RESEARCH AND TEACHING: COMPLEMENTS OR SUBSTITUTES?

John Quiggin
Australian Research Council Federation Fellow
The University of Queensland
Email address: j.quiggin@uq.edu.au

ABSTRACT

The idea that research and teaching are complementary activities is central to the idea of the modern university, developed in the 19th century and still dominant today. Despite its importance, the belief that research and teaching should be undertaken jointly has rarely been subject to close examination. The object of this paper is to clarify some of the conceptual issues surrounding the relationship between research and teaching. The main contribution of the paper is the development of an analytical model within which concepts of complementarity and substitutability can be explored.

1. INTRODUCTION

The idea that research and teaching are complementary activities is central to the idea of the modern university, developed in the 19th century and still dominant today. The concept of research as an organised activity, rather than a private hobby can be traced back to the founding of the University of Berlin by Humboldt in 1810, the world's first university of the modern kind, dedicated to the scientific approach to knowledge, to the combination of research and teaching, and to the proliferation of academic pursuits. The German model, rapidly emulated in the United States, displaced the older notion of the university as an essentially clerical institution, devoted to the transmission of existing knowledge and religious belief, rather than to the generation of new knowledge. The establishment, over the course of the 20th century, of the research doctorate as the entry ticket to academic appointments, tightened the link between teaching and research. With this development, teaching appointments at top-ranked universities were open only to those with a demonstrated capacity to undertake research.

Despite its importance, the belief that research and teaching should be undertaken jointly has rarely been subject to close examination. Popular
stereotypes, with their customary ambiguity provide suitable examples for both believers and sceptics. The sceptics can contrast brilliant and dedicated teachers who have forsaken their own research ambitions to pass the light of learning on to their students with invisible research demons who use locked doors and teaching assistants to keep students at bay. Believers can contrast the enthusiasm of those working at the frontiers of the discipline with the tired hacks whose yellowing lecture notes are recycled from their own undergraduate days. However convincing some individual examples may be, it seems unlikely that issues can be resolved in this fashion.

It seems unlikely, however, that the issue can remain one of purely speculative interest. Changes to the structure and financing of universities, particularly in countries influenced by the reform movements variously known as Thatcherism, economic rationalism and neoliberalism\(^1\), have generated increasing pressures on universities to account for the funds they receive through verifiable measures of teaching and research output. In many cases, the funding arrangements give rise to pressures on academics to reduce time commitments to research (at least in the absence of grant funding). There has also been some growth in the number of research-only positions.

The object of this paper is to clarify some of the conceptual issues surrounding the relationship between research and teaching. The paper begins with a survey of the (very limited) empirical evidence. The main contribution of the paper is the development of an analytical model within which concepts of complementarity and substitutability can be explored.

2. **EMPIRICAL EVIDENCE**

Empirical evidence on the complementarity or substitutability of research and teaching is limited. Freedman (2003), examines the case of George Stigler, one of the most eminent of 20th century economic researchers and concludes that Stigler provides a less-than-optimal model for undergraduate teaching. Good, even great, economists do not necessarily make successful teachers at that level. In contrast, graduate work and particularly locating an appropriate supervisor is more like a search for a congenial marriage partner.

While these conclusions appear quite reasonable, the fact that

---

1 Among the developed countries, the members of the British Commonwealth, including Australia, Canada, New Zealand, and the United Kingdom, have been most affected by this movement. By comparison, there have been only modest changes in the financing and accountability structures applied to universities in the Unites States.
individual case studies are considered as a useful source of information is indicative of the limited availability of systematic quantitative, or even qualitative, evidence.

The most useful empirical studies are those of M. Fox (1992) and K. Fox and Milbourne (1999). Fox (1992) finds evidence of mutual complementarity between research and teaching at the graduate level, a finding consistent with the conclusions drawn by Freedman. Fox and Milbourne find that a 10 per cent increase in the number of teaching hours may reduce research output by as much as 20 per cent.

The paucity of substantial evidence on complementarity between research and teaching at the undergraduate level should not be taken as evidence of the absence of such complementarity. The most likely reason for the absence of empirical research on a topic of such obvious interest is the lack of good measures of teaching performance. Over the period since the 1960s, it has become common practice to elicit evaluations of teaching from students, and such evaluations are used, to some extent, in promotion and tenure decisions. However, they are of limited value in relation to the current problem, for several reasons.

First, the results of teaching evaluations are rarely available to researchers in the form that would be required, namely one that allowed identification of individual academics whose teaching performance could be matched to research output. Second, instruments and procedures for teaching evaluation differ between universities, while measures of research output are not, in general, comparable across disciplines. Third, and most seriously, there are doubts about the usefulness of teaching evaluations. A common belief among teachers is that favorable student evaluations can be promoted by soft grading, light workloads and easy course material, pejoratively referred to as 'dumbing down'.

Bosshardt and Watts (2001) observe that the correlation between evaluations by students and teachers' evaluation of their own performance is low, and note that this low correlation raises some questions about administrative use of SET data from principles classes - in effect, choosing to use such data means that, at the least, student ratings are accepted as important even when they differ from faculty self-assessments and, at most, that student ratings of teaching are accepted as more reliable and valid than self-assessments. Those arguments are defensible, certainly, but they are rarely made explicit. These problems become more difficult when our regression results are taken into account. ... on the preparation item, student and instructor ratings are only in weak agreement. It seems likely that students are in a better position to judge an instructor's ability to speak clear English and teach with enthusiasm than they are to judge grading rigor or
how well prepared an instructor is in a subject that most students have not seen before; but things perceived to be true are true in their consequences.

The results of Bosshardt and Watts are supported by a case study undertaken by Brickley and Zimmerman (1998) of William E. Simon Graduate Business School where teachers' ratings on the basis of students' appreciation were the main criterion of teaching performance evaluation. Brickley and Zimmerman found that faced with this incentive structure, teachers adopted low-cost strategies to increase their popularity with students in preference to high-cost strategies aimed at improving the quality of teaching.

Another approach to the problem is derived from Stigler's work rather than from his somewhat unsatisfactory example. In the analysis of industrial structure, Stigler proposed the 'survivor principle', namely that the survival of a firm in the face of competition was prima facie evidence that the firm was technically and allocatively efficient. Applying Stigler's principle, the fact that the modern university, with its combination of teaching and research, has survived in a wide range of institutional and regulatory settings, and in competition with enterprises more narrowly focused either on teaching or on research, seems to indicate that, over a wide range of teaching and research activities, joint provision is efficient.

3. MODEL

The simplest possible model of complementarity between research and teaching is one with a single input, \( l \), denoting hours of academic labour, and two aggregate outputs, \( r \) and \( e \), denoting research and education (teaching). In this simplified model, differences in the quantity and quality of research and teaching are subsumed in the aggregate outputs.

For any individual academic \( i \), there is a technology set

\[
T^i = \{(l, r, e) : \text{Output } (r, e) \text{ is feasible for } i \text{ given Input } l\}
\]

which may be represented by an input requirements function

\[
l^i (r, e) = \inf \{l : (l, r, e) \in T^i\}
\]

Research and teaching are complements (for \( i \)) if

\[
l^i_{rt} \leq 0 \quad \text{(3.1)}
\]

and substitutes if
Even this simple definition raises issues that are important in reviewing the literature on this topic. The proposed definition says that the hours required to deliver a given output of teaching are lower, the higher the number of hours allocated to research, and *vice versa*. This is the standard economic definition of complementarity in production.

But for any given input of labour, the more hours are allocated to research, the fewer can be allocated to teaching. In some discussions, this fact is taken to mean that research and teaching are not complementary activities. Complementarity would arise only if time spent on research allowed an increase in teaching output (presumably in the form of higher quality teaching that more than offsets reduced teaching hours) with no additional effort.

Implicitly, on this view, research and teaching are complements only if, for at least some \((r, e)\)

\[
\frac{\partial l}{\partial e} < 0
\]

This would arise, for example, with a Leontief (fixed-output proportions) technology in which teaching activity automatically produces a proportionate amount of research.

As noted above Fox and Milbourne (1999) show that, other things equal, an increase in teaching effort results in a reduction in research output. Hence, research is not a free by-product of teaching. However, the results of Fox and Milbourne do not necessarily imply that research and teaching are substitutes in the standard economic sense.

For most of the purposes under discussion, the characterization of complementarity and substitution given in (3.1) and (3.2) is appropriate. In particular, suppose that there exists a set \(I = \{1, \ldots, I\}\) of academics, each with available labour input \(\tilde{t}^i\). The total available input is therefore

\[
l = \sum_{i} \tilde{t}^i
\]

Suppose further that the technology displays constant returns to scale and is the same for each \(i\). Now consider, the problem of maximizing some welfare function \(W(R, E)\) where \(W\) is strictly increasing in both arguments.
The following proposition, the proof of which is left as an exercise, gives the rationale for the standard university arrangement under which academics engage in both research and teaching:

**Proposition 1.** Suppose \( 3.1 \) holds for all \( r,t \), and the technology is the same for each. Then the optimal solution satisfies

\[
\frac{r^i}{R} = \frac{e^i}{E} = \frac{T^i}{T} \forall i
\]

Furthermore, there exists a shadow price of research \( p \) such that

\[
\frac{\partial l^i}{\partial e} = p \frac{\partial l^i}{\partial r} \forall i
\]

This result may easily be extended in a variety of directions. The assumption of a fixed quantity of available labour may be relaxed by incorporating the disutility of labour in the objective function, provided this is convex and increasing over the relevant range. The assumption of constant returns to scale may be generalized to allow for homothetic labour costs, and a range of results from aggregation theory may be applied. Differences in the skills of academics may be captured by interpreting the labour input \( l^i \) as effective hours of labour. Under this range of conditions, the shadow price \( p \) will, in general, play no explicit role in the allocation of teaching and research responsibilities, which can be dealt with simply by an equal sharing rule.

On the other hand, the result will not necessarily hold if academics are heterogeneous, with some having a comparative advantage in research and others in teaching. As in the theory of international trade, if differences in comparative advantage are sufficiently great, they will outweigh gains from complementarity in production and lead to specialisation as an optimal outcome.

More generally, differences in comparative advantage imply that the optimal proportions of research and teaching will be different for different academics. This difference may be formalized, as is often the case in the United States, with academic positions being explicitly characterized by the
proportions of time to be allocated to teaching, research and service activities. In other contexts, there may be an informal allocation of teaching responsibilities that takes account of research output.

In Australian universities, traditional practice has been to require roughly equal outputs of teaching from all academics, but to offer promotion primarily on the basis of research outputs. The resulting equilibrium is one in which academics with a comparative advantage in research supply more hours and receive a higher reward (both in aggregate terms and per hour) than those with a comparative disadvantage. Assuming that the incentives are calibrated appropriately, such a structure can yield an approximately optimal allocation of resources to research, at the cost of a sub-optimal allocation of resources to teaching. Such a solution may be adopted as a result of problems in measuring teaching output. This issue is addressed below.

3.1 Quality and quantity

The simple model of complementarity presented above illuminates some important issues. However, the model is unsatisfactory in two important respects, which turn out to be related.

The first is the aggregation of quality and quantity into a single index. This is not a crucial problem in relation to research, at least as quality is normally measured. There is no particular reason to suppose that the kind of research typically regarded as 'high quality', such as path-breaking theoretical research, is more complementary with teaching than, say, applied research. This is particularly true in relation to undergraduate research. In addition, the required input of research effort per page of output is typically higher for high-quality research. Hence, in implementing a model like that above, it would seem reasonable to use any of the standard quality-adjusted measures of aggregate research output, rather than attempting to treat the quantity and quality of research separately.

On the other hand, the issue of quality is crucial in relation to teaching. To the extent that the role of the teacher is to present the contents of a standard text in a stimulating and engaging fashion, there is no obvious reason why research on topics going beyond the scope of the text should be particularly beneficial or why time devoted to such teaching should be helpful in research. Hence, any complementarity between research and teaching will arise only if the aims of teaching are more ambitious. In particular, complementarity seems most likely to arise when courses are prepared individually by teachers and reflect their personal knowledge of the subject, derived at least in part from research work.

The second problem is that the two-output model presented above is
inherently symmetrical. If research is complementary with teaching (the cross-derivative of the cost function is negative), then teaching is necessarily complementary with research. However, most discussion of the relationship between teaching and research implies an asymmetrical relationship, in which most of the benefit goes from research to teaching. This is particularly true in systems, such as that in Australia, where there is little graduate coursework undertaken by research students.

It is useful, therefore, to consider a dynamic generalization of the model of the previous section, designed to address the problems described above in a fairly simple fashion. In each period $t$, there are four outputs $(r_t, c_t, q_t, h_t)$ where $r_t$ denotes research output as before, $c_t$ (for classes taught) is a quantitative measure of teaching output, $q_t$ denotes the quality of teaching, and $h_t$ denotes human capital in period $t$. The inputs are $(l_t, h_{t-1})$ so that previously accumulated human capital is used in teaching. The crucial feature of the technology is that while $l_t$ is allocable between class time and other activities, $h_{t-1}$ acts as a quasi-fixed factor, which contributes to the production of $r_t, q_t$ and $h_t$, but not $c_t$. The output of educational services is given by a function $e(c, q)$ which is assumed to be the same for all individuals $i$ and date $t$.

We will simplify further by assuming that the labour input requirement function takes a quasilinear form with a unit coefficient on class time so that the labour supply constraint becomes

$$\bar{T}_t = c_t + l'(r_t, q_t, h_t, h_{t-1})$$ (3.3)

The problem of choosing the optimal allocation of resources in cases of this kind is one of dynamic programming and the solution may be characterized by a valuation function $v_t(h_t)$ reflecting the future flow of returns to the stock of human capital $h_t$, so that the value of output for academic $i$ at time $t$ may be expressed as

$$z_t = e(c'_t, q'_t) + p_tr'_t + v_t(h_t)$$ (3.4)

where $p_t$ as before, is a shadow price for research output. Maximising (3.4) subject to (3.3) using a Lagrangian yields the first-order conditions (with subscripts and superscripts dropped except where needed for clarity)

$$\frac{\partial e}{\partial c} = \lambda = 0$$ (3.5)
\[ \frac{\partial e}{\partial q} - \lambda \frac{\partial l}{\partial q} = 0 \]

\[ p - \lambda \frac{\partial l}{\partial r} = 0 \]

\[ \frac{\partial v}{\partial h_t} - \lambda \frac{\partial l}{\partial h_t} = 0 \]

which describe the trade-off between the objectives of research, teaching and investment in human capital within a given period. Under plausible conditions, these activities will be substitutes within a given period when account is taken of the labour supply constraint.

The key issues of interest for the present discussion relate to the optimal path for \( h \), for a given initial level of human capital \( h_0 \) and labour input requirement function \( l(r_t, q_t, h_t; h_{t-1}) \), which in turn depends on the properties of \( v(h) \). If \( v \) is concave, there will exist a unique steady-state optimum \( h^* \) and the optimal time path for \( h \), following standard 'turnpike' arguments will be one which moves from \( h_0 \) to a value near \( h^* \), remains there until the individual approaches retirement, then declines. If \( v \) is strongly concave (which will be true if \( l(r_t, q_t, h_t; h_{t-1}) \) is strongly convex with respect to radial expansions in output), the optimal value \( h^* \) will be robust to variations in \( p \), the relative price of teaching and research output and to changes in the discount rate used in the optimisation process. Further, provided that the value \( h^* \) is consistent with the existence of an interior optimum to the problem of maximising (3.4) subject to (3.3), the steady-state optimum will involve a mixture of research and teaching. By contrast, if there exist increasing returns to effort devoted to research, \( v \) may be convex over some range, and the optimal solution is likely to be a corner solution in which individuals with high productivity in the research task specialise in research, offering only a small number of high-quality courses, while others do little or no research, allowing human capital to converge to a fairly low-level steady state.

Consideration of this model leads to a slightly different view of complementarity between research and teaching. Supposing, as seems plausible, that there are decreasing returns to research effort beyond some effort level \( l < \bar{l} \), it is desirable that research should be combined with some other activity. If research is directly complementary with the acquisition of human capital, and human capital is complementary with high teaching quality, then a combination of research and teaching is optimal. The view that no more than four hours of research work per day is desirable has been...
widely attributed both to Poincare and GH Hardy (Nielsen 2004).

3.2 Agency theory

Thus far, attention has been focused on the optimal allocation of time between research and teaching that would arise under conditions of contracting with perfect information. However, many aspects of research and teaching are subject to asymmetric information. Considering the variables in the dynamic model described above, human capital $h$ is not directly observable, but must be inferred from final outputs, particularly research output and the quality of teaching. Of the final outputs, it seems reasonable to assert that the problems of asymmetric information are least in relation to classroom hours $h_t$, and greatest in relation to teaching quality $q_t$. Research output $r$ is harder to measure than the provision of hours of teaching, but more easily measurable and quantifiable than the quality of teaching. We are thus presented with a multi-task agency problem, of the kind first analyzed in detail by Holmstrom and Milgrom (1991) and addressed by a large subsequent literature.

A notable topic in this literature, related to the question at issue in this paper, is that of the desirability or otherwise of rewarding school teachers on the basis of the test scores achieved by their students. The general, though not universal conclusion of the literature is that extensive reliance on such rewards is likely to produce adverse outcomes, of which the most important is 'teaching to the test'. That is, teachers are likely to focus on the subset of skills required to score well on tests, rather than on the broad set of skills for which the test is supposed to be a proxy. In particular, it is suggested, skills such as critical and creative thinking, which are hard to test in the multiple choice format normally favored in universal testing programs, will be neglected in favor of relatively narrow competencies (including the basic requirement of using a pencil to fill in a circle quickly and accurately).

The problem of incentive structures for teachers is an instance of the more general question of whether it is desirable, in a multitask concept, to provide sharp or 'high-powered' incentives for the performance of a subset of tasks that are measurable. As would be expected, the degree of technical complementarity is important, but so is the information structure. To the extent that observable components of research output are correlated with unobservable components of teaching quality in an effort allocation that is privately optimal for academics, universities will find it beneficial to propose contracts involving a mixture of teaching and research.

This issue has received some attention in the literature. Dzagourova and Smirnova (2003) examine the problem in relation to the contractual
structures in Russian universities. They argue that contractual requirements for a high minimum number of teaching hours, combined with relatively low pay, provide incentives to shirk in research. Dzagourova and Smirnova present and analyze an agency-theoretic model of the problem, drawing on that of Holmstrom and Milgrom.

In the contracts examined by Dzagourova and Smirnova, payment is determined by the most easily observable output variable, classroom hours. In Western countries, it is more common to observe a payment structure in which payment and promotion prospects depend on research output as well as, and sometimes instead of, classroom hours. Although, as has been argued, this contractual structure may work well on average, it does not provide any direct incentive to improve teaching quality. Moreover, even if teaching quality and research outputs are complements on average, there will be many individual academics who are excellent researchers and poor teachers, or vice versa.

4. CONCLUDING COMMENTS

The modern university is founded on the assumption that research and teaching are complementary activities. Although the success of the university as an institution over the last two centuries may be taken as strong evidence in support of this assumption, there is little direct evidence for or against it. Discussion of the issue has been confused by the absence of a clear notion of complementarity and of the relationship between complementarity and the optimal organisation of university teaching and research.

The object of this paper has been to clarify some of the conceptual issues, using an explicit and dynamic representation of the technology of teaching and research. From considerations of agency theory, it has been argued that the difficulty of observing, and contracting on, the quality of teaching, is a crucial reason for the joint provision of teaching and research activities. Paradoxically, perhaps, it is precisely this measurement difficulty that accounts for the lack of direct evidence on complementarity between teaching and research.

REFERENCES


Brickley, J. and Zimmerman, J. (1998), *Changing Incentives in a Multitask*
Environment: Evidence from a Top-Tier Business School, William E. Simon Graduate School of Business Administration, University of Rochester.

Dzagourova, N. and Smirnova, M. (2003), Inefficiency of the Basic Contract in the Russian Economic Universities: the problem through lens of multitasking and common agency approaches, High School of Economics, Public University, Moscow, Russia,


