The apparent objectivity of the formula-based measurement of merit is convenient because annual debates about allocation priorities will no more be needed (Johnes, 1996). Questions of equity appear to be solved by the legislature although the legislative intent may be buried under the complex details (Spencer, 1980, 1982; Caudle *et al.*, 1987). Efficiency may be enhanced by formula design that incorporates market incentives (Massy, 1994; Johnes, 1996). However, if the measurement of merit is inaccurate, formula-based allocations may lead to lack of effectiveness and inequity among the recipients, even if they share the overall mission of the system.

Our case study deals with recent proposals to reorganise the funding of the Finnish universities. The country has a regional university system consisting of 21 (mostly) small units spread throughout the country. Spending on universities accounts for 1 per cent of Finland's gross domestic product. Approximately two thirds of it consists of state funding and one third of external private funding. State funding is allocated by the Ministry of Education based on the target number of degrees the universities are expected to produce the following year (cf. Hölttä 1995).

The faculty of a Finnish university consists of professorial positions, lecturers, and assistants. There is a great diversity among universities, but it is common to have two to four faculty of each type in a university department. According to a rough rule of thumb, the number of degrees awarded per department is one tenth of the size of the total

enrolled population, or it is typically in the range 5-20. A major technical problem in formula-based funding of the departments is that their “products” (degrees, research papers) are subject to high random variation.

We will first pose the problem in abstract terms, as a problem of the division of a fixed pie according to criteria subject to randomness. After that we will present empirical evidence on the uncertainty in the case of the Finnish university system.

Allocation with random, zero-sum errors

Suppose first that a national university budget is to be divided to universities, schools within universities, or to departments within schools. Concentrating on one level of the hierarchy, suppose that we have to divide a fixed pie into shares $D_i \geq 0$, $D_1 + \dots + D_I = 1$, according to the merit Z_i of the recipients $i = 1, \dots, I$. To simplify the notation, we will assume that the recipients are of equal “size” in the sense that an even division would be just if their merits were equal. (This does not restrict the generality of the discussion because aggregates of arbitrary sizes can always be built from the recipients.) Then, we can define the shares to be of the form

$$D_i = (I + bZ_i)/I. \quad (1)$$

The shares deviate from the average linearly according to the merit Z_i , with a slope given by b/I , where b is a policy variable that controls the overall spread of the shares around the mean. For definiteness, here and in the sequel the merits Z_i are normalised so their standard deviation is one.

Should we have two merits, one merit $Z_i(1)$ might be the number of degrees produced by department i per full time position (cf. Roy, 1985). Then, the other merit $Z_i(2)$ might be the number of refereed publications; it could be some scientific citation index; or it could be a qualitative peer evaluation determined by negotiation (cf. Huffman *et al.*, 1994). In this case we may take the shares to be of the form

$$D_i = (1 + b\{c_1Z_i(1) + c_2Z_i(2)\})/I. \quad (2)$$

where $c_j > 0$ with $c_1 + c_2 = I$. This yields zero-sum deviations from an even split in which b regulates the magnitude of the deviations, and c_1 and c_2 determine the relative importance of the two merits. An extension to more than two merits is immediate.

The allocation with random, zero-sum errors will be illustrated below in the section “simulations under different division criteria”.

Data

We think of a university department as receiving four types of funds:

- 1 salaries of permanent personnel;
- 2 funds for basic operating costs such as supplies, computers, services, travel, books, and teaching not given by permanent personnel;
- 3 internal support for research such as graduate assistantships, university funded projects, etc., and
- 4 outside funding from private or public sources for research, and possible income from services provided to outsiders.

We use the University of Joensuu as an example of how the share of the university out of the total university system can vary in real life. With its 90 professors and 5,000 students its share of all professors of Finland is 4.6 per cent and its share of all university students of Finland is 4.5 per cent. We then use two departments, one in social sciences (A) that does a considerable amount of service teaching but relatively little funded research, and a science department (B) that has a heavy research orientation, to show how the share of a department may vary within university[1].

Our estimates of the complete budgeted allocations to the Finnish university system (in 1994 prices) decreased from FIM 5.0 billion to FIM 4.6 billion during 1991-1995 (USD 1 = FIM 5.6 approximately, see Table I.). The decrease was primarily due to the severe economic depression that started in the early 1990s.

Table I shows that governmental funding of the University of Joensuu decreased in real terms from about FIM 200 million to about FIM 160 million. In 1991-1994, its share of the state university budget decreased from 3.95 per cent to 3.55 per cent but rose to 4.15 per cent in 1995.

Table I shows that the share of department A has increased from about 1.5 per cent to 1.7 per cent, whereas B has increased its share from 4.9 per cent to 5.6 per cent in 1992-1995. In relative terms, both departments have experienced similar relative improvement inside the university.

To put the state appropriations into perspective, in Table I we also present rough estimates of outside funding concerning the total university system and the University of Joensuu based on data from a national database, and estimates of outside funding supplied by department B. For department A the role of outside funding was not important, so those figures are not included. The figures are presented relative to the 1994

Table I

Descriptive statistics of funding of universities in Finland in 1991–1995 (at 1994 prices)

	1991	1992	Year 1993	1994	1995
Estimated public expenditures (FIM, millions)[2]					
State budget	4,966	5,000	4,776	4,391	3,924
University of Joensuu budget	196	191	179	156	163
Share or budgeted allocations					
University of Joensuu share of state budget (%)	3.95	3.82	3.75	3.55	4.15
Department A share of University of Joensuu budget (%)	–	1.52	1.47	1.66	1.70
Department B share of University of Joensuu budget (%)	–	4.94	4.80	5.58	5.59
Ratio of outside funding to the 1994 budgeted allocations					
All universities	0.12	0.14	0.16	0.20	–
University of Joensuu	0.24	0.27	0.28	0.33	–
Department B	–	0.59	0.84	0.87	1.50

levels of Table I so both relative magnitude and time trends can be studied.

We find that the department B has an unusually high relative income from funded research both within the University of Joensuu and the whole system. We also see that the University of Joensuu has a clearly higher income from outside sources than the average.

Estimation of the variability in 1991–1994

We will build simplified statistical models to describe the availability of funds to a university department. The first model represents the annual variability observed during 1991–1994. The second model shows how the situation would change if pure merit-based criteria were used. The goal is to quantify the statistical variability in each case.

Let $C(t)$ be the overall state allocation to universities during year t and define $c(t) = \log(C(t))$. Past data show that $c(t)$ has been nonstationary, following the development of the economy. The downward trend visible in Table I might continue, but it might turn upwards, as well. In fact, during the previous 15 years the university sector expanded considerably. A simple model that captures this idea is that $c(t)$ is a (normally distributed) random walk with

$$c(t+s) - c(s) \sim N(0, t\sigma_c^2) \quad (3)$$

for all s and $t > 0$. Based on Table I, we estimate that $\sigma_c = 0.074$.

Define $0 < U(t) < 1$ as the share of the University of Joensuu during year t from $C(t)$, and let $u(t) = \text{logit}(U(t))$. The share seems rather stable. A reasonable model assumes that the $u(t)$ are independent and identically distributed (*i.i.d.*) with

$$u(t) \sim N(\mu_u, \sigma_u^2). \quad (4)$$

Based on Table I we estimate that $\mu_u = -3.22$ (corresponding to a median share of 0.038, or 3.8 per cent out of $C(t)$) and $\sigma_u = 0.060$.

Similarly, define the shares of departments A and B as $D_i(t)$ with $d_i(t) = \text{logit}(D_i(t))$, $i = A, B$, and assume that the $d_i(t)$ are *i.i.d.* with

$$d_i(t) \sim (\mu_{d,i}, \sigma_{d,i}^2). \quad (5)$$

Based on Table I we estimate that $\mu_{d,A}(t) = -4.123$ (corresponding to about 1.6 per cent); $\sigma_{d,A} = 0.071$; $\mu_{d,B}(t) = -2.90$ (corresponding to about 5.2 per cent); and $\sigma_{d,B} = 0.084$.

Now, the actual allocations $W_i(t)$ of the two departments can be defined as

$$W_i(t) = C(t)U(t)D_i(t) \quad (6)$$

for $i = A, B$. In the past, the major cause of variability has been the nonstationary $C(t)$. The shares $U(t)$ and the $D_i(t)$ s have been roughly stationary and equally variable. When we consider formula-based allocation schemes, their effect would primarily involve the $D_i(t)$ s, and to a lesser extent $U(t)$. In particular, the variability of $C(t)$ would not be influenced.

Simulations under different division criteria

Consider now the variability of the measures of merit. The first measure is the number of degrees produced in a department. In the Finnish system there is no tuition, so the students do not have a similar incentive to graduate as US students do, for example. Similarly, changes of major are relatively common. The fact that about one in ten among the enrolled gets a degree each year, although the degrees are planned to take less than five years of full time study, also suggests that there is a great deal of

unpredictability in the time it takes a student to complete a degree. All these factors point to Poisson distribution as being a reasonable descriptor of the variability in the number of degrees completed annually in a department (cf. Feller, 1968, p. 282). In Alho (1996) we provided additional evidence for the Poisson assumption.

Therefore, we have used a Poisson model

$$X \sim Po(\lambda_X) \quad (7)$$

for the number of degrees of a department, and the model

$$Y \sim Po(\lambda_Y) \quad (8)$$

for the number of degrees produced in all other departments of the university. To make the results of the formula-based allocation schemes comparable to the current situation of the University of Joensuu we assume that the expected number of degrees at the university is $\lambda_X + \lambda_Y = 550$. If the share of, say, department A should be 1.6 per cent, the expected number of degrees producing this would be $\lambda_X = 8.8$. Accordingly, the 5.2 per cent share of department B would correspond to $\lambda_X = 28.7$ degrees of the total of 550.

Recall that for a Poisson distribution with expectation λ , the variance is also λ . For example, if we expect nine degrees annually in a department, then the standard deviation is three, and the ratio of the standard deviation to the expectation is one third. In other words, we expect the number of degrees to vary by about 33 per cent annually.

If the number of degrees produced directly determines the allocations, the share of the department $i = A$ would be calculated according to (1) as $D_i = (1 + bZ_i)/I$, where $I = 550/8.8 = 62.5$ (i.e. the university is implicitly thought to consist of I departments that are identical to the one we are considering) and

$$Z = \{X - \hat{\lambda}_X\}/\hat{\lambda}_X^{1/2} \quad (9)$$

is the empirical merit in which we have used the estimate

$$\hat{\lambda}_X = (X + Y)/I \quad (10)$$

for the expectation and variance of X . For example, if in a given year the rest of the university produces $Y = 550$ degrees and department A produces $X = 5$ degrees, then $\hat{\lambda}_X = 555/62.5 = 8.88$. It follows that $Z_i = (5 - 8.88)/2.98 = -1.30$. If $b = 0.1$, for example, then $D_i = (1 - 0.13)/62.5 = 0.0139$ as opposed to $1/62.5 = 0.16$ that would have been their share, had they produced their expected share.

Now the only difference between simulations corresponding to the current system (as given by (5)) and the

formula-based system with a single merit is the relative variability.

We will use the number of publications as the second type of merit. Their number is harder to characterise, because they may include various types of reports, refereed journal articles, monographs, computer programs etc. For the purpose of studying the possible stabilising effect of the second merit, we will simplify the matter and assume that some standard can be chosen, such that the number of publications has the same expectation as the number of degrees in a department.

Results

We used the empirical estimates of the previous section to establish a baseline level of variation for the departments. Since Finland underwent a major economic depression during 1991–1995 past five years, it is informative to include a second set of simulations in which the variability in (3), (4) and (5) is one half of the past empirical level, and a third set of simulations in the variability in (3) and (4) is eliminated altogether. Since the policy variable b controls the level of variability in the system, we experimented with three values ($b = 0.05, 0.10, 0.20$). The higher the value of b , the larger the potential incentive of the formula-based allocation system. Since large values of b entail heavy variability, it is natural to try to reduce it by aggregating merit data across several years. We will show what effect this has. Similarly, since the use of a single merit leads to problems of stability, we will illustrate the case of two sources of merit.

Table II has results from simulations that match the situation of department A, with $\lambda_X = 8.8$ degrees in a university with $\lambda_X + \lambda_Y = 550$, each year. The X s and Y s are assumed to be independent over time. This ignores the negative autocorrelation we expect because graduation during one year prevents it from occurring during another. The autocorrelation is small, however (Alho 1996).

Table II assumes the past level of variability in the state budget and the shares of university and department are as estimated in (3), (4) and (5). The standard deviation of the annual relative change is the measure of variability we have used. i.e. if $C(t)$ is the allocation of year t , then we have used $\text{Var}(\log(C(t)/C(t-1)))^{1/2}$ as the measure of year to year change.

The calculations were carried out using Minitab. The results shown are based on 1,000 simulations for each parameter

combination. An indication of the amount of sampling variability remaining in the estimates can be seen in the bottom rows (current) that show four replications of the same simulation set.

We see that if five-year data can be used instead of single-year data, variability can be reduced. Similarly, two measures of merit is preferable to one. We will comment on the interpretation of this aspect later. Consider the role of the policy variable b now. For small values of b variability is small it can be reduced to about 0.11 in Table II by taking $b = 0.0$. However, small values of b mean that very little weight is given to merit in the allocation of the funds.

One policy implication of Table II is that if we do not want the variability of a department budget to increase, then the amount of incentive permissible depends on the way merit is measured. Under the past level of variability, a maximum value of a bit over $b = 0.2$ can be used if two measures of merit can be aggregated over a five-year period. If only single year, single merit data are available, then we see by interpolation from Table II that b cannot exceed 0.07, approximately.

By interpolation we draw another policy implication of Table II, namely that the availability of two similar measures of merit stabilises the budget in approximately the same way as aggregating a single measure over three or four years.

Further simulations were carried out in the case in which there were no external sources of randomness. The current level of variability in the share of department A induces a standard deviation of annual relative change of about 0.10 on the allocations. The value of b that gives the same level of variability is approximately $b = 0.07 - 0.08$ for a single year, one or two merit systems; and $b = 0.17$ for five-year, one or two merit systems. These agree fairly closely with the estimates one obtains by simulation from Table II.

The above figures refer to a department with $\lambda_X = 8.8$ expected degrees per year. The 5.2 per cent share of department B would

correspond to $\lambda_X = 28.7$ degrees out of the total of 550. With a larger number of degrees expected than for department A we would expect the measures of merit to be more stable. However, although the expected number of degrees is three times as large as for department A, the results are virtually indistinguishable from those of Table II. Since these values of expected numbers of degrees cover the bulk of departments in Finnish universities, we conclude that Table II can be generalised to most units in the system.

Conclusion

We have used data from two university departments to demonstrate what would happen if their public funding was made to depend directly on their share of academic degrees and/or scientific publications produced by the university. Both departments are small, and their output has a Poisson-like character. If the level of variability in the state budget persists on a high level, then the choice between formula-based or other allocation systems within the university will not have a major effect at the department level (cf. Johnes, 1996). However, if the economic development stabilises, then the choice made within university matters, and formula-based Poisson-like measures of merit will lead to high annual variability in allocations that cannot be controlled by the department. This can be demoralising; or what is intended to act as an incentive, may act as a disincentive instead.

This is in sharp conflict with the aims of an allocation system based on quantitative measures of merits. We have discussed several options to control the variability.

First, there is the option of using formulas that depend on more than one merit, as suggested by Roy (1985), for example. Our models and simulations support the use of a combination of several complementary measures. However, unless they are clearly less variable than the Poisson-like measures we have considered, the improvement is likely to be small.

Second, we can ask whether more reliable merit ratings could be achieved by aggregating data over several years, e.g. five years (cf. Roy, 1985). This does increase the stability of the allocations but in small units considerable variability remains. Moreover, the effect of innovations becomes only slowly visible when data aggregated over time are used. Hence, aggregation over time reduces the incentive value of the formula-based allocations.

Table II

Standard deviation of annual relative change in revenues of department A assuming that the variability in state, university and department budgets is at current level

Policy	Single-year data		Five-year data	
	1 merit	2 merits	1 merit	2 merits
$b = 0.05$	0.129	0.114	0.116	0.109
$b = 0.10$	0.168	0.139	0.122	0.108
$b = 0.20$	0.310	0.226	0.170	0.135
Current	0.143	0.143	0.147	0.145

Third, small units could be combined into bigger units (Huffman *et al.*, 1994). Our simulations demonstrated that this did not induce real year-to-year stability. In fact, combining five similar units has a much smaller effect than aggregating data over five years for a single unit.

Thus, we conclude that formula-based allocation systems do not work as expected in the context of small units, if only Poisson distributed output measures are available. An indirect recognition of this fact is that formula-based allocation schemes often include ad hoc rules that bound the maximum change in the allocations. The stability problems may be aggravated by the potentially corrupting effect of formula-based schemes, for example, if a university department is rewarded by the number of academic degrees it awards, or by the number of research papers it publishes, it is not hard to predict that the quality of both will decline (Huffman *et al.*, 1994).

Hierarchical allocation systems may provide a way to reduce the variability of the formula-based allocations, provided that accepted measures of merit exist that are more stable than the Poisson-like measures we have considered (cf. Johnes, 1996). The potential usefulness of the hierarchy is due to the fact that different criteria may be used at different levels of the hierarchy, and in different parts of the organisation.

Another potential way of handling the random variation in the measurement of merit is to permit the establishment of departmental "savings accounts". Having the right to open their own savings accounts and to accumulate a buffer fund would protect the units against the uncertainties of state budgets or stochastic allocations.

More generally, our results suggest that allocation systems should be based on criteria whose robustness is assured. It may not be desirable to go after efficiency, if one cannot realistically measure it.

Note

- 1 The Ministry of Education has established a database (KOTA) that provides budgetary and performance information concerning Finnish universities. However, there have been several technical changes in budget categories. Therefore, we could not use the KOTA figures, as such. A solution we found feasible was to take the year 1991 as a baseline, and estimate the total annual budgeted funds to the university system that it would have had, had the 1991 budget structure been in force. All figures have been inflated to the 1994

price level using the consumer price index (for 1995 the index of July was used).

References

- Alho, J. (1996), "Stokastiset tutkintomarkkinat", in Ahponen, P. (Ed.), *Yhteisellä asialla - monella äänellä*, Studies in Social Policy, No. 4, University of Joensuu.
- Caudle, S.L. and Newcomer, K.E. (1987), "Grant allocation decision rules: expectations versus realities", *Evaluation Review*, Vol. 10 No. 6, pp. 860-79.
- Feller, W. (1968), *An Introduction to Probability Theory and Its Applications*, Wiley, New York, NY.
- Fisher, R.C. (1996), *State and Local Public Finance*, Irwin, Chicago, IL, pp. 503-17.
- Hölttä, S. (1995), *Towards the Self-Regulative University*, University of Joensuu Publications in Social Sciences, No. 23, University of Joensuu, Joensuu.
- Huffman, W.E. and Just, R.E. (1994), "Funding, structure, and management of public agricultural research in the United States", *American Journal of Agricultural Economics*, Vol. 76 No. 4, pp. 744-59.
- Johnes, G. (1996), "The British experience of reform in education", *Journal of Higher Education Policy and Management*, Vol. 18 No. 2, pp. 125-37.
- Massy, W.F. (1994), *Resource Allocation Reform in Higher Education*, National Association of College and University Business Officers, Washington, DC.
- Mayston, D.J. (1998), "Devolved budgeting, formula funding and equity", *Management Accounting Research*, Vol. 9, pp. 37-54.
- Musgrave, R.A. and Musgrave, P.B. (1989), *Public Finance in Theory and Practice*, McGraw-Hill, New York, NY, pp. 29-31.
- Nagel, S.S. (1988), *Policy Studies - Integration and Evaluation*, Greenwood Press, New York, NY.
- Roy, R. (1985), "Funding science: the real defects of peer review and an alternative to it", *Science, Technology, Human Values*, Vol. 10 No. 3, pp. 73-81.
- Serban, A.M. and Burke, J.C. (1998), "Meeting the performance funding challenge", *Public Productivity and Management Review*, Vol. 22 No. 2, pp. 157-76.
- Spencer, B.D. (1980), *Benefit-Cost Analysis of Data Used to Allocate Funds*, Springer, New York, NY.
- Spencer, B.D. (1982), "Technical issues in allocation formula design", *Public Administration Review*, Vol. 42, pp. 524-9.
- Young, H.P. (1994), *Equity in Theory and Practice*, Princeton University Press, Princeton, NJ.

Further reading

- Box, G.E.P. and Jenkins, G.M. (1976), *Time Series Analysis*, Holden-Day, San Francisco, CA.