Incentive Funding Meets Incentive-Based Budgeting: Can They Coexist?

Daniel W. Lang
University of Toronto

Abstract

Two major developments in the financial management of higher education have occurred more or less contemporaneously: incentive or performance funding on the part of government and incentive-based budgeting on the part of institutions. Both are based on fiscal incentives. Despite their several inherent and interconnected similarities, incentive funding and incentive-based budgeting have been viewed and appraised on parallel tracks. This study investigates their convergence. In doing so, it sharpens the definitions of both, identifies their respective track records, and discusses problems that are chronic to both. The study concludes that although incentive funding and incentive-based budgeting are sometimes at cross-purposes, they are functionally interconnected. The study uses Canada as an example because it is the jurisdiction that so far has seen the most extensive mutual deployment of performance funding and incentive-based budgeting.

Résumé

Deux changements importants sont survenus plus ou moins simultanément dans la gestion financière de l’enseignement supérieur : le financement incitatif ou basé sur le rendement, pour ce qui est du gouvernement, et le budget basé sur des mesures incitatives, pour ce qui sont des institutions. Tous deux sont basés sur des incitatifs fiscaux. Malgré plusieurs similitudes inhérentes et inter-reliées, le financement incitatif et le budget basé sur des mesures incitatives ont été considérés et évalués en parallèle. Cette étude se penche sur leur convergence. Ce faisant, elle en affine les définitions, identifie leurs résultats respectifs et traite des problèmes chroniques qui s’appliquent à tous les deux. L’étude conclut que, même si le financement incitatif et le budget
basé sur des mesures incitatives travaillent parfois à contre-courant, ils sont inter-reliés dans leurs fonctions. L'étude utilise le Canada comme exemple parce que, jusqu'à présent, c'est la juridiction qui a connu le plus important déploiement mutuel en matière de financement basé sur le rendement et en matière de budget basé sur des mesures incitatives.

Introduction

The last 25 years has seen the introduction and evolution of two practices in the financing of public universities that are based on incentives: performance funding and incentive-based budgeting. Both are known by other names—for example, “incentive funding,” “set-aside” funding, “matching” funding, “value-centred management,” “responsibility centre budgeting,” “value-centered management,” and even “every tub on its own bottom.” Despite contemporaneous timing and similar nomenclature, the two practices are not usually associated with one another. Performance funding is an instrument of public policy and is exercised “top down” by government. Incentive-based budgeting is a matter of institutional choice and strategy. One is mainly about revenue in the form of public subsidies, while the other is about all revenue and expense, in particular net revenue or expense.

On closer examination, however, we see underlying organizational principles that are shared by both. Both address principal–agent relationships. Both assume that resource dependence determines much institutional behaviour. Both assume certain implicit relationships between patterns of revenue and patterns of expense. Both assume that all institutional cost functions are linear. Few assume that unit costs vary by institutional size, complexity, or mission. The problem is that governments and universities rarely share the same assumptions. This leads to some as yet unexamined questions. How can or should the two practices interact? Are they on a course to collision or a course to mutual benefit?

Canada presents a useful context in which to examine this question. Eight of the 10 provinces employ some form of performance indicators. Some are directly attached to funding and some indirectly attached. Many of these arrangements have been in place for at least a decade, sometimes longer. The arrangements vary in structure and the amounts of funding. Almost all Canadian universities are public. More than a dozen Canadian universities—particularly large, research-intensive ones—have deployed incentive-based budgeting in one variant or another. The combination of performance funding and incentive-based budgeting in Canada is fortuitous. Although performance funding is in use in several American states, and nearly 50 American universities use some version of incentive-based budgeting, the overlap is small: in only a few of the American states that use performance funding are there also public universities that use incentive-based budgeting. The overlap is much larger in Canada.

Performance Funding

Performance funding is not a new idea. Nearly two decades ago, Guy Neave (1988) introduced the phrase “the evaluative state.” At that time, Neave was reflecting on a variety of practices and policies that had been installed to assist universities and, more often, the states that supported them, to cut the higher educational suit to fit the public purse cloth by quantitative measurement. A decade later, Einhard Rau (1999) presented a small
but important paper that asked: “Performance Funding in Higher Education: Everybody Seems to Love It But Does Anybody Really Know What It Is?” The title of Rau’s paper was telling. By the turn of the 20th century, practices that previously had been tolerated on an assumption that they were ephemeral and would go away were not only still in use but were also more popular, at least among governments and other agencies that provided public subsidies to higher education. Moreover, and perhaps more importantly, Rau’s research indicated that despite a decade of experience, mainly of the trial-and-error variety, performance funding was poorly understood and, in the views of many, ineffective or flawed. Even the language of performance funding is problematic. Performance funding, performance indicators, benchmarking, best practice, incentive or “set-aside” funding, performance budgeting, performance reporting, performance agreements or contracts—they all seem at once to be different and the same. In addition to not knowing exactly what performance funding is, we are not certain that it works, and why or why not.

It is not possible to discuss performance funding as if it were a single-cell public-policy organism. Not all are based on incentives. There are several subsets, the most common of which are performance set-asides or earmarks that reserve small proportions of public subsidies for higher education to be paid out on the basis of pre-determined and purpose-built metric targets—hence, performance indicators. Funding thus reserved is an entitlement and potentially open-ended. From the government’s fiscal perspective, the set-aside amount may be overspent or underspent. In most cases, institutions do not compete with one another for these funds in a zero-sum contest. The ultimate public policy objective is to influence or modify institutional behaviour by means of financial incentives. The incentives are exactly that: they are fiscal inducements that only coincidentally correspond to institutional costs. In certain cases, primarily in Europe, this form of performance funding is called payment for results.

There is, however, a competitive version of performance funding. This looks a lot like the set-aside model but with an important difference: it is a zero-sum contest. Depending on the performance of individual competing institutions, the final value per measured “performance” may rise or fall, but the government never spends more than the budgeted amount. The World Bank promotes and underwrites this type of performance funding in countries with relatively limited discretionary resources to direct to the development of colleges and universities (Salmi & Hauptman, 2006).

As expressions of fiscal policy, these two versions of competitive performance funding serve different purposes. The first offers benefit advantages. The state promotes and, hopefully, secures institutional performances that are desirable as public policy. A frequent example in the United States is expanding access for under-represented social groups. By providing open-ended performance funding, the state is indicating a willingness to accept and pay for as much of a given desired performance as institutions provide. The second, because the funding is a fixed sum, offers cost advantages. As performances expand in response to the incentive within the fixed sum, unit costs are either contained or reduced, and, as is the case in Alberta and Switzerland, institutions are assumed to be more efficient. This is the contemporary version of performance funding that comes closest to what Neave had in mind 25 years ago.

Some jurisdictions, the State of Texas, for example (Ashworth, 1994), have used bundled performance set-asides. Under this arrangement, incentive funding is accessible
by universities in response to a collection or “bundle” of several indicators. This allows each university to use the performance indicators and consequent funding for purposes of strategy and planning as well as budgeting because the financial outcomes of responding to the various indicators can be modeled.

Switzerland deploys what might be called “negative” performance funding that is used more as a stick than as a carrot, to promote efficiency and private fundraising on the part of universities. Public subsidies are limited by indicators, in exchange for an incentive to replace the lost public revenue by private funding and by improved efficiency. The results so far are mixed (Schenker-Wicki & Hurlimann, 2006).

Although technical, there are two fundamental aspects of performance funding—especially set-asides and payments for results—that affect their effectiveness in terms of the institutional behaviours that they engender. The first aspect is not so much about performance indicator algorithms as it is about the source of the funds that the algorithms allocate. If the funds available for allocation are new or additive, the incentive is truly a carrot that institutions may, literally, take or leave according to their autonomous judgment. The experience of the Excellence Initiative in Germany bears this out (Melnyk, 2014). If, however, the funds available for allocation come from existing public grants to colleges and universities, the incentive may be as much a stick as a carrot, and as such will be harder for institutions to ignore, regardless of their missions, as appears to be the case in Alberta (Barnetson, 1999; Barnetson & Cutright, 2000).

The second factor that affects the effectiveness of performance funding in modifying institutional behaviour is the match between the amount of funding that is set aside and the “performance” or other behaviour that any given incentive is put in place to engender. If the match is inaccurate or deficient, performance funding will fail. Let’s use rates of graduation. To improve rates of graduation, a college or university might take several steps that involve additional expense—for example, more academic counselling, writing labs, math labs, teaching assistants, and financial aid. The list could be longer, but the length of the list of measures that might be taken to improve rates of graduation is not the point. The point is the cost of the list. If the amount of funding set aside does not reflect, at least approximately, the marginal cost of the institutional performance for which the formula calls, the incentive will be ignored, as it often is (Chan, 2014; El-Khawas, 1998; McCollm, 2002; Miao, 2012; Rau, 1999; Schmidt, 2002; Schmidtlein, 1999). Measured in terms of effective cost ratios, the incentives should be ignored, even from the point of view of government (Harris, 2013).

Matching performance funding is an arrangement somewhat similar to performance funding in which the funding is not public. Governments, in order to leverage private funding, offer to match charitable gifts that as de facto endowments are restricted to purposes designated by the state instead of donors. Funding is set aside for each purpose and not released until matching private gifts are actually received. The funding thus set aside can be either a fixed amount (hence the competitive version of performance funding) or open ended. The consequent performance funding is thus a mixture of public and private funding. Matching funding fits the basic inducement or incentive definition because the public portion is never enough to meet the total cost (Brooks, 2000). In Canada, the federal government, through the Canada Foundation for Innovation, used matching funding as a device to finance research infrastructure (Canada Foundation for Innovation, 2013).
The Track Record of Performance Funding

There have been two iterations of performance funding. The first began in the early 1980s and extended to, approximately, 2010, during which time the number of jurisdictions deploying performance funding rose to a peak around 2006 and then began to decline. In some cases, the decline was permanent and in others temporary (Dougherty & Reddy, 2013; McKeown-Moak, 2013; Ziskin, 2014). Since the mid-1990s, the Rockefeller Institute of Government has conducted a series of surveys of the use of performance funding in the United States. The deployment of performance funding grew rapidly from 1979 to 2006, at which time it was in place in some form in 27 states. Two-thirds of those states, however, at one time or another discontinued the practice or held it in abeyance (Dougherty & Reddy, 2013; Midwestern Higher Education Compact, 2009). Also, of the remaining states where performance funding remained in place, two used it for two-year colleges only. Thus, as of 2012, performance funding for universities was in use in a dozen American states, or in less than half of the historical high. There were, however, at that time some signs that its use was increasing for community colleges (Dougherty, Natow, Bork, Jones, & Vega, 2011).

In approximately the same period in Canada, two provinces—Alberta and Ontario—introduced performance funding. Three others followed. In both of the initial Canadian cases, although performance funding remained in place, the amounts of funding allocated on the basis of performance were reduced to nearly negligible levels.

The Rockefeller Institute, in speculating about the levelling off in the use of performance funding in the United States, stated: “The volatility of performance funding confirms the previous conclusion that its desirability in theory is matched by its difficulty in practice. It is easier to adopt than implement and easier to start than to sustain (Burke, Rosen, Minassians, & Lessard, 2000, p. 7).

What makes performance funding volatile? One explanation has already been mentioned: the amounts of funding associated either with performance funding generally or with specific performance indicators usually do not correspond with the cost structures of the performances that are being measured and putatively rewarded. For example, given the efforts that a university would have to exert in order to raise rates of graduation—smaller classes, enhanced academic services, supplementary financial aid—the net costs that the university would have to incur might be greater than the additional income that those efforts would generate.

Also in terms of cost structures, performance funding often fails to take into account the fact that universities have long production cycles and variable economies of scale. For example, the typical undergraduate program takes four years to complete; many programs take longer. For that reason, universities are something like super-tankers: it takes a long time to change their direction, even when they are willing to change in response to financial incentives. Let us again take the rate of graduation as an example. First, the rate of graduation is not a simple sum of annual retention rates. Most graduation rate performance indicators are not calculated until one or two years after the normal program length—for example, after the sixth year for a four-year program (National Center for Education Statistics, 2013). This allows for the inclusion of students who “stop out” or temporarily switch from full-time to part-time status, but who nevertheless eventually
graduate. Thus, even if a university makes every possible authentic effort to increase its rate of graduation, the results of those efforts will not be seen until several years later. But performance funding universally operates annually. This means that a university must incur costs long before it receives supplementary “performance” revenue to cover those costs, and even then usually partially instead of fully.

Even the delayed recovery of costs is problematic. One of the reasons most often cited for the disinclination of some universities to take performance funding seriously is uncertainty about the future (Burke & Modarresi, 2000; Callahan, 2006; Dougherty & Natow, 2010; Hearn, Lewis, Kallsen, Holdsworth, & Jones, 2006; McColm, 2002). Will the definition and calculation of performance indicators change over time? Will the amount of funding attached to performance change? Will new indicators be introduced that offset older indicators? These concerns about stability are not unfounded (Dougherty & Natow, 2010; Hearn et al., 2006). In Ontario, for example, the performance funding cum performance indicators program changed four times in eight years.

Some jurisdictions deal with the problem of cost by limiting the number of indicators so that the performance funding available to each indicator will be higher and therefore closer to a reflection of the actual costs of the performances that it measures. This, however, creates a Catch-22 problem. As the array of performance indicators narrows, the indicators cover less of each university’s total performance, which in turn makes the measurement of institutional performance less reliable and performance funding less influential. Context is crucial in appreciating the complexity of this problem. With one exception, no Canadian province or American state has ever allocated more than six percent of its total funding for post-secondary education through performance funding. The one exception—South Carolina—suspended its performance funding program in 2003. It is difficult to imagine any manipulation of an array of performance indicators that could realistically have matched the performance measured with the actual costs of that performance. These facts, however, do not necessarily mean that the amounts were ineffective as incentives. If these allocations were the only truly additive funding available, they still could have been large enough to serve as incentives, particularly in cases where knowledge of attendant costs was problematic. Not one of the stakeholder groups surveyed by the Rockefeller Institute—from state governors and legislators to deans and chairs of faculty senates—thought that the amount of funding allocated by performance funding was too large. The almost unanimous consensus was that funding was too small (Burke & Minassians, 2003).

Whether the amounts are “large” or “small” is in the eye of the beholder. The metrics of performance funding are more than a technical footnote. The conventional metric in almost all the relevant research literature is performance funding as a percentage of state funding, less capital funding. These amounts, jurisdiction by jurisdiction, are then reported to range from as low as one percent to as high as 100 percent, as is evident in the case of Tennessee (Jones, 2103). As straightforward as these data are, they nevertheless mask three major questions, each of which has a lot to do with incentives and impact.

Performance funding so far has essentially been a system of incentives or bonuses. The policy and fiscal “performance” objectives of the incentives have varied over time from jurisdiction to jurisdiction and from “first iteration” to “second iteration,” but the modality of an incentive has not changed. Incentives are not intended or expected to meet
all the costs of the “performances” that they promote. In other words, to colleges and universities, they are marginal revenue. To government, they are the costs of leverage. This exposes the first question: as percentages, are the two—the marginal revenue and the cost—the same? The answer is either no or not necessarily.

Unless a college or university receives all its funding from the state, the conventional metric will always overstate the arithmetical influence or leverage of performance funding. For colleges and universities that are approaching PINO status, the arithmetic effect could be almost negligible. For example, in Ontario, using the conventional metric, two percent of operating funding was originally allocated on the basis of performance. However, as marginal revenue to the University of Toronto, the percentage is less than one-quarter of one percent of total revenue (Council of Ontario Universities, 2014). In terms of whether or not performance funding is successful, the conventional metric can be misleading. What is a cost to a state or provincial treasurer is not necessarily an equivalent incentive to a college or university president. This leads to a second question. Is the median percentage of performance funding revenue across a system the same as the mean? If it is not, as is often the case when funding formulae are based on averages (Lang, 2005), what may be an incentive to one institution in the system may be a disincentive to another.

The third question is also arithmetic. The percentages in the literature appear to be expressions of policy. In other words, as coefficients they are fixed amounts. What if the larger amounts of funding of which performance funding is a part have changed over time, or there have been no adjustments for price inflation? The consequent real-dollar effect as an incentive to institutions may have gained or lost leverage.

What lessons can we learn from trial and error? An examination of the experience of the United States came to the notable but mixed conclusion that performance funding worked in jurisdictions in which the performance indicators emphasized quality or outcomes and did not work where the emphasis was on efficiency (Burke & Modarresi, 2000). But even the track record in terms of quality is not promising. Of the 38 public universities that are members of the elite American Association of Universities, only seven are in jurisdictions that at one time or another deployed performance funding. In the 2013 Times Higher Education Supplement ranking of the world’s top universities, only one university from a North American performance funding jurisdiction ranked in the top 50.

Efficiency is particularly problematic in terms of the measurement of instructional and administrative costs. Most notably, all the universities that in the latest Rockefeller Institute survey of American states reported deployment of performance funding in the future as being “unlikely” or “highly unlikely” cited lack of funding as the reason. In total, 65 percent of all responding states were in the “unlikely” or “highly unlikely” categories. Only six percent were in the “likely” category. Four states that had had performance funding in place for several years reported plans to suspend it due to fiscal constraint. If performance funding really reduced cost and improved efficiency, it would be counterproductive to hold it in abeyance in times of fiscal constraint. This is not simply an empirical coincidence. Performance funding in the public sector is a monopsony. There is only one “buyer”: the state.

When public funds are set aside to finance performance funding, the amounts either are added to the funds already available to institutions or supplant them by redirection or reduction. In the latter case, the result for the institutions is a zero-sum game. Zero-sums
in public finance are often assumed to be beneficial because they stimulate competition. But monopsonies are inherently inefficient (Cooke & Lang, 2009). When underfunding is cited as a cause of performance funding failure, the discussion does not go far enough to uncover a more basic problem. An inference is still possible: that a zero-sum approach might be made to work if more funding were allocated on the basis of performance. That is not so. Monopsonies are always inefficient. Consider, too, that virtually all the metrics of performance funding apply to government as a single financer or nominal buyer. No performance funding program has yet to differentiate between incentives or invite competitive bidding for them (Lundsgaard, 2002). That is monopsony behaviour.

The track records of some “bundled” or “composite” incentive funding schemes provide another lesson about incentive funding. “Bundled” or “composite” performance funding runs the risk of “pooling.” Institutions, sometimes for good reasons and sometimes not, offset bad performances in certain areas with good performances in others. Some “weighted” institutional ranking schemes, like USN&WR and Maclean’s, implicitly encourage such behaviour. The public policy antidote is incentive funding in which funds are set aside by performance and indicator by indicator.

This type of funding has another dimension that sometimes, despite its effectiveness, makes it less attractive to government. In some jurisdictions, governments and funding agencies are becoming wary of performance funding. There are two reasons for this: one political and one financial. The political reason is that this form of funding, some governments are beginning to realize, can work in two directions. If a specific performance target is set, is benchmarked, is visibly measurable by a performance indicator, and is financed by earmarked funding, the effects of inadequate funding can be measured as well institutional performance. In other words, the performance of government as a funding agent becomes visibly measurable, too, and may just as easily become a political liability as an asset.

The other reason is that a tight, realistic, and predictable fit between performance indicators and performance funding generates what amounts to entitlement funding. In other words, the more successful performance funding is in terms of raised institutional performance, the more it costs. Open-ended funding schemes make governments nervous, especially those in tight fiscal circumstances (Blakeney & Borins, 1998; Wildavsky, 1975).

The final lesson learned explains the second iteration of performance funding. Six American states have announced plans to reintroduce performance funding (Dougherty & Reddy, 2013), and Ontario has commissioned a study of the prospect of an expanded performance funding program (Ziskin, 2014). Governments have recognized the importance of an at least approximate match between the amount of performance funding that is made available and the performances that they wish, as a matter of policy, to promote. On the other side, institutions are paying more attention to the net relationship between marginal revenue and marginal cost.

**Incentive-Based Budgeting**

By the end of the 1980s—coincidentally at the same time that performance funding was being introduced—a number of large, research-intensive universities in North America had begun experimenting with an organizational and budgetary concept the principal objectives of which were to relocate and enhance responsibility for planning and budgeting, usually by decentralization, and in turn improve institutional performance in the
allocation and generation of resources and in delivering services. Three decades later, between 50 and 60 universities in the United States and Canada follow the practice, albeit using several different but similar names. For example, what Indiana University calls responsibility center budgeting the University of Michigan calls value center management and the University of Toronto calls simply the new budget model.

Whatever nomenclature is used, it is an expression of the total cost and total revenue attributable to a university academic division. It gives a campus, college, faculty, or department control over the income that it generates and the expenses that it incurs, including indirect and overhead costs. Control over income may include the determination as well as the receipt of fees. Control over expense includes local options for securing goods and services that otherwise would be available only through central university service units. This ineluctably has a highly decentralizing effect by locating many decisions involving the generation and management of resources at different locations in the organizational structure of the university, locations at which, in theory, there is greater familiarity and knowledge about the connections between budgets and programs.

A major difference between the nomenclature of performance funding and that of incentive-based budgeting is the meaning of “cost.” Cost in terms of incentive funding usually means the cost to government and means only the cost of inducing a particular behaviour or performance on the part of institutions. Cost in terms of incentive budgeting means all costs—direct, indirect, and overhead or infrastructure—and, because of the inclusion of revenue, also means net revenue or cost. This is a major and fundamental difference between activity-based costing (ABC) and incentive-based budgeting. ABC is only about cost; it never gets to net cost. “Activity” in ABC is a generic; it applies, as the term implies, to institution-wide classes of expense—for example, library acquisitions or undergraduate instruction. Incentive-based budgeting, on the other hand, is organizational; it applies to academic budgetary units—for example, a faculty of engineering or a department of anthropology.

**Track Record of Incentive-Based Budgeting**

Incentive-based budgeting emphasizes and exposes costs that are often known but not recognized or are deliberately not known because of their strategic implications (Gillen, Denhart, & Robe, 2011). While this demands accuracy and a sound methodology for attributing indirect and overhead costs, its ultimate purpose is not to account for costs. There are other reasons for an institution’s wanting to know about its cost and income structures. The most obvious of these reasons are to account fully for the costs of research and to ensure that auxiliary or ancillary services that are supposed to be self-funding really are. Less obvious but perhaps ultimately more important is to understand better the dynamics of marginal costs and marginal revenues. This is exactly the type of decision that universities have to make about responding to performance funding incentives. It is also the type of decision that governments, as designers and proponents of performance funding, often do not, in Scott’s (1998) terms, “see.”

In terms of budget planning, incentive-based budgeting has a salutary but often upsetting “nowhere-to-hide” effect. Here, incentive-based budgeting shares some of the pooling problems that affect performance funding. Because the arithmetic of performance funding operates at the institutional level, below-average performances of some faculties
can offset above-average performances of other faculties, thus generating less or even no performance funding on a net basis institution wide. When we consider that the basic political economy of any university is to optimize the intersection of quality and cost for every program and service, we see a necessary and almost automatic connection to performance funding. The costs thus identified are the costs that the university can connect to the marginal income generated from performance funding. Having made that connection, the university can make an informed decision whether or not to respond to the performance funding incentive.

This in turn motivates entrepreneurial behaviour and the generation of revenue. In most other institutional planning and budget regimes, the generation of revenue is regarded mainly as the responsibility of the university’s administration. That, as well, is how governments envision incentive funding working. To academic divisions, most services—for example, libraries, media centres, and campus security—are free goods. Because income as well as cost is attributed to campus, colleges, faculties, or departments under incentive-based budgeting, the effect on principals, deans, or chairs is virtually immediate: the generation of revenue (and the reduction of cost) counts. This is the level at which performance funding enters the equation. Mistaken decisions or even wishful thinking about costs versus benefits under incentive funding make real differences close to home.

Institutional plans become anthologies of academic unit plans, the practicality of which within the context of performance funding depends on the array “performances” to which performance funding applies. For example, the rate of graduation is usually calculated at the institutional level. Thus, even though incentive-based budgeting promotes the development of “bottom-up” planning, universities still have to plan centrally for performance funding and be held accountable for that “performance,” even if the practice implicitly invites the pooling of programs that perform weakly with programs that perform strongly. Similarly, if we look at typical indicators to which performance is tied—graduation, retention, time to degree, improved “value-added” student performance—we see that these performances in many universities depend as much or more on shared central student services as on individual faculty services. Here we learn an important lesson: although the momentum of incentive-based budgeting is in the direction of decentralization, the effect of incentive funding is in the direction of centralization. Governments that deploy performance funding may not intend this, but it has the effect nonetheless.

In theory, incentive-based budgeting encourages and rewards the reduction of costs as well as the generation of new revenue. The track record of revenue generation is much better than that of cost reduction (Curry, Laws, & Strauss, 2013). This phenomenon invites a discussion of comparison, best practice, and benchmarking, which collectively are sometimes categorized as a form of performance indicators. Benchmarking in higher education is an import from business in the for-profit sector. In the view of some, although benchmarking did not originate in higher education, it has become a virtually mandatory practice for colleges and universities (Alstete, 1995). Benchmarking is the subset of comparison that focuses mainly on process (Birnbaum, 2000).

When benchmarks are drawn from true peers, their financial effect is primarily on costs and efficiency. In other words, institutions find ways to reduce cost. This, then, connects to incentive budgeting. It also connects to incentive funding when the objective is to promote institutional efficiency. Benchmarking for best practice and ultimately cost
reduction, because it focuses on process, is very laborious and demands sophisticated financial information systems (Gaither, Nedwick, & Neal, 1994; Lang, 2002). This is more than some universities can afford or more than some governments are willing to underwrite. It can also be risky because, in the absence of a corollary effort to insure that best practices are drawn from institutions that are peers, there can be no assurance that what is a best practice in one institution can be a best practice in another (Lang, 2000). As Robert Birnbaum (2000) observed, when that happens, the conversion of a benchmark into a performance indicator is in practice useless for funding purposes. This is why incentive funding is least effective in undifferentiated systems of higher education, and incentive-based budgeting is more effective.

**Chronic Problems**

**Aggregation.** Finding the right level of aggregation is as essential as it is difficult in the successful deployment of performance funding. Michael Porter said that “diversified companies do not compete; only their business units do” (Porter, 1996). This applies to universities. They are very diversified. Porter’s proposition is fundamental to most forms of incentive-based budgeting, which in effect push planning and budgeting down to the level of faculties as “business units.” If we examine individual performance indicators carefully, we see that most of the “performances” that the indicators measure do not really operate at the institutional level. In Ontario, for example, one has to look only at the results of annual surveys of graduates that have been conducted for nearly a decade to see the great extent to which indicator performances vary by program. The variability statistically is greater than it is when measured by institution (Ontario Graduate Surveys, 1999–2009. But it is at the institutional level that the arithmetic of performance funding operates.

Is this a problem to be solved or a lesson to be learned? As a problem, it is unsolvable, at least by any currently known form of performance funding. Programs are diversified for good reasons. In the case of professional programs, third-party regulators (of which government often is one) have powerful influences on the structure and content of programs. There is plenty of evidence that program structure and anticipated employment have strong effects on retention and graduation (Adams & Becker, 1990; Angrist, Lang, & Oreopoulos, 2006; Lang et al., 2009).

Let’s say that the absence of institutional differentiation is an institutional behavioural problem that a system could solve by deploying performance funding. Should it be? Here we enter an unfortunate and fundamentally untenable middle ground between system performance and institutional performance. Performance funding can have externalities. In simple economic terms, an externality is a consequence of an activity between two parties—for example, a government as a principal and a university as an agent—that has an unintended effect on other parties or “performances.” In this case, using rate of graduation as an example, if program diversification were reversed by the incentive of performance funding, students might end up with less curricular and program delivery choice, and employers might end up with graduates whom they regard as less prepared.

**Matching performance funding with performance.** Performance funding as an incentive to change institutional behaviour works when performance funding matches, at least approximately, the cost of performing. That sounds like common sense, but it is the shoal on which performance funding most often founders. It founders for three reasons.
The first is that governments confuse the outputs and outcomes that they hope performance funding will achieve. Let’s take the graduation rate again as an example. There are three reasons for the state to desire higher rates of graduation. The economic objective is to expand the supply of human capital. The social objective is equity through access to higher wages and, in some countries, higher social standing. The budgetary or cost objective is to realize a cost advantage by producing graduates at a lower unit cost. (The benefit advantage would have been to produce higher quality or more graduates at the same unit cost.) Each of these objectives requires a different standard of measurement. More significantly, each requires a different amount of funding. “Mix and match” will not work. Pooling does not work either (Martin, 2011). In some jurisdictions in which this problem is recognized, governments rationalize the mix-and-match practice by assuming that institutional autonomy and “block grants” will enable individual institutions to offset negative mismatches between performance and the cost of performing according to one performance indicator with a positive mismatch according to another indicator. This is indeed a rationalization. It becomes even more so in undifferentiated systems, like most in Canada, in which institutions with different missions are expected to conform to the same indicators.

The second problem is the notion that performance funding can be an incentive. The idea itself is not unsound. In execution, however, there is difficulty in funding an incentive as a true incentive. By definition, an incentive should generate new funding. In other words, performance funding should be truly additive or supplementary. It should not be the result of reallocation through which one source of funds supplants another. Algorithms of performance funding can be highly complex, but the difference between performance funding that supplements and performance funding that only supplants is instantly apparent to institutions that deploy incentive-based budgeting. Although Australia foreclosed its performance funding scheme, it realistically recognized this propensity by dividing performance funding into two parts: one for “facilitation,” which, at least nominally, addressed actual costs, and one frankly called “rewards,” which was an incentive unrelated to cost (Massy, 2003; Australian Government, Department of Education Employment, and Workplace Relations, 2009).

The third problem is cost, which to some degree is an amalgam of the problem of confused objectives and the problem of incentives that fail to generate truly new funds. Logically, performance funding as an incentive can be less than the average unit costs or even less than the marginal unit costs of a given behaviour or “performance” if it generates truly new funds. In other words, it is “extra.” This is the logic that most states apply in determining the scale of funding to be allocated by means of performance indicators. It makes performance funding seem affordable by displacing average costs with marginal costs. This is a major disconnection between performance funding and incentive-based budgeting.

For a time in the history of incentive funding, this worked for government and for institutions. This can be explained in two ways. The first is that public subsidies as a proportion of total funding for universities was relatively high when the deployment of performance funding was rising towards its apogee (Derochers, Linehan, & Wellman, 2010). Howard Bowen (1980) was right when he said that cost in higher education is elusive because institutions spend all the revenue that they generate. They do not seek and cannot identify inherent costs. Costs rise to meet revenue, hence the unfortunate but appropriate term “cost disease.” Thus it was possible, although neither certain nor admitted, that for
a period of time a university could achieve a performance funding objective by spending marginal amounts that were equal to or even less than the marginal performance income.

The second, but not mutually exclusive, explanation is that it is relatively recently that universities have begun to understand their costs fully. Although ABC was in use in private firms in the early 1980s, it was not deployed in post-secondary education until the late 1990s. The Lumina Foundation’s Delta Cost Project began in 2008. Its first report spanned the years between 1998 and 2008. The Center for College Affordability produced its first report on costs in 2011 (Gillen et al., 2011). Incentive-based budgeting, which analyzes costs more precisely and systematically than ABC, was in wide practice in public universities by the latter half of the 1990s (Dougherty & Reddy, 2013; Lang 2002). Most of these dates coincide with the levelling off or first iteration of the use of performance funding. Thus, when we now talk about performance funding matching (or not) the costs of performing, universities know a lot more than they previously did about the costs of the various performances for which performance funding indicators call. In other words, they now can “do the math,” which in many if not most cases means a realization that marginal performance funding is less than the marginal cost of performing. When universities “do the math” and in turn either respond or not to the performance funding incentives, they send a clear signal to government about the adequacy of the funding. Thus, a significant difference between first- and second-generation performance funding is a financially larger commitment.

**Cost functions, equitability, and adequacy.** Although the concept of elasticity is normally associated with prices and markets, it has an application to performance funding, too. Performance funding is almost always linear. It doesn’t have to be, but it is. Each percentage point rise in a performance indicator generates the same funding. Because colleges and universities now know more about costs, they know that not all performance increases are equal in terms of cost. All other things being equal, an institution starting below the average—for example, again, the rate of graduation of its peers—will find the marginal cost of a percentage point increase in the rate lower than would an institution that started above the average. For the first institution, the performance funding incentive would be elastic. For the latter, it would be inelastic. Thus, governments should not be surprised when performance funding produces diminishing returns at higher unit costs.

This can lead to an inequity versus adequacy problem. A putative advantage of incentive funding is that it can be equitable. The German Excellence Initiative is an example (Melynk, 2014). Any given institution within a system or jurisdiction can attract funding by improving its performance. More to the point, that institution will generate the same performance funding as will another institution that improves its performance by the same amount and according to the same measurement. That is equitable. But the marginal revenue/marginal cost equation may be different for each of the two institutions. It may be adequate for one and inadequate for the other. Incentive-based budgeting is based essentially on a marginal revenue/marginal cost equation. This is a problem for jurisdictions that aim to improve system performance among institutions of widely different sizes and missions.

**Multiple principals/multiple agents.** A reasonable case can be made that performance funding could have been invented to address a principal–agent problem between states as principals and universities as agents. Principal–agent relationships become problematic when the following conditions are present:
Agent and principal have different objectives or at least construe the same objectives in different ways. Principals have conflicting or incompatible objectives, as might occur when outcomes are confused with outputs. Information is asymmetrical, in which case, the principal lacks information about the agent’s behaviour or the outcomes of that behaviour. Information is asymmetrical, in which case, the agent lacks information about the principal’s objective, including asymmetry caused by the principal underfunding the nominal objective.

When performance funding was introduced, much of the theory behind the principal–agent problem was theoretical insofar as higher education was concerned. Government, as a principal, provided or otherwise controlled nearly all funding received by public colleges and universities. Universities, as agents, were managed centrally or “top down.” There was one principal and one agent.

Today, many public universities are “public” only in the sense that they are eligible for state funding. As governments for various reasons cut back funding for higher education, they became minor shareholders and created a financial vacuum into which other principals were drawn, sometimes as a matter of public policy that encouraged universities to seek alternative sources of income. Different principals have different objectives. If they have different objectives, they will, for good reason, expect different “performances” from universities as their agents and devise different performance funding incentives and indicators. Universities as agents are forced to trade off among principals or, more particularly and problematically, among their principals’ performance indicators. This of course blunts the effect of performance funding. As performance funding become less powerful for these reasons, incentive-based budgeting becomes more powerful because it encourages and rewards efforts to diversify and expand revenue to replace reductions in public subsidies.

Universities have also changed in the ways they perform as agents. They have become decentralized in budgeting and planning and have brought more stakeholders into governance. Some stakeholders—for example, fee-paying students—are in practical effect principals. Agency as measured by several commonly used performance indicators has moved from the institutional level to the faculty level. Deans instead of presidents thus are becoming the real respondents to performance funding incentives. Some that have introduced incentive-based budgeting already reflect this by attributing enrolment-driven costs and revenue, including performance funding, proportionately into various categories of cost—for example, registrants by program, registrants by course (actual instruction), and graduates by program—each of which could be measured by a different performance indicator. For research, agents are principal investigators, organizationally even more distant from the central administration.

Donors are becoming more frequent principals, often with the encouragement of government. This in turn engenders further confusion. While institutions see donors as principals, governments may see them as agents whose private wealth may be leveraged to replace public subsidies as incentives. This is the public policy concept that underpins government “matching” programs that function as de facto performance funding.

Cost. Incentive-based budgeting may assume more knowledge of costs than an institution might actually have. If the implementation of incentive-based budgeting at the several universities that have deployed it were to reveal only one thing, it would be that the accurate determination and attribution of indirect costs and overhead is, on the one hand, essential and, on the other hand, very demanding and expensive.
What this means, however, for governments that want to install incentive funding is that the funding equation may be calculated differently from university to university within their respective jurisdictions. We know, for example, that some universities and colleges in Ontario have paid little or no attention to performance funding. That could be explained by mismatches between the amount of performance funding and the cost of the respective performances. But it also could be explained by the incapacity of some institutions to match costs and funding. The lesson here is not that to ensure the success of incentive funding, governments should promote certain institutional budget models. It is that there is virtue in simplicity, consistency, and flexibility in designing incentive funding systems. The case of South Carolina is instructive here. That performance funding system—one of the oldest and largest in America—collapsed because it was too complicated for colleges and universities to incorporate into their planning and budgeting models (Burke, 2002). The South Carolina system probably could have been manageable by institutions that were using incentive-based budgeting. In other words, one incentive-based system would have matched up with another incentive-based system.

**Reducing unit costs of overhead.** In terms of incentives to operate more efficiently and effectively, incentive-based budgeting has been more successful on the income side than on the cost side. It has generated interest in finding new sources of revenue, which often meant responding to governments’ interests in expanding access and developing new programs in response to student and employer demand. It has done less well on the cost side (Curry et al., 2013). It has encouraged universities to reduce costs, where practical, by reducing volume—for example, by occupying less space. This, however, is different from reducing the unit costs of operating space. The fact that universities now know more about their costs does not necessarily mean that they can or will reduce costs accordingly. The problem seems chronic but could be irrelevant to incentive funding. It would be were it not for the almost universal practice of measuring performance by indicators of academic performance. Incentive-based budgeting functions almost exclusively in faculties where in most universities about 60 percent of costs reside. Thus, both incentives—performance funding and incentive-based budgeting—to the consternation of governments interested in efficiency and of universities interested in cost reduction—do not directly affect a large percentage of total institutional cost.

**The Future: Collision or Symbiosis?**

There are several possible scenarios. Each begins with two reasonable assumptions. First, governments will continue to be interested in accountability, with which incentive funding is often closely associated. Second, universities will continue to expand their knowledge of costs and increasingly plan and budget on the basis of net costs.

**Scenario 1**

Given the track record of incentive funding, governments might recognize that its cost-effectiveness is problematic and in political terms a liability. In terms of results, funding that is installed to change institutional behaviour by incentive is expensive (Sanford & Hunter, 2011; Shin & Milton, 2004). Although perceived as expensive from the point of the view of the state, performance funding has so far been perceived by colleges and universities as too small, in which case they often ignore the incentives or find them too
costly to comply with (Callahan, 2006; Chan, 2014; Cooke & Lang, 2009; McColm, 2002; Miao, 2012; Rabovsky, 2012). Ironically, incentive-based budgeting partly explains the apparent ineffectiveness of incentive funding. Universities that deploy incentive-based budgeting know that the income–expense equation of incentive funding rarely balances. Not only is this known as a matter of fact, but incentive-based budgeting inherently forces universities to take it into account on their bottom lines. Under incentive-based budgeting, there is no way of hiding an imbalance of performance funding income and the cost of attracting it. If the second generation of performance funding is not different from the first, the outcome will be a collision or at least a parting of ways.

Scenario 2

Here we can draw some generalizations from the experience in Canada. In some respects, this has already happened in two provinces. Performance funding in Alberta and Ontario is still in place, but both of those provinces in different ways have moved on to prescriptive measures that are more compliance sticks than incentive carrots. Additionally, in Alberta, as in Switzerland, the view seems to be that the most effective way to force universities to operate more efficiently is to reduce their funding. This coincides with Martin’s (2012) view that as long as additive revenue is available to universities, they will not reallocate existing resources in response to public policy preferences.

Scenario 3

Declines in public funding for higher education will further weaken the impact of public performance funding on university behaviour as resource dependence shifts to other sectors: corporate and private philanthropy, students and parents, foundations, and “private partners”—all of whom will seek “performances” that advance their interests. Performance funding will cease to be a monopsony, as there will be multiple “buyers” of performance. In that case, efficiencies might result, which may be to the indirect or direct financial benefit of the state. In terms of affordability and competing demands on public funding, such a transition in the role of government in financing higher education might be desirable or at least tolerable in terms of public policy. Some American states are beginning to include private philanthropy as a metric for performance funding (Jones, 2103). Whether desirable or not, it is a transition that universities can better manage by incentive-based budgeting. In that case, the outcome will be symbiotic.

Scenario 4

In virtually every Canadian province, due either to deliberate policy or to fiscal necessity, universities have turned to other sectors for financial support, conditions that may invite symbiosis. Some voices are beginning to argue that public systems of higher education are too big, too centralized, and too complex to be managed “top down” successfully (Berdahl, 2000; Callan, 1994; Gaither, 1999; MacTaggart, 1996). March (1978) used the phrase “limited rationality” to describe the inability of large, centralized organizations to make universally competent decisions. Public universities, expressly because they are public, are typical of what Scott (1998) called “complex inter-dependencies” that cannot easily be reduced to the schematic, system-wide visions that performance funding
often represents. There is considerable evidence that allowing greater autonomy may be a more powerful carrot than performance funding (Altbach, 2004; Clark, 1998; MacTaggart, 1998; Maxwell et al., 2000). Incentive-based budgeting systematically promotes institutional discretion and efficiency (Hearn et al., 2006). In that case, governments may continue to use incentive funding but will allow more permutations and combinations among performance indicators in order to encourage and promote diversity over isomorphism (Jones, 2013; Weingarten & Deller, 2014), as already appears to be happening in Saskatchewan and some American states where performance funding, instead of functioning outside funding formulas, is being brought within them to fund specific performance outcomes (Dougherty et al., 2011; Miao, 2012). In this case, the second generation of performance funding is not a bonus. There is no set-aside. There is still an incentive, but it is an incentive not to lose funding rather than to gain new funding. Initiatives such as these will side-step a collision by allowing more institutional strategic choice. Universities that adopt incentive-based budgeting in some form will be capable of exercising that choice rationally. The result will be symbiotic, as the interests of the institutions and the state will be better served. But in smaller institutions, particularly ones with homogeneity of programs, incentive-based budgeting may not be suitable (Curry et al., 2013; Lang, 2001); in such cases, there might not be a collision, but there will be no symbiotic mutual benefit.

**Scenario 5**

Finally, there is the possibility of a shift in the name of efficiency from supply-side subsidies to demand-side subsidies, as is already being discussed in the UK, where the government speaks about “funding students instead of institutions,” which is how the British refer to performance funding as “payment for results.” This coincides with the views of some economists (Krugman, 2011; Wolf, 2002) that economic growth “pulls” education instead of being “pushed” by it. A switch to demand-side subsidies could make performance indicators less necessary and incentive-based budgeting more so. Two American states—Texas and Nebraska—that were early adopters of performance funding are reconsidering it in light of evidence that the correlation between investments in higher education and economic growth is dubious (Gillen et al., 2011; Lindsay, Vedder, Bishirjian, & Stille, 2012; Vedder, Robe, & Denhart, 2012). The Higher Education Quality Council of Ontario at the beginning of 2013 began a review of returns on investment in higher education. Demand-side public subsidization may logically follow and evoke a different view of incentive funding and redefine performance, with the topsy-turvy result, perhaps, that institutions informed by incentive-based budgeting will lobby governments to earmark funding for “performances” that they can—they will argue—uniquely provide. Some versions of “second-generation” performance funding look like this (Ziskin, 2014). This is exactly the kind of strategic behaviour that proponents of incentive-based budgeting predict (Hearn, 2006; Lang, 2001; Massy, 2003; Whalen, 1991). This may be what some governments have in mind when they provide matching performance funding to encourage universities to be more entrepreneurial (Caruana, Ramaseshan, & Ewing 2006; Clark, 1998; Lazzersoni & Piccaluga, 2003). Here the result will be symbiosis through what will amount to performance contracts. Institutional knowledge of net costs will make this possible. Larger “non-trivial” public investments in performance funding will make it necessary (Clark, Trick, & Van Loon, 2001; Jones, 2013).
References


**Contact Information**

Daniel W. Lang  
OISE  
University of Toronto  
dan.lang@utoronto.ca

Now an emeritus professor at the University of Toronto, Mr. Lang was Senior Policy Advisor to the President, the Vice Provost, Planning and Budget, and the Vice-President, Computing and Communications. He received a BA and MA from Wesleyan University and a PhD from the University of Toronto. His principal areas of interest are institutional planning and management, finance, accountability, and quality assurance. His current research investigates how, when, and why community college students decide to transfer, the effects of fiscal incentives, the role of informal knowledge in the formation of human capital, and the performance of consortia.