The Design of the University System*

Gianni De Fraja†
University of Leicester
Università di Roma "Tor Vergata"
and C.E.P.R.

Paola Valbonesi‡
University of Padova

July 12, 2011

Abstract

This paper compares the organisation of the university sector under unregulated private provision with the structure which would be chosen by a welfare maximising government. It studies a general equilibrium model where students attend university to earn higher incomes in the labour market, and universities teach them and carry out research. Each university chooses its tuition fee to maximise the amount of resources it can devote to research. Research bestows an externality on society. Government intervention needs to balance labour market efficiency considerations – which would tend to equalise the number of students attending each university – with considerations of efficiency on the production side, which suggest that the most productive universities should teach more students and carry out more research. We find that government concentrates research more that the private market would, but less than it would like to do if it had perfect information about the productivity of universities. It also allows fewer universities than would operate in a private system.

JEL Numbers: I210, I280, H420.

Keywords: Higher education; The organisation of the university sector.

---

*We are grateful to Subir Bose, Giorgio Brunello, Ottorino Chillemi, Vincenzo Denicolò, Loretti Dobrescu, Dennis Epple, Benedetto Gui, Eric Hanushek, Riccardo Leoni, Riccardo Martina, Claudio Mezzetti, Ludovic Renou, Richard Romano, Francesca Sgobbi, Roberto Tamborini, Jian Tong, Tommaso Valletti and especially two referees of this *Journal* for helpful suggestions. Earlier versions of this paper were presented in Lancaster, Bloomington Indiana, Gainesville Florida, Leicester, Turin, Athens, Brescia, Padua, Trento, Southampton, Bologna, Florence, Milan, Bath, Oxford, and at the 2009 Royal Economic Society, European Economic Association and Italian Economic Society conferences in Guildford, Barcelona and Rome. Financial support from the Italian Ministry of Education (projects 60A15-0093-09 and 20089PYFHY-002) is gratefully acknowledged.

†University of Leicester, Department of Economics Leicester, LE1 7RH, UK, Università di Roma "Tor Vergata", Dip. SEFEMEQ, Via Columbia 2, I-00133 Rome, Italy, and C.E.P.R., 90-98 Goswell Street, London EC1V 7DB, UK; email: `defraja@economia.uniroma2.it`.

‡University of Padova, Department of Economics, Via del Santo 33, 35100 Padova, Italy, e-mail: `paola.valbonesi@unipd.it`.

1 Introduction

This paper studies the optimal organisation of the university sector. Theoretical analyses devoted to this topic are scarce, relative, for example, to public sector procurement, health care, and primary and secondary school systems. This is surprising, in view of the many peculiar features characterising the university sector, its importance for society, and of course its immediate relevance to the daily life of many researchers.

In the model of the paper, universities supply teaching and research. Students attend university to increase their labour market earnings. They differ in the overall utility they derive from attending university, and have preferences regarding which university they attend, for example because of mobility costs. Universities are the only institutions which can impart tertiary education, and their payoff is the amount of research that they do. We compare a private university system unencumbered by government intervention with one where the government[1] chooses the funding and the student number of each university to maximise a standard utilitarian welfare function.

Government intervention may be justified by the externality bestowed on society by university research. Our approach studies the *structure* of the funding of the sector, and therefore sits between the macroeconomic study of the optimal aggregate amount of research and the microeconomic study of the optimal internal organisation of each individual university. In other words, we do not ask whether too much or too little research is done overall; we concentrate instead on the *distribution* of a given amount of total research and teaching among different universities, under the assumption that each individual university is carrying out its activities as efficiently as its resources and technology permit.

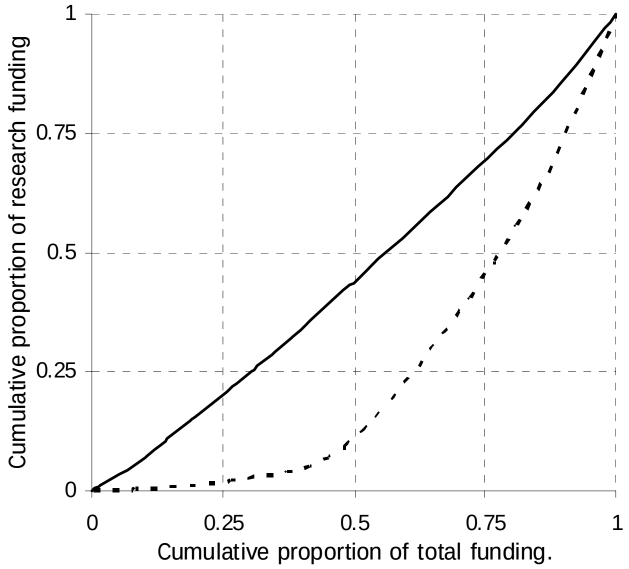
The gist of our results can be summarized by stating that the unfettered private market spreads research too thinly, and therefore the government would like to concentrate research and teaching in the most productive universities. A tension emerges. While the location of research is a matter of indifference, and so there is no harm in concentrating it in few institutions, the students' different preferences regarding the location of their place of study, due for example to imperfect geographical mobility implies that the same is not true of teaching: concentrating teaching in some institutions prevents some students from

[1] We refer to the financing or regulatory agency, as "the government", although, conceivably, the views of an international organisation could also be influent for a developing country.

attending university who would benefit from doing so. The tension is between concentrating teaching and research in the most productive universities only, which is efficient from the *cost* viewpoint, and ensuring that the students who benefit most attend university, irrespective of their preferred location, which is efficient from the *benefit* viewpoint. This tension is due to the link between teaching and research created by the university's budget constraint: teaching raises tuition fees and alumni donations with which universities fund their research. The government would like to sever this link. That is, it would like to allocate the total amount of research exclusively to the most productive institutions, to allocate teaching according to the trade-off between benefits and costs in teaching, and to make all students contribute to research funding, including those taught at universities where no research is done. The government is however unable to do so unless it has perfect information about universities' productivity. Instead the government offers universities a direct research funding linked to the number of students they enrol: the more productive universities "can afford" to select a larger grant and more students, because for them it is less costly to teach students, and, in consequence, they do more research than they would in an unfettered private market. Less productive universities opt instead for fewer students and a smaller research block grant. It is worth underlining that competition plays no role in our model of the university sector; one would expect that in a richer set-up, the pressure towards concentration we identify would be tempered by the positive effect of competition on productive and allocative efficiency.

The link established by the government between teaching and research mirrors what happens in the unfettered private market, where research activities must be funded by tuition fees. Therefore, while our model does not rule out technological complementarities between teaching and research, it suggests a possible additional explanation of the universally observed regularity that higher education and research are supplied jointly; this is the interaction between the preferences of the universities and the need of the government to provide them with an incentive to teach students. In other words, complementarity between teaching and research would be observed even in the absence of specific exogenous assumptions on the technology of the university production function.

In practice, the structure of funding and organisational structure varies widely across different university systems. As an example, Figure 1 shows the



State sector universities. UK 2009, dotted line; Italy 2005, incentive proportion only, solid line.
Source: HEFCE (UK), and MIUR (Italy).

Figure 1: Concentration of research funding in Italy and in the UK.

different concentration of research funding in UK and in Italy, two countries comparable in size and economic development. Research is clearly much more concentrated in the UK.[2] These huge differences call for a theoretical analysis of the organisation of the sector, to identify desirable direction for reform and explain difference in overall performance. Note that, while the UK system is generally considered "better" than the Italian one, it is not necessarily the case that concentration is *per se* preferable. It may affect the number of universities in the top 200 in the Shanghai Jiao Tong University ranking (which includes 19 British and four Italian universities), but, if the marginal cost of research is increasing, then it may be efficient to spread resources evenly across institutions, and so concentration would not be the best funding structure.

Both in the UK and in Italy, like in the rest of Europe, universities are government funded and regulated (for example in the tuition fees they can charge) to a very considerable extent. US states have instead a substantial

[2]For Italy, we plot only the "incentive" allocation. At the moment, only a small proportion of total funding is allocated in this way, the rest according to historic funding, and the aim is gradually to increase the allocation based on teaching and research quality. In the same vein, Drèze and Estevan (2007, Figures 1 and 2, pp 280-281) show that, in economics, research is more concentrated in the US than in Europe.

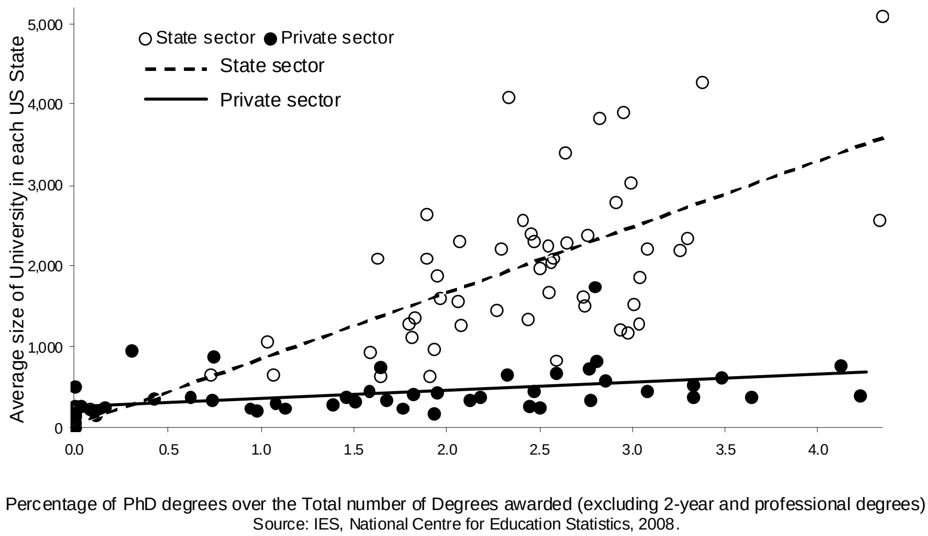


Figure 2: Research intensity and number of students in US states.

relatively unregulated private sector alongside the state university system. Here too we see different patterns; Figure 2 plots one measure of research intensity (the share of degrees awarded which are PhD's) against the average size of universities in each US state, separately for the private sector and for the state university system. States where universities are more research intensive tend also to have universities with more students, and the effect is clearly stronger in the public sector universities, as the analysis of the paper suggests should be the case.

We find that the private and government designed systems differ not just in concentration of teaching and research, but also in the number and nature of universities. There are fewer universities in the government designed system than in an unfettered private system. While first best efficiency requires "teaching only" universities, the government is prevented in practice from establishing them by its information disadvantage: the system it designs shares with that emerging under unfettered private provision the feature that all universities do both teaching and research. Finally, "research only" universities can only exist in a private system where the government offers all institutions a lump sum subsidy, unrelated to teaching or research performance.

The paper is organised as follows. We present the model in Section 2: the universities in Section 2.1, the students and the labour market in 2.2. In

Section 3 we study a private university system unencumbered by government intervention, and in Section 4 we derive the government optimal policy. Section 5 shows that the government policy can be implemented by offering a lump-sum grant which is higher the lower is the fee charged by universities to students. In Section 6 these two systems are compared against the common benchmark of the system a perfectly informed government would design. Section 7 is a brief conclusion. The proofs of all mathematical results are gathered in the Appendix.

2 The model

2.1 Universities

We study an economy with a continuum of separated local education markets and a single economy-wide "global" labour market. The local education markets can be thought of as different towns or counties. In the global labour market there are two types of jobs, skilled and unskilled: skilled jobs require a university education. This is obtained in the local education markets, in each of which there is a single potential university, which has monopoly power:[3] it is available to all local residents, and only to them. A cost of mobility is clearly just one way to capture differences among potential students regarding their desired place of study. Different amenities, subject mix, or entry requirements are others.This is a simple way of capturing the assumption that students are not infinitely mobile: if they were, given constant returns to scale in production, there would be only one university, teaching *all* students, and carrying out *all* research. [4]

[3] We do not consider competition for students among universities. Among the few theoretical contributions to the topic, in Del Rey (2001), universities choose the amount of research, and funding is positively related to the number of students. De Fraja and Iossa (2002) show that, if students are sufficiently mobile, competition among universities causes the emergence of "elite" institutions, which carry out more research and teach the best students. Empirical analyses focused on the US have studied the effects of competition for students on tuition fees, peer effects, students' future earnings and financial aid policies (see Hoxby 1997, Dale and Kruger 2002, Epple *et al* 2003, and Epple *et al* 2006, and references therein). The link between competition and governance is analysed empirically in Aghion *et al* (2008).

[4] Concentration of all teaching activities in a single centre may be optimal in some cases, such as sporting academies, but clearly not for tertiary education. Empirical work suggests that indeed students are not infinitely mobile: for the US, Long (2004) finds that "distance negatively impact the likelihood of an individual choosing a college" (p 284). In the Flanders, while "proximity hardly influences the participation decisions", it affects the choice of where

Potential universities differ in the value of an exogenously given[5] productivity parameter, $\theta \in \left(0, \bar{\theta}\right]$, with $\bar{\theta} > 1$. The distribution of θ in the economy follows a differentiable function $F(\theta)$, with measure $F\left(\bar{\theta}\right)$ normalised without loss of generality to 1, and density $f(\theta) = F'(\theta) > 0$ for $\theta \in \left(0, \bar{\theta}\right)$. As standard, $F(\theta)$ is assumed to have a monotonic hazard rate: $\frac{d}{d\theta}\left(\frac{1-F(\theta)}{f(\theta)}\right) < 0$.

Universities can engage in research and teaching. To do so they must build lecture theatres, laboratories, libraries and so on, and employ "professors". The production function of a university of type θ is given by:

$$n = \frac{t + r}{\theta}, \tag{1}$$

where $n > 0$ is the number of professors employed, $r \geqslant 0$ the amount of research carried out, and $t \geqslant 0$, the number of students taught. (1) implies that research and teaching both require professors, and that θ is a positive measure of productivity: a university with a higher θ needs fewer professors for a given amount of research and teaching.[6] The linear relationship with the number of professors implies that there are no economies of scale or scope, and that teaching and research are perfect substitutes as outputs.[7] All non-staff resources are assumed to be proportional to staff numbers, and therefore are fully captured by the parameter θ.

and what to study (Kelchtermans and Verboven 2010, p 356). Universities may respond to limited students' mobility by running multiple campuses. Some indeed do so, but they remain the exception rather than the rule, suggesting that institutions are in fact less mobile than students. The longer earlier version of the paper (De Fraja and Valbonesi 2008) introduces some student mobility, and shows that this does not alter the qualitative features of the analysis.

[5] The assumption of an exogenously given productivity parameter may seem artificial, and is clearly more plausible for the regulation and procurement literature (Baron and Myerson 1982, Laffont and Tirole 1986) which inspires our paper. The earlier version of the paper, De Fraja and Valbonesi (2008) sketches how natural assumptions on the hiring process together with the assumption that the ability of its staff determines a university's productivity create a link between the exogenously given population distribution of academics' ability and the endogenously determined distribution of universities' productivity. Work in progress incorporates this idea into a rigorous model, where graduates have preferences over the quality of the research institution they join as professors.

[6] The normalisation of $\frac{1}{\theta}$ as the number of staff needed to produce one combined unit of outputs will prove convenient when solving the government optimisation problem.

[7] The highly specialised nature of scientific research suggests that complementarities nowadays are at best limited: resources devoted to teaching are in fact resources subtracted to research.

Universities receive income from students, who pay a tuition fee of $p \in \mathbb{R}$, which can be negative if students are subsidised, and possibly from the government, in the form of a grant $g \in \mathbb{R}$, which again can be negative, and therefore a tax. A university's budget constraint[8] is

$$pt + g - yn = 0, \tag{2}$$

where y is the salary paid to a professor, determined by a competitive labour market for skilled workers (see below).

We posit that the objective function of each university is the maximisation of the amount of research it does, interpreted here as "blue sky" research. This is in line with the view that universities strive to maximise prestige which in turn is closely linked to research success (e.g. Raines and Leathers 2003, p 194, Garvin 1980, p 41).[9] Universities view teaching as a source of income, used to fund research, not an activity that increases their utility directly. This tallies with the empirical regularities that universities are by and large managed by academics whose vocation is research, and that academics' rewards and progress are more closely linked to success in research than in teaching.[10] Of course this does not mean that universities do not care about teaching: a university may take great pride and care passionately about the quality of its teaching, and indeed most do, but this is exactly in the same sense as a firm cares passionately about the quality of its products: both are means of furthering the organisation's objective, not ends in themselves.[11] Note also that, in our

[8]Other components of a university's budget constraint such as income from endowment, alumni donations, or state subsidies to tuition fees are included implicitly in (2). For example, p can be interpreted as including the net present values of future alumni donations. (2) captures the plain fact that if a university with a given endowment income and alumni donations wishes to increase its research expenditure, it must do so by increasing tuition revenues. This is an accepted explanation for the steep increase in tuition costs of the past decades in the US university sector (Ehrenberg 2007).

[9]Externally funded research or consultancy fees can be treated as separate activities and we disregard them in what follows.

[10]Tuckman *et al* (1977) noted a long time ago that, for individual academics, "outstanding teaching appears to yield a low rate of return" (p 697) and that "teaching and public service yield low compensation; publishing and administration carry much larger returns" (p 701); see also Hammond *et al* (1969). More recently, self-selection and the endogeneity of teaching loads make it more difficult to ascertain an independent impact of teaching on rewards (Euwals and Ward 2005, p 1663, and Golden *et al* 2009), and indeed, in his empirical survey of the academic labour market, Ehrenberg (2004) does not report studies of the relationship between teaching and research and lifetime earnings.

[11]Similarly, "non-profit" hospitals are often assumed to strive to maximise profit, which they

set-up, the concept of research can be plausibly extended: define "research" as any academic activity which bestows an externality on society by benefitting individuals or organisations who cannot be made to pay for it.[12]

We also assume that there is an upper bound to the amount of research that can be carried out in each university: let $r_{\max}$ be this bound. This constraint is only necessary when the government has perfect information, and, to leave it out of the main analysis of the paper, we think of it as "high", in a sense made precise in Assumption 2.

2.2 Students and the labour market

Each local education market serves a population of potential students, with measure normalised to 1. Every potential student can attend university, obtain a degree, and subsequently work in the skilled labour market (which includes working as a university professor).

Assumption 1 *A university graduate working in the skilled labour market receives income y; an individual with basic education only can only work in an unskilled job and obtain income $y - \Delta$.*

$\Delta > 0$ denotes therefore the salary premium earned by graduates. A previous version of the paper shows that Assumption 1 can be relaxed in several directions. At the level of the individual university, Section 6.3 in De Fraja and Valbonesi (2008) allows the students' utility, and hence their willingness to pay for a university education, to depend on the type of the university attended: higher θ universities are "better" for students. As they show, the realism of

do not distribute to owners as dividends, but use instead to further their aims, from excellence in research, to treating patients who are unable to pay (Danzon 1982, Dranove and White 1994). The link is in fact arguably stronger in universities, as typically good teaching attracts good students who both reduce the cost of teaching and may attract good staff improving research prowess (Rothschild and White 1995).

[12] Thus for example universities may subsidise doctoral supervision, or offer scholarship and financial aid to students from deprived backgrounds: these activities are undertaken by universities because they increase their payoff, even though – by definition – they do not generate enough revenue to cover their cost. The externality they bestow generates benefits to the rest of society, for example by increasing future research activities or enhancing diversity and offering role-models to able individuals in deprived neighbourhoods, and so they fit this extended definition of "research". Further outputs which have been suggested, such as the transfer of knowledge (Johnes *et al* 2005), or the production of human capital (Rothschild and White 1995), could be included simply by extending the definition of teaching.

the assumption is reflected in the realistic feature of the resulting tuition fee structure that universities with higher θ can also charge higher fees. In practice of course labour market incomes, y and $y - \Delta$, also depend on the values of "global" variables such as the aggregate number of graduates, which affects demand and supply in the skilled and unskilled labour markets, and the total amount of research, which affects productivity, and hence earnings. Explicitly adding these effects increases the complexity of the analysis, again with limited added insight (De Fraja and Valbonesi, 2008).

Students incur an effort cost if they go to university. This is denoted[13] by $a \in [a_{\min}, a_{\max})$, and its value is distributed among the potential students in each local market according to a distribution function $\Phi(a)$ with density $\phi(a) = \Phi'(a)$ and monotonic hazard rate $\frac{d}{da}\left(\frac{\Phi(a)}{\phi(a)}\right) > 0$.

While basic education is available in each local labour market at no cost, a student can go to university only if the potential university in their local market becomes active. If the type θ university does so, it sets a tuition fee $p(\theta)$ which students must pay to attend university. An individual of type a chooses to attend university and subsequently work in the skilled labour market if and only if[14] her earnings, net of tuition fees, and reduced by the cost of effort while at university, exceed the income obtained from the unskilled labour market: this implies that only individuals with a equal or lower than $\Delta - p(\theta)$ attend university and work in the skilled labour market.

To ensure that only some of the potential students attend university, we posit $\Delta \in (a_{\min}, a_{\max})$: with a zero tuition fee, some students would like to go to university, others would not.

Note that, for simplicity, a denotes only the cost of attending university. The observation that labour market rewards are higher for higher "ability" individuals can be easily captured by positing a correlation between the effort cost of attending university and the effort cost of working: low a individuals receive the same salary but enjoy higher utility than high a individuals.

[13] The restriction to linear cost is simply a normalisation of the measure of the cost relative to the distribution function $\Phi(a)$. As shown in De Fraja and Valbonesi (2008) including a function $c(a)$ changes nothing in the analysis, provided that $-\frac{\phi(a)}{\Phi(a)} < \frac{c''(a)}{c'(a)}$ for every $a \in (a_{\min}, a_{\max})$.

[14] We maintain throughout, for the sake of definiteness, the assumption that indifferent students attend university: since they have measure 0 in $[a_{\min}, a_{\max}]$, this entails no loss of generality.

We study the steady-state of a dynamic model where new research carried out in a period balances the reduction of research stock due to obsolescence, and where the students' tuition is financed either by parents or by loans secured on their future income. In the absence of perfect capital markets, differential borrowing cost among households can be included in the cost of attending university, a, which would then measure a combination of the utility cost of effort and the interest payments on any loan used to finance university attendance. Interesting on its own right, the steady-state analysis is also a necessary first step in the study of a fuller dynamic model, where the accumulated stock of knowledge, not just current research, affects labour market earnings, and where intergenerational transfers are possible.

3 Equilibrium with no government

In this section we study the benchmark where the university sector is private, and government intervention is limited to at most a lump sum subsidy or tax, independent of anything a university does. Universities charge the same price to all students, and do not exclude students who are willing to pay the enrolment fee. This must be the case if they cannot observe a student's type, and they have no instruments to provide incentives for truth-telling.[15] A university therefore acts as a local monopolist, and chooses the number of students – or the tuition fee – to maximise its objective function. This is obtained by substituting (1) into (2), and then using the demand function, $p = \Delta - \Phi^{-1}(t)$. The maximisation problem of a type θ university is therefore the following:

$$\max_t \left[\frac{\theta}{y} \left(\left(\Delta - \Phi^{-1}(t) \right) t + g \right) - t \right].$$

If an internal solution exists, it satisfies the first order condition:

$$\frac{\theta}{y} \left(\frac{-t}{\phi\left(\Phi^{-1}(t)\right)} + \Delta - \Phi^{-1}(t) \right) - 1 = 0, \tag{3}$$

which can be solved in t to obtain the optimal number of students. The comparison of this outcome with the structure determined by government intervention

[15] Gary-Bobo and Trannoy (2008) show that a research maximising university would select students and charge a tuition fee if and only if students have imperfect information about their own ability. The empirical trade-off between tuition fees and admission standards is explored in Gilboa and Justman (2005).

is the focus of the rest of the paper. The function Z_k, defined next, is a helpful device in this comparison. For $k \in [0,1]$, let $Z_k : [0,1] \longrightarrow \mathbb{R}$ be defined by:

$$Z_k(t) = \Phi^{-1}(t) + \frac{kt}{\phi(\Phi^{-1}(t))}. \tag{4}$$

Lemma 1 *For every $k \in [0,1]$, (i) $Z_k(t)$ is strictly increasing, and (ii) $Z_k(0) = a_{\min}$. (iii) Let $k_1 > k_0$; then, for every $t > 0$, we have $Z_{k_1}(t) > Z_{k_0}(t)$ and $Z'_{k_1}(t) > Z'_{k_0}(t)$.*

The proof of all the results in the paper is in the Appendix.

Z_k can be called the adjusted marginal teaching cost. To interpret it, note that if a university wants to add marginal students in measure dt to those already enrolled, it needs to enrol students whose cost of attendance is $\Phi^{-1}(t)$, and so the fee must be adjusted to entice them. $k \in [0,1]$ is an inverse measure of the social benefit of lower education costs. When $k = 1$, there is no social benefit, and the second component of the RHS of (4) is simply the lost fee income due to the fact that already enrolled students also pay the lower tuition fee. The number of the inframarginal students who enjoy a lower fee, relative to the number of newly enrolled type $\Phi^{-1}(t)$ students, is the hazard rate, $\frac{\Phi(\Phi^{-1}(t))}{\phi(\Phi^{-1}(t))}$, which equals $\frac{t}{\phi(\Phi^{-1}(t))}$ in the second term of the RHS of (4). This component of the marginal cost is offset when there is a social benefit to the fact that attendance to university becomes cheaper. When $k = 0$, the lower revenues of the university are exactly compensated by the lower fees paid by the students, and the marginal cost of an extra student is unaffected by the fact that the university suffers a revenue loss. In the intermediate case, $k \in (0,1)$, the university's revenue loss is only partly compensated by gains elsewhere; as we see in the next section, this happens when the decision maker strictly prefers the universities to be funded by students fees than by general taxation, for example because of distributional, deadweight or administrative costs of taxation.

Using (4), the optimal choice of a type θ university can be presented formally.

Proposition 1 *If the university of type θ is active, it sets the following tuition fee:*

$$p(\theta) = \Delta - \Phi^{-1}\left(Z_1^{-1}\left(\Delta - \frac{y}{\theta}\right)\right), \tag{5}$$

and enrols

$$t(\theta) = Z_1^{-1}\left(\Delta - \frac{y}{\theta}\right) \tag{6}$$

students.

The number of students is set in (6) at a level such that the revenue which can be extracted from an additional student – her additional income Δ –, reduced by $Z_1(t(\theta))$ – the lost revenue due to the fee reduction necessary to induce this additional student –, equals the university's cost of teaching this student, $\frac{y}{\theta}$. The amount of research that the university can carry out is the "profit" made from fee paying students, and is obtained substituting $p(\theta)$ and $t(\theta)$ given in (5) and (6) into the budget constraint (2):

$$r(\theta) = \frac{\theta}{y}\left(\left(\Delta - \Phi^{-1}\left(Z_1^{-1}\left(\Delta - \tfrac{y}{\theta}\right)\right)\right) Z_1^{-1}\left(\Delta - \tfrac{y}{\theta}\right) + g\right) - Z_1^{-1}\left(\Delta - \tfrac{y}{\theta}\right). \quad (7)$$

Expressions (5)-(7) fully describe the choice of each active university as a function of exogenously given parameters only.

The next Corollary gives the relationship between productivity, size and fees in the absence of government intervention.

Corollary 1 $\frac{dt}{d\theta} > 0$, $\frac{dp}{d\theta} < 0$, $\frac{dr}{d\theta} > 0$, $\frac{d^2r}{d\theta^2} > 0$, $\frac{dn}{d\theta} > 0$, *and* $\frac{d}{d\theta}\left(\frac{yr(\theta)}{\theta}\right) > 0$.

More productive universities do therefore teach more students, charge a lower price to attract them,[16] employ more academics and carry out more research. They also spend more on research, as shown in the last inequality. These relationships between size, output and efficiency are often observed in empirical studies.[17] They are sometime attributed to complementarities between teaching and research (Becker 1975 and 1979): teachers (respectively researchers) are thought to be more productive if they also do some research (respectively teaching). The empirical evidence for this is flimsy (Hattie and Marsh 1996). Corollary 1 shows that the observed relationships between size, efficiency and research output need not necessarily imply that there are economies of scale and scope in the technology, but they could be due instead to an underlying unobserved parameter: universities which employ more staff have lower measured

[16] This possibly counterintuitive result is reversed if students respond to "quality" as well as price, see De Fraja and Valbonesi (2008).

[17] Cohn *et al* (1989) study a sample of 1887 US colleges and universities in 1981-82, through a multi-product cost approach: they find economies of scale and scope. De Groot *et al* (1991) Glass *et al* (1995) and Dundar and Lewis (1995) extended Cohn's research focusing respectively on the sensitivity of the cost functions estimates to different output measures, on flexible cost function, and on the departmental production function. For stochastic and non-stochastic frontier analyses, see for example Izadi *et al* (2002) and McMillan and Chan (2006) and the references cited therein.

unit costs, both in teaching and in research, even though each university employs the same linear technology with no economies of scale and scope. Clearly, the relationships given in Corollary 1 would be strengthened by technological economies of scale. Finally, research is convex in productivity.

The above analysis clearly holds if the university of type θ is able to operate. We turn next to the question of which universities are in fact active. Clearly, for university of type θ to be active, it must be that, at its optimal number of students, it can make positive revenues to pay for its research. This is the case if $r(\theta)$ given in (7) is positive. Note that, for an active university, the optimal number of students is independent of the grant g. This has the following immediate consequence.

Corollary 2 *If $g \geqslant 0$, a university of type θ enrols students if and only if*

$$\theta > \frac{y}{\Delta - a_{\min}}. \tag{8}$$

The interpretation of (8) is natural: the teaching cost of a student is $\frac{y}{\theta}$. For the university to want to teach a strictly positive measure of students, it must be worth for at least the students with the lowest a to pay for this cost, and the willingness to pay for tuition of this student is the increase in her labour market earnings as a consequence of her having a degree, Δ, reduced by the utility cost of attending university, $a_{\min}$.

Consider research now. Clearly, if $g = 0$, a university can do research if and only if it can make positive revenues from its teaching. If $g > 0$, universities with type lower than the RHS in (8) can do some research by enrolling no students and spending all their grant on research: they are research only institutions, but they are the least productive among the active universities. Conversely, if $g < 0$, then some universities are prevented from becoming active even though they could raise enough tuition fees to pay their teaching costs, and the smallest universities teach a strictly positive number of students and do no research.

4 Government intervention

We now assume that the government intervenes actively in the higher education sector. We do not address the issue of actual *ownership* of universities. Anecdotal evidence suggests that there is little systematic difference in efficiency or in objective function between private and public universities.[18] We simply assume that the government imposes regulatory constraints on the university sector, and that these constraints are the same for public and for private universities.

Formally, the government maximises total utility in society, and universities know their own θ and how much r they carry out, while the government cannot observe either. Some research is of course *ex-ante* observable or *ex-post* measurable: but the very fact that governments all over the world expend considerable, and costly, effort to measure research output indicates that universities enjoy an information advantage relative to their funders and regulators. All our model needs is that at least some research effort is not observable. In the presence of uncertainty in research activities, so that a large unobserved research effort may well lead to no results, and conversely, given that serendipity and luck may yield huge returns at little cost, and in the presence of a potentially extremely long time before effort determines output, it is very difficult to infer and reward effort from the results obtained.[19] In other words, we assume, in line with our perception of universities, public or private alike, that they have both a trade-off between teaching and research which leans more towards research than the government's, and an informational advantage over the government.

The government designs the university policy.[20] In our model, this is simply a pair of functions $\{p(t), g(t)\}$, offered to all potential universities: the government commits to linking the number of students enrolled, t, with the tuition fee a university is allowed to charge, $p(t)$, and with the block funding grant awarded to the university, $g(t)$. Both $p(t)$ and $g(t)$ can be negative: the government could subsidise students and tax universities. Faced with this policy, each university can choose the number of students it enrols, receiving the corresponding government grant $g(t)$, and charging the corresponding tuition fee $p(t)$. The government's grant to universities is funded by general taxation; to keep things simple, we model taxation as a lump sum tax h, the same for all individuals: since income affects utility linearly, this is optimal. Naturally,

[18] Three in the top ten universities in the 2009 Shangai ranking and four in the top six in the 2009 Times Higher Education ranking are public.

[19] And even when output is observed, it might be difficult to asses: is a certain profound mathematical theorem a fundamental discovery, or a "useless" fascinating intellectual game? The mathematicians who could answer this question have no incentive to do so truthfully.

[20] Beath *et al* (2005) take the funding mechanism as given and concentrate on the effects on incentives for teaching and research quality of different public funding schemes.

the tax is constrained not to exceed the income from the unskilled labour market: $h \leqslant y - \Delta$. In the rest of the paper we assume that this constraint is in fact slack; that is, plausibly, the total tax needed to finance the preferred level of tertiary education is not so high as to require more than the aggregate income that would be obtained with no university sector, when all workers are unskilled. Raising one unit of resources in tax has an exogenously given cost $(1+\lambda) > 1$. As De Fraja and Valbonesi (2008) shows, in addition to the standard administrative and distortionary costs of taxation, λ also captures the government's preference for redistribution.

To determine the government's optimal policy, we take the standard revelation approach. The government asks each university to report its own type, and commits to imposing a vector of variables as functions of the reported type, which the university must adhere to: by the revelation principle, the government cannot improve on the payoff it can obtain by restricting its choices to the set of mechanisms such that no university has an incentive to mis-report its type. With this perspective, a policy is a triple, $\{t(\theta), p(\theta), g(\theta)\}_{\theta \in (0,\bar{\theta}]}$, the number of students, the tuition fee and the government grant as functions of the reported type. The employment at university of type θ is given by $n(\theta) = \frac{p(\theta)t(\theta)+g(\theta)}{y}$. We include *both* t and p as policy variables, thus allowing, potentially, the number of students enrolled in a university to be different from the number of individuals who, given the tuition fee, would prefer to graduate. Clearly, it cannot exceed it, and so we must impose the constraint

$$\Phi^{-1}(t(\theta)) \leqslant \Delta - p(\theta), \qquad \theta \in \left[\underline{\theta}, \bar{\theta}\right]. \tag{9}$$

(9) says that the type of the marginal student must be no greater than the type of the student who is indifferent between going and not going to university. As we show below, (9) is in fact binding at the government's optimal policy: it cannot happen that the number of university places needs to be rationed by non-price methods. Intuitively, this is so because the shadow cost of public funds exceeds 1, and so it is always preferable for the government to raise funds through tuition fees than through taxes.

To set up the government's problem as an optimal control in a suitable way, we introduce the auxiliary variable R, the total amount of research:

$$R = \int_0^{\bar{\theta}} r(\theta) f(\theta) \, \mathrm{d}\theta = \int_{\underline{\theta}}^{\bar{\theta}} r(\theta) f(\theta) \, \mathrm{d}\theta. \tag{10}$$

In the above, $\underline{\theta}$ is the cut-off point type of university, such that those above operate, those below do not, and hence $r(\theta) = 0$ for $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$. $\underline{\theta}$ is determined endogenously, as the "initial time" (Leonard and van Long 1992, p 222 ff). Finally, $r(\theta)$ too is treated as a variable chosen by the government, subject, as explained above, to the incentive compatibility constraint that all universities prefer to reveal their type truthfully. We derive this constraint in Proposition 2. Note first that the utility of a type θ university who has reported type θ is

$$r(\theta) = [p(\theta)\, t(\theta) + g(\theta)]\, \frac{\theta}{y} - t(\theta), \qquad \theta \in \left[\underline{\theta}, \bar{\theta}\right]. \tag{11}$$

Proposition 2 *Let $\underline{\theta}$ be the least productive active university. The following are jointly necessary and sufficient conditions for incentive compatibility. For $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$:*

$$r'(\theta) = \frac{p(\theta)\, t(\theta) + g(\theta)}{y}, \qquad r(\underline{\theta}) = 0, \qquad r\left(\bar{\theta}\right) \text{ free,} \tag{12}$$

$$t'(\theta) \geqslant 0. \tag{13}$$

The participation, or individual rationality, constraint is $r(\theta) \geqslant 0$ for $\theta \geqslant \underline{\theta}$. Given that universities have a stronger preference for research than the government and that $r'(\theta) \geqslant 0$, the participation constraint simplifies to $r(\underline{\theta}) = 0$, as stated in (12).

In Proposition 2, we are assuming that the government can induce any incentive compatible non-negative research level. This would be the case if a university must accept the government's proposal (for example if it needs a licence) in order to operate at all. If universities could "go it alone", then they would be able to carry out at least the amount of research determined in Section 3, and so the participation constraint would be for $r(\theta)$ to be greater than or equal to the RHS of (7).

Further, the number of students in a local market must be non-negative:

$$t(\theta) \geqslant 0, \qquad \theta \in \left[\underline{\theta}, \bar{\theta}\right]. \tag{14}$$

$t(\theta)$ cannot exceed 1 either: since $\Delta < a_{\max}$, if the number of students from a local education market were 1, then the total utility from individuals in that market could be increased simply by stopping the students with the highest cost of effort from attending university, and so a situation were there are some $\theta \in \left[0, \bar{\theta}\right]$ were $t(\theta) = 1$ cannot happen at the optimum, and the constraint $t(\theta) \leqslant 1$ can be omitted.

In the jargon of optimal control analysis, the problem can be written as a free initial time optimal control problem with R as a parameter; the integral constraint (10) is re-written as a state constraint (Leonard and van Long, 1992, p 190), with $r_0(\theta)$ as an auxiliary variable:

$$r_0'(\theta) = r(\theta) f(\theta), \qquad r_0(\underline{\theta}) = 0, \qquad r_0(\bar{\theta}) = R. \tag{15}$$

The instruments described and the constraints derived, we can finally present the government's problem. This is the maximisation of a welfare function made up of three components. First, the total after tax income of the population. Second, the disutility costs borne by those who attend university. And third, the overall externality bestowed by the aggregate amount of research, which we measure by ωR, with $\omega \geqslant 0$.[21] ω captures the positive effect of research on total factor productivity, as well as non-monetary benefits of research such as the national pride at the award of Nobel prizes, Fields Medals, and other formal or informal recognition. We assume, naturally, that the social value of research does not exceed the salary cost of the most efficient research institutions:

$$\omega < (1+\lambda)\frac{y}{\underline{\theta}}. \tag{16}$$

Otherwise the government would have a stronger desire to do research than these institutions.

Proposition 3 *The government's problem is:*

$$\max_{\substack{p(\theta),t(\theta),r(\theta),\\ g(\theta),R,\underline{\theta}}} \int_{\underline{\theta}}^{\bar{\theta}} \left((\Delta - p(\theta))\, t(\theta) - \int_{a_{\min}}^{\Phi^{-1}(t(\theta))} a\phi(a)\,\mathrm{d}a - (1+\lambda)\, g(\theta) \right) f(\theta)\,\mathrm{d}\theta$$
$$+ y - \Delta F(\underline{\theta}) + \omega R, \tag{17}$$

s.t. (9), (11), (12), (13), (14) and (15).

To derive the payoff function intuitively, note that the total utility of the potential students in a local education market where a type θ university is active is given by

$$(y - \Delta - h) + (\Delta - p(\theta))\, t(\theta) - \int_{a_{\min}}^{\Phi^{-1}(t(\theta))} a\phi(a)\,\mathrm{d}a. \tag{18}$$

[21] Formally this is identical to including a proportion ω of the universities payoff in the computation of social welfare, just as a share of a regulated firm's profit is typically included. An earlier version of the paper included explicitly a link between labour market incomes and the aggregate amount of research: $y(R)$. As a referee pointed out, this link can be included in the catchall term ω, with a considerable gain in simplicity.

This has a natural interpretation: all potential students receive after tax income $y - \Delta - h$ at least; of the potential students, $t(\theta)$ do go to university, and receive an *additional* income, net of tuition fee, equal to $\Delta - p(\theta)$: this is the second term in (18). The aggregate disutility cost of attending university, given by the last term in (18), must then be subtracted. Integrate over θ and add ωR, the direct benefit of research, to obtain (17).

Proposition 4 describes the government's optimal policy. Define the function $\sigma(\theta)$ as

$$\sigma(\theta) = \Delta - \frac{y}{\theta} + \frac{(1 - F(\theta))\, y + \frac{\omega}{1+\lambda} \int_{\theta}^{\bar{\theta}} \tilde{\theta} f(\tilde{\theta}) \mathrm{d}\tilde{\theta}}{f(\theta)\, \theta^2}. \tag{19}$$

We said before that we require $r_{\max}$ to be "high". Formally we now posit:

Assumption 2 $r_{\max} > \bar{\theta} \int_{\underline{\theta}}^{\bar{\theta}} \tilde{\theta}^{-2} Z^{-1}_{\frac{\lambda}{1+\lambda}} \left(\sigma(\tilde{\theta}) \right) \mathrm{d}\tilde{\theta}.$

This implies that at the solution, no university is constrained by the requirement $r(\theta) \leqslant r_{\max}$, and this constraint can be omitted from Problem (17).

Proposition 4 *Let (16) and Assumption 2 hold. If a solution to problem (17) exists, then it satisfies:*

$$t(\theta) = Z^{-1}_{\frac{\lambda}{1+\lambda}} (\sigma(\theta)), \tag{20}$$

$$p(\theta) = \Delta - \Phi^{-1}(t(\theta)), \tag{21}$$

$$r(\theta) = \theta \int_{\underline{\theta}}^{\theta} \tilde{\theta}^{-2} t(\tilde{\theta}) \mathrm{d}\tilde{\theta}. \tag{22}$$

(20) gives the number of students in university θ, which (13) requires to be increasing. As shown in the Appendix, (20) is indeed increasing: more productive universities have more students. (22) shows that they also do more research. As we saw in Section 3, this was also the case with unfettered private provision. The explanation, however, is slightly different in this case: with unfettered private provision, a more productive university needs to teach more students to be able to carry out more research. On the other hand, when the government controls the sector, it asks more productive universities to teach more students because they are more productive: precisely for the same reason it also asks them to carry out more research. This different angle is brought in starker relief in Section 6, which presents the case in which the government

is not constrained by its information disadvantage, and is therefore able to separate fully the allocation of teaching and research. Notice also that the research *expenditure*, as well as the *amount* of research done, is increasing in θ. To see this, use (12) and (11) to show that $\frac{d}{d\theta}\left(\frac{yr(\theta)}{\theta}\right) = \frac{yt(\theta)}{\theta^2} > 0$.

The global variables are simply obtained from substitution of the values obtained from Proposition 4. For example, the least productive active university, and the total amount of research are given by:

$$\underline{\theta} = \sigma^{-1}\left(a_{\min}\right), \tag{23}$$

$$R = \int_{\underline{\theta}}^{\bar{\theta}} \theta \int_{\underline{\theta}}^{\theta} \tilde{\theta}^{-2} Z^{-1}_{\frac{\lambda}{1+\lambda}}\left(\sigma(\tilde{\theta})\right) \mathrm{d}\tilde{\theta} f\left(\theta\right) \mathrm{d}\theta. \tag{24}$$

(23) corresponds to (8) and follows from setting the RHS of (20) to 0, and using the fact that $Z_k\left(0\right) = a_{\min}$. (24) is simply the sum of every university's research.

5 Implementation: The design of the funding scheme

In this section we show how the above mechanism can be implemented in practice. Begin by noting that (11) must hold for every θ. Differentiate it with respect to θ:

$$\frac{y}{\theta}\left(r'\left(\theta\right) + t'\left(\theta\right)\right) - \frac{y}{\theta^2}\left(r\left(\theta\right) + t\left(\theta\right)\right) - g'\left(\theta\right) - \frac{d\left(p\left(\theta\right) t\left(\theta\right)\right)}{d\theta} = 0.$$

Denote by $T_F\left(\theta\right)$ the total funding available to a type θ university. Substitute (12) and (11) into the above, to obtain:

$$T_F'\left(\theta\right) = g'\left(\theta\right) + \frac{d\left(p\left(\theta\right) t\left(\theta\right)\right)}{d\theta} = \frac{y}{\theta} t'\left(\theta\right) > 0. \tag{25}$$

The north-east quadrant of Figure 3 depicts the relationship between t and T_F. This is derived from (25) in the following way: the locus in the south-west quadrant illustrates the relationships between θ and the number of students $t\left(\theta\right)$ taught by a type θ university, given in (20). This is increasing. Also increasing, by (25), is the locus showing the relationship between θ and the total funding available to a type θ university, $T_F\left(\theta\right)$, drawn in the north-west quadrant. Joining the two loci via the 45 degree line in the south-east quadrant, we obtain, in the north-east quadrant, the relationship between the number of students and the total funding available to a university. If the government offers this relationship to all universities, that is if the government allows universities to

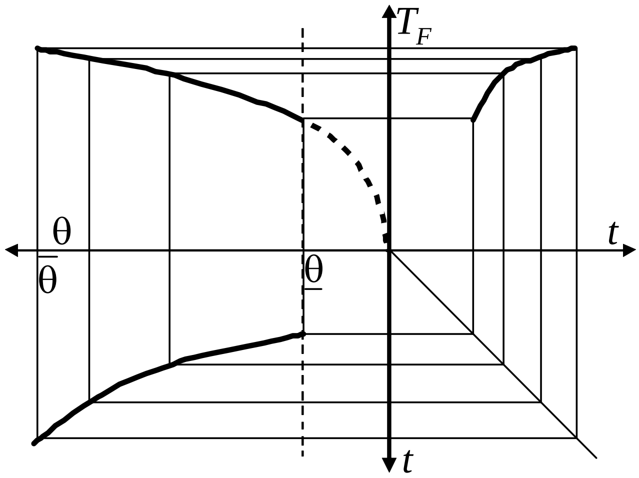


Figure 3: The implementation of the optimal policy.

choose any point on the curve, then each university will select the combination of students given by (20) and total funding which will allow it carry out the amount of research given by (22). Note that, since both $t(\theta)$ and $T_F(\theta)$ are increasing, then so is the relationship between number of students and total funding, as depicted.

The following Corollary illustrates this relationship.

Corollary 3 *Taken as a function of t, T_F is concave.*

Therefore, as in Laffont and Tirole's (1986 pp 630-631) analysis of procurement contracts, the policy can be implemented by offering universities a menu of "subsidy per student"-"research grant" combination. This implies that the government can simply offer all universities a menu of linear contracts, $g_T(p)$, where p, the tuition fee per student, is given by the slope of the tangent of the curve in the north-east quadrant, and g_T, the lump sum grant, is the ordinate of this tangent: this is such that the lower the tuition fee the higher the grant. Faced with this menu, each university will simply select the combination of funding and fee per student that corresponds to its own type. Concavity of the curve implies that universities which charge students less are "rewarded" with a larger grant.

6 The systems compared

To interpret the solution obtained in Proposition 4, it is helpful to compare it with the policy the government would choose if, *ex-ante*, it knew perfectly the type of each university, or, equivalently, if it could *ex-post* measure precisely each university's research effort. This is referred to as the first best and is described it in the next proposition.

Proposition 5 *If the government had perfect information it would choose:*

$$t(\theta) = Z^{-1}_{\frac{\lambda}{1+\lambda}}\left(\Delta - \frac{y}{\theta}\right), \tag{26}$$

$$p(\theta) = \Delta - \Phi^{-1}(t(\theta)), \tag{27}$$

$$r(\theta) = \begin{cases} r_{\max} & \textit{for} \quad \theta \geqslant \frac{1+\lambda}{\omega}y \\ 0 & \textit{for} \quad \theta \in \left[\frac{y}{\Delta - a_{\min}}, \frac{1+\lambda}{\omega}y\right) \end{cases}. \tag{28}$$

The number of students, given in (26), has a similar expression as for the unfettered market case, (6), and for the case of asymmetric information, given in (20). From the former it differs as $Z_{\frac{\lambda}{1+\lambda}}$ replaces Z_1, from the latter in that the argument of $Z^{-1}_{\frac{\lambda}{1+\lambda}}$ does not include the information distortion terms. Since $\sigma\left(\bar{\theta}\right) = 0$, the two expressions are identical when $\theta = \bar{\theta}$, that is we have efficiency at the top, and the price reflects this. On the other hand, the expression for a university's amount of research, (28), is radically different from the corresponding expression for the case of asymmetric information, (22). While Assumption 2 ensures that the upper boundary on research is not binding in the asymmetric information case, this is not possible in this case, since the government does not need to provide incentives for research, but can simply command and control the activities of each university. It therefore allocates research to the most productive universities, its only constraint the technological upper bound.[22] Unlike the private market case and the case of imperfect information, with perfect information there are "teaching only" universities: universities with θ in $\left[\frac{y}{\Delta - a_{\min}}, \frac{1+\lambda}{\omega}y\right)$ enrol students but carry out no research.

[22] In Proposition 5, research can take two values only, $r_{\max}$ or 0, as shown by the dotted line in Figure 4. This follows from the hypothesis that marginal cost is 0 up to the exogenously given upper bound, and is $+\infty$ beyond it. With less extreme forms of decreasing returns to research expenditure, the "bang-bang" nature of the research policy would be tempered: the concentration of research in the most productive universities would remain, but different high productivity universities would do different amount of research.

Next, from Proposition 5, we can determine the government subsidy $g(\theta)$:

$$g(\theta) = \begin{cases} \frac{r_{\max}}{\theta} y - \frac{t(\theta)^2}{\phi(\Phi^{-1}(t))} & \text{for} \quad \theta \geqslant \frac{1+\lambda}{\omega} y \\ -\frac{t(\theta)^2}{\phi(\Phi^{-1}(t))} & \text{for} \quad \theta \in \left[\frac{y}{\Delta - a_{\min}}, \frac{1+\lambda}{\omega} y\right) \end{cases}.$$

The subsidy is *negative* for the "teaching only" universities: they receive more in fees that they pay out in salaries, and their surplus is transferred to the high θ universities which do research. In words, if the government had perfect information about the productivity of the universities, it would choose a policy whereby students attending low θ universities pay for research carried out elsewhere.

Notice that in our model there is no natural welfare ranking of the total amount of research in the three regimes. This is because even though research does bestow an externality, it is not necessarily underprovided by an unfettered private university sector: whether or not it is depends in general on the balance between technology – how much the nation's scientific and cultural state affects the production of goods and services –, the direction research takes – how much of what researchers do spills over to the rest of society –, and the subjective preferences of the government.[23] In other words, the first best value of R (the preferred amount of research when the government has perfect information) may well be lower than the amount of research carried out in an unfettered private market. To see this, think of a situation where ω is small, indicating that research has little social benefit. If Δ is sufficiently large, then students are quite willing to pay for university tuition. Indeed so willing that universities raise enough tuition fee income to pay an amount of research which exceeds that which a welfare maximising government would want to choose.

The following assumption allows us to abstract from the effects of the *aggregate* amount of research and to concentrate on the more microeconomic aspects of the *distribution* of teaching and research across institutions.

Assumption 3 $r_{\max}$ *and* ω *are such that the equilibrium total amount of research,* R*, is the same in the three regimes considered.*

[23] While a mathematical theorem may eventually help improve computer software used in designing robots, a chemical discovery may allow the development of more effective drugs, reducing the number of days lost due to illness, and advances in game theory may lead to improved incentive mechanisms used by organizations to select and motivate staff, other research activities could instead be viewed as an end in itself, academics indulging in their intellectual hobbies, with no expected current or long term benefit to society.

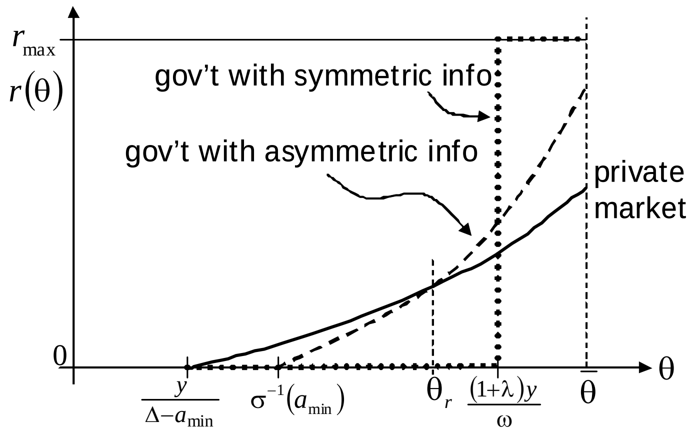


Figure 4: The amount of research with private and government provision.

It is in general possible for the parameters $r_{\max}$ and ω to satisfy Assumption 3. To see this, let $\hat{R}$ be the amount of research in the unfettered private market. $\hat{R}$ is independent of both $r_{\max}$ and ω. Therefore there exists a value of ω such that the aggregate amount of research chosen by the government under asymmetric information equals $\hat{R}$ (this is the value of ω such that the RHS in (24) equals $\hat{R}$). Finally, note that $r_{\max}$ appears only in the solution for the first best case. The total amount of research in this case is $\left(1 - F\left(\frac{1+\lambda}{\omega}y\right)\right) r_{\max}$, and, when $r_{\max}$ is such that this equals $\hat{R}$, then Assumption 3 holds.

Figure 4 illustrates the distribution of the total amount of research across universities in the three regimes, under Assumption 3, so that the total research is the same in all three. The three curves depict the amount of research as a function of the productivity of the university, θ: the solid one with private provision, the dashed curve in the case of government intervention with imperfect information, the dotted curve when the government can perfectly observe each university's productivity. The integral, with measure $f(\theta)$, of the three curves is the same, and so the dashed line is above the solid one for high θ and vice versa, as drawn. Both the solid (private market), and the dashed (government intervention in asymmetric information) line are convex. The former from $\frac{d^2 r}{d\theta^2} > 0$ in Corollary 1, the latter because the slope of $r(\theta)$ is proportional to total funding (12), which, by (25), in increasing in θ. The dotted line is 0 below the threshold value of θ given in (28) (the "teaching only" institutions), and at its maximum above this threshold. The intuition is that, with perfect information, the government can separate teaching and research: the former is chosen on the basis of efficiency and equity considerations in teaching only, in each local market to the point where marginal benefits equal marginal costs. The total amount of research, R, is chosen at the optimally global level, given by the condition that *global* marginal benefits, ω, equal *global* marginal costs. This total amount of research is allocated to the university sector in the most cost effective way, by asking the most productive universities to do as much research as they can.

Figure 4 also illustrates the role of the assumption that universities must either accept the government's offer, or remain inactive. If instead they could also choose to operate independently, then $r(\theta)$ could not be reduced below the RHS of (7), and so the part of the dashed line to the left of their intersection θ_r would move up to overlap the solid one.

Consider next the distribution and the number of students in the three regimes. The following summarises the comparison.

Corollary 4 *Let Assumptions 2 and 3 hold. Then:*

1. *The universities active with unfettered private provision are the same a perfectly informed government would allow to operate; conversely fewer universities are active with asymmetric information.*

2. *With unfettered private provision, each active university has fewer students than it would have with a perfectly informed government.*

3. *Relative to perfect information, the government information disadvantage reduces the number of students at each university except the most productive.*

The diagrams of Figure 5 illustrate this. The Corollary implies that, compared with private provision, government intervention concentrates students in the most productive institutions: the higher (lower) productivity institutions have more (fewer) students than they would in a private system. The horizontal axis measures the number of students. This, for university θ, is given by the intersection of the increasing line $Z_{\frac{\lambda}{1+\lambda}}(t)$, under government provision, and $Z_1(t)$ for the private market, with the appropriate horizontal line:

$\Delta - \dfrac{y}{\theta}$	for the private market or with perfect information,
$\Delta - \dfrac{y}{\theta} + \sigma(\theta)$	with imperfect government information.

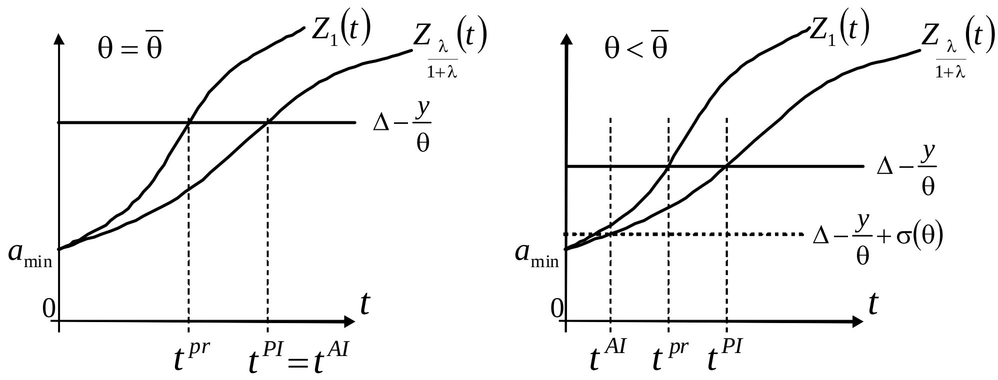


Figure 5: The determination of the number of students. AI: asymmetric information. PI: perfect information. pr: unfettered private provision.

The LHS of Figure 5 shows the determination of the number of students at the highest possible value of θ. Assumption 3 holds, and so the horizontal line $\Delta - \frac{y}{\theta}$ is the same in all three regimes. The number of students is lower with private provision than with government intervention. This is so because, as Lemma 1 implies, when k increases from 0 to 1, the curve Z swings anticlockwise around the point $(0, a_{\min})$, and since $1 = \lim_{\lambda \to +\infty} \frac{\lambda}{1+\lambda}$, the curve Z is lower under government provision for every finite value of λ. In this case we have $t^{pr} < t^{AI} = t^{PI}$. This has a natural explanation: unlike private universities, the government receives some benefit from the fact that students pay lower fees. It therefore will want to push the number of students recruited beyond what a private university sector would do. The strength of this effect depends on the social cost of raising taxes to pay for students' tuition: if this is very high, then the overall government cost of enticing students to attend university becomes similar to the private universities' and the curve $Z_{\frac{\lambda}{1+\lambda}}(t)$ draws closer to $Z_1(t)$. When $\theta = \bar{\theta}$, we have $t^{AI} = t^{PI}$: this is the standard "efficiency at the top" result. The RHS of Figure 5 considers a value of θ lower than $\bar{\theta}$. For such θ, the horizontal curve is lower, for all three regimes, than with $\theta = \bar{\theta}$, but is "more lower" when the government has imperfect information, as depicted by the dotted horizontal line, because $-\frac{y}{\theta} + \sigma(\theta)$ is increasing at $\theta = \bar{\theta}$. As the RHS diagram shows, for sufficiently low θ, we have that $t^{AI} < t^{pr}$ and both of them are smaller than t^{PI}: productive universities do more research than less productive ones, both in an unfettered market and with the optimal government policy, but they recruit more students in the latter regime. This is roughly suggested by the data presented in Figure 2 in the Introduction.

The intuition is the following. Research is cheaper in more productive universities, and so the government wants them to carry out more research. To give them the incentive to do so, it rewards them with a combination of a larger total income (grant plus student fees) and a bigger number of students. A less productive university, which would like to receive the higher total income promised to a productive one, is thus deterred from claiming to have high productivity: if it did so to receive a bigger grant, it would also have to recruit an increased number of students. This is costly, as they can only be recruited by charging them a lower fee. Since it is less productive, extra students are more expensive than they would be for a higher θ university, and the extra total income (the higher lump-sum grant and the increased fee revenue) received for teaching more students is not sufficient to cover the cost of teaching them: research would have to be sacrificed to meet the teaching obligation, which would defy the point of claiming to be of higher type in the first place. The term $\sigma(\theta)$ can naturally be interpreted as the information cost incurred by the government: except for the most productive university, the balance between student fees and lump-sum grant is inefficiently skewed towards student fees, so as to make less productive universities less willing to expand their student intake.

As θ decreases further, the intersection of the horizontal lines with the curves $Z_1(t)$ and $Z_{\frac{\lambda}{1+\lambda}}(t)$ move towards the vertical axis. The value of θ for which this intersection reaches the axis is the type of the least productive university that teaches any students: clearly this happens for the same value when provision is via an unfettered private market and in the first best, and for a higher value of θ when the government has imperfect information: fewer universities are active in this case.

Figure 6 sketches the relationship between θ and the number of students, in the three regimes, analogously to Figure 4 for research. The dotted line, depicting the perfect information case, coincides with the dashed one, the asymmetric information case, at $\theta = \bar{\theta}$ (efficiency at the top). Instead it coincides with the solid one, the private market case, at $\theta = \frac{y}{\Delta - a_{\min}}$, because $Z_1(0) = Z_{\frac{\lambda}{1+\lambda}}(0)$.

Notice the information externality among universities in different local education markets: whether some students attend university or not depends on the cost conditions in the rest of the university sector, even though the cost of providing university education is fully determined at the local level.

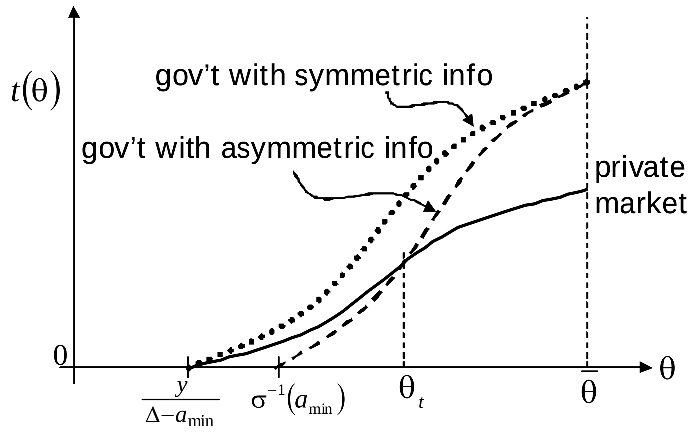


Figure 6: The number of students with private and government provision.

As shown above, the expression for the slope of the total research expenditure is $\frac{yt(\theta)}{\theta^2}$ both for government intervention with asymmetric information and in the unfettered private market. This implies that, for $\theta < \theta_t$, that is for θ lower than the intersection of the dashed and the solid line in Figure 6 – those university types that would teach more students in the unfettered private market than with government intervention under asymmetric information –, the total research would grow with θ slower in the latter regime, and vice versa for $\theta > \theta_t$. Therefore, since at $\theta = \theta_t$ the slope of the total research is the same in the two regimes, and since at $\theta = \theta_r$ the total expenditure on research is the same in the two regimes, then $\theta_r > \theta_t$. It follows that low (high) type universities, those with $\theta < \theta_t$ (those with $\theta > \theta_r$), teach more (fewer) students, and spend more (less) on research with private provision than with government intervention; intermediate types, $\theta \in (\theta_t, \theta_r)$, teach fewer students and spend more on research with private provision.

Combining Figures 4 and 6 gives the relationship between the number of students and the research carried out, or the expenditure on research. The comparison between unfettered market provision and government intervention with asymmetric information indicates that this relationship follows a similar locus in the two regimes, but it stretches out more with government intervention, as the highest θ universities have fewer students and spend less on research under unfettered private provision: the largest universities in the private sector should be smaller, that is have fewer students and spend less on research, than the largest universities in the state sector. Figure 7 roughly bears this out. It

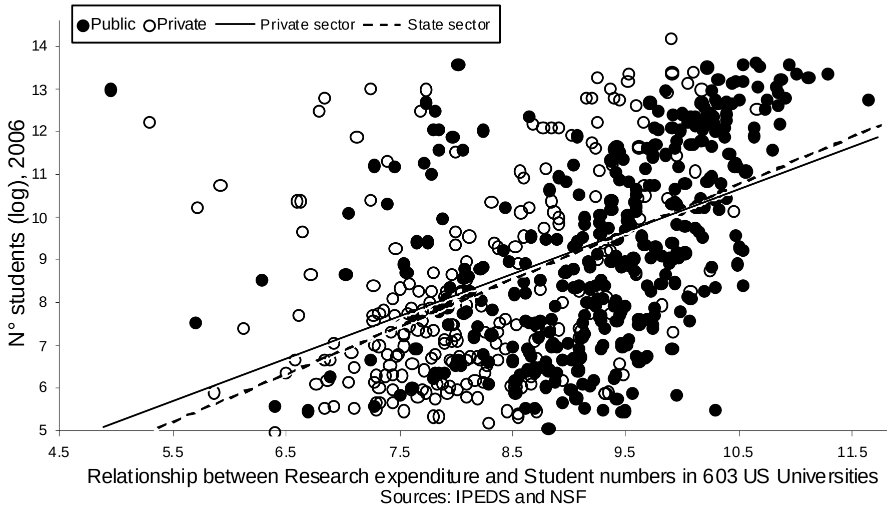


Figure 7: Research expenditure and student numbers.

plots the number of students against the expenditure on research in 243 private and 360 state universities in the US: these are all the universities in the 2006 Survey of Research and Development Expenditures at Universities and Colleges (NSF) for which data on student enrollement in the corresponding period is recorded by the IES. The two regression lines are not significantly different (in intercept or slope), whereas the distribution of public university is shifted towards north-east.[24]

Recall that Figure 6 is drawn for the special case when Assumption 3 holds and the total amount of research R is the same in the three regimes. When this is not the case, the horizontal curve in Figure 5 would have a different position in each regime, and the comparison would have to be made taking into account of the different position of this curve. In general, however, note that this horizontal curve shifts *down* when R is higher: this naturally reflects the trade-off between teaching and research.

[24]The mean of the log of the expenditure on research is 8.381 for private universities and 9.396 for state university, with a value of 5.345 for the t-statistics of their difference being non-zero; the figures for the log of the number of students are 8.245 and 9.302, with t-statistics of 13.04.

7 Concluding remarks

This paper studies how a utilitarian government should intervene in the university sector. Intervention can be beneficial because of an externality in research, and because, even though the total *amount* of research which is carried out in an unfettered private might be the optimal level, its *distribution* across universities is not. Specifically, we show that the private market spreads research too thinly: if the government could freely determine who does how much research, it would concentrate it in the most productive universities, and allow less productive universities as "teaching only" institutions, whose students subsidise research carried out elsewhere. The information disadvantage of the government vis-à-vis the universities implies that this is not possible, and the government must allow all teaching universities to do at least some research. This is inefficient, and it increases the overall cost of university provision. In response to this increase, the government reduces the number of universities relative to private provision. The overall effect on the number of students is ambiguous, because while there are fewer universities, the more productive ones are given stronger incentives to admit more students.

The model we utilised to carry out our analysis has a very rich set-up: it is a general equilibrium model, with taxation and higher education, with a continuum of different local markets, a continuum of different students in each market, and information, teaching and research interactions among all the local markets mediated by a global labour market. To keep the complexity at a manageable level, we have introduced a number of simplifying assumptions. An earlier version of the paper (De Fraja and Valbonesi 2008) shows that the analysis is robust to relaxation of several of the assumptions introduced here. For example, whilst the model presented here model takes a strictly utilitarian set-up, with social welfare given by the sum of individual utilities, De Fraja and Valbonesi (2008) shows how distributional concerns can be taken into account without affecting the results. Similarly, for some student mobility and for the cases where their willingness to pay for education depends on the research output of the university they attend. Research in progress studies the endogenous determination of the productivity of each university.

References

Aghion, Philippe, Mathias Dewatripont, Caroline Hoxby, Andreu Mas-Colell, and Andre Sapir (2010), "The governance and performance of research universities: Evidence from Europe and the US." *Economic Policy*, 25/61, 7–59.

Baron, David P and Roger B Myerson (1982), "Regulating a monopolist with unknown cost." *Econometrica*, 50, 911–930.

Beath, John, Joanna Poyago-Theotoky, and David Ulph (2005), "University funding systems and their impact on research and teaching: A general framework." Loughborough University Discussion Paper.

Becker, William (1975), "The university professor as a utility maximiser and producer of learning, research and income." *Journal of Human Resources*, 10, 109–115.

Becker, William (1979), "Professorial behavior given a stochastic reward structure." *American Economic Review*, 69(5), 1010–1017.

Cohn, Elchanan, Sherrie L W Rhine, and Maria C Santos (1989), "Institutions of higher education as multi-product firms." *The Review of Economics and Statistics*, 4, 281–329.

Dale, Stacy Berg and Alan B Krueger (2002), "Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables." *Quarterly Journal of Economics*, 117, 1491–1527.

Danzon, P (1998), "Hospital profits: The effect of reimbursement policies." *Journal of Health Economics*, 1, 29–52.

De Fraja, Gianni and Elisabetta Iossa (2002), "Competition among institutions and the emergence of the elite university." *Bulletin of Economic Research*, 54, 275–293.

De Fraja, Gianni and Paola Valbonesi (2008), "The design of the university system." CEPR Discussion Paper.

de Groot, Hans, Walter W McMahon, and J Fredericks Volkwein (1991), "The cost structure of American research universities." *The Review of Economics and Statistics*, 43, 121–133.

Del Rey, Elena (2001), "Teaching versus research: A model of state university competition." *Journal of Urban Economics*, 49, 356–373.

Dranove, David and William D White (1994), "Recent theory and evidence on competition in hospital markets." *Journal of Economics and Management Strategy*, 3, 169–209.

Drèze, Jacques H and Fernanda Estevan (2007), "Research and higher education in economics: Can we deliver the Lisbon objectives?" *Journal of the European Economic Association*, 5, 271–304.

Dundar, Halil and Darrell R Lewis (1995), "Departmental productivity in American universities: Economies of scale and scope." *Economics of Education Review*, 14, 119–144.

Ehrenberg, Ronald G (2004), "Econometric studies of higher education." *Journal of Econometrics*, 121, 19–37.

Ehrenberg, Ronald G (2007), "The economics of tuition and fees in American higher education." In *International Encyclopedia of Education*, forthcoming.

Epple, Dennis, Richard E Romano, and Holger Sieg (2003), "Peer effects financial aid and selection of students to colleges and universities: An empirical analysis." *Journal of Applied Econometrics*, 18, 501–525.

Epple, Dennis, Richard E Romano, and Holger Sieg (2006), "Admission, tuition, and financial aid policies in the market for higher education." *Econometrica*, 18, 885–928.

Euwals, Rob and Melanie E Ward (2005), "What matters most: Teaching or research? empirical evidence on the remuneration of British academics." *Applied Economics*, 37, 1655–1672.

Garvin, David A (1980), *The Economics of University Behavior*. Academic Press, New York.

Gary-Bobo, Robert and Alain Trannoy (2008), "Efficient tuition fees and examinations." *Journal of the European Economic Association*, 6, 1211–1243.

Gilboa, Yaakov and Moshe Justman (2005), "Academic admissions standards: Implications for output, distribution, and mobility." *Journal of the European Economic Association*, 3, 1105–1133.

Glass, J Colin, Donal G McKillop, and Noel Hyndman (1995), "The achievement of scale efficiency in UK universities: A multiple-input multiple-output analysis." *Education Economics*, 3(3), 249–263.

Golden, Bill B, Leah J Tsoodle, Oluwarotimi O Odeh, and Allen M Featherstone (2009), "Determinants of agricultural economic faculty salaries: A quarter of a century later." *Review of Agricultural Economics*, 28, 254–261.

Hammond, Phillip E, John W Meyer, and David Miller (1969), "Teaching versus research: Sources of misperceptions." *Journal of Higher Education*, 40, 682–692.

Hattie, John and H W Marsh (1996), "The relationship between research and teaching: A meta-analysis." *Review of Educational Research*, 66(4), 507–542.

Hoxby, Caroline M (1997), "How changing the market structure of the US higher education explains college tuition." National Bureau of Economic Research, New York.

Izadi, Hooshang, Geraint Johnes, Reza Oskrochi, and Robert Crouchley (2002), "Stochastic frontier estimation of a CES cost function: The case of higher education in Britain." *Economics of Education Review*, 21, 63–71.

Johnes, Geraint, Jill Johnes, Emmanuel Thanassoulis, Pam Lenton, and Ali Emrouznejad (2005), "An exploratory analsysis of the cost structure of higher education in England." UK Departmentfor Education and Skills.

Kelchtermans, Stijn and Frank Verboven (2010), "Participation and study decisions in a public system of higher education." *Journal of Applied Econometrics*, 25, 355–391.

Laffont, Jean-Jacques and Jean Tirole (1986), "Using cost observation to regulate firms." *Journal of Political Economy*, 94, 614–41.

Laffont, Jean-Jacques and Jean Tirole (1993), *A Theory of Incentives in Procurement and Regulation*. MIT Press, Cambridge, Massachusetts.

Leonard, Daniel and Ngo van Long (1992), *Optimal Control Theory and Static Optimization in Economics*. Cambridge University Press, Cambridge, UK.

Long, Bridget Terry (2004), "How have college decisions changed over time? an application of the conditional logistic choice model." *Journal of Econometrics*, 121, 271–296.

MacMillan, Melville L and Wing H Chan (2006), "University efficiency: A comparison and consolidation of results from stochastic and non-stochastic methods." *Education Economics*, 14, 1–30.

Raines, J Patrick and Charles G Leathers (2003), *The Economic Institutions of Higher Education*. Edward Elgar, Cheltenham.

Rothschild, Michael and Lawrence J White (1995), "The analytics of pricing in higher education and other services in which customers are inputs." *Journal of Political Economy*, 103, 573–86.

Tuckman, Howard P, James H Gapinski, and Robert P Hagemann (1977), "Faculty skills and the salary structure in academe: A market perspective." *American Economic Review*, 67, 692–702.

Appendix

Proof of Lemma 1. (i) Differentiate (4), writing $(\cdot)$ for $\left(\Phi^{-1}(t)\right)$:

$$Z_k'(t) = \frac{1}{\phi(\cdot)} - \frac{\frac{\phi'(\cdot)}{\phi(\cdot)}kt}{\phi(\cdot)^2} + \frac{k}{\phi(\cdot)} = \frac{kt}{\phi(\cdot)^2}\left(\frac{\phi(\cdot)}{kt} + \frac{\phi(\cdot)}{t} - \frac{\phi'(\cdot)}{\phi(\cdot)}\right) . \tag{A1}$$

Next expand the derivative of the hazard rate:

$$\frac{d}{dx}\left(\frac{\Phi(x)}{\phi(x)}\right) = 1 - \frac{\Phi(x)\,\phi'(x)}{\phi(x)^2} = \frac{\Phi(x)}{\phi(x)}\left(\frac{\phi(x)}{\Phi(x)} - \frac{\phi'(x)}{\phi(x)}\right) > 0 .$$

Evaluate (A1) at $x = \Phi^{-1}(t)$, and substitute the above into (A1) to obtain:

$$Z_k'(t) = \frac{kt}{\phi(\cdot)}\left(\frac{1}{kt} + \frac{1}{\Phi(\cdot)}\frac{d}{dx}\left(\frac{\Phi(\cdot)}{\phi(\cdot)}\right)\right) .$$

This is positive because of the assumption that the hazard rate is monotonically increasing, and (i) is established. (ii) follows immediately from (4). Consider (iii) next. $Z_{k_1}(t) - Z_{k_0}(t) = \frac{(k_1-k_0)t}{\phi(\Phi^{-1}(t))} > 0$, and

$$\frac{dZ_k'(t)}{dk} = \frac{1}{\phi(\cdot)}\frac{d}{dx}\left(\frac{\Phi(\cdot)}{\phi(\cdot)}\right) > 0 .$$

This ends the proof of the Lemma. ■

Proof of Proposition 1. This is obtained immediately substituting the definition of Z_k, (4) with $k = 1$, into the first order condition (3). ■

Proof of Corollary 1. The first assertion follows from (6), noting that Z_k^{-1} is increasing. The second from the first and $\frac{dp}{dt} = -\frac{1}{\phi(\Phi^{-1}(t))} < 0$. For the third, develop $\frac{dr}{d\theta}$:

$$\frac{dr}{d\theta} = \left((\Delta - Z_1(t))\frac{\theta}{y} - 1\right)\frac{dt}{d\theta} + \frac{p(t(\theta))\,t(\theta) + g}{y} > 0, \tag{A2}$$

since the first term vanishes by the first order condition on the choice of r, (3). Differentiate the second term to obtain, again using (3):

$$\frac{d^2r}{d\theta^2} = \frac{1}{y}\frac{dt}{d\theta}\left(\frac{dp}{dt}t(\theta) + p(t(\theta))\right) = \frac{1}{\theta}\frac{dt}{d\theta} > 0$$

Finally, for the last two assertions:

$$\frac{dn}{d\theta} = \frac{\frac{dt}{d\theta} + \frac{dr}{d\theta}}{\theta} - \frac{t(\theta) + r(\theta)}{\theta^2} ,$$

$$\frac{d}{d\theta}\left(\frac{yr(\theta)}{\theta}\right) = \frac{y}{\theta^2}\left(\frac{dr}{d\theta}\theta - r(\theta)\right) .$$

Using (A2) and $t(\theta)+r(\theta)=\frac{p(t(\theta))t(\theta)+g}{y}\theta$, they simplify, respectively, to $\frac{1}{\theta}\frac{dt}{d\theta}>0$ and $\frac{yt(\theta)}{\theta^2}>0$. ■

Proof of Corollary 2. The preferred value of t is given by the intersection of the increasing function $Z_1(t)$ with the horizontal line $\Delta-\frac{y}{\theta}$ (see (6)). This intersection occurs for a positive value of t if $Z_1(0)=\Phi^{-1}(0)=a_{\min}<\Delta-\frac{y}{\theta}$. This establishes the Corollary. ■

Proof of Proposition 2. Let the government policy be $\{t(\theta), p(\theta), g(\theta)\}$. By choosing to report type $\hat{\theta}\in[0,\overline{\theta}]$, university of type θ is allowed to set a price for tuition $p(\hat{\theta})$, receives a grant $g(\hat{\theta})$, and is required to teach $t(\hat{\theta})$ students. Given the market salary for its staff, y, it employs:

$$\frac{p(\hat{\theta})t(\hat{\theta})+g(\hat{\theta})}{y} \tag{A3}$$

academics, who will enable it to carry out an amount of research x such that:

$$\frac{p(\hat{\theta})t(\hat{\theta})+g(\hat{\theta})}{y}=\frac{x+t(\hat{\theta})}{\theta}\ .$$

Hence the utility of a university of type θ for reporting $\hat{\theta}$ is

$$\xi(\theta,\hat{\theta})=\frac{p(\hat{\theta})t(\hat{\theta})+g(\hat{\theta})}{y}\theta-t(\hat{\theta})\ .$$

The revelation principle requires that the above is maximised at $\hat{\theta}=\theta$. The first order condition for the choice of $\hat{\theta}$ is:

$$\left.\frac{\partial\xi(\theta,\hat{\theta})}{\partial\hat{\theta}}\right|_{\hat{\theta}=\theta}=\left.\frac{\partial\left(\frac{p(\hat{\theta})t(\hat{\theta})+g(\hat{\theta})}{y}\theta-t(\hat{\theta})\right)}{\partial\hat{\theta}}\right|_{\hat{\theta}=\theta}=0\ , \tag{A4}$$

which gives:

$$\left(p(\theta)t'(\theta)+p'(\theta)t(\theta)+g'(\theta)\right)\frac{\theta}{y}-t'(\theta)=0\ . \tag{A5}$$

Next differentiate $r(\theta)$ given in (11),

$$r'(\theta)=\left[p(\theta)t'(\theta)+p'(\theta)t(\theta)+g'(\theta)\right]\frac{\theta}{y}+\frac{p(\theta)t(\theta)+g(\theta)}{y}-t'(\theta)\ ,$$

and substitute (A5) into it to obtain (12). Now (13): following Laffont and Tirole (1993, p 121), a sufficient condition for a policy to be incentive compatible is that:

$$\frac{\partial^2\xi(\theta,\hat{\theta})}{\partial\theta\partial\hat{\theta}}\geqslant 0\ . \tag{A6}$$

We have

$$\frac{\partial^2 \xi(\theta, \hat{\theta})}{\partial \theta \partial \hat{\theta}} = \frac{\partial \left(\frac{p(\hat{\theta}) t(\hat{\theta}) + g(\hat{\theta})}{y} \right)}{\partial \hat{\theta}} = \frac{\partial \left(\frac{r(\hat{\theta}) + t(\hat{\theta})}{\hat{\theta}} \right)}{\partial \hat{\theta}} = \frac{1}{\hat{\theta}} \left(r'(\hat{\theta}) + t'(\hat{\theta}) - \frac{r(\hat{\theta}) + t(\hat{\theta})}{\hat{\theta}} \right);$$

substitute $\frac{r(\hat{\theta}) + t(\hat{\theta})}{\hat{\theta}} = \frac{p(\hat{\theta}) t(\hat{\theta}) + g(\hat{\theta})}{y} = r'(\hat{\theta})$ from (12), to obtain that (13) is required for the second order condition (A6) to hold. ■

Proof of Proposition 3. Consider local labour market θ. The total pre-tax utility of the potential students is:

$$\int_{a_{\min}}^{\Phi^{-1}(t(\theta))} (y - p(\theta) - a - h) \phi(a) \, \mathrm{d}a + (1 - t(\theta)) (y - \Delta - h) \ ,$$

where the first term is the total utility of the individuals who go to university, and the second the total utility of those who work in the unskilled labour market. Rearrange to obtain (18). Integrating for $\theta > \underline{\theta}$, using the fact that $(1 + \lambda) \int_{\underline{\theta}}^{\bar{\theta}} g(\theta) f(\theta) \, \mathrm{d}\theta = h$ (the total tax paid equals the total value of the subsidies given by the government to the university sector increased by the deadweight loss costs of taxation per unit of tax raised), adding the direct benefit of research, ωR, and rearranging gives (17). ■

Proof of Proposition 4. Begin by constructing the Lagrangean for (17):

$$\begin{aligned} \mathcal{L} = & \left((\Delta - p(\theta)) t(\theta) - \int_{a_{\min}}^{\Phi^{-1}(t(\theta))} a\phi(a) \, \mathrm{d}a - (1 + \lambda) g(\theta) \right) f(\theta) \\ & + \mu(\theta) \delta \frac{p(\theta) t(\theta) + g(\theta)}{y} + \beta(\theta) \left[\frac{y}{\theta} (r(\theta) + t(\theta)) - g(\theta) - p(\theta) t(\theta) \right] \\ & + \tau(\theta) \left[\Delta - \Phi^{-1}(t(\theta)) - p(\theta) \right] + \eta(\theta) t(\theta) + \rho r(\theta) f(\theta) \ , \end{aligned} \tag{A7}$$

where $\beta(\theta)$, $\tau(\theta)$, $\eta(\theta)$, are the Lagrange multipliers for constraints (11), (9), (14), respectively and $\mu(\theta)$ and ρ are the Pontagryin multipliers for the state variables in constraints (12) and (15), respectively. To simplify the analysis of the perfect information case, we have multiplied the incentive compatibility constraint (12) by an indicator $\delta \in \{0, 1\}$, with $\delta = 1$ for the imperfect information case, and $\delta = 0$ for the case in which the government can costlessly observe the type of the university and therefore is not subject to (12) and (13). The first order conditions are:

$$-\frac{\partial \mathcal{L}}{\partial r(\theta)} = \delta\mu'(\theta) = -\beta(\theta)\frac{y}{\theta} - \rho f(\theta) \ ; \tag{A8}$$

$$\frac{\partial \mathcal{L}}{\partial g(\theta)} = -(1+\lambda) f(\theta) + \frac{\delta\mu(\theta)}{y} - \beta(\theta) = 0 \ ; \tag{A9}$$

$$\frac{\partial \mathcal{L}}{\partial p(\theta)} = -t(\theta) f(\theta) + \frac{\delta\mu(\theta) t(\theta)}{y} - \beta(\theta) t(\theta) - \tau(\theta) = 0 \ ; \tag{A10}$$

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial t(\theta)} &= \left[\Delta - p(\theta) - \Phi^{-1}(t(\theta))\right] f(\theta) + \frac{\delta\mu(\theta) p(\theta)}{y} \\ &\quad + \beta(\theta)\left(\frac{y}{\theta} - p(\theta)\right) - \frac{\tau(\theta)}{\phi(\Phi^{-1}(t(\theta)))} + \eta(\theta) = 0 \ ,\end{aligned} \tag{A11}$$

and, for R and $\underline{\theta}$ (Leonard and van Long, 1992, Theorem 7.11.1, p 255):

$$\rho = \omega + \int_{\underline{\theta}}^{\bar{\theta}} \frac{\partial \mathcal{L}(\theta)}{\partial R} \mathrm{d}\theta = \omega, \tag{A12}$$

$$\mathcal{L}(\underline{\theta}) = 0 \ . \tag{A13}$$

Derive $\beta(\theta)$ from (A9):

$$\beta(\theta) = \frac{\delta\mu(\theta)}{y} - (1+\lambda) f(\theta) \ , \tag{A14}$$

and substitute it into (A8), using (A12), $\rho = \omega$:

$$\delta\mu'(\theta) = -\frac{\delta\mu(\theta)}{\theta} + (1+\lambda) y \frac{f(\theta)}{\theta} - \omega f(\theta) \ .$$

When $\delta = 1$, the two differential equations:

$$\begin{aligned}\mu'(\theta) &= -\frac{\mu(\theta)}{\theta} + (1+\lambda) y \frac{f(\theta)}{\theta} - \omega f(\theta) \qquad && \mu(\underline{\theta}) \text{ free} \quad \mu(\bar{\theta}) = 0 \ ; \\ r'(\theta) &= \frac{p(\theta) t(\theta) + g(\theta)}{y} && r(\bar{\theta}) \text{ free} \quad r(\underline{\theta}) = 0,\end{aligned}$$

determine the state variable $r(\theta)$ and the multiplier $\mu(\theta)$:

$$\mu(\theta) = \frac{\rho \int_{\theta}^{\bar{\theta}} \tilde{\theta} f(\tilde{\theta}) \mathrm{d}\tilde{\theta} - (1 - F(\theta))(1+\lambda) y}{\theta} \ . \tag{A15}$$

Next substitute (A14) into (A10), to obtain:

$$\tau(\theta) = \lambda t(\theta) f(\theta) \ . \tag{A16}$$

(A16) implies that $\tau(\theta) > 0$ if $\lambda > 0$ and $t(\theta) > 0$, and so (9) holds as an equality: $p(\theta) = \Delta - \Phi^{-1}(t(\theta))$. Substitute this, $\beta(\theta)$ from (A14), $\tau(\theta)$ from (A16) and $\eta(\theta) = 0$ (because $t(\theta) > 0$) into (A11) and re-arrange:

$$\frac{\partial \mathcal{L}}{\partial t(\theta)} = \frac{\delta \mu(\theta) p(\theta)}{y} + \left(\frac{\delta \mu(\theta)}{y} - (1+\lambda) f(\theta)\right) \frac{y}{\theta} - \left(\frac{\delta \mu(\theta)}{y} - (1+\lambda) f(\theta)\right) p(\theta) - \frac{\lambda t(\theta) f(\theta)}{\phi(\Phi^{-1}(t(\theta)))} = 0 \ ,$$

that is

$$\Phi^{-1}(t(\theta)) = \frac{\delta \mu(\theta)}{(1+\lambda) f(\theta) \theta} - \frac{y}{\theta} + \Delta - \frac{\frac{\lambda}{1+\lambda} t(\theta)}{\phi(\Phi^{-1}(t(\theta)))} \ . \tag{A17}$$

Substitute $\mu(\theta)$ from (A15) to obtain:

$$Z_{\frac{\lambda}{1+\lambda}}(t(\theta)) = \Delta - \frac{y}{\theta} + \delta \frac{\omega \int_{\theta}^{\bar{\theta}} \tilde{\theta} f(\tilde{\theta}) \mathrm{d}\tilde{\theta} - (1 - F(\theta))(1+\lambda) y}{(1+\lambda) f(\theta) \theta^2} \ . \tag{A18}$$

From the definition of $\sigma(\theta)$, (20) is obtained. Note that $\sigma(\theta)$ is increasing: write $\sigma(\theta)$ as

$$\sigma(\theta) = \Delta - \frac{y}{\theta} + \frac{\theta}{1+\lambda} \mu(\theta),$$

and differentiate, to obtain

$$\sigma'(\theta) = \frac{y}{\theta^2} + \left(y - \frac{\theta}{1+\lambda} \omega\right) f(\theta) > 0.$$

The inequality follows from (16).

Since $\sigma(\theta)$ and $Z_k(t)$ are increasing, (20) implies that $t(\theta)$ also is and so (13) is satisfied. Since $t(\theta)$ is 0 for some $\theta \in [0, \bar{\theta}]$ (or, more precisely, for some $\theta \in [1, \bar{\theta}]$), there is a threshold value of θ, call it $\underline{\theta}$, such that $t(\theta) > 0$ if and only if $\theta > \underline{\theta}$.

Now we want to establish that the lowest θ determined in (23), $\underline{\theta}$, is also the value of θ such that $t(\theta) = 0$ and $t(\theta) > 0$ in a right neighbourhood. Expand (A13). At $\underline{\theta}$, the terms in the square brackets in (A7) and the term $\eta(\underline{\theta}) t(\underline{\theta})$ are all 0 because of the slackness complementarity constraints. Also 0 is the term $\rho r(\underline{\theta}) f(\underline{\theta})$, because $r(\underline{\theta}) = 0$, and so:

$$\mathcal{L}(\underline{\theta}) = \left((\Delta - p(\underline{\theta})) t(\underline{\theta}) - \int_{a_{\min}}^{\Phi^{-1}(t(\underline{\theta}))} a \phi(a) \mathrm{d}a - (1+\lambda) g(\underline{\theta})\right) f(\underline{\theta}) + \delta \mu(\underline{\theta}) \frac{p(\underline{\theta}) t(\underline{\theta}) + g(\underline{\theta})}{y} t(\underline{\theta}) f(\underline{\theta}) = 0 \ .$$

Since $r(\underline{\theta}) = 0$,

$$g(\underline{\theta}) = \left(\frac{y}{\underline{\theta}} - p(\underline{\theta})\right) t(\underline{\theta}) \ ,$$

and so $\mathcal{L}(\underline{\theta}) = 0$ implies:

$$\mathcal{L}(\underline{\theta}) = \Bigg\{ \left[\Delta - p(\underline{\theta}) - (1+\lambda)\left(\tfrac{y}{\underline{\theta}} - p(\underline{\theta})\right) + \tfrac{\delta\mu(\underline{\theta})}{f(\underline{\theta})\underline{\theta}}\right] t(\underline{\theta}) - \int_{a_{\min}}^{\Phi^{-1}(t(\underline{\theta}))} a\phi(a)\,\mathrm{d}a \Bigg\} f(\underline{\theta}) = 0 \ . \tag{A19}$$

Write (A17) (with $\eta(\theta) = 0$) as

$$\frac{\delta\mu(\theta)}{f(\theta)\theta} = \left(Z_{\frac{\lambda}{1+\lambda}}(t(\theta)) + \frac{y}{\theta} - \Delta\right)(1+\lambda) \ ,$$

and (A19) becomes:

$$\mathcal{L}(\underline{\theta}) = \left\{ \left[-\lambda\Phi^{-1}(t(\underline{\theta})) + Z_{\frac{\lambda}{1+\lambda}}(t(\underline{\theta}))(1+\lambda)\right] t(\underline{\theta}) - \int_{a_{\min}}^{\Phi^{-1}(t(\underline{\theta}))} a\phi(a)\,\mathrm{d}a \right\} f(\underline{\theta}) = 0.$$

This is 0 at $t(\underline{\theta}) = 0$; moreover,

$$\frac{\partial\mathcal{L}}{\partial t(\underline{\theta})} = f(\underline{\theta})\left[\frac{-\frac{\lambda}{1+\lambda}t(\theta)}{\phi(\Phi^{-1}(t(\underline{\theta})))} + t(\underline{\theta})\,Z'_{\frac{\lambda}{1+\lambda}}(t(\underline{\theta})) - \Phi^{-1}(t(\underline{\theta})) + Z_{\frac{\lambda}{1+\lambda}}(t(\underline{\theta}))\right] =$$
$$= t(\underline{\theta})\,Z'_{\frac{\lambda}{1+\lambda}}(t(\underline{\theta}))\,f(\underline{\theta}) \, .$$

This is strictly positive for $t(\underline{\theta}) > 0$. Therefore $\mathcal{L}(\underline{\theta}) = 0$ is increasing at $t(\underline{\theta}) = 0$, making 0 the only value of t where $\mathcal{L}(\underline{\theta}) = 0$, and so $t(\underline{\theta}) = 0$. What remains to be established are (21) and (22). The first follows from (9). To derive (22), start from the following:

$$r(\theta) = [p(\theta)\,t(\theta) + g(\theta)]\,\frac{\theta}{y} - t(\theta) = \int_{\underline{\theta}}^{\theta} \frac{p(\tilde{\theta})t(\tilde{\theta}) + g(\tilde{\theta})}{y}\mathrm{d}\tilde{\theta} \ ,$$

and write the second equality as:

$$g(\theta)\,\theta = \int_{\underline{\theta}}^{\theta} g(\tilde{\theta})\mathrm{d}\tilde{\theta} + \int_{\underline{\theta}}^{\theta} p(\tilde{\theta})t(\tilde{\theta})\mathrm{d}\tilde{\theta} + t(\theta)\,y - p(\theta)\,t(\theta)\,\theta \ .$$

Differentiate both sides with respect to θ, and divide by θ:

$$g'(\theta) = \frac{t'(\theta)}{\theta}y - \frac{d\,(p(\theta)\,t(\theta))}{d\theta} \ .$$

Integrate both sides in the above to get

$$g(\theta) = \left(\frac{t(\theta)}{\theta} + \int_{\underline{\theta}}^{\theta} \tilde{\theta}^{-2}t(\tilde{\theta})\mathrm{d}\tilde{\theta}\right) y - p(\theta)\,t(\theta) \ ,$$

and therefore:

$$r(\theta) = \left[p(\theta) t(\theta) + \left(\frac{t(\theta)}{\theta} + \int_{\underline{\theta}}^{\theta} \tilde{\theta}^{-2} t(\tilde{\theta}) \mathrm{d}\tilde{\theta} \right) y - p(\theta) t(\theta) \right] \frac{\theta}{y} - t(\theta) \ ,$$
$$= \theta \int_{\underline{\theta}}^{\theta} \tilde{\theta}^{-2} t(\tilde{\theta}) \mathrm{d}\tilde{\theta} \ ,$$

from which (22) is derived. In addition, it is easy to see that integration of the above gives (24). ■

Proof of Corollary 3. Take a given θ and $\varepsilon > 0$. We have $t'(\theta_t) = \frac{t(\theta+\varepsilon)-t(\theta)}{\varepsilon}$ for some $\theta_t \in [\theta, \theta+\varepsilon]$, from which we can write $\varepsilon = \frac{t(\theta+\varepsilon)-t(\theta)}{t'(\theta_t)}$ and $T_F'(\theta_T) = \frac{T_F(\theta+\varepsilon)-T_F(\theta)}{\varepsilon}$ for some $\theta_T \in [\theta, \theta+\varepsilon]$, that is

$$T_F'(\theta_T) = \frac{T_F(\theta+\varepsilon) - T_F(\theta)}{t(\theta+\varepsilon) - t(\theta)} t'(\theta_t) = \frac{y}{\theta} t'(\theta_T) \ .$$

The second equality is (25). Take the limit $\varepsilon \to 0$, which implies $(\theta_t - \theta_T) \to 0$ and, with a slight abuse of notation, the above is

$$\frac{dT_F(\theta)}{dt(\theta)} = \frac{y}{\theta}.$$

The LHS is decreasing in θ and therefore in t, which shows that the slope of the curve in the north-east quadrant is decreasing, establishing the Corollary. ■

Proof of Proposition 5. Impose $\delta = 0$ in the proof of Proposition 4. This eliminates the constraint given by the information disadvantage of the government. (A9) becomes:

$$-\beta(\theta) - (1+\lambda) f(\theta) = 0 \ . \tag{A20}$$

Because of the possibility that the optimum is a corner solution, (A8) must be replaced by

$$\begin{aligned} r = 0 \quad \text{if} \quad & \frac{\partial \mathcal{L}}{\partial r(\theta)} = (1+\lambda) f(\theta) \frac{y}{\theta} - \rho f(\theta) < 0 \ ; \\ & \frac{\partial \mathcal{L}}{\partial r(\theta)} = (1+\lambda) f(\theta) \frac{y}{\theta} - \rho f(\theta) = 0 \quad \text{for} \quad r \in (0, r_{\max}) \ ; \\ r = r_{\max} \quad \text{if} \quad & \frac{\partial \mathcal{L}}{\partial r(\theta)} = (1+\lambda) f(\theta) \frac{y}{\theta} - \rho f(\theta) > 0 \ . \end{aligned}$$

Using (A20), this implies that

$$r = 0 \text{ if } \theta < \frac{1+\lambda}{\rho} y;$$
$$r = r_{\max} \text{ if } \theta > \frac{1+\lambda}{\rho} y.$$

That is, the solution is "bang-bang". (A10) becomes

$$-t(\theta) f(\theta) - \beta(\theta) t(\theta) - \tau(\theta) = 0 \ ,$$
$$\tau(\theta) = \lambda f(\theta) t(\theta) \ ,$$

as before. (A18) in turn becomes:

$$Z_{\frac{\lambda}{1+\lambda}}(t(\theta)) = \Delta - \frac{y}{\theta} \ ,$$

and the last university active is given by the solution in θ of

$$a_{\min} = \Delta - \frac{y}{\theta} \ .$$

As before, $\rho = \omega$, and the Proposition is obtained. ■