ECONOMIC POLICY

OESTERREICHISCHE NATIONALBANK
EUROSYSTEM

68th Economic Policy Panel Meeting

4-5 October 2018
Vienna

Hosted by the Oesterreichische Nationalbank

Academic Salaries and Public Evaluation of University Research: Evidence from the UK Research Excellence Framework

Gianni De Fraja (University of Nottingham)
Giovanni Facchini (University of Nottingham)
John Gathergood (University of Nottingham)

The organisers would like to thank the Oesterreichische Nationalbank for their support.

Centre for Economic Policy Research
Research Excellence, Policy Relevance
CESifo
SciencesPo.

Academic Salaries and
Public Evaluation of University Research:
Evidence from the UK Research Excellence Framework*

Gianni De Fraja† Giovanni Facchini‡ John Gathergood§

September 9, 2018

Abstract

We study the effects of public evaluation of university research on the pay structures of academic departments. A simple model of university pay determination shows how the pay-performance relationship can be explained by the incentives inherent in the research evaluation process. We then analyse the pay-performance relationship using data on the salary of all UK university full professors, matched to the performance of their departments from the 2014 UK government evaluation of research, the Research Excellence Framework (REF). A cross sectional empirical analysis shows that both average pay level and pay inequality in a department are positively related to performance. It also shows that the pay-performance relationship is driven by a feature of the research evaluation that allows academics to transfer the affiliation of published research across universities. To assess the causal effect of the REF on pay structure, we take advantage of the time dimension of our data and of quasi-experimental variation in the performance of academic departments generated by the research evaluation rules. Our results indicate that higher achieving departments benefit from increased future hiring and higher professorial salaries with the salary benefits of REF performance concentrated among the highest paid professors.

Keywords: Research evaluation, Research funding, Research Excellence Framework.
JEL Codes: D47, H42, I28, L30.

---

*We are grateful to Andrea Ichino, Jacques Maraisse, Paula Stephan, Todd Zangler, three referees of this journal and the audiences at the 2017 RES Conference in Bristol, at the 2017 BRICK Workshop on the Economics of Science (Turin), at the 2017 Barcelona GSE Forum, and at the 2017 AISUK conference for helpful comments.

†University of Nottingham, School of Economics, Università di Roma "Tor Vergata", and C.E.P.R. Email: gianni.defraja@nottingham.ac.uk.

‡University of Nottingham, School of Economics, Università degli Studi di Milano, C.E.P.R. and GEP. Email: giovanni.facchini@nottingham.ac.uk.

§University of Nottingham, School of Economics and Network for Integrated Behavioural Science. Email: john.gathergood@nottingham.ac.uk.

1 Introduction

Many countries, particularly in Europe, have in recent years introduced public evaluation of the scientific research carried out by universities. The outcome of these evaluations affects directly the allocation of public funding to the university sector and is an important input to produce rankings used by students, industry, the media, and other users to measure university quality. In this paper we study the effects of a national research evaluation of academic departments on departmental wage structure using a rich dataset covering all UK universities. Two features of the UK setting make it ideally suited to study this question. First, the UK has a systematic and comprehensive assessment of research provided by the Research Excellence Framework (REF). This determines the government "block" research funding, a significant source of research income for UK institutions. The importance of REF performance is also leveraged by its contribution to the ranking published in several university league tables. As a result, it also affects indirectly other sources of research income, post-graduate student recruitment,[1] and prestige. Second, unlike many other European countries, the size of academic departments and full professors' salaries are not subject to national regulation, other than an agreed minimum salary. Hence universities are free to determine professorial hirings and set professorial pay. Indeed our data exhibits large observed salary differences, with the highest paid professors in some of the elite institutions earning as much as seven times the national minimum. A fundamental question is then how performance in the REF translates to wage structures at the department level.

Our analysis draws upon a dataset comprising details of salaries paid to full professors at all UK departments between 2013 and 2016 and the performance of their department in the REF carried out in 2014. Together, these data allow us to evaluate the association between the department's pay structure and its performance in the REF and the effects of departmental performance on subsequent outcomes. The university sector in the UK is an example of a quasi-market, where individual institutions compete, often fiercely, according to rules designed by the government to mimic the incentive system operating in the private markets. In the latter, the literature (e.g. Nickell and Wadhwani 1990, Nickell et al. 1994, Hildreth and Oswald 1997, Abowd et al. 1999, and Lazear 2000) has long firmly established a positive correlation between firm performance and average pay. Some evidence suggests also that firms with higher within-firm pay *inequality* might improve performance (Grund and Westergaard-Nielsen 2008, Edmans and Gabaix 2015, Mueller et al. 2016). Our first contribution is to analyse whether this also holds true for universities. A subject of amused or heated discussions among academics, there is surprisingly little systematic evidence on this important question. Moreover, the limited existing literature has focused on broad national differences in university pay (Altbach et al. 2012) or the internal labor markets of universities (Oyer 2007, Haeck and Verboven 2012) rather than variation among institutions within a country.

To frame our analysis, we present a theoretical model of university pay determination where universities face a national competitive evaluation of their research, and aim at maximising an aggregate measure of research success. This in turn determines the research funding they will receive

[1] This can be an important source of funding for some departments; income from undergraduate students, on the other hand, is unlikely to be affected by a good REF performance. This so both because fees are capped, and because students choose on the basis of word of mouth information from their teachers, or from newspaper rankings which give a negligible weight to research.

from the government. Research is produced using elastically supplied capital and different kinds of labour, to capture the diverse attributes of the academics employed. Among the various models that can be built to capture the structure of the university sector, and the internal organisations of individual universities, our model allows us to highlight the key role played by the internal wage structure and predicts a positive correlation between the research performance of a department, the average salary of its staff, and inequality in departmental pay.

Consistent with the theoretical model, our empirical analysis uncovers a positive cross-sectional relationship between professorial pay in UK universities and their REF performance. This finding is very robust: it holds when we control for a range of departmental characteristics and for academic discipline and university type fixed effects. It also holds across the whole range of academic disciplines. Interestingly, we find that the pay-performance relationship is weaker, though still statistically significant, in the most well-known research intensive universities, and stronger among those established more recently. We also find a positive relationship between professorial pay inequality within a department, and REF performance at the department level. Unlike for the mean salary, this finding is strongest in the most prestigious universities, and is statistically significant for disciplines in the sciences and engineering, but not in medicine and biology, the social sciences, the arts and humanities.

The cross–sectional results in Section 4 can be interpreted as capturing the steady-state, static equilibrium of the theoretical model. That is, they suggest how a pay-performance relationship, exogenously given technology, and individual departments' idiosyncratic productivity combine to determine the equilibrium values of both the department's pay structure, and its research performance. As such they cannot establish causality. In fact the relationship between a department's wage structure and its research evaluation performance may of course be two-way: while better performing departments may obtain greater resources to increase pay and hirings, the causal link could also run in the opposite direction, with more resources improving performance. To address this identification problem we exploit quasi-experimental variation in REF performance arising from a feature of the research evaluation mechanism, which results in two departments with the same average research quality receiving differential levels of funding allocations (and vice–versa in two departments with the same level of funding being characterized by different average research qualities). These differences can be interpreted as perturbations from the pre-REF steady-state, and their highly unpredictable nature likens them to a random allocation of additional funding to some departments, in preference to others, observationally very similar. Using this identification strategy, explained in greater detail in Section 5, we estimate the causal effects of a better research performance on the *subsequent* wage structure of the department.

Our findings show that departments which obtain an "above expectation" REF result see faster subsequent growth in wages and professorial headcount compared with less performing, comparable departments. Importantly, this effect can be observed only when performance is measured by GPA, but not when it is measured through research funding allocations. One plausible interpretation of these results is that the REF affects departments primarily through establishing an objective measure of reputation (GPA score) and not as much through funding. Thus the vaunted government policy of rewarding excellence, by steering funding heavily where "world leading" research is carried out is partially reversed by the universities' central administrations, which appear to allocate more senior posts according the simpler GPA measure. In other words, the research evaluation exercise seems to create a measure of research quality, and departments which perform strongly in terms of research quality are rewarded with higher wages and new senior posts.

Both the theory model and the empirical analysis link average pay and research performance at the *department* level. As a result, the focus our study differs from that in the relatively more established literature linking *individual's* compensation and their research productivity, where great attention has been dedicated to the challenging task of measuring an academic's output. The early work by Diamond (1986) uses citations as an indicator of a researcher's impact, and finds that the marginal effect of an additional citation on individual income is positive. Other contributions distinguish between the number of citations, used as a proxy for "quality", and the measure of "quantity" given by the number of papers published. Most analyses study a small sample of departments (e.g. Hamermesh et al. 1982, Moore et al. 1998 and Bratsberg et al. 2010). In a recent paper, Hamermesh and Pfann (2012) consider instead the members of a larger group of 43 economics department at public institutions in the United States, and find a positive association between output and salaries. This holds both when output is measured by quality, proxied by citations, or by quantity, the number of papers. As far as we know, Sgroi and Oswald (2013) is the only paper which provides a solid theoretical foundation to the balance between quality and quantity. The paucity of information on individual pay has constrained the analysis of the previous incarnations of the REF.[2]

A small recent strand of this literature studies the determinants of individuals' research output in continental Europe: among these, Bosquet and Combes (2017), Zinovyeva and Bagues (2010, 2015) and Checchi et al. (2014) in France, Spain, and Italy, respectively. The first of these studies shows that the characteristics of colleagues matter for research, while the last two focus on the link between research performance and the chances of promotion. Kwiek's recent contribution (2017), on the other hand, shows how in continental Europe, salary increases are associated with increases in administrative and managerial duties.

The rest of the paper proceeds as follows. To frame our analysis, a simple theoretical model of resource allocation within universities that we use to interpret our results is presented in Section 2. The main features of the REF and the data used in the analysis are described in Section 3. Our empirical results are presented in Sections 4, 5 and 6. Section 7 concludes. Additional results and more information on the UK university sector are available in the Appendix.

[2]Early comprehensive studies (e.g. Johnes et al. 1993, Taylor 1995, Sharp and Coleman 2005) have emphasised the role played by systematic biases in the panels' quality assessment, based on characteristics of the institutions: new universities vs. more established ones, institutions based in England vs. those based in other parts of the country, units of assessment that had a panel member vs. those which did not, and so on. Controlling for the quality of the submission in the 1996 and 2001 assessments of economics and econometrics departments, Clerides et al. (2011) do not find systematic evidence of biases in favour of specific institutions. The exception is membership in the assessment panel, which has a positive and significant impact on the ranking of the department in the 1996 exercise. This is in line, as well as with this paper, with Butler and McAllister's (2009) study of the evaluation of the political science panel in the 2001 exercise. The important role played by the panel composition on the evaluation process of academics has been emphasised also by Zinovyeva and Bagues (2015) for the case of Spain.

2 A Model of Research Evaluation and University Competition

Boxes 1 and 2 present a simple theoretical model which illustrates how the incentives created by the research evaluation exercise may shape the pay structure within a university's academic departments. While other models of the university sector are of course possible, we build one that allows us to focus on the the effects of the competition induced by the REF, and so we concentrate on the production of research, abstracting, in particular, from explicitly including teaching. Any constraint imposed by teaching, such as the requirement to recruit a given number of students, is implicitly captured in the production function or in the budget constraint.

In our simple model, research output is generated by different units within a university, the departments. All units employ capital and two types of academics, and they may differ in their "technology", captured by the extent of the complementarity between the three inputs, the two types of labour and capital. Universities allocate their budget across different departments in order to maximise a weighted average of the research outputs of all their departments[3] and receive transfers from the government which depend on the research performance of each unit. Capital is elastically supplied, and we capture the scarcity of academic labour, with a simple reduced form supply function, inelastic for both types of labour, but, naturally, more so for more skilled academics. With this model, we determine the wage structure of each department and the allocation of funds across departments in each state financed university.

We begin, in Box 1, by studying the problem faced by a department which needs to allocate the budget it receives from the central administration of the university. Lemma 1 in Box 1 indicates that, in the long run, the industry equilibrium is such that the amount of both capital and labour employed by a department increase with the budget allocated to it, whereas the amount of labour (capital) employed declines (increases) with the importance of capital in the production process. This latter parameter influences research output in unexpected ways: small departments become smaller still as capital intensity increases, whereas large ones instead increase further in size. This tallies with the anecdotal observation that capital intensive departments tends to be large.[4]

Expression (9) in Lemma 2 is standard for models where labour supply is inelastic, and expression (10) shows that the mean salary and the dispersion of salaries within a department are collinear, and Proposition 1 shows that they both increase with the budget allocated to the department, and with the importance of labour in the production process of research. Intuitively the mean increases because of the realistic assumption that the labour supply is not perfectly elastic. As labour becomes more valuable, for example because the university allocates more funding to the department, or because capital becomes less important, the department will want to hire more (academics of both types), and supply being inelastic, will need to pay more for it.

[3] While a large body of literature emphasises the role played by conflicts of interest within large institutions (Milgrom and Roberts, 1992), in the case of universities it is plausible to assume that individual academics and the heads of both universities and departments all share the same goal with regards to research, namely the maximisation of its quality. For this reason we posit a complete information setting.

[4] This follows from the fact that the sign of the derivative of output with respect to the parameter β_i is the same as the sign of $\ln \frac{2\beta_i B_i}{c_i r}$: therefore it is negative when the budget is low, but it becomes positive for a large enough budget.

BOX 1 The department optimisation problem

We model the higher education sector as a quasi-market comprising K universities, indexed by $k = 1, \ldots, K$. They aim to maximise an aggregate measure of their research output in the n academic disciplines, indexed by $i = 1, \ldots, n$.

Research is produced using three inputs: capital and two types of labour, for example, good professors and superstar academics. Let w_ℓ be the salary of an academic of type ℓ and assume that the supply of this type of professors is given by [a]

$$L_\ell = \mu_\ell w_\ell, \qquad \ell = 1, 2, \tag{1}$$

where L_ℓ is the amount of labour of type ℓ, $\ell = 1, 2$. The parameter μ_ℓ in the supply function of the two types of labour captures plausibly different job market opportunities for the two types of academics, which depend on their research potential. The research output[b] of university k in discipline i, $k = 1, \ldots, K$, $i = 1, \ldots, n$, is denoted by $\rho_{k,i}$, and obeys a Cobb-Douglas technology, a simplified special case of the functions typically used in empirical analyses of universities' production function (for example Thanassoulis et al. (2011)):

$$\rho_{k,i} = \theta_k L_1^{\alpha_1} L_2^{\alpha_2} K^{\beta_i}, \tag{2}$$

where K is the amount of capital, given by labs, equipment, technical personnel, and so on: this can be purchased in a competitive market at a price r.

The parameters θ_k and β_i correspond to fixed effects in our empirical specification. They characterise respectively the overall research productivity of an institution, due for example to different research environments and international connections,[c] and the importance of capital in a given discipline, which depends on factors such as laboratory costs and the like.

If a given department i receives a fixed budget B_i from university k central administration, its budget constraint is given by:

$$rK + w_1 L_1 + w_2 L_2 = B_i. \tag{3}$$

Thus department i in institution k chooses K, L_1 and L_2 to maximise (2) subject to (1) and (3). To lighten notation, let

[a] We therefore ignore any oligopsonistic interaction among institutions: taking them into account would change the absolute levels of academic employment and salaries, but would not alter their *relative* values across institutions and disciplines, which is the focus of our paper.

[b] We do not specify how research output is measured. It could be one of the REF measures considered below, (19) or (20), but the model could be applied to a world without REF, and research output is the less mechanically defined prestige and reputation that is fed and maintained by prizes, accolades, publications, policy influence, and any distinction that enhances academic esteem.

[c] We take θ_k to be exogenously given: it may depend on reputation or history, and in particular, it is not affected by changes in the quality of other departments. Thus our analysis is based on the idea that the correlation between the quality of the various departments in a given university is not a necessary consequence of technological spillovers, but may be caused by an unobserved factor, common to all departments. A similar set-up emerges if θ_k is interpreted as a measure of the cost of doing research, and if the plausible assumption is made that academics are willing to trade-off a university's prestige and overall research environment for a lower salary (see De Fraja and Valbonesi (2012), or De Fraja (2016)). If this is the case a prestigious university would find it easier to hire and retain high quality academics and for this reason enjoy a higher productivity.

$$A_i = \alpha_1^{\frac{\alpha_1}{2}} \mu_1^{\frac{\alpha_1}{2}} \alpha_2^{\frac{\alpha_2}{2}} \mu_2^{\frac{\alpha_2}{2}} \left(\frac{2\beta_i}{r} \right)^{\beta_i}, \tag{4}$$

$$c_i = \alpha_1 + \alpha_2 + 2\beta_i. \tag{5}$$

We can now establish the following:

Lemma 1. *The solution of the maximisation problem of department i in institution k satisfies:*

$$L_\ell = \sqrt{\frac{\alpha_\ell \mu_\ell}{c_i} B_i}, \qquad \ell = 1, 2, \tag{6}$$

$$K = \frac{2\beta_i}{c_i} \frac{B_i}{r}, \tag{7}$$

and the research output is given by

$$\rho^*_{k,i}(B_i) = \theta_k A_i \left(\frac{B_i}{c_i} \right)^{\frac{c_i}{2}}. \tag{8}$$

The wage structure of the department is instead characterized in the following:

Lemma 2. *At the solution of department i's maximisation problem the salaries for the two types of academics are given by:*

$$w_\ell = \sqrt{\frac{\alpha_\ell B_i}{c_i \mu_\ell}}, \qquad \ell = 1, 2. \tag{9}$$

Therefore the mean salary and its standard deviation are given by

$$\bar{w} = \frac{\alpha_1 + \alpha_2}{\sqrt{\alpha_1 \mu_1} + \sqrt{\alpha_2 \mu_2}} \sqrt{\frac{B_i}{c_i}}, \qquad \sigma_w = \frac{\left| \sqrt{\frac{\alpha_1}{\mu_1}} - \sqrt{\frac{\alpha_2}{\mu_2}} \right|}{\sqrt{\alpha_1 \mu_1} + \sqrt{\alpha_2 \mu_2}} \sqrt[4]{\alpha_1 \mu_1 \alpha_2 \mu_2} \sqrt{\frac{B_i}{c_i}}, \tag{10}$$

and the Gini coefficient of the department members' pay by:

$$G_i = \frac{\sqrt{\alpha_1 \alpha_2} \left(\sqrt{\mu_1 \alpha_2} - \sqrt{\mu_2 \alpha_1} \right)}{(\alpha_1 + \alpha_2) \left(\sqrt{\mu_1 \alpha_1} + \sqrt{\mu_2 \alpha_2} \right)}. \tag{11}$$

The proof of this and the other results can be found in the Appendix. The following is an immediate consequence of the previous result.

Proposition 1. *The derivatives of the mean salary and of the standard deviation of salary in department i in institution k are proportional to the derivative of* $\sqrt{\frac{B_i}{c_i}}$*. This is* $-c_i^{-\frac{3}{2}} B_i^{\frac{1}{2}} < 0$*, when differentiating with respect to* β_i*, and* $\frac{1}{2} c_i^{-\frac{1}{2}} B_i^{-\frac{1}{2}} > 0$ *when differentiating with respect to* B_i*.*

The increase in the standard deviation requires the additional hypothesis that the ratio $\frac{\alpha_i}{\mu_i}$ is different for the two types of labour, so that the $\left|\sqrt{\frac{\alpha_1}{\mu_1}} - \sqrt{\frac{\alpha_2}{\mu_2}}\right|$ in (10) is non-zero.[5] As long as the two types of labor have different efficiency adjusted costs, as the size of the department increases, so does the wedge between the total compensation of the two groups of workers, and hence the measured standard deviation of departmental salaries. While the exact values of individual wages, aggregate wages and of the Gini coefficient in (9)-(11) does depend on the assumption of a Cobb-Douglas technology, the informal discussion in this paragraph suggests that the relationships highlighted in Proposition 1 will hold more generally, when at least some workers are paid above their marginal productivity.[6]

Given that, *ceteris paribus*, a department's research output increases with its budget, the model predicts a positive association in the long run between average salary and research output and between inequality in salary and research output, which is consistent with the main findings of our empirical analysis. Conversely, the mean and the standard deviation of the department salaries varies in the opposite direction to capital intensity. The intuition for this effect is the same as for size: an increase in β reduces the importance of labour in production, so for a given budget fewer workers will be employed.[7]

Corollary 1 in Box 2 implies that, in the steady state, universities with a higher ability to do research, measured by θ_k in the model, will be able to devote more resources to all their departments, which will also produce higher output. In other words, in equilibrium, research output and the pay structure of each department in every institution are simultaneously determined by the common technology, captured by the parameters α_1, α_2, and β_i, with a ranking of institutions determined instead by the unobservable idiosyncratic parameter θ_k. This ranking suggests that some institutions will tend to perform better in all disciplines, and pay their professors more.[8] As a result, our model can explain the empirical pattern that some groups of universities tend to perform better in the research evaluation across (nearly) all disciplines, – for example the Russell Group in the UK context.

[5] Note that it is theoretically conceivable that the lower productivity workers are paid *more*. For this to happen, however, their supply must be sufficiently more inelastic than that of higher productivity workers to compensate for their lower productivity. This seems unrealistic though, as in all plausible situations the supply of superstar professors is likely to be less elastic than that of the good professors.

[6] Given the simplified Cobb-Douglas technology we consider here, the Gini coefficient (11) is independent of the department budget and hence size, but depends only on the relative productivities and the relative elasticities of the two types of academics. This result would not hold in a richer model.

[7] To the extent that STEM subject are more capital intensive than social sciences and humanities, this constitutes a testable implication of this model.

[8] The model in the box, with a Cobb-Douglas production function, and the condition on the productivity of capital, rules out corner solutions where some universities close down some departments. A more flexible production function, and explicit inclusion of general equilibrium effects might generate specialisation of different universities in different disciplines.

BOX 2 The university maximisation problem

We now consider the university's allocation problem. We make the following assumptions regarding the objective function and the resources a university has at its disposal.

Assumption 1. *The objective function of university k is*

$$U_k = \sum_{i=1}^{n} u_i \rho^*_{k,i}\left(B_i\right), \qquad k = 1, \ldots, K. \tag{12}$$

That is, university k aims at maximising the total weighted research output of all its departments, with exogenously given weights, u_i. This is a catch-all simplifying assumption to capture the idea that universities care about research success.[a]

The next assumption establishes a link between research success and the overall budget made available by the funding agency to university k, which is denoted by $\bar{B}_k$. While these budgets are in practice allocated each year on the basis of past success, we can think of the simultaneous set-up presented here as the steady state.

Assumption 2. *The overall budget allocated by the government to university k is*

$$\bar{B}_k = \sum_{i=1}^{n} \gamma_i \rho^*_{k,i}\left(B_i\right). \tag{13}$$

The weights γ_i are exogenously given, fixed by the government agency in charge of university funding. A linear formulation is a very natural starting point for the analysis, and was used in the 2014 REF (see (20) in Box 3 below), when the government rewarded excellence by skewing the measure of performance strongly towards high quality outputs, but the sum of the funding of two departments would not be altered by their merging. This seems a desirable property. Incorporating external sources of revenues, such as sponsorships, grant funding, income from patents or donations from alumni, would not alter the analytical set-up, as all these are positively related to prestige. Note furthermore that the funding weights γ_i depend on institutional differences, and will not in general be proportional to the utility weights in (12), u_i.[b]

[a] A more complex model could modify (12) replacing the weighted sum with a more general function of the research performance to fit better the details of the REF, for example the GPA score and the funding formula in (19) and (20) in Box 3 in Section 3.1. The idea of (12) is that the university's management aims at maximising overall prestige, given by a weighted average of the prestige of its activities, and that funding raised in any way, including the government research allocation, is devoted to enhance research prowess. The simple formulation in (12) conveys the main idea of the model. It could be extended, with no conceptually important changes, for example, by making the payoff depending on an institution's rank in each discipline, rather than the level of its output, or including an exponent for the output. The latter would capture an institution's preference for equality or inequality, according to whether the exponent is smaller or greater than 1.

[b] Note also that in the special case where the ratio between u_i and γ_i is constant in i, that is when the relative "prestige" of any two disciplines equals their relative funding, the Lagrange multiplier disappears from the budget allocation (16). That is, in this case, and only in this case, all departments in a given university grow and shrink proportionally according to its funding.

Recall definitions (4) and (5) to write university k's problem as:

$$\max_{\{B_i\}_{i=1}^n} \sum_{i=1}^{n} u_i \theta_k A_i c_i^{-\frac{c_i}{2}} B_i^{\frac{c_i}{2}} \tag{14}$$

$$\text{s.t.: } \sum_{i=1}^{n} B_i = \sum_{i=1}^{n} \gamma_i \theta_k A_i c_i^{-\frac{c_i}{2}} B_i^{\frac{c_i}{2}}. \tag{15}$$

We can now determine the allocation of funds to the departments.

Corollary 1. *Let $\beta_i < 1 - \frac{\alpha_1+\alpha_2}{2}$. Then there exists a $\lambda_k > 0$, such that the solution of university k's problem is given by:*

$$B_i = c_i \left[\frac{(u_i + \lambda_k \gamma_i)\, A_i \theta_k}{2\lambda_k} \right]^{1-\frac{c_i}{2}}, \qquad i = 1, \ldots, n. \tag{16}$$

The condition in the statement of the corollary, ensures that all departments receive a positive share of the total funds, and it avoids the need to consider corner solutions, where some departments are shut down. Finally, note that to close the model, (16) is substituted into (15) to obtain λ_k as a function of the β_i's and θ_k, and the other parameters, which are constant across disciplines and institutions. Writing this as $\lambda(\theta_k; \boldsymbol{\beta})$, where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$, we can determine the research output of each discipline as a function of the exogenous parameters:[a]

$$\rho_{k,i}^* = \theta_k A_i c_i^{\frac{c_i}{2}} \left[\frac{\left(\frac{u_i}{\lambda(\theta_k;\boldsymbol{\beta})} + \gamma_i \right) A_i \theta_k}{2} \right]^{c_i\left(1-\frac{c_i}{2}\right)}. \tag{17}$$

[a]Note that it is not practical to obtain explicit expressions for $\rho_{k,i}^*$, as it is highly non-linear in the parameters. For example, an increase in the capital-intensity of a discipline, measured by β_i, first increases the research performance then decreases it, due to the increase in cost and the beneficial effect of diverting resources to other "less expensive" disciplines.

3 Data

Our dataset combines public information on the submissions and results from the REF, available on the REF 2014 website, with repeated cross-section data containing information on the characteristics of all full professors in UK academic departments. With these data, we can identify the characteristics of the professorial wage structure at the point of the REF exercise (October 2013) and then track the evolution of the wage structure over subsequent years.[9] In this section we start by presenting the institutional environment of the REF, we discuss next the sample construction, and we report some summary statistics.

[9]Note though that the data do not contain individual level identifiers and as a result we cannot construct an individual professor level panel.

3.1 Research Excellence Framework (REF) Outcome Data

The REF 2014 was a government run evaluation to assess the quality of research in UK higher education institutions.[10] As well as ensuring accountability for public investment in research and producing evidence of its benefits, the assessment informs the selective allocation of the annual "block" budget for research to institutions. This funding is the so-called QR (quality related) allocation, and, unlike the funds distributed by the research councils, which pay for specific projects, universities are free to choose how to allocate them across projects, and indeed disciplines. The funding allocated on the basis of REF results is approximately one quarter of all taxpayer money awarded to higher education institutions.[11]

Do UK universities place greater emphasis on their GPA or funding scores? Institutions are not required to submit all their academics; instead they may choose whom to submit for assessment[12]. The presence of N_i in (20), but not in (19), thus creates for them an important trade-off. GPA, for its immediacy and simplicity is a good measure of prestige, used in many league tables. If institutions only care about GPA, then they should submit very few researchers, in the limit only their very best ones. This however would harm their funding, which is proportional to the number of staff submitted for assessment. While we report results for funding scores in the main text we repeat all our estimations using GPA as the dependent variable and report them in the Appendix.

It is important to note that the funding determined by formula (20) in Box 3 is calculated by department, but allocated to the university and that the outcome of the REF exercise determines the funding received by the university in every year until the subsequent assessment is carried out. That is while the contribution to the university budget of each of its departments can be measured down to the last penny, universities are *not* required to allocate these funds to the departments responsible for obtaining them.

The three components of the quality indicator generate different incentives in recruitment and retention of academics. These differences can be used to understand how does the "transfer market" for academics operate, given the different incentives provided to departments to hire academics with an excellent publication record or who have carried out research with very strong impact (see Box 3). As we explain, the value to the institution of an academic's outputs travels with her, but the value of her contribution to the environment and especially her impact does not. This suggests that, when hiring or responding to outside offers prior to the REF census date, institutions should value more a researcher with a stellar publication record, even though she has no demonstrable impact outside academia, than a researcher whose less prestigious recent publications can however be shown to have influenced a certain Act of Parliament, an EU directive, or industry practice.

[10] Similar exercises have been carried out at regular intervals since 1992, with early runs in 1986 and 1989, as explained on the REF website.

[11] Detailed information of how public funds are allocated to UK universities can be found at www.hesa.ac.uk/stats-finance. The full set of REF rules, the identity of the reviewers, and the outcomes are all available at www.ref.ac.uk.

[12] For the next exercise, REF 2021, this element is being removed. Departments will be required to submit all of their professors with responsibility for research

BOX 3: The 2014 Research Excellence Framework

As a result of the evaluation, each academic department is assigned a numerical 'quality' profile which describes the percentage of the department's output, environment and impact rated on a 5-point "star" scale from 4* to 0*, where 4* is defined as "Quality that is world-leading in terms of originality, significance and rigour" and 0* is "Quality that falls below the standard of nationally recognised work."[a] The profiles of the three components (output, impact and environment) are aggregated into a single quality profile, given by a weighted average of the three components, with weights 65%, 20%, and 15% for the three components.

Formally, let π_{ik}^{s} be the proportion of the submission department i' in university k judged to be of quality $s^{\star}$, with $s = 0, 1, \ldots, 4$. π_{ik}^{s} is then given by

$$\pi_{ik}^{s} = 0.65 O_{ik}^{s} + 0.2 I_{ik}^{s} + 0.15 E_{ik}^{s} \qquad s = 0, 1, \ldots, 4, \tag{18}$$

where O_{ik}^{s}, I_{ik}^{s}, and E_{ik}^{s} are the shares of department i in university k's research output, impact and environment which has been attributed a grade $s^{\star}$ by the panel. Clearly $\sum_{s=0}^{4} O_{ik}^{s} = 1$, and similarly for I_{ik}^{s} and E_{ik}^{s} and hence for π_{ik}^{s}.

The REF involves peer-review assessment by 36 subject-specific expert panels of the "reach and significance" of the research carried out by the academics submitted for assessment. The 36 panels are grouped into four "Main Panels", corresponding to very broad disciplinary areas: medicine and biology, the other sciences and engineering, the social sciences, and the arts and humanities. Universities are not obliged to submit all their departments for evaluation, nor are they compelled to submit all the academic members of each department submitted. By not doing so though they might forgo part of their funding, which is based on a formula weakly increasing in the number of academics submitted. This is an important difference with the previous exercises, where the funding was proportional to the product of the number of FTE staff submitted and the average quality of their research: thus submitting an additional, weak, researcher could have lowered the department average and hence the funding as well as the prestige.[b]

As already mentioned, panels assess academic departments in three areas: research output, environment, and impact. Output is assessed through the evaluation of scholarly work (such as books or journal articles), published in the period since the previous assessment. Each academic submitted is required to submit four different items.[c] An academic's outputs are attributed to the institution where she is employed on 31 October 2013, the REF census date, even if they were produced while the faculty member was employed by a different institution.

[a] A tongue-in-cheek analysis of the accuracy of these definitions is carried out in the intriguing paper by Règibeau and Rockett (2016).

[b] The change to the funding formula for the 2014 exercise described in detail in (20) was intended to soften the trade-off, and induce universities to submit all their research staff. Anecdotal evidence suggests however that the desired effect was not achieved, and rules have changed again for the next exercise when all staff involved in research will have to be submitted.

[c] Hamermesh and Pfann (2012) find a negative correlation between the number of citations and the number of papers published by the members of a sample of top US economics departments. Thus the small number of items individuals are required to submit for the REF, might indicate that the UK policy maker preferences are skewed towards the "quality" of research, measured by citations, rather than the sheer publication count.

The environment component is a written submission describing the achievements of the academic department, together with data on research grant income and PhD completions. Similarly, impact is assessed by considering written 'case studies', one for every eight academics submitted, accompanied by supporting evidence which shows how the research of the department has brought benefits *outside of academia* through, for example, influence on government policy or industry practice. Unlike output, impact is attributed to the institution where it was carried out irrespective of which institution is currently employing the researcher responsible for it at the census date.[a] Similarly, it would be hard for a department to argue that someone with a very short tenure could have had the opportunity to affect its research "environment". That is, the "output" component is transferable, while the "environment" and "impact" components are not.

The quality profiles of individual departments presented in equation (18) are typically used to construct two indicators. The first is the grade point average, GPA, which has a limited official status, but is widely used by the media and in the public discourse to rank departments in national league tables. GPA is calculated as a weighted average of the scores, with the proportion in each category as weight: department i's GPA is calculated simply as:

$$GPA_i = \sum_{s=0}^{4} \pi_i^s s. \tag{19}$$

The second indicator is a funding score formula, FS, which is used by the government as the basis to determine research funding allocations. This formula is intended to provide incentives towards high quality research, and so it gives high weight to $4^\star$ output, and no weight to output judged less than $3^\star$.[b] With the above notation, and N_i denoting the number of full–time equivalent researchers submitted by institution i, its yearly funding until the following evaluation exercise is given by

$$FS_i = \Phi_t \Gamma_i \left(4\pi_i^4 + \pi_i^3 \right) N_i, \tag{20}$$

where Φ_t is the QR unit funding, determined every year t depending on the overall public funding for universities, and Γ_i is a subject specific weight which takes the value 1.6 for STEM subjects, 1.3 for intermediate cost subjects such as geography, architecture, sport sciences, design, music and 1 for all other subjects.

[a] To use a fictitious example, suppose Professor Lapping publishes important papers while he is employed by Poppleton University. He then moves to Porterhouse College *before* the REF census date. Then his publications will be included in the "output" submission of Porterhouse College and in the "impact" submission of Poppleton University.

[b] While institutions did not know the exact details of the formula, which were determined after the publication of the results (Else 2015), they knew the principles which would underpin it.

3.2 Sample Construction

To construct our sample we combine data on academic department–level results from the REF 2014 exercise with repeated cross-sections of salary data containing details of professorial pay within departments. The salary data is at the individual level as a repeated cross-section dataset provided by the UK Higher Education Statistical Agency (HESA) for the years 2013-14, 2014-15 and 2015-16. It covers the universe of all individuals employed with the academic rank of full professor by a higher education institution.[13] This allows us to observe pay within departments at the point of the REF census date (October 2013) and track it subsequently over the following two years; however, as we do not have an individual identifier we cannot construct an individual-level panel. In total there are approximately 17,000 full-time equivalent professorial positions in the UK which are filled by approximately 19,000 individuals, some of whom work part-time. HESA matches each individual to one of the 36 REF panels, and therefore we use the composition of departments determined by this match, even if there may be instances in which an individual in a given department is submitted to a different panel for assessment (for example, an economics member of staff might have been submitted to the management panel). Information about the average pay and its dispersion within a department is obtained from this data, which also reports details of the age structure of departments' professoriate: for each department, we know the fraction of professors whose age falls in each ten-year band.

From these matched data we calculate various departmental level characteristics of the wage structure, such as mean wages and measures of wage inequality. When constructing these measures we make some small sample restrictions. First, we include only professors paid a full-time equivalent of more than a threshold value of £50,500 in 2013.[14] To reduce the possibility of identifying individuals, the sample is limited to units with more than three full-time equivalent professors, and we exclude units which were not submitted to the REF. We also omit the only department of the London Business School, which has very low reliance on government funding and is an extreme outlier, in regards of salaries, as it pays on average considerably more than the national average. Together these restrictions reduce our sample from approximately 17,000 full-time equivalent full professors to 16,300. The final piece of information we add is the total remuneration of the universities' heads, typically known as Vice Chancellors, which is published every year in the Times Higher Education newspaper.

Departments are partitioned into the four "Main Panels" defined by the REF exercise. We also divide universities into five groups, according to their institutional characteristics. These are the most established universities, which include the original Russell group, labelled "Russell" – Oxbridge, LSE, and the authors' institution among them; the founding members of the recently disbanded "1994 group", which comprised younger and smaller research-intensive universities, – York, Essex, Queen Mary among them; "New Universities" mostly created from locally controlled vocational institutions;

[13]The UK academic hierarchy comprises three grades. In the past they were confusedly given four names, lecturer, senior lecturer or reader and professor. Nowadays, many institutions are converting to a US style three name hierarchy, with the conventional titles of assistant, associate and full professor. With few exceptions, all institutions pay staff on the first two levels according to nationally agreed, relatively short, scales, and so salary and tenure in the post are highly correlated. Individual negotiation, on the other hand, is the norm for full professors, and for this reason we focus on the pay structure among them only.

[14]This is to account for the fact that in some institutions there are academics who are paid a very low full-time equivalent annual pay, and are employed for a very small fraction of the time (a typical figure is 10%). Our understanding is that some institutions classify as professorial staff collaborators (such as external examiners) who would be considered external payees in other institutions, and whose research cannot be submitted to the REF evaluation. All our results are robust if we include also professors paid less than the threshold.

"specialist" institutions, such as the Royal College of Arts, whose focus is exclusively on a single discipline; and the rest, mostly universities with historically less emphasis on research (such as Hull, Bradford), labelled "Others". A full listing of the groups is provided in the Appendix A.

3.3 Summary Statistics

Summary data on the characteristics of the 1,093 academic departments that comprise our final dataset are reported in Table 1. The average department has approximately 14 full professors (Full-Time Equivalent, FTE), with an average pay at around £74,000. The number of professorial FTE in a department ranges from 3 to over 300 – the latter figure applying to a very large medical school. Average salary ranges from just above £50,000 to just below £130,000. The Gini coefficient for the department's professorial pay is on average low at 8 percent, but varies between zero (completely equal pay) to 36 percent.[15] Notice that, given the minimum pay constraint, the maximum theoretical value for the department with the country's average membership and pay is around 30 percent. Over two thirds of professors fall in the 41-60 age range. The table also includes summary data for the total number of full-time equivalent staff, including non-professors, submitted to the REF, the pay of the university Vice Chancellor, and a dummy indicating whether a member of academic staff from the department sat on the REF peer-review panel. We use these variables as controls in our analysis. Table 2 summarizes the same time-varying variables from Table 1 two years after the REF exercise. In Table 3 we show summary statistics for the changes in these variables across the two waves of data. Table 4 summarises the performance of departments in the REF exercise. A breakdown of average scores across the full quality profile is reported in Table A1 in the Appendix.

The distribution of average departmental salary and of REF funding scores is shown in Figure 1. Departments are grouped according to the REF main panel which evaluated them (on the LHS plots), and by the type of their university (on the RHS plots). The top plots in Figure 1 illustrate a right-tail of high paying departments across panels and university types, with more pronounced skewness in the social sciences and specialist universities. The distributions of funding scores, shown in the bottom plots, are similar across panels, with a higher average among the medicine panel due to the typically large size of medical schools. For this and other reasons, we repeat in the appendix the analysis excluding all the department in the "Clinical Medicine" unit of assessment. Plot D shows variation in performance across university types, suggesting a hierarchical ranking with the Russell group of universities on average the strongest performers, followed by the "1994" group, the "Others" and the "New Universities". Table 5 shows the correlation matrix across the different measures of research performance we consider in our study. Note that the correlations between the GPA score and its components are well below 1, indicating that the panels do judge each component separately.

Figure 2 illustrates the unconditional correlation between our main variables of interest: average departmental pay and funding score. It shows a positive average pay-performance gradient across subject areas and university types. The slopes of the fitted regression lines are similar across main panels, but less so across university types: the fitted line has a lower gradient in the Russell group

[15] Academic pay dispersion has grabbed little attention; exception are studies of inequalities due to sex and race. See for example Porter et al. (2008). In our analysis we use the Gini coefficient as the measure of inequality. Other measures, such as the natural log of the standard deviation of salaries in the unit, or the variance of the natural log of salaries in the unit, produce very similar results. By construction, these alternative measures of inequality are highly correlated with the Gini coefficient. Results are available from the authors on request.

universities and is steepest in the "New Universities".

4 Results: Equilibrium prior to the REF

Universities knew the REF rules well in advance of the submission, and therefore had the opportunity to allocate funding to departments according to their objective function. The relationship between departmental pay structures at the point of the REF evaluation (October 2013) and departmental REF performance presented in this section therefore corresponds loosely to the steady-state of the model introduced in Section 2. The aim is to uncover the association between pay and performance. To this end, we estimate a series of econometric models taking the following general form:

$$RP_{ik} = \beta_0 + \beta_1 AvSalary_{ik} + \beta_2 Gini_{ik} + \beta_3 \mathbf{X}_{ik} + \phi_i + \psi_t + \epsilon_{ik}, \tag{21}$$

where RP_{ik} is a measure of REF performance for the submission made by university k, which is of university type t (i.e. "Russell", "1994 group" etc.), to the panel assessing disciplinary field i. In the main analysis the REF outcome is the natural log of the funding score, and in the appendix we report also results for the determinants of the overall GPA. The results obtained using the two dependent variables are qualitatively very similar (see Figures A1 and A2, Tables A2, and A3 in the Appendix).

$AvSalary_{ik}$ and $Gini_{ik}$ are the average salary in October 2013 of the professoriate in department (i, k), in logs, and inequality in department (i, k), measured by the Gini coefficient of the professors' salaries.[16] The matrix $\mathbf{X}_{ik}$ contains additional controls including the total number of professorial full time equivalents (FTE), the total number of FTE members of staff submitted to the REF (both in logs), an indicator for whether the department had a member of staff serving on the corresponding REF panel, the total remuneration of the university's head (in logs), and the share of individuals in the professoriate who are respectively below 40 years of age, between 41 and 50 years of age, and between 51-60 years of age, with the professors older than 60 as the reference group. In some specifications we also include discipline (ϕ_i, $i = 1, \ldots, 36$) and institution type (ψ_t, $t = 1, \ldots, 5$) fixed effects to account respectively for unobserved characteristics common to all departments in the same subject and to departments in similar institutions. The large number of institutions, and the fact that many of them submitted very few departments, often only one or two, prevents us from including institution fixed effects.

Table 6 presents the main results of this section. Column (1) shows estimates from a specification where regressors are average salary and inequality and where we control for a basic set of covariates. Results indicate that the size of the submission, measured by the *total* number of academic staff, thus including non-professors, improves REF performance.[17] At the same time, the additional effect of submitting professors rather than less senior staff is only weakly statistically significant. Moreover, we find that having a member of staff on the corresponding REF panel has a positive and significant

[16] An alternative measure of pay inequality within a department is given by the variance of log salary, which has a correlation coefficient with the Gini measure of 0.93. Results using this alternative measure are very similar and available upon request.

[17] One feature of the REF is that departments are allowed to hire academics on a part-time basis (subject to a minimum threshold of 0.2 FTE) and include the academic's papers in their REF submission. We cannot perfectly identify the number of part-time staff in the HESA data, which is recoded at the FTE level. For example, two professors in the same age bracket who are both paid the same salary would be recorded in the HESA data a $1 \times$ FTE. However, when we include the proportion of observable part-time staff as a control variable, results are unchanged. See Tables A7 and A8 for details.

effect on the REF funding score. There is also a positive association between REF performance and the university head's total compensation (see Figure A3 in the Appendix for more details).

In Columns (2) and (3) we additionally include unit of assessment and institution type fixed effects. This improves further the fit of the models, from an already very high value of 0.87, to over 90%. In both specifications the average wage and the measure of inequality keep a robust link with REF performance, though their coefficient is lowered in value and inequality is less precisely estimated when we include the institution type fixed effect.

The magnitude of the effects we have uncovered is substantial. The results in our preferred specification, Column (3), indicate that a 10% increase in average salary is associated with a 5.1% increase in the REF funding. Though less precisely estimated, the coefficient for the Gini coefficient of the professorial salary is also sizeable: if the coefficient grows from 8.2 (the sample average) to 9, a 10% increase, the funding score increases by 5.7%.

There is also a strong size effect: a 10% increase in the size of the total REF submission is associated to a 10.5% increase in REF funding: if two identical departments (with all co-variates equal to the sample mean) were to merge, their aggregate REF funding would increase by 6.4%. Note that if departments where solely concerned with total funding, they would submit as many staff as possible, given that – at the census date – an academic's salary is effectively a sunk cost. A positive coefficient might therefore indicate that some departments are constrained in the number of people they submit, as they already submit everyone, or because it is not the case that they care only about funding.[18] On the other hand, the analysis of Section 6 suggests that, indeed, universities are willing to trade-off prestige against funding. The additional effect of a 10% increase in the number of professorial FTE employed, keeping the overall size constant, is a modest 1.1% increase in the REF funding. Having a member of the department on the evaluation panel increases instead the funding score by 6%: arguably a non-negligible effect. Finally, age seems to matter little: while the coefficient for under 40 is significant[19] only few professors are under 40. We find very similar results when GPA score is used as the dependent variable; see Table A2 in the Appendix.

The analysis of the unit of assessment fixed effect coefficients offers us an insight on systematic differences across fields that are not captured by our observables. Figure 3 displays plots of these fixed effects with 95% confidence intervals, taking as baseline the Economics and Econometrics panel. To gain insights on the magnitude of these effects, a department in the discipline with the highest fixed effect (Sports Science or Communications and Media Studies) would receive approximately *twice*[20] the annual funding than an otherwise identical department in the discipline with the lowest estimated fixed effect (Economics and Econometrics).

This lower REF success on average of the Economics and Econometrics UK departments could be due either to a lower "quality" of the average submission in the field, or to a more "demanding" assessment of research by this panel's members, and our data are unable to shed any light on which of these alternative explanation is more likely. Using a methodology which measures quality as the

[18]Direct analysis of the first possibility is not possible as HESA does not collect data on submittable FTE.

[19]The average department has 14.4 members, so replacing an over 40 professor with a younger one increases the number of under 40 professors by 6.94%. Given a coefficient of 0.276, *ceteris paribus* this swap increases the funding score by 1.9%.

[20]In a regression of $\ln Y$ on covariates, if a dummy variable switches from 0 to 1, the percentage impact on Y is $100\,(e^c - 1)$, where c is the estimated coefficient of the dummy variable. See Halvorsen and Palmquist (1980) and Giles (1982) for details.

number of citations attracted by papers published in high quality journals, Oswald (2015) argues however that the quality of UK economics is high.

4.1 Results by Subject Groups and University Types

So far our analysis has highlighted the existence of a strong positive relationship between average professorial wage and REF performance. One important question is whether the shape of this relationship varies across fields. We therefore estimate the specification of Column (1) in Table 6, separately for sub-samples corresponding to the four main REF panels, including university type fixed effects. As Table 7 shows, the effect of average salary is positive and statistically significant for all panels. It is considerably larger in the main panel A (medicine and biology), than in the other subject areas: all the coefficients are pairwise statistically significantly different, except the difference between main panels B and D, whose equality is only weakly rejected, with a p-value of 0.0914.[21]

The difference across main panels conceals some heterogeneity among the disciplines that make up the four groups. Figure 4 plots the coefficient estimates for the average professorial wage, with 95% confidence intervals shown in bars, from the same model specification as in Column (1) of Table 6, run separately in each of the 36 subjects corresponding to the units of assessment. Note the very high coefficient for Clinical Medicine. This, and the fact that a large proportion of the academics employed in these departments work also for, and are separately paid by, the National Health Service, and their salary structure might well be driven by different considerations. Given the large number of professors in this unit of assessment (15.5% of the total), we are re-assured that the results are unaffected when we exclude all the departments submitted to the Clinical Medicine unit of assessment (Table A4 in the Appendix).

As suggested by Panel (D) of Figure 1, the average REF performance differs across university types, with Russell group universities ahead of the other groups. To what extent is this result affected by the shape of the pay–performance relationship? To answer this question, in Table 8 we run the same specification as in Column (3) of Table 6 on four different subgroups of institutions: the "Russell group", the "1994 group", the "New Universities" and the "Others" (we omit specialist universities as they comprise a total of only eight departments). We find that the relationship between average professorial wage and REF performance becomes progressively stronger as we move from the "Russell group" institutions to the "New universities", roughly following the quality of research, as depicted in Figure 2, Panel (B).[22] This result holds true also when we exclude from our analysis Cambridge and Oxford, two institutions that offer substantial non-monetary compensation to many senior academics, for example in the form of subsidised accommodation.[23]

[21] Note that the independent role of the inequality in wages in the overall sample appears to be driven by the disciplines in main panel B, science and engineering and A, medicine. The coefficients for the other main panels are both smaller and not estimated precisely. Furthermore, the effect of having a panel member uncovered in Table 6 is less statistically significant and smaller in value for medicine and biology than for the social sciences and the arts and humanities. Again the GPA score as the measure of research performance yields qualitatively very similar results (Table A3 in the Appendix).

[22] T-tests for the equivalence of means strongly reject equality between any two pairs of the average salary coefficients of Table 8.

[23] Of course a very important component of an academic's non-monetary compensation is the quality of the research environment in which she operates. The latter is fully described and quantified by the relevant REF panel, whose judgement on this aspect of professional life is likely to be very well in line with that of the professors' in our dataset. This variable, however, enters linearly in both the GPA and the funding scores and including it, for example together with average pay as a component of a composite "compensation" variable, would introduce a mechanical correlation. Use of the research

4.2 Results by Score Components

We have described in Section 3.1 how the overall research profile of a unit is obtained as a weighted average of the profiles in each of the three components of the assessment: outputs, environment, and impact. As we explained, while output can easily be transferred across departments by hiring the academic who has produced it, this is not the case for environment and impact. Thus we expect that if universities use higher salaries to improve their REF performance, the effect of salaries should be stronger on the measure of output than on the other components of the overall funding score. To assess this implication, in Table 9 we run the specification from Column (3) of Table 6, which includes unit and institution type fixed effects, and is reported for convenience in Column (1) of Table 9, three more times, each with only one of the separate components of the overall REF funding score as dependent variables.

These results are presented in Columns (2)-(4). The overall positive association between average salary and REF performance is driven primarily by the relationships between salary and output and between salary and environment funding score. There is weaker evidence for a positive relationship between average salary and impact funding score, which is consistent with the rules of the REF which are such that institutions cannot "buy-in" impact success.[24] The results also indicate some interesting details in the role played by our controls. We find that having a member of staff on the panel has a positive and statistically significant effect on the funding score obtained for research environment and impact. There is no significant effect instead on the output funding score. These results are consistent with the idea that panel membership might be more important for the elements of the REF evaluation that are arguably more subjective, rather than for those which are based on more objective criteria such as the reputation of the outlet where a scholarly work has been published, its impact factor or the number of citations received. Finally, results in these regressions also indicate a strong relationship between the age profile of the academic department and performance in outputs funding score, but not for environment and impact funding score, though the coefficient on the proportion of under-40 professors is positive and weakly significant in the model for outputs funding score.

5 Responses to the REF: Empirical Strategy

Did the *outcome* of the REF exercise affect the subsequent wage structure of individual departments? Given the possibility of reverse or contemporaneous causality, and of omitted variables, such as location, co-determining both pay and performance, we tackle this question by adopting an identification strategy which exploits the quasi-experimental variation in REF performance across units brought about by some of the specific features of the exercise.

As pointed out in Box 3, the funding formula (20) translates a department's quality profile – the percentage of activity evaluated to be of quality at 4*, 3*, 2*, 1*, and 0* – into the level of funding

environment in the previous research assessment exercise is severely hampered both by the time interval, and especially by the drastic reduction from 67 to 36 in the number of units of assessment, which would make it very hard to match units across the exercises. Inclusion of fixed effects for discipline and institution type should capture at least part of the research environment component of an academic compensation.

[24] Cross-model tests of the equality of the coefficients on average salary in the model for output (2) against the model for impact (4), and also the model for environment (3) against the model for impact (4), in both cases reject the equality of coefficients at the 5% level. Note also that larger departments, those with more professors, especially those aged under 40, and those where pay inequality is higher do better in the environment component.

received. Importantly, the funding formula adopts a different set of weights compared to the headline measure of success given by the GPA used in the media. In particular, the former heavily over-weights 4* research relative to the GPA, as it can be evinced by comparing equations (20) and (19). As a result, two departments of the same average quality (GPA) submitting the same number of academics can well receive different levels of funding. More precisely, a department which achieves a given GPA score with more research evaluated as 4* will receive higher funding than a department achieving the same GPA with a lower percentage of research evaluated as 4*.[25] The non–linearity in the funding formula implies also that symmetrically, two departments receiving the same level of funding could well be characterized by different GPA scores.

Our identification strategy will thus rely on exploiting (1) quasi-experimental variation in the returns to the REF evaluation for departments with matched GPA and FTE and (2) quasi-experimental variation in GPA for departments with the same funding score. Bluntly, we ask the following question: Do universities care more about GPA or funding? To answer it, we compare two mirror approaches: the first matches GPAs (and size and broad subject area) to determine whether quasi-random shocks in funding generate corresponding changes in the pay structure, and the second approach matches funding (and broad subject area only, as the funding score already accounts for size) to study if quasi-random shocks in GPA determine corresponding changes in the pay structure. Our identification strategy rests on the assumption that differences in funding score given GPA and FTE, and conversely differences in GPA given funding, are as good as random.[26] Irrespective of the strategy pursued by the department, which is in fact often determined at the university level, there are good reasons to think that this assumption is satisfied, as the inherent uncertainties in the REF peer-review, which resembles the peer-review process by academic journals, generate noise in the evaluation of research quality, which we can exploit as a source of quasi-experimental variation.

Importantly, the two exercises will allow us to shed light on different implications of the REF for each university. Strong funding score performance contributes directly to university budgets, relaxing financial constraints at the margin. A strong GPA score indicates instead that the department's research is of high quality, with positive reputational effects and potentially indirect financial benefits (including increased awards by competitive funding bodies and increased Master level student recruitment). This analysis therefore helps us to understand the tradeoff between financial and reputational performance.

More formally, using the samples of matched departments we estimate the effects of university department performance in the REF on the department level wage structure using the following "difference-in-difference" model:[27]

$$\Delta Y_{i,t,t-1} - \Delta Y_{j,t,t-1} = \alpha + \beta_1 \Delta RP_{i,j,t-1} + \beta_2 \Delta \mathbf{X}_{i,j,t-1} + \tau_p + \epsilon_{i,j} \tag{22}$$

In (22) the dependent variable is the difference in the difference at dates t and $t-1$ (denoted by $t,t-1$)

[25]For example, two departments of the same size in FTE might both achieve a GPA score of 3 but the first department achieves this through a 100% $\times$ 3* profile, and the second department through a 50% $\times$ 4*, 50% $\times$ 2* profile. While the GPA scores are identical, the second department would receive £2.50 for each £1 received by the first.

[26]Editor Andrea Ichino invited us to develop this identification strategy: we wish to thank him for this suggestion.

[27]Using a difference-in-difference approach is preferable to using a difference-in-levels one as level differences may reflect pre-REF heterogeneity across departments.

in outcome Y between department i and department j in main panel p, where i and j form a matched pair by (i) GPA and FTE, or (ii) funding score. The dates we consider are $t =$ October 2015, and $t-1 =$ October 2013. The latter is the exact REF census date. The REF results were published in December 2014, and therefore, given the internal administrative times in funding a post, the search process, negotiations, and delays prior to an appointee taking up a post, October 2015 is the earliest date where a response to the REF can be detected. As in (21), RP is the measure of research performance, and $\Delta RP_{i,j,t-1}$ is the difference in research performance between departments i and j at the time of the REF exercise $t-1$. As in the earlier analysis we measure research performance using (i) funding score and in (ii) GPA. $\Delta \mathbf{X}_{i,j,t-1}$ refers to a set of control variables, defined as the difference in the variable between departments i and j at the time of the REF exercise $t-1$.[28] τ_p is a set of main panel fixed effects and ϵ is the error term. Hence the model estimates a difference-in-difference in an outcome variable within matched pairs. In other words, we exploit differences across departments in the *change* in outcome variable (e.g. change in mean salary), estimating whether the department within a matched pair that received (i) a higher funding score or (ii) a higher GPA did in fact experience a larger increase in the outcome variable over subsequent years of the data.

Practically, our identification strategy rests on the availability of sufficiently close matches between department-pairs both on GPA and FTE, and on funding score. We limit matched pairs to be within the same main panel to avoid the risk of mis-matches on important unobservables (e.g. wage flexibility or capital requirement). To achieve the closest matches between departments we use a simple algorithm. In the first matching exercise, we pair departments such that we minimise the sum of distances in standardised GPA and FTE. We use an additive quintic loss function, because we want to avoid being very close on one dimension, say GPA, while being far away on the other, FTE in this case.

Figure 5 illustrates the distribution of GPA and FTE across departments in the top two diagrams and the within-cell differences in GPA and FTE between department pairs in the bottom two diagrams: within-cell differences are very small for the majority of departments. Hence in many cases we are able to match pairs of departments with very similar GPA and FTE.[29] This gives variation in the research funding dimension of the assessment, holding research quality constant.

In the second matching exercise, we pair departments by funding score using a nearest-neighbour match, on the single matching target variable. We do this in order to obtain quasi-experimental variation in average quality (GPA) across departments achieving the same funding score. This gives variation in the research quality dimension of the assessment, holding research income constant. Analogously to Figure 5, Figure 6 illustrates the distribution of log funding score across departments in the top figure and the within-cell differences in log funding score between department pairs in the bottom figure. Again, within-cell differences are very small for most of the distribution. In our main analysis we use the full distribution of matched pairs. In sensitivity analysis we further restrict the sample to the closest matches, a restriction which leaves the results qualitatively unchanged (see Tables A9 and A10).

[28] In particular $\mathbf{X}$ includes the percent of professors aged under 40, the percent of professors aged between 41-50, the percent of professors aged between 51-60 and the log of total FTE staff submitted to the REF.

[29] The closest pair are Geography, Environmental Studies and Archaeology at the University of Hull (GPA 2.96, FTE 34.5) and Social Work Policy at the University of Birmingham (GPA 2.95, FTE 34.5).

6 Responses to the REF: Results

We examine four outcomes. First, the natural log of the department mean salary, which answers the question of the effect on average salaries. Second, the natural log of the total wage bill paid to the top quartile of professors within the department. This measures the effect of REF performance on 'superstar' salaries within the department: these individuals could be hired from outside on the strength of the increased reputation, or be the consequence of counteroffers to professors already in the department whose visibility has increased because of the REF success. Third, the inequality in salaries within the department. Fourth, the total number of professors in the department: departments performing strongly in the exercise may be rewarded with new senior posts and/or internal promotion of non-professorial staff.

The funding score determines the direct monetary research funding allocation given by the government to universities as a result of the REF performance of its academic departments. Table 10 shows results for the sample of departments that are closely matched by GPA and FTE.[30] Results reveal no statistically significant coefficients on the funding score variable in any of the estimated models. These results indicate that the receipt of higher funding by the university in which the department sits (for a given research quality) appears not to cause any changes in the wage structure of the department. In other words, the financial benefits of REF performance appear not to influence the wage outcomes of the individuals within the department.

This picture however changes when we estimate the effect of GPA score on department wage structure. GPA is a measure of the average research quality of a department, commonly employed as the headline measure of research quality by the media and the public. Table 11 reports results for the sample of departments that are closely matched by Funding Score.[31] While there is no statistically significant effect from higher GPA on mean salaries, the estimates reveal a positive effect on the top quartile wage bill which is statistically significant at the 1% level. The coefficient value of 0.072 implies that a department achieving a one standard deviation higher GPA score, one for example that passes from the sample mean to 87-th percentile, would see growth in the wage bill of top paid staff which is 2% higher than that in a matched department with the same funding score, but achieved with a lower GPA. In the sample the mean growth in salaries (in real terms) over the period is 4.2%. Hence departments achieving a one standard deviation higher GPA score see more than pay growth faster by about 50%.It should be noted that there is no contradiction in principle between an increase in top-level wages and lack of change of the mean salary.[32]

[30] Specifically, we restrict the sample to the interquartile range of the difference in the error term from the additive quintic loss function employed to match departments. We show results for the full sample in Table A5. Results in the full sample are very similar to those in the restricted sample.

[31] Specifically, we restrict the sample to the interquartile range of the difference in the error term from the additive quintic loss function employed to match departments. We show results for the full sample in Table A6. Results in the full sample are very similar to those in the restricted sample. Tables A9 and A10 in the Appendix explore the possible presence of spillovers between departments within the same faculty. It does so by replacing the difference in performance variable at the departmental level with the same variable at main panel level. This evidence is suggestive that spillovers might be present.

[32] If new hires or promotions are made in the middle of the wage distribution as well as the very top, then the overall effect on inequality may be small. For example, consider a department with 12 professors, four of them paid 60, four paid 70 and four paid 100. Suppose that, as a consequence of the REF results, this department ends up with four extra professors, for a total of 16, three paid 60, two paid 65, seven paid 70, 3 paid 100, and a superstar paid 130. Easy calculations show that the Gini and the average pay are roughly unchanged, but the average pay of the top quartile increases from 100 to 107.5.

In the final model shown in Column 4 for the change in total FTE of professors within the department, we see a positive and statistically significant coefficient. The coefficient value of 6 implies that a department with one standard deviation higher GPA score sees approximately an additional two staff added to the unit. In the sample the change in professorial FTE over the period at the 75th percentile is 1.10 (hence on average departments grow by approximately one person in the two-year period after the REF). Hence higher GPA performance by one unit causes departments to grow approximately twice as fast than this rate of growth in the sample. In short, the response to REF success seems to be in the guise of more senior posts and increases in the wage at the top of the distribution.

From a policy perspective, a general conclusion from the exercise carried out in this section is that the research evaluation creates a measure of research quality, the GPA, and that departments which perform strongly in terms of this indicator are rewarded in higher wages and expansion in the size of their professorial body. Even when the increase in GPA does *not* result in a higher contribution to the university coffer, universities internal funding allocation processes undo the desire of the government to reward disproportionately "world leading" research, with the intuitive GPA formula driving instead the rewards to centres of excellence within institutions. In other words, prestige appears to offer departments a stronger claim to internal funds to be devoted to salaries than direct financial contribution. This result could well be due to a stronger multiplier effect on grant funding and post-graduate student recruitment by the headline measure of the GPA than by the rather ad-hoc funding formula.

7 Conclusion

This paper studies the relationship between pay and research performance in UK universities. The UK setting is especially interesting because universities research quality is periodically assessed in a national exercise, now labelled REF, and because they can freely compete on the salaries they offer to senior academics.

Our empirical analysis, guided by a simple theoretical model where academics differ in ability and are inputs into the production of research, and universities seek to maximise the weighted average research quality, shows a positive pay-performance relationship in all disciplines. This is true both in subjects areas which anecdotal evidence suggests to be more competitive, such as business and management, economics, engineering, but also in subjects where there appears to be far less cross-institution movement of staff and possibly less competition, such as disciplines within the arts and humanities. Inequality of pay is also associated with higher research performance, especially in the more research intensive institutions.

That universities respond to REF rules by pursuing academics who will contribute to the measured research performance is suggested also by our analysis of the three components that make up the aggregate research score. The positive salary-performance gradient is due mainly to the relationship between salary and scholarly publications: when an academic moves, this is the component of her recent record that can be transferred from one institution to another, whereas any "impact" that her research may have had contributes to the score of her previous institution.

Those whose task is to design the details of the evaluation process should also consider our consistent finding that, after controlling for other potential covariates, panel membership is associated with stronger performance, and this result in turn is driven by the effect of panel membership on the arguably more subjective "environment" and "impact" components of the evaluation, which, incidentally, will see their weight increased in the next assessment exercise, the REF2021.

The salary structure snapshot taken in October 2013, can of course only determine the presence of correlations, which, while interesting on their own, lack a casual interpretation. To study the causal effect of REF performance we have employed a matching strategy to set up a quasi experimental design, exploiting the unavoidable elements of uncertainty in the peer-review based evaluation process. The idea is that a favourable surprise in the REF outcome relatively to an otherwise identical department generates a random unexpected increase in the allocated funds, or in the departmental "prestige", just as a controlled experiment would. Our results in Section 6 suggest that while bringing additional funding to the institution does not affect the professorial pay structure in a department, additional prestige does, in the form of increases in the top pay, and in the number of professors. One realistic explanation is that the extra cash simply ends up in the university budget, whereas media attention and prestige give departments additional power in the negotiations for the allocation of the budget, or easier access to senior potential appointees, or, finally, increase the pressure on senior staff retention budgets by making current professor more visible externally.

While individual UK academics and administrators will no doubt find these results of much interest, they warrant wider attention, as they contain important lessons on the effects of liberalising pay and introducing competition for resources in a largely publicly funded system. These lessons may be useful for other European countries, which are in the process of developing and strengthening lively quasi-market systems in the university sector.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*, 251–333.

Altbach, P. G., L. Reisberg, M. Yudkevich, G. Androushchak, and I. F. Pacheco (2012). *Paying the Professoriate: A Global Comparison of Compensation and Contracts*. New York: Routledge.

Bosquet, C. and P.-P. Combes (2017). Sorting and agglomeration economies in french economics departments. *Journal of Urban Economics 101*, 27–44.

Bratsberg, B., J. Ragan, and J. Warren (2010). Does raiding explain the negative returns to faculty seniority? *Economic Inquiry 48*, 704–721.

Butler, L. and I. McAllister (2009). Metrics or peer review? Evaluating the 2001 UK research assessment exercise in political science. *Political Studies Review 7*, 3–17.

Checchi, D., G. De Fraja, and S. Verzillo (2014). Publish or perish? An analysis of the academic job market in Italy, 1991-2012. CEPR Discussion Paper 10084, London, UK.

Clerides, S., P. Pashardes, and A. Polycarpou (2011). Peer review vs metric-based assessment: Testing for bias in the RAE ratings of UK economics departments. *Economica 78*, 565–583.

De Fraja, G. (2016). A theoretical analysis of public funding for research. *Rand Journal of Economics 47*, 498–528.

De Fraja, G. and P. Valbonesi (2012). The design of the university system. *Journal of Public Economics 96*, 317–330.

Diamond, A. M. (1986). What is a citation worth? *Journal of Human Resources 21*, 200–215.

Edmans, A. and X. Gabaix (2015). Executive compensation: A modern primer. Technical Report 21131, NBER Working Paper Series.

Else, H. (2015). Research funding formula tweaked after REF 2014 results. *Times Higher Education 20 February*.

Giles, D. E. (1982). The interpretation of dummy variables in semilogarithmic equations: Unbiased estimation. *Economics Letters 10*, 77–79.

Grund, C. and N. Westergaard-Nielsen (2008). The dispersion of employees' wage increases and firm performance. *ILR Review 61*, 485–501.

Haeck, C. and F. Verboven (2012). The internal economics of a university: Evidence from personnel data. *Journal of Labor Economics 30*, 591–626.

Halvorsen, R. and R. Palmquist (1980). The interpretation of dummy variables in semilogarithmic equations. *American Economic Review 70*, 474–475.

Hamermesh, D. and G. A. Pfann (2012). Reputation and earnings: The roles of quality and quantity in academe. *Economic Inquiry 50*, 1–16.

Hamermesh, D. S., G. E. Johnson, and B. A. Weisbrod (1982). Scholarship, citations and salaries: Economic rewards in economics. *Southern Economic Journal 49*, 472–481.

Hildreth, A. K. G. and A. J. Oswald (1997). Rent-sharing and wages: Evidence from company and establishment panels. *Journal of Labor Economics 15*, 318–337.

Johnes, J., J. Taylor, and B. Francis (1993). The research performance of UK universities: A statistical analysis of the results of the 1989 research selectivity exercise. *Journal of the Royal Statistical Society, Series A 156*, 271–286.

Lazear, E. (2000). Performance pay and productivity. *American Economic Review 90*, 1346–1361.

Milgrom, P. R. and J. Roberts (1992). *Economics, Organization and Management*. New York: Prentice Hall.

Moore, W. J., R. J. Newman, and G. K. Turnbull (1998). Do academic salaries decline with seniority? *Journal of Labor Economics 16*, 352–366.

Mueller, H. M., P. P. Ouimet, and E. Simintzi (2017). Wage inequality and firm growth. *American Economic Review 107*, 379–383.

Nickell, S., J. Vainiomaki, and S. Wadhwani (1994). Wages and product market power. *Economica 61*, 457–473.

Nickell, S. and S. Wadhwani (1990). Insider forces and wage determination. *The Economic Journal 100*, 496–509.

Oyer, P. (2007). Is there an insider advtange in getting tenure? *American Economic Review 97*, 501–205.

Porter, S. R., R. K. Toutkoushian, and J. V. Moore III (2008). Pay inequities for recently hired faculty, 1988-2004. *The Review of Higher Education 31*, 465–487.

Règibeau, P. and K. E. Rockett (2016). Research assessment and recognized excellence: Simple bibliometrics for more efficient academic research evaluations. *Economic Policy 31*, 611–652.

Sgroi, D. and A. J. Oswald (2013). How should peer-review panels behave. *Economic Journal 123*, F255–F278.

Sharp, S. and S. Coleman (2005). Ratings in the research assessment exercise 2001 – the patterns of university status and panel membership. *Higher Education Quarterly 59*, 153–171.

Taylor, J. (1995). A statistical analysis of the 1992 research assessment exercise. *Journal of the Royal Statistical Society, Series A 158*, 241–261.

Thanassoulis, E., M. Kortelainen, G. Johnes, and J. Johnes (2011). Costs and efficiency of higher education institutions in England: A DEA analysis. *Journal of the Operational Research Society 62*, 1282.

Zinovyeva, N. and M. Bagues (2010). Does gender matter for academic promotion? Evidence from a randomized natural experiment. FEDEA Documento de Trabajo 2010-15, Madrid.

Zinovyeva, N. and M. Bagues (2015). The role of connections in academic promotions. *American Economic Journal: Applied Economics 7*, 264–292.

Table 1: Department-Level Pay Summary Statistics, 2013-14

	Mean	Min	Median	SD	Max
Professorial FTE (N)	13.98	3.00	9.00	19.48	311.00
Average Salary (£000s)	73.67	50.87	72.89	9.85	128.46
Gini Coefficient Salary (x100)	8.22	0.25	7.81	4.06	35.62
% Age under 40	5.26	0.00	0.00	9.11	66.67
% Age 41-50	31.08	0.00	30.77	18.47	100.00
% Age 51-60	38.27	0.00	36.36	19.54	100.00
% Age over 60	25.39	0.00	23.08	18.50	100.00
Vice Chancellor Pay (£000s)	300.73	143.00	284.00	63.48	623.00
Department has a REF Panel Member	0.39	0.00	0.00	0.49	1.00
Academic FTE Submitted to REF (N)	34.01	2.00	24.72	32.83	449.74

Notes: Sample size = 1093 academic departments submitted to REF 2014 which remain in existence in 2015/16 academic year. Professorial FTE is the total FTE of professors in the department. FTE submitted to REF 2014 measures the total number of FTE (including non-professorial researchers) submitted to the REF evaluation. Vice Chancellor pay is total remuneration (including salary and discretionary payments). Department has a REF Panel Member is a dummy variable indicating whether the REF 2014 main panel or sub-panel included a member of the department.

Table 2: Department-Level Pay Summary Statistics, 2015-16

	Mean	Min	Median	SD	Max
Professorial FTE (N)	14.14	1.02	8.67	20.10	315.50
Average Salary (£000s)	76.62	54.79	75.47	10.41	173.78
Gini Coefficient Salary (x100)	8.15	0.00	7.80	4.14	29.81
% Age under 40	4.07	0.00	0.00	7.51	50.00
% Age 41-50	30.16	0.00	30.00	18.55	100.00
% Age 51-60	40.92	0.00	40.00	19.67	100.00
% Age over 60	24.85	0.00	23.58	18.20	100.00

Notes: Sample size = 1093 academic departments submitted to REF 2014 which remain in existence in 2015/16 academic year. Professorial FTE is the total FTE of professors in the department. Vice Chancellor pay is total remuneration (including salary and discretionary payments).

Table 3: Changes in Department-Level Pay Summary Statistics, 2013-2016

	Mean	Median	SD	Min	25th	75th	Max
Change in Professorial FTE (N)	0.16	0.00	3.09	-24.50	-1.00	1.12	19.99
Change in Average Salary (£000s)	2.94	2.84	4.77	-21.79	0.72	4.99	46.72
Change in Gini Coefficient Salary (x100)	-0.00	-0.00	0.03	-0.20	-0.01	0.01	0.13
Change in Top Quartile Wage Bill (£000s)	22.63	2.60	186.40	-1469.35	-46.40	72.73	2620.19
Change in Professorial FTE (%)	1.26	0.00	28.03	-80.71	-14.62	15.15	175.00
Change in Average Salary (%)	4.20	3.94	6.41	-20.61	0.98	6.79	45.22
Change in Top Quartile Wage Bill (%)	12.25	3.02	49.48	-70.85	-14.20	24.29	215.94

Notes: Sample size = 1093 academic departments submitted to REF 2014 which remain in existence in 2015/16 academic year. Professorial FTE is the total FTE of professors in the department. FTE submitted to REF 2014 measures the total number of FTE (including non-professorial researchers) submitted to the REF evaluation. Vice Chancellor pay is total remuneration (including salary and discretionary payments). Department has a REF Panel Member is a dummy variable indicating whether the REF 2014 main panel or sub-panel included a member of the department. Total salary bill is sum of departmental professorial pay.

Table 4: Department-Level REF Outcome Summary Statistics

	Mean	Min	Median	SD	Max
Funding Score	59.48	0.36	38.63	74.90	939.96
Overall Grade Point Average	2.97	1.51	3.02	0.33	3.78
Outputs Grade Point Average	2.87	1.39	2.91	0.29	3.68
Environment Grade Point Average	3.14	0.75	3.20	0.58	4.00
Impact Grade Point Average	3.16	0.70	3.27	0.53	4.00

Notes: Sample size = 1093 academic departments submitted to REF 2014 which remain in existence in 2015/16 academic year. Definitions of Funding Score and other REF Grade Point Average variables are provided in the main text.

Table 5: Correlation Between REF Performance Measures

	Funding Score	GPA Score	GPA Outputs	GPA Environment	GPA Impact
Funding Score	1				
GPA Score	0.473***	1			
GPA Outputs	0.369***	0.901***	1		
GPA Environment	0.506***	0.845***	0.635***	1	
GPA Impact	0.379***	0.777***	0.477***	0.644***	1

Note: Sample size = 1093 departments submitted to REF 2014. For explanation of REF performance measures see main text. Denotes significance at 10% level, ** 5% level, *** at 1% level.

Table 6: OLS Regression Estimates: Department Pay Characteristics and REF 2014 Performance. Dependent Variable: REF Funding Score

	(1)	(2)	(3)
Log Average Salary	0.493***	0.842***	0.510***
	(0.096)	(0.092)	(0.091)
Gini Coefficient Salary	0.914***	1.030***	0.568*
	(0.317)	(0.290)	(0.271)
Log Professorial FTE	0.037*	0.149***	0.110***
	(0.021)	(0.022)	(0.020)
Log REF FTE	1.118***	1.092***	1.050***
	(0.023)	(0.022)	(0.021)
Panel Member = 1	0.142***	0.090***	0.060***
	(0.023)	(0.021)	(0.019)
Log Vice Chancellor Pay	0.280***	0.179***	0.117**
	(0.055)	(0.050)	(0.047)
% Age under 40	-0.015	0.318***	0.222*
	(0.118)	(0.111)	(0.104)
% Age 41-50	-0.120*	0.096	0.059
	(0.067)	(0.063)	(0.059)
% Age 51-60	-0.082	0.024	-0.052
	(0.064)	(0.060)	(0.056)
1994 Group Uni. = 1			-0.024
			(0.024)
Other Type Uni. = 1			-0.205***
			(0.028)
New Uni = 1			-0.414***
			(0.034)
Specialist Uni = 1			-0.072
			(0.112)
Unit of Assessment FE	NO	YES	YES
University Type FE	NO	NO	YES
R-squared	0.866	0.898	0.913
Observations	1093	1093	1093

Notes: Sample size = 1093 academic departments submitted to REF 2014. the dependent variable in the OLS regression is the natural log of research funding score. Regressors are shown in table rows. Additional columns add control variables as described in table header. Standard errors in parenthesis. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

Table 7: OLS Regression Estimates By REF Main Panel

	(1) Panel A Medicine	(2) Panel B Science & Engineering	(3) Panel C Social Sciences	(4) Panel D Arts & Humanities
Log Average Salary	1.082***	0.611***	0.401**	0.809***
	(0.257)	(0.206)	(0.163)	(0.149)
Gini Coefficient Salary	1.465*	1.566***	0.544	0.794
	(0.722)	(0.578)	(0.571)	(0.540)
Log Professorial FTE	-0.042	0.200***	0.027	0.130***
	(0.039)	(0.049)	(0.041)	(0.038)
Log REF FTE	1.141***	1.030***	1.226***	1.000***
	(0.048)	(0.051)	(0.043)	(0.037)
Panel Member = 1	0.092*	0.095*	0.165***	0.129***
	(0.053)	(0.044)	(0.040)	(0.035)
Log Vice Chancellor Pay	0.237*	0.235*	0.281***	0.134*
	(0.135)	(0.109)	(0.098)	(0.081)
% Age under 40	0.053	0.438	0.144	0.098
	(0.327)	(0.284)	(0.183)	(0.188)
% Age 41-50	0.018	0.011	0.041	-0.030
	(0.182)	(0.145)	(0.119)	(0.093)
% Age 51-60	0.053	-0.133	0.046	-0.059
	(0.164)	(0.151)	(0.115)	(0.087)
University Type FE	YES	YES	YES	YES
R-squared	0.921	0.894	0.849	0.876
Observations	183	263	369	278

Note: Dependent variable is natural log of research funding score. OLS regression estimated on four mutually exclusive samples of academic departments categorised by REF Main Panel. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

Table 8: OLS Regression Estimates By University Type

	(1) Russell	(2) 1994	(3) Others	(4) New
Log Average Salary	0.170*	0.322*	1.125***	1.800***
	(0.102)	(0.171)	(0.264)	(0.400)
Gini Coefficient Salary	0.763**	1.056*	-0.909	-0.665
	(0.298)	(0.498)	(0.768)	(1.098)
Log Professorial FTE	0.092***	0.138***	0.297***	0.078
	(0.025)	(0.040)	(0.057)	(0.073)
Log REF FTE	1.077***	0.969***	0.986***	1.062***
	(0.026)	(0.044)	(0.055)	(0.062)
Panel Member = 1	0.043*	0.077*	0.051	0.152*
	(0.023)	(0.034)	(0.053)	(0.086)
Log Vice Chancellor Pay	0.153***	0.023	-0.064	0.007
	(0.056)	(0.092)	(0.139)	(0.144)
% Age under 40	-0.124	0.571***	0.145	0.136
	(0.144)	(0.198)	(0.315)	(0.263)
% Age 41-50	0.003	0.304**	0.074	0.036
	(0.086)	(0.122)	(0.145)	(0.156)
% Age 51-60	-0.039	0.195*	-0.006	-0.109
	(0.083)	(0.117)	(0.131)	(0.144)
Unit of Assessment FE	YES	YES	YES	YES
R-squared	0.939	0.882	0.882	0.820
Observations	421	256	211	198

Note: Dependent variable is natural log of research funding score. OLS regression estimated on four mutually exclusive samples of academic departments categorised by University Type. Sample of 'specialist' universities not included as it contains only 8 observations. See Appendix A for details of classifications of universities. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

Table 9: OLS Regression Estimates: Department Pay Characteristics and REF Component Funding Scores

	(1) Overall Funding	(2) Outputs	(3) Environment	(4) Impact
Log Average Salary	0.510***	0.547***	0.540***	0.416*
	(0.091)	(0.092)	(0.183)	(0.191)
Gini Coefficient Salary	0.568*	0.398	1.923***	0.100
	(0.271)	(0.276)	(0.547)	(0.572)
Log Professorial FTE	0.110***	0.082***	0.166***	0.161***
	(0.020)	(0.021)	(0.041)	(0.043)
Log REF FTE	1.050***	-0.000	0.359***	0.145***
	(0.021)	(0.021)	(0.042)	(0.044)
Panel Member = 1	0.060***	0.024	0.148***	0.124***
	(0.019)	(0.020)	(0.039)	(0.041)
Log Vice Chancellor Pay	0.117**	0.203***	0.107	-0.024
	(0.047)	(0.048)	(0.096)	(0.100)
% Age under 40	0.222*	0.203*	0.473*	0.210
	(0.104)	(0.106)	(0.209)	(0.219)
% Age 41-50	0.059	0.093	0.063	-0.058
	(0.059)	(0.060)	(0.119)	(0.125)
% Age 51-60	-0.052	-0.053	0.032	-0.121
	(0.056)	(0.057)	(0.113)	(0.118)
1994 Group Uni. = 1	-0.024	-0.017	-0.007	-0.028
	(0.024)	(0.024)	(0.048)	(0.050)
Other Type Uni. = 1	-0.205***	-0.233***	-0.316***	-0.182***
	(0.028)	(0.028)	(0.056)	(0.059)
New Uni = 1	-0.414***	-0.413***	-0.572***	-0.386***
	(0.034)	(0.035)	(0.069)	(0.072)
Specialist Uni = 1	-0.072	-0.044	-0.507*	0.089
	(0.112)	(0.114)	(0.225)	(0.236)
Unit of Assessment FE	YES	YES	YES	YES
University Type FE	YES	YES	YES	YES
R-squared	0.913	0.492	0.526	0.294
Observations	1093	1093	1093	1093

Note: Sample size = 1093 departments submitted to REF 2014. The dependent variable is obtained from formula (20). Column (1) is the same as Column (4) in Table 6. Column (2) the funding score for the department output component, Column (3) environment component and Column (4) impact component. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

Table 10: Difference-in-Difference Estimates of Effects of REF 2014 Funding Score on Changes in Departmental Wage Structure, 2013 - 2016, Sample of Closest Matches Only

	(1) Change Log Mean Salary	(2) Change Log Top Quartile Wage Bill	(3) Change Gini Coeff.	(4) Change Total FTE Professors
Funding Score	-0.012	0.025	0.005	1.127
	(0.020)	(0.446)	(0.008)	(0.934)
% Age under 40	-0.017	1.832	0.017	7.285***
	(0.056)	(1.282)	(0.023)	(2.684)
% Age 41-50	-0.048*	0.692	-0.007	0.448
	(0.027)	(0.616)	(0.011)	(1.289)
% Age 51-60	0.002	0.403	-0.002	-0.530
	(0.024)	(0.546)	(0.010)	(1.142)
Log REF FTE	0.005	0.237	-0.001	-0.878*
	(0.011)	(0.241)	(0.004)	(0.503)
Main Panel FE	YES	YES	YES	YES
R-squared	0.020	0.013	0.005	0.051
Observations	280	280	280	280

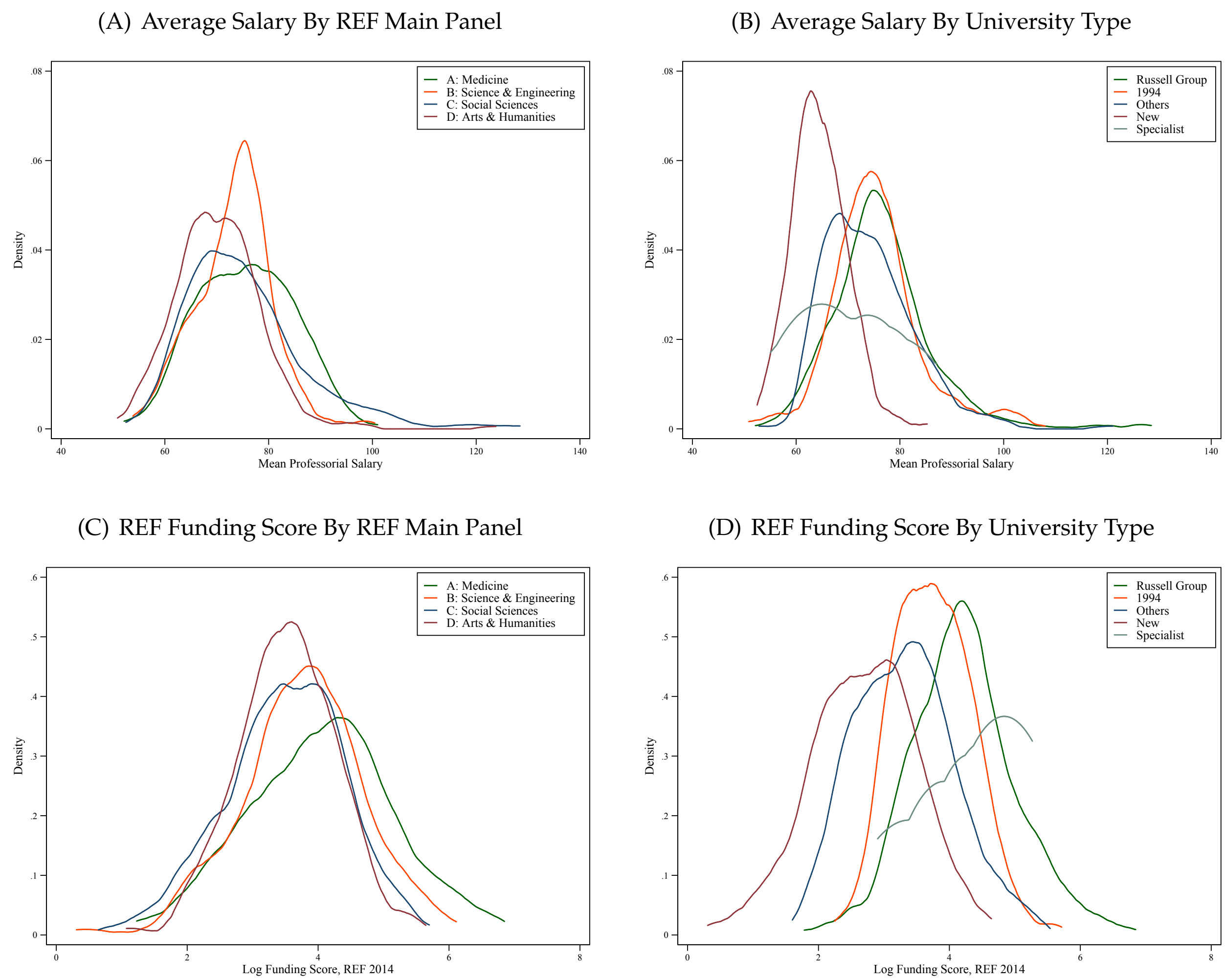
Note: Table shows OLS estimates from difference-in-difference models based on Equation 22. The dependent variable and funding score variables are the differenced variables, calculated as the differences between matched pairs (by GPA and FTE) in the first-differences in the variables between 2013 and 2015. Other control variables are differences between the matched departments in levels measured in the 2013 data. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

Table 11: Difference-in-Difference Estimates of Effects of REF 2014 GPA on Changes in Departmental Wage Structure, 2013 - 2016, Sample of Closest Matches Only

	(1) Change Log Mean Salary	(2) Change Log Top Quartile Wage Bill	(3) Change Gini Coeff.	(4) Change Total FTE Professors
GPA Score	-1.502	0.072**	-0.021	5.988***
	(2.528)	(0.019)	(0.044)	(1.228)
% Age under 40	-0.270	-0.074	0.025	2.002
	(1.309)	(0.051)	(0.023)	(2.707)
% Age 41-50	-0.598	-0.012	-0.004	0.088
	(0.644)	(0.025)	(0.011)	(1.332)
% Age 51-60	0.854	0.002	-0.002	-0.975
	(0.634)	(0.025)	(0.011)	(1.312)
Log REF FTE	-0.128	0.012	-0.002	0.193
	(0.216)	(0.008)	(0.004)	(0.446)
Main Panel FE	YES	YES	YES	YES
R-squared	0.029	0.017	0.007	0.010
Observations	272	272	272	272

Note: Table shows OLS estimates from difference-in-difference models based on Equation 22. The dependent variable and funding score variables are the differenced variables, calculated as the differences between matched pairs (by Funding Score) in the first-differences in the variables between 2013 and 2015. Other control variables are differences between the matched departments in levels measured in the 2013 data. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

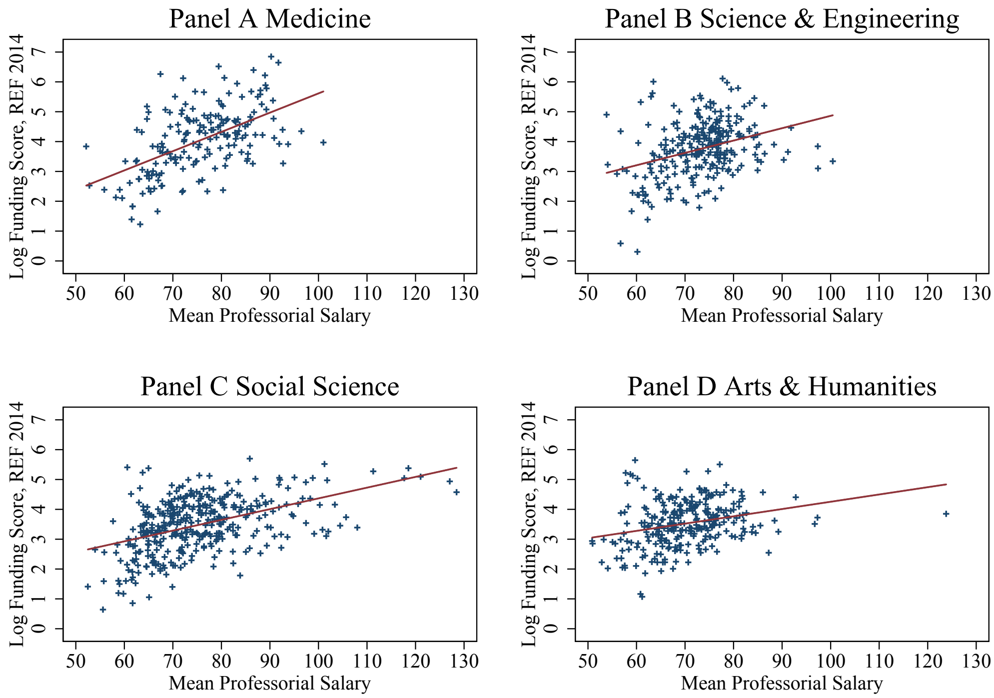
Figure 1: Distribution of Average Salary and REF Performance (Funding Score) Among Academic Departments by REF Main Panel and University Type, 2013



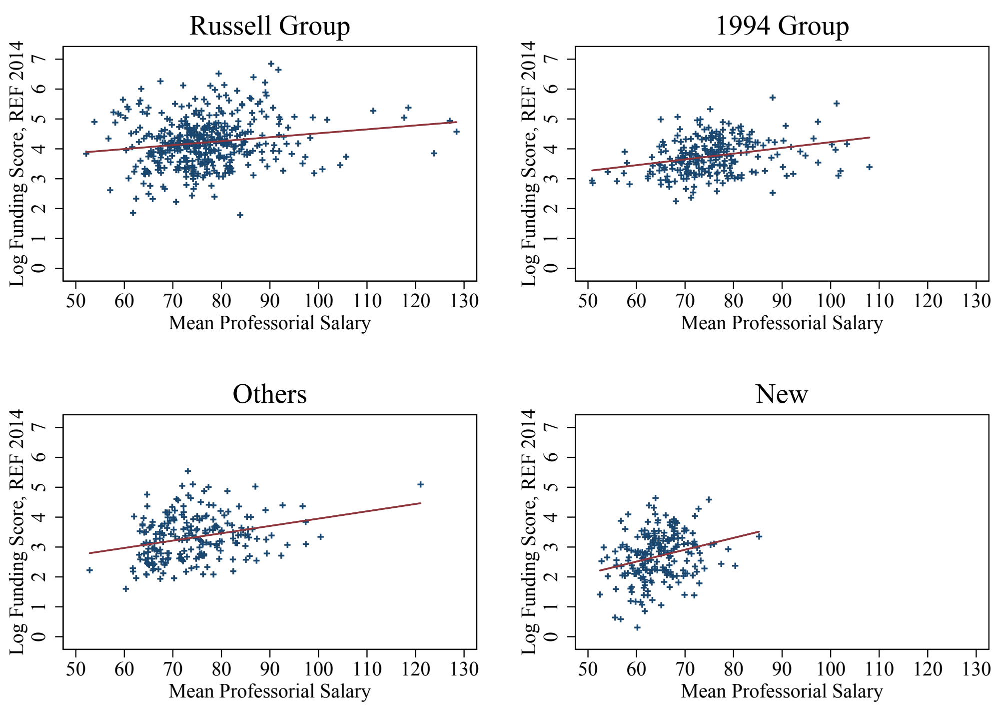
Note: The upper part of the Figure illustrates distribution of average departmental salary among academic departments classified by REF Main Panel and university type. The lower part shows the distribution of REF Funding Score in academic departments, again classified by REF Main Panel and university type. Kernel density functions, epanechnikov kernel.

Figure 2: Correlation Between Mean Pay and REF Funding Score

(A) By REF Main Panel

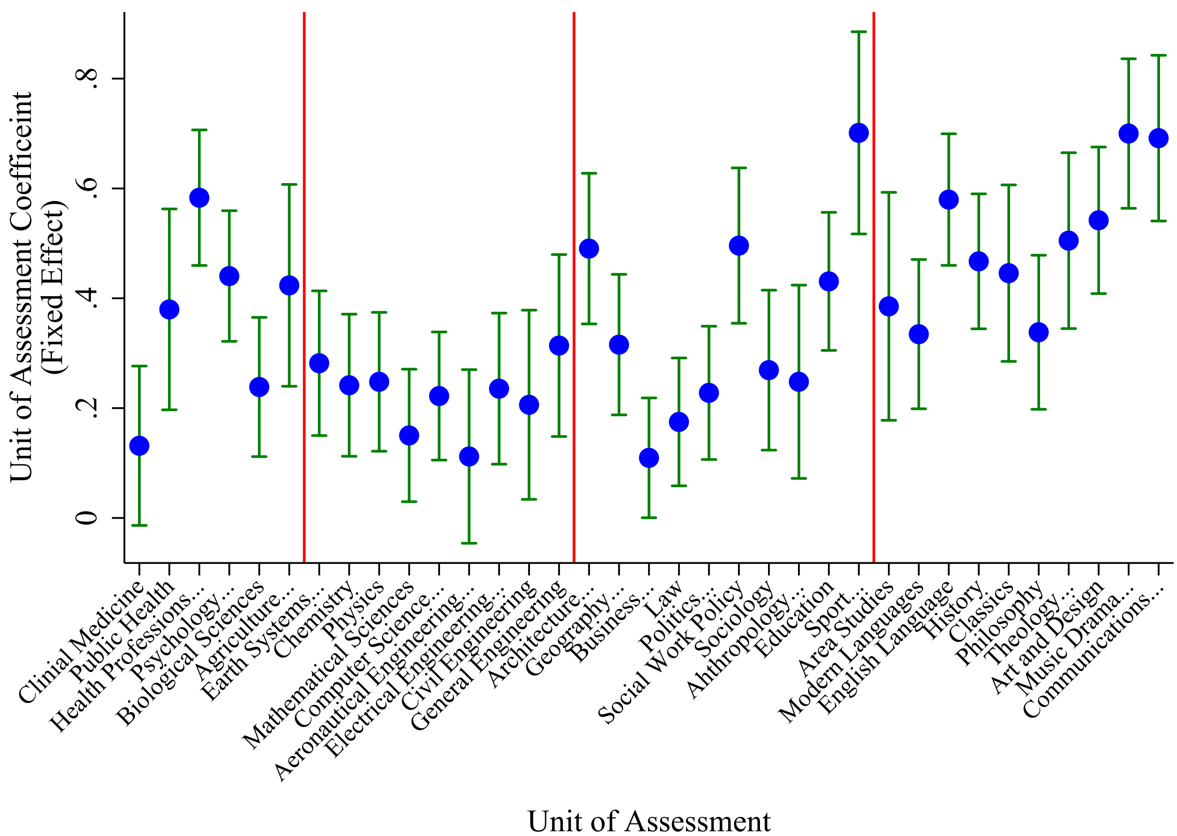


(B) By University Type



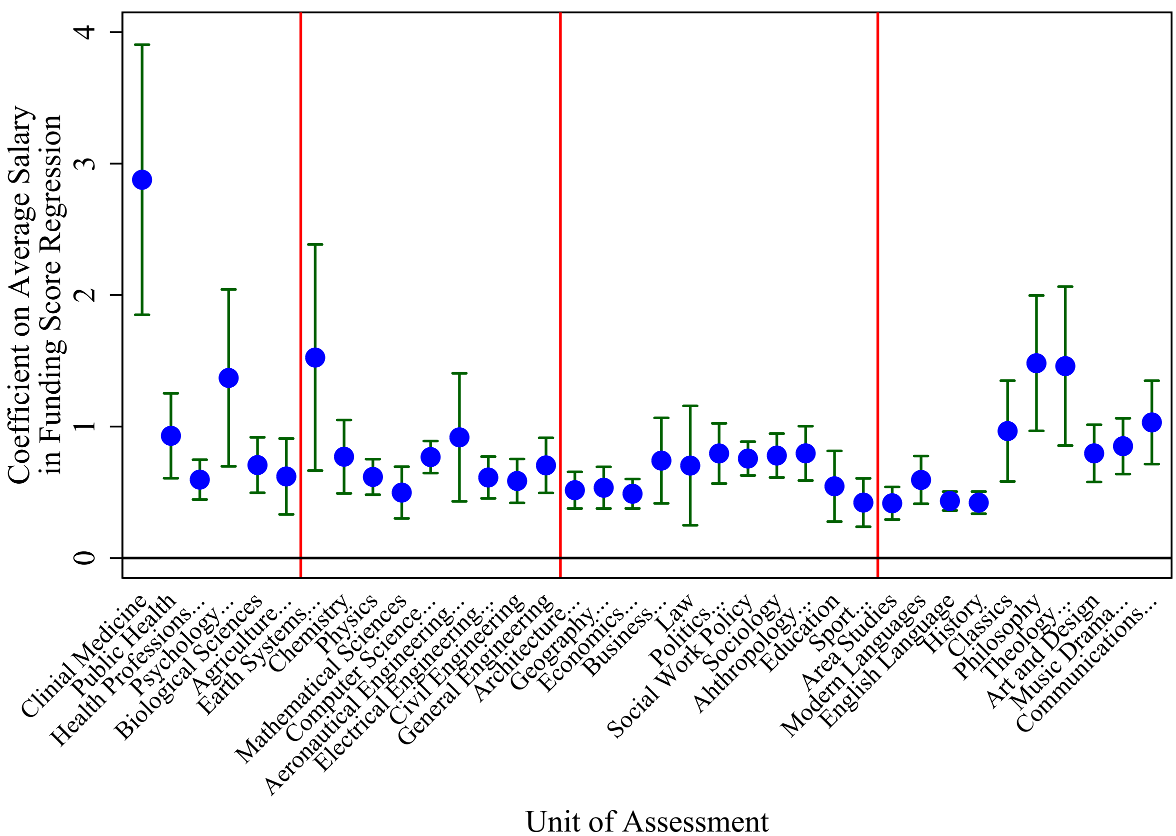
Note: Each observation is an individual academic department. Figure show scatter plots and fitted regression lines. Observations grouped by REF Main Panel (Plot group A) and University Type (Plot group B)

Figure 3: Plot of Estimated Unit of Assessment Fixed Effects from Regression Model (Omitted Unit: Economics and Econometrics)



Note: Figure shows a plot of the estimated unit of assessment fixed effects from Table 6, Column 4. Omitted group is Unit of Assessment 18, Economics and Econometrics. 95% confidence intervals shown in whiskers.

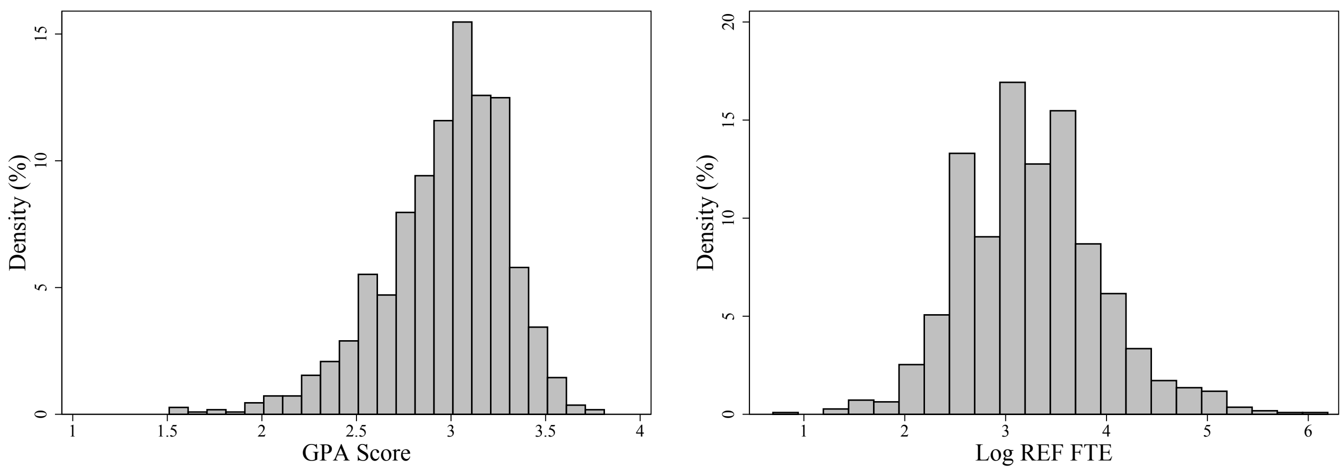
Figure 4: Coefficient Plots From Unit of Assessment Sample OLS Regressions



Note: Figure shows regression coefficient values and 95% confidence intervals (shown by vertical whisker bars) for coefficient estimates on average pay variable in OLS regression of funding score against average pay and controls (control variables as in Table 6, Column 3). Separate regressions estimated for each unit of assessment sample.

Figure 5: Distributions of Grade Point Average and Research Power Across Units and Within Matched Pairs

(A) Distributions in the Main Sample



(B) Distributions of Within-Cell Differences

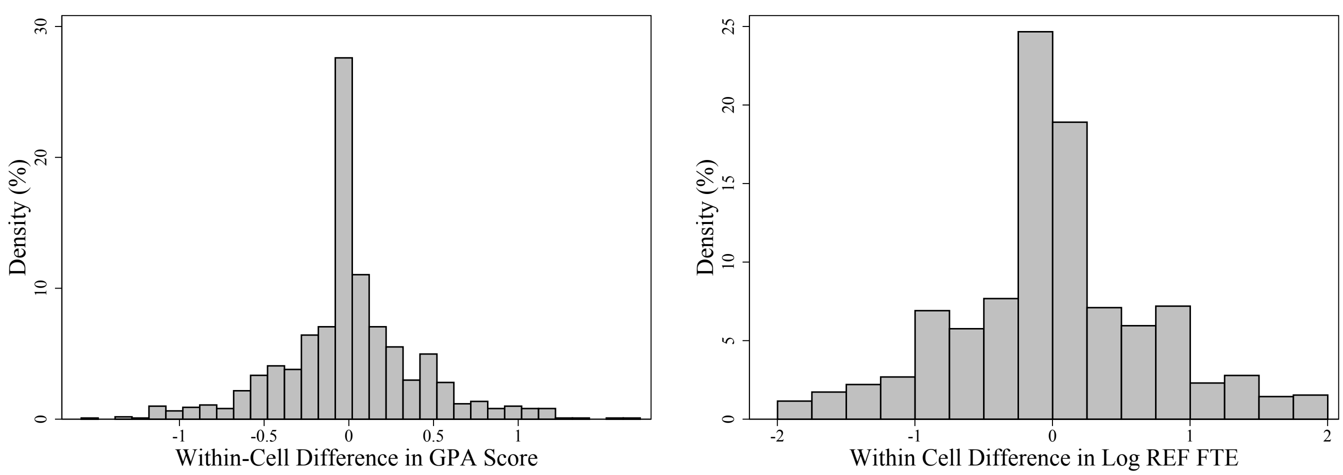
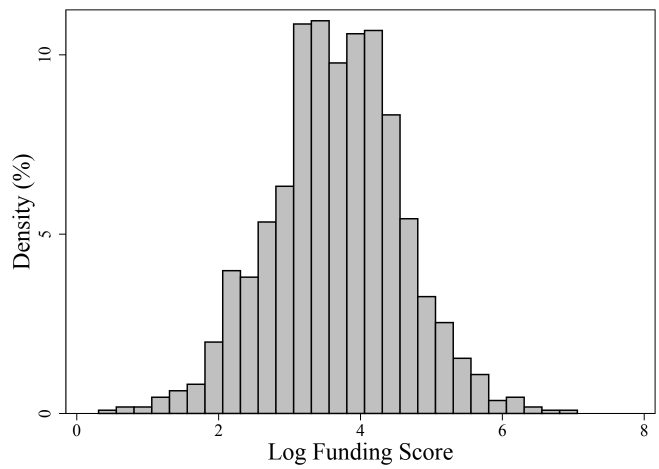
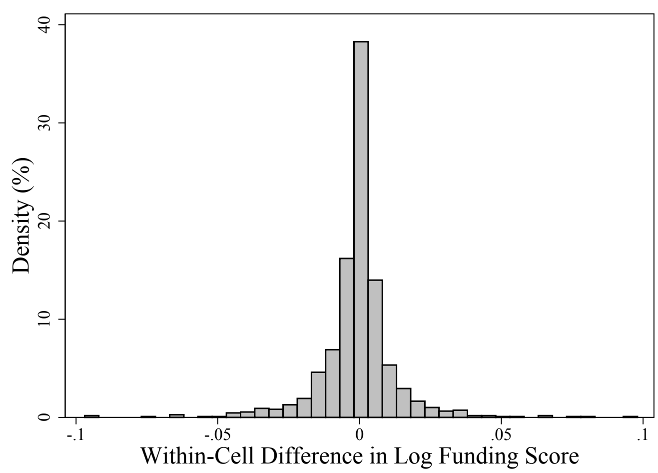


Figure 6: Distributions of Log Funding Score Across Units and Within Matched Pairs

(A) Distributions in the Main Sample



(B) Distributions of Within-Cell Differences



Appendix

A Classification of UK Universities by Types

This appendix lists the members of the University 'Type' Groups used in the analysis (excluding therefore institutions with fewer than three professors, or which did not make a submission to the REF, and the London Business School). These groupings are based the membership of University associations during the relevant period. For each institution, we give in brackets the number of panels to which a submission was made, the fulltime equivalent number of staff submitted, and the fulltime equivalent number of full professors in post in October 2013.

Russell Group: In 1994 a group of 17 'research intensive' UK universities formed an association known as the 'Russell Group', which grew to 20 Universities by 2006. In 2012 four additional universities joined from the newly-defunct '1994' group. Our classification of 'Russell Group' uses the 20 members from 2006 onwards on the basis that this group represents long-running core members.

University of Birmingham (29, 990, 411)
University of Bristol (26, 981, 487)
University of Cambridge (28, 1874, 711)
University of Leeds (26, 1015, 428)
University of Liverpool (16, 584, 253)
Imperial College, London (10, 1071, 612)
King's College London (19, 861, 496)
London School of Economics (12, 490, 249)
University College London (29, 2059, 902)
University of Newcastle-upon-Tyne (24, 809, 374)
University of Nottingham (23, 1053, 430)
University of Oxford (28, 2264, 905)
University of Sheffield (24, 770, 394)
University of Southampton (23, 878, 392)
University of Warwick (21, 907, 474)
University of Edinburgh (28, 1603, 558)
University of Glasgow (25, 887, 400)
Cardiff University (24, 679, 590)
The Queen's University of Belfast (21, 729, 193)
University of Manchester (30, 1426, 780)

The 1994 Group: The '1994' Group was also formed in 1994, its membership comprising smaller research-intensive universities that had not been invited to join the Russell Group. This group disbanded in 2012.

University of Bath (11, 414, 137)
University of Durham (20, 629, 250)
University of East Anglia (16, 353, 150)
University of Essex (10, 304, 131)
University of Exeter (19, 621, 230)
University of Lancaster (14, 503, 173)
University of Leicester (19, 576, 237)
Birkbeck College (11, 254, 92)
Queen Mary University of London (16, 556, 326)
Royal Holloway and Bedford (16, 360, 182)
School of Oriental and African Studies (8, 134, 78)
Loughborough University (13, 544, 154)
University of Reading (20, 537, 223)
University of Surrey (11, 362, 103)
University of Sussex (18, 443, 175)
University of York (23, 628, 264)
University of St Andrews (18, 498, 205)

"New" Universities: This group comprises institutions which were given status as universities from 1992 onwards. Prior to that time most of the members of this group were known as 'polytechnics' and delivered mainly post-high school technical education.

Buckinghamshire New University (1, 7, 6)
University of Chester (4, 45, 19)
Canterbury Christ Church University (2, 36, 8)
Edge Hill University (4, 55, 13)
Falmouth University (2, 50, 7)
Harper Adams University (1, 17, 3)
University of Winchester (1, 12, 4)
Liverpool Hope University (3, 38, 18)
University of Bedfordshire (7, 111, 47)
University of Northampton (1, 13, 3)
Roehampton University (6, 74, 34)
University of Worcester (1, 10, 3)
Anglia Ruskin University (5, 55, 35)
Bath Spa University (2, 26, 19)
Bournemouth University (4, 80, 35)
University of Brighton (6, 145, 29)
Birmingham City University (8, 98, 42)
University of Gloucestershire (4, 37, 15)
Coventry University (3, 44, 16)
University of East London (6, 78, 27)
University of Greenwich (3, 41, 12)
University of Hertfordshire (8, 138, 39)
University of Lincoln (9, 92, 34)
Kingston University (7, 115, 46)
Leeds Beckett University (5, 116, 31)
Liverpool John Moores University (9, 138, 47)
Manchester Metropolitan University (7, 220, 35)
Middlesex University (9, 257, 80)
De Montfort University (7, 136, 49)
Nottingham Trent University (7, 118, 58)
Oxford Brookes University (12, 204, 66)
University of Plymouth (13, 295, 101)
University of Portsmouth (5, 109, 25)
Sheffield Hallam University (7, 139, 37)
London South Bank University (1, 34, 4)
Teesside University (1, 16, 6)
University of West London (3, 29, 13)
University of the West of England (8, 192, 72)
University of Chichester (1, 8, 3)
University of Wolverhampton (7, 119, 37)
Cardiff Metropolitan University (2, 23, 9)
University of South Wales (4, 37, 16)
University of Abertay Dundee (1, 14, 3)
Queen Margaret University, Edinburgh (1, 8, 5)
Robert Gordon University (2, 20, 10)
Glasgow Caledonian University (5, 108, 41)
Edinburgh Napier University (5, 65, 28)
University of Ulster (14, 311, 123)
London Metropolitan University (2, 17, 9)

Specialists: This group comprises a set of high specialised universities offering a limited range of subjects including, in some cases, universities offering only a single subject.

Royal College of Art (1, 60, 7)
University of the Arts, London (1, 110, 27)
Royal Academy of Music (1, 14, 4)
Royal Northern College of Music (1, 11, 5)
Royal Veterinary College (1, 103, 33)
St George's Hospital Medical School (1, 44, 41)
London School of Hygiene and Tropical Medicine (1, 57, 66)
University for the Creative Arts (1, 21, 7)
Guildhall School of Music and Drama (1, 16, 105)

Others: Universities not included in any of the above groups are assigned to this 'other' group.

The Open University (12, 325, 96)
Cranfield University (3, 224, 50)
University of Central Lancashire (11, 201, 52)
University of Huddersfield (10, 144, 61)
University of Westminster (8, 138, 53)
Aston University (5, 164, 45)
University of Bradford (4, 93, 44)
Brunel University London (15, 452, 97)
City University (10, 316, 141)
University of Hull (10, 244, 54)
University of Keele (13, 239, 71)
University of Kent (18, 500, 135)
Goldsmiths College (10, 210, 82)
University of Salford (8, 185, 71)
University of Strathclyde (14, 508, 174)
University of Aberdeen (19, 475, 236)

Heriot-Watt University (8, 212, 75)
University of Dundee (11, 287, 110)
University of Stirling (12, 239, 83)
Aberystwyth University (11, 202, 71)
Bangor University (2, 70, 17)
Swansea University (14, 330, 149)

B Proofs

Proof of Lemma 1. The problem of department i is:

$$\max_{L_1,L_2} \ln\theta_k + \alpha_1 \ln L_1 + \alpha_2 \ln L_2 + \beta_i \ln K,$$
$$\text{s.t.: } rK = B_i - w_1 L_1 - w_2 L_2,$$
$$L_\ell = \mu_\ell w_\ell, \qquad \ell = 1, 2.$$

Substituting the constraints into the maximand, we can write this problem as:

$$\max_{L_1,L_2} \left\{ \ln\theta_k + \alpha_1 \ln L_1 + \alpha_2 \ln L_2 + \beta_i \ln\left(B_i - \frac{L_1^2}{\mu_1} - \frac{L_2^2}{\mu_2}\right) - \beta_i \ln r \right\}.$$

Solving the first order conditions of the above gives

$$L_\ell = \sqrt{\frac{\alpha_\ell \mu_\ell}{c_i} B_i}, \qquad \ell = 1, 2. \tag{A1}$$

It is easily checked that the second order conditions are satisfied. Substitute (A1) into (3) and using definitions (4) and (5), gives

$$K = \frac{2\beta_i}{c_i}\frac{B_i}{r}.$$

the expression for the level of capital. The total research output is also obtained by direct substitution as

$$\rho^*_{k,i}(B_i) = \theta_k A_i \left(\frac{B_i}{c_i}\right)^{\frac{c_i}{2}}. \tag{A2}$$

□

Proof of Lemma 2. We need to substitute (A1) into (1), and carry out simple calculations to find that salaries for the two types of academics in department i in institution k are given by (10), and the Gini coefficent by (11). □

Proof of Corollary 1. The first order conditions for the Lagrangian of problem (14)-(15) are

$$\frac{1}{2} u_i \theta_k A_i c_i^{1-\frac{1}{2}c_i} B_i^{\frac{1}{2}c_i - 1} - \lambda_k \left(1 - \frac{1}{2}\gamma_i \theta_k A_i c_i^{1-\frac{1}{2}c_i} B_i^{\frac{1}{2}c_i - 1}\right) = 0, \qquad i = 1, \ldots, n.$$

Rearranging, we derive (16). For this condition to identify a maximum, $\frac{1}{4} B_i^{\frac{1}{2}c_i - 2}(c_i - 2)\, c_i^{1-\frac{1}{2}c_i}$ must be negative, which is the case if $c_i < 2$, that is if $\beta_i < 1 - \frac{\alpha_1+\alpha_2}{2}$, as assumed. □

C Additional tables and figures

Table A1: Summary Statistics Departmental Level REF Performance by Component

	mean	sd	min	max
Overall % 4*	26.88	14.66	0.00	79.00
Overall % 3*	47.46	11.89	3.00	83.00
Overall % 2*	21.95	12.08	0.00	75.00
Overall % 1*	3.23	5.23	0.00	55.00
Overall % 0*	0.48	1.29	0.00	10.00
Outputs % 4*	21.79	11.24	0.00	69.70
Outputs % 3*	48.44	12.23	0.00	100.00
Outputs % 2*	25.44	12.05	0.00	72.90
Outputs % 1*	3.82	5.68	0.00	60.60
Outputs % 0*	0.50	1.15	0.00	10.30
Environment % 4*	34.03	33.46	0.00	100.00
Environment % 3*	47.83	27.50	0.00	100.00
Environment % 2*	16.39	23.39	0.00	100.00
Environment % 1*	1.73	8.32	0.00	90.00
Environment % 0*	0.03	0.81	0.00	25.00
Impact % 4*	38.00	28.04	0.00	100.00
Impact % 3*	43.98	22.94	0.00	100.00
Impact % 2*	14.93	19.48	0.00	100.00
Impact % 1*	2.32	8.26	0.00	90.00
Impact % 0*	0.77	4.82	0.00	40.00

Note: Sample size = 1093 departments submitted to REF 2014. For explanation of REF performance measures see main text.

Table A2: OLS Regression Estimates: Department Pay Characteristics and REF 2014 Performance. Dependent Variable: Grade Point Average Score

	(1)	(2)	(3)
Log Average Salary	0.430***	0.715***	0.433***
	(0.074)	(0.073)	(0.071)
Gini Coefficient Salary	0.698***	0.581**	0.194
	(0.246)	(0.231)	(0.214)
Log Professorial FTE	0.032*	0.112***	0.080***
	(0.017)	(0.017)	(0.016)
Log REF FTE	0.117***	0.093***	0.059***
	(0.018)	(0.018)	(0.016)
Panel Member = 1	0.109***	0.072***	0.047***
	(0.018)	(0.016)	(0.015)
Log Vice Chancellor Pay	0.222***	0.147***	0.096**
	(0.043)	(0.040)	(0.037)
% Age under 40	0.040	0.245***	0.163*
	(0.092)	(0.089)	(0.082)
% Age 41-50	-0.032	0.071	0.037
	(0.052)	(0.050)	(0.047)
% Age 51-60	-0.041	0.009	-0.056
	(0.050)	(0.048)	(0.044)
1994 Group Uni. = 1			-0.013
			(0.019)
Other Type Uni. = 1			-0.171***
			(0.022)
New Uni = 1			-0.347***
			(0.027)
Specialist Uni = 1			-0.083
			(0.088)
Unit of Assessment FE	NO	YES	YES
University Type FE	NO	NO	YES
R-squared	0.343	0.477	0.563
Observations	1093	1093	1093

Notes: Dependent variable is GPA score. Column 1 includes log average salary and log sd salary only. Additional columns add control variables as described in table header. Vice Chancellor pay variable omitted from Column 5 as it is collinear with fixed effects. Standard errors in parenthesis. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

Table A3: OLS Regression Estimates by Main Panel (Grade Point Average Score)

	(1) Panel A Medicine	(2) Panel B Science & Engineering	(3) Panel C Social Sciences	(4) Panel D Arts & Humanities
Log Average Salary	0.753***	0.482***	0.422***	0.756***
	(0.201)	(0.154)	(0.126)	(0.134)
Gini Coefficient Salary	1.204*	1.066**	0.090	0.451
	(0.566)	(0.432)	(0.440)	(0.483)
Log Professorial FTE	-0.038	0.163***	0.019	0.115***
	(0.030)	(0.036)	(0.031)	(0.034)
Log REF FTE	0.122***	0.032	0.208***	0.015
	(0.037)	(0.038)	(0.033)	(0.033)
Panel Member = 1	0.080*	0.080**	0.124***	0.106***
	(0.042)	(0.033)	(0.031)	(0.031)
Log Vice Chancellor Pay	0.219*	0.177*	0.237***	0.115
	(0.106)	(0.082)	(0.076)	(0.072)
% Age under 40	-0.074	0.417*	0.167	0.068
	(0.256)	(0.212)	(0.141)	(0.168)
% Age 41-50	0.056	-0.006	0.098	-0.052
	(0.143)	(0.109)	(0.091)	(0.083)
% Age 51-60	0.046	-0.121	0.074	-0.073
	(0.128)	(0.113)	(0.089)	(0.078)
University Type FE	YES	YES	YES	YES
R-squared	0.418	0.497	0.398	0.320
Observations	183	263	369	278

Note: Dependent variable is GPA score. OLS regression estimated on four mutually exclusive samples of academic departments categorised by REF Main Panel. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

Table A4: OLS Regression Estimates Omitting Medicine UOA

	(1) Funding Score	(2) GPA Score
Log Average Salary	0.497***	0.426***
	(0.092)	(0.072)
Gini Coefficient Salary	0.530*	0.155
	(0.276)	(0.217)
Log Professorial FTE	0.111***	0.080***
	(0.021)	(0.016)
Log REF FTE	1.058***	0.064***
	(0.021)	(0.017)
Panel Member = 1	0.061***	0.049***
	(0.020)	(0.016)
Log Vice Chancellor Pay	0.109*	0.090**
	(0.048)	(0.038)
% Age under 40	0.222*	0.163*
	(0.104)	(0.082)
% Age 41-50	0.050	0.032
	(0.060)	(0.047)
% Age 51-60	-0.059	-0.061
	(0.056)	(0.044)
1994 Group Uni. = 1	-0.027	-0.015
	(0.024)	(0.019)
Other Type Uni. = 1	-0.207***	-0.172***
	(0.028)	(0.022)
New Uni = 1	-0.420***	-0.351***
	(0.034)	(0.027)
Specialist Uni = 1	-0.118	-0.108
	(0.133)	(0.104)
Unit of Assessment FE	YES	YES
University Type FE	YES	YES
R-squared	0.906	0.565
Observations	1067	1067

Note: * Denotes significance at 10% level, ** 5% level, *** at 1% level.

Table A5: Difference-in-Difference Estimates of Effects of REF 2014 Funding Score on Changes in Departmental Wage Structure, 2013 - 2016

	(1) Change Log Mean Salary	(2) Change Log Top Quartile Wage Bill	(3) Change Gini Coeff.	(4) Change Total FTE Professors
Funding Score	0.002	-0.028	0.005	1.001
	(0.005)	(0.119)	(0.004)	(0.758)
% Age under 40	-0.053	2.642***	0.021	6.970***
	(0.037)	(0.895)	(0.017)	(1.950)
% Age 41-50	-0.032*	0.245	-0.011	1.144
	(0.017)	(0.423)	(0.008)	(0.921)
% Age 51-60	-0.001	0.403	-0.007	-0.087
	(0.016)	(0.393)	(0.007)	(0.855)
Log REF FTE	-0.002	0.050	-0.003	-0.525
	(0.007)	(0.169)	(0.003)	(0.368)
Main Panel FE	YES	YES	YES	YES
R-squared	0.013	0.016	0.019	0.068
Observations	562	562	562	562

Note: Table shows OLS estimates from difference-in-difference models based on Equation 22. The dependent variable and funding score variables are the differenced variables, calculated as the differences between matched pairs (by GPA and FTE) in the first-differences in the variables between 2013 and 2015. Other control variables are differences between the matched departments in levels measured in the 2013 data. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

Table A6: Difference-in-Difference Estimates of Effects of REF 2014 GPA on Changes in Departmental Wage Structure, 2013 - 2016

	(1) Change Log Mean Salary	(2) Change Log Top Quartile Wage Bill	(3) Change Gini Coeff.	(4) Change Total FTE Professors
GPA Score	-0.073	0.067***	0.011	6.226***
	(0.098)	(0.023)	(0.011)	(1.173)
% Age under 40	0.859	-0.091***	0.012	4.540***
	(0.895)	(0.034)	(0.016)	(1.756)
% Age 41-50	0.294	-0.007	0.002	1.656*
	(0.437)	(0.017)	(0.008)	(0.857)
% Age 51-60	0.647	0.005	0.002	-0.614
	(0.431)	(0.016)	(0.008)	(0.845)
Log REF FTE	0.073	0.012*	0.002	-0.604*
	(0.140)	(0.005)	(0.002)	(0.276)
Main Panel FE	YES	YES	YES	YES
R-squared	0.008	0.031	0.004	0.039
Observations	571	571	571	571

Note: Table shows OLS estimates from difference-in-difference models based on Equation 22. The dependent variable and funding score variables are the differenced variables, calculated as the differences between matched pairs (by Funding Score) in the first-differences in the variables between 2013 and 2015. Other control variables are differences between the matched departments in levels measured in the 2013 data. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

Table A7: OLS Regression Estimates: Department Pay Characteristics and REF 2014 Performance. Dependent Variable: Grade Point Average Score

	(1)	(2)	(3)
Log Average Salary	0.435***	0.728***	0.444***
	(0.074)	(0.073)	(0.071)
Gini Coefficient Salary	0.678***	0.552**	0.176
	(0.246)	(0.230)	(0.213)
Log Professorial FTE	0.034*	0.116***	0.083***
	(0.017)	(0.017)	(0.016)
Log REF FTE	0.116***	0.091***	0.057***
	(0.018)	(0.017)	(0.016)
Panel Member = 1	0.107***	0.068***	0.044***
	(0.018)	(0.016)	(0.015)
Log Vice Chancellor Pay	0.226***	0.155***	0.099***
	(0.043)	(0.040)	(0.037)
% Age under 40	0.042	0.251***	0.168*
	(0.092)	(0.088)	(0.081)
% Age 41-50	-0.028	0.078	0.043
	(0.052)	(0.050)	(0.046)
% Age 51-60	-0.044	0.007	-0.057
	(0.050)	(0.047)	(0.044)
Proportion PT Staff	0.100*	0.174***	0.122**
	(0.055)	(0.051)	(0.048)
1994 Group Uni. = 1			-0.020
			(0.019)
Other Type Uni. = 1			-0.169***
			(0.022)
New Uni = 1			-0.346***
			(0.027)
Specialist Uni = 1			-0.097
			(0.088)
Unit of Assessment FE	NO	YES	YES
University Type FE	NO	NO	YES
R-squared	0.345	0.482	0.565
Observations	1093	1093	1093

Notes: Dependent variable is GPA score. Column 1 includes log average salary and log sd salary only. Additional columns add control variables as described in table header. Vice Chancellor pay variable omitted from Column 5 as it is collinear with fixed effects. Standard errors in parenthesis. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

Table A8: OLS Regression Estimates: Department Pay Characteristics and REF 2014 Performance. Dependent Variable: REF Funding Score

	(1)	(2)	(3)
Log Average Salary	0.500***	0.858***	0.524***
	(0.096)	(0.092)	(0.091)
Gini Coefficient Salary	0.891***	0.996***	0.546*
	(0.317)	(0.289)	(0.271)
Log Professorial FTE	0.039*	0.154***	0.114***
	(0.021)	(0.022)	(0.020)
Log REF FTE	1.115***	1.089***	1.048***
	(0.023)	(0.022)	(0.021)
Panel Member = 1	0.140***	0.085***	0.056***
	(0.023)	(0.021)	(0.019)
Log Vice Chancellor Pay	0.285***	0.188***	0.121**
	(0.055)	(0.050)	(0.047)
% Age under 40	-0.012	0.326***	0.229*
	(0.118)	(0.111)	(0.104)
% Age 41-50	-0.115*	0.105*	0.065
	(0.067)	(0.063)	(0.059)
% Age 51-60	-0.085	0.021	-0.054
	(0.064)	(0.060)	(0.056)
Proportion PT Staff	0.114	0.209***	0.150**
	(0.071)	(0.064)	(0.061)
1994 Group Uni. = 1			-0.032
			(0.024)
Other Type Uni. = 1			-0.203***
			(0.028)
New Uni = 1			-0.414***
			(0.034)
Specialist Uni = 1			-0.089
			(0.112)
Unit of Assessment FE	NO	YES	YES
University Type FE	NO	NO	YES
R-squared	0.867	0.899	0.913
Observations	1093	1093	1093

Notes: Dependent variable is natural log of the research funding score. Column 1 includes log average salary and log sd salary only. Additional columns add control variables as described in table header. Vice Chancellor pay variable omitted from Column 5 as it is collinear with fixed effects. Standard errors in parenthesis. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

Table A9: Difference-in-Difference Estimates of Effects of REF 2014 Funding Score at Main Panel Level on Changes in Departmental Wage Structure, 2013 - 2016, Sample of Closest Matches Only

	(1) Change Log Mean Salary	(2) Change Log Top Quartile Wage Bill	(3) Change Gini Coeff.	(4) Change Total FTE Professors
Main Panel FS (Incl. Dept)	0.000	-0.010	0.000	-0.025
	(0.000)	(0.008)	(0.000)	(0.017)
% Age under 40	-0.010	1.680	0.023	8.009***
	(0.059)	(1.317)	(0.023)	(2.766)
% Age 41-50	-0.049*	0.776	-0.006	0.856
	(0.028)	(0.622)	(0.011)	(1.306)
% Age 51-60	0.004	0.283	-0.003	-0.503
	(0.025)	(0.553)	(0.010)	(1.161)
Log REF FTE	-0.005	0.499	-0.002	0.040
	(0.015)	(0.330)	(0.006)	(0.692)
Main Panel FE	YES	YES	YES	YES
R-squared	0.022	0.018	0.007	0.056
Observations	269	269	269	269

Note: Table shows OLS estimates from difference-in-difference models based on Equation 22. The dependent variable and funding score variables are the differenced variables, calculated as the differences between matched pairs (by GPA and FTE) in the first-differences in the variables between 2013 and 2015. In this regression the Funding Score is differences at the level of the Main Panel. Other control variables are differences between the matched departments in levels measured in the 2013 data. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

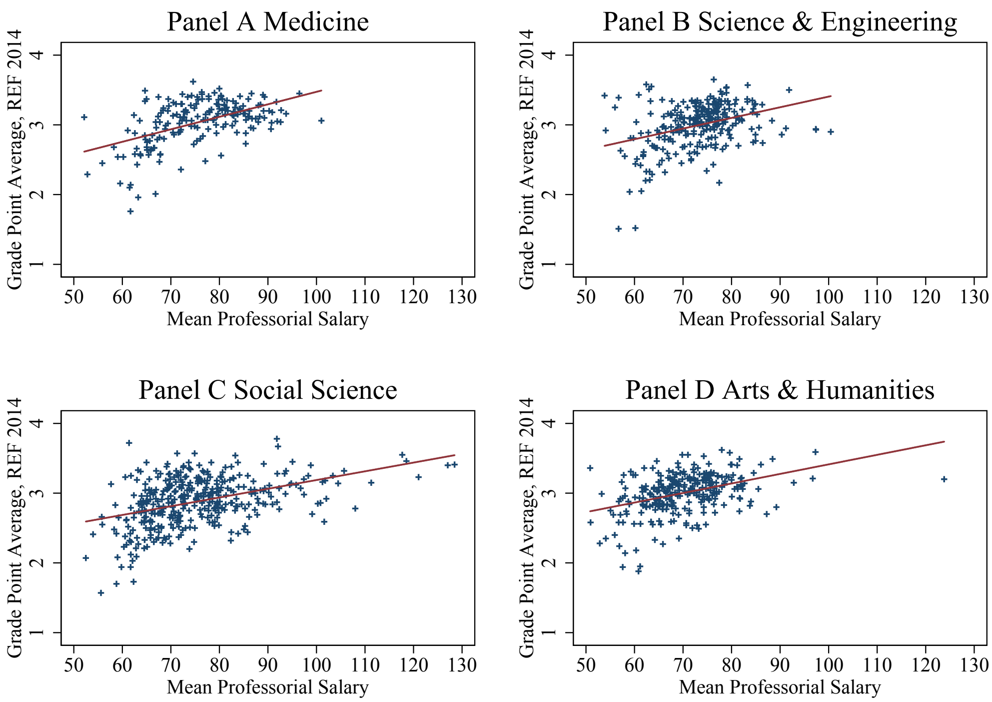
Table A10: Difference-in-Difference Estimates of Effects of REF 2014 GPA at Main Panel Level on Changes in Departmental Wage Structure, 2013 - 2016, Sample of Closest Matches Only

	(1) Change Log Mean Salary	(2) Change Log Top Quartile Wage Bill	(3) Change Gini Coeff.	(4) Change Total FTE Professors
Main Panel GPA Score (Incl. Dept)	-2.222	0.053*	-0.011	3.740**
	(2.544)	(0.022)	(0.046)	(1.436)
% Age under 40	-0.158	-0.034	0.029	2.150
	(1.324)	(0.053)	(0.024)	(2.831)
% Age 41-50	-0.744	-0.003	-0.003	0.211
	(0.649)	(0.026)	(0.012)	(1.387)
% Age 51-60	0.786	0.013	-0.002	-0.971
	(0.638)	(0.026)	(0.011)	(1.363)
Log REF FTE	-0.176	0.015*	-0.002	0.186
	(0.217)	(0.009)	(0.004)	(0.464)
Main Panel FE	YES	YES	YES	YES
R-squared	0.036	0.014	0.008	0.011
Observations	259	259	259	259

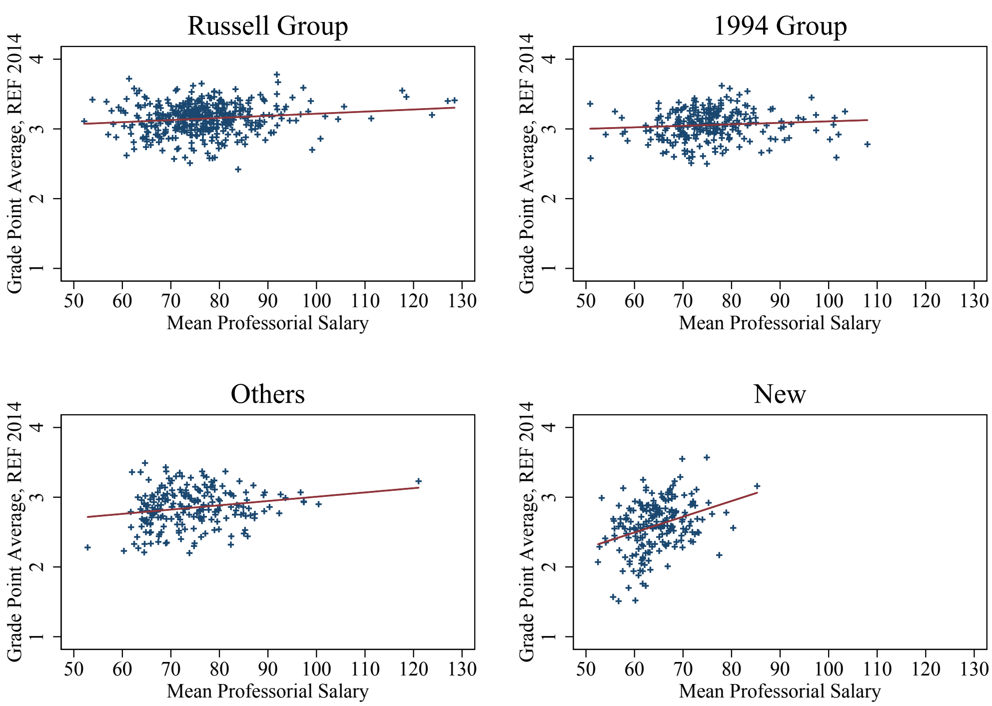
Note: Table shows OLS estimates from difference-in-difference models based on Equation 22. The dependent variable and funding score variables are the differenced variables, calculated as the differences between matched pairs (by Funding Score) in the first-differences in the variables between 2013 and 2015. In this regression the GPA Score is differences at the level of the Main Panel. Other control variables are differences between the matched departments in levels measured in the 2013 data. * Denotes significance at 10% level, ** 5% level, *** at 1% level.

Figure A1: Correlation Between Mean Pay and GPA Score

(A) By REF Main Panel



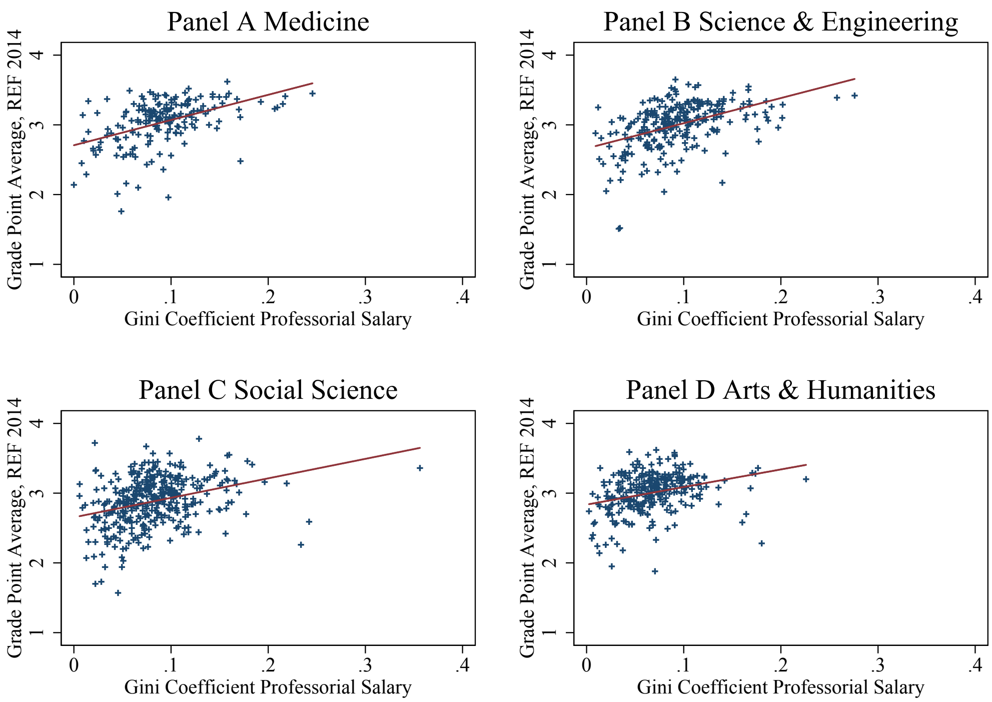
(B) By University Type



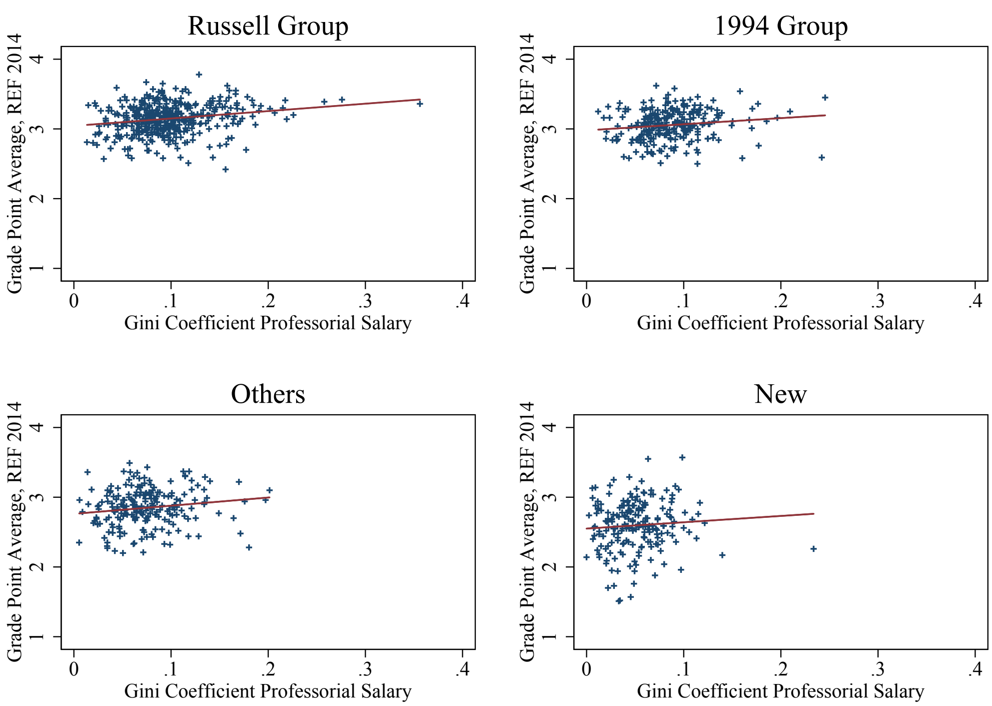
Note: Each observation represents an individual academic department. Figure show scatter plots and fitted regression lines. Observations grouped by REF Main Panel (Plot group A) and University Type (Plot group B).

Figure A2: Correlation Between Gini Pay and GPA Score

(A) By REF Main Panel

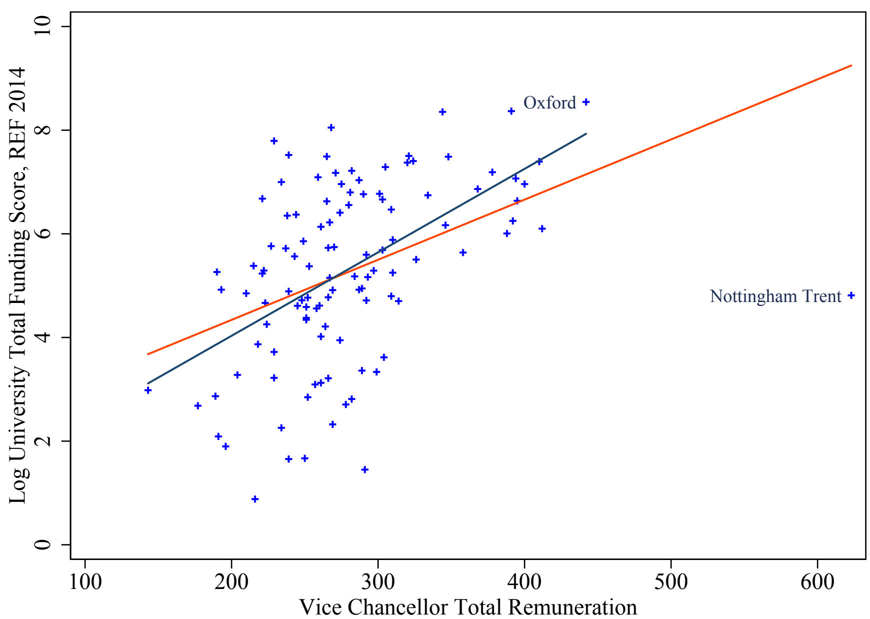


(B) By University Type



Note: Each observation represents an individual academic department. Figure show scatter plots and fitted regression lines. Observations grouped by REF Main Panel (Plot group A) and University Type (Plot group B).

Figure A3: Vice Chancellor Pay and Funding Score Performance



Note: Each observation corresponds to an individual university. Figure shows a scatter plot of total remuneration of university vice chancellors (x-axis, including pension contributions and discretionary payments) and log funding score (y-axis). The red fitted regression line is estimated on all observations; the blue fitted regression line is fitted on observations excluding the far-right outlier value.