

(How) Do research and administrative duties affect university professors' teaching?*

Aurora García-Gallego

LEE & Economics Department-Universitat Jaume I, Castellón, Spain

Nikolaos Georgantzís

Economics Department, Universidad de Granada & LEE-Universitat Jaume I, Castellón, Spain

Joan Martín-Montaner

IEI & Economics Department-Universitat Jaume I, Castellón, Spain

Teodosio Pérez-Amaral

Quantitative Economics Department-Universidad Complutense, Madrid, Spain

Abstract

We analyze the interaction between university professors' teaching quality and their research and administrative activities. Our sample is a high-quality individual panel data set from a medium size public Spanish university. Although, researchers teach roughly 20% more than non-researchers, their teaching quality is also 20% higher. Over much of the relevant range, we find a nonlinear and positive effect of research output and teaching quantity on teaching quality. Instructors with no research are 5 times more likely than the rest to be among the worst teachers and up to two-thirds of the professors could improve their teaching by increasing research.

* We would like to express our sincere thanks to the Universitat Jaume I for granting us access to the data. Without their careful and systematic collection of information on faculty duties and performance, this study would not have been possible. The authors want to thank Michael McAleer for his detailed comments and proofreading. Thanks are also due to Covadonga Gijón and Iñaki Iriondo for helpful suggestions. Financial support by the *Spanish Ministry of Economics and Perspective* (projects ECO2011-23634, ECO2008-06191 and ECO2011-27619), *Junta de Andalucía* (P07-SEJ-03155), and *UJI-Bancaixa* (project P1-1B2010-17) is gratefully acknowledged.

I. Introduction

Universities exist since the 12th century. Over this period, they have experienced great changes regarding both their functions and the way they are run, as recently summarized in Martin (2012). For the last century and a half, there has been a widespread idea that universities encompass functions which, following modern theories of the firm¹, should exhibit some kind of complementarity: (a) to increase current knowledge through research and (b) to spread it to the new generations through teaching. This agreement on the functions of the university as an institution does not mean that there is only one way of running it. In fact, in some relevant cases a strategy of teaching specialization has been adopted², while in other cases research has been undertaken in research oriented institutes. However, most universities world-wide seem to adopt a clear cut approach based on the complementarity between teaching and research, motivating in most of the cases their faculty members to perform a mix of the aforementioned tasks together with a certain amount of administrative duties.

Taking this observation as a positive indication on the perceived synergic relation between different types of academic duties, the normative question is whether the actual mix of the corresponding inputs and outputs is the optimal, or, alternatively, whether it could be improved by a redefinition of the obligations and objectives set by governments and university authorities. From an individual academic's point of view, a professor's choice to do research and a researcher's choice to teach might simply be dictated by intrinsic synergies arising from the joint dedication to both duties resulting in the provision of higher quality output. From the interplay among intrinsic (individual, supply-driven) and extrinsic (institutional, demand-driven) motivators and the natural time constraint binding an academic's dedication to both activities, two opposing effects emerge regarding the complementarity or substitutability among tasks: a quantity of available time effect on teaching, versus a quality effect on teaching. This latter statement leads to an alternative approach based on individual specialization. In this scenario, although each department should provide high-quality teaching and high-quality research, this does not necessarily mean that each faculty member should implement both. A proper answer to this dilemma requires confirming whether these perceived synergies really hold.

¹ See, for example, Spiller and Zelner (1997) and Lindbeck and Snower (2003).

² In various countries, we can find higher education institutions that have focused only on the teaching side, such as the *Grandes Écoles* in France, the liberal arts colleges in the United Kingdom or the *Fachhochschulen* in Germany.

It is often assumed that research may have positive spillovers on teaching because it facilitates an up-to-date choice, a deeper understanding of topics and a more rigorous approach to the subjects taught. Furthermore, the time devoted by researchers to their students may have a higher quality than that devoted to them by non-researchers. Research on the true nature of this statement can be generally qualified as inconclusive. On one hand, a number of studies share this common aim but differ in “the variables investigated, their measurement, as well as the investigated population” as pointed by Verburgh, Elen and Lindblom-Ylänne (2007), thus lacking a common framework. On the other hand, and despite this lack of consensus, most studies conclude that a very weak relation between research and teaching can be found, or even no relation at all, see Marsh and Hattie (2002).

In this paper, we turn to a particularly rich annual panel data set of 604 individual university professors, available over the period 2002-2006. The panel was extracted from the staff files of the University. The sample of professors covers a variety of different disciplines including humanities, social sciences, economics, management, natural sciences and engineering. We use several indicators of the quantity and quality of teaching, research output, and the amount of administrative duties performed.

Summarizing our results, we find that professors with a typical research output are somewhat better teachers than professors with less research. Moreover, non researchers are five times more likely than researchers to be poor teachers. In general, the quality of university-level teaching is positively affected by published research across most levels of research output. The remainder of the paper is organized as follows. Section II consists in a review of the previous literature and the current situation of the debate. Section III proposes a theoretical benchmark. Section IV presents the database. The econometric estimations and discussion of the main results are contained in section V. Section VI concludes.

II. Previous studies

The existence and the nature of a relation between research productivity and teaching performance of faculty has been for years at the core of a strong debate concerning the design of universities all over the world. To which extent and how are these activities linked? Although the existing literature offers reasonable justifications for both a positive and a negative relation between them, to the best of our knowledge, there is little, if any, evidence on the links among all the professorial activities as required by the multidimensional description of the input and the output spaces.

On the one hand, common wisdom usually leads us to think that the best researchers should also excel in teaching, provided they are enlarging the bulk of knowledge on a given discipline that should be later taught at the classroom. Successful research also implies several individual characteristics such as higher interest in learning and more ability to motivate students. Besides, research requires being more organized, which in turns helps to be more effective to hand down knowledge. On the other hand, it could be argued that good research requires a high degree of specialization which is not compatible with the broad view that students expect to find in their instructors. In fact, Friedrich and Michalak Jr. (1983) conclude that the higher organizational capability of researchers does not compensate the fact that students perceive them as less knowledgeable than instructors that do not research. A more extensive discussion on this topic can be found in Friedrich and Michalak Jr. (1983) or in Marsh and Hattie (2002).

The answer to this dilemma, therefore, becomes an empirical issue. However, empirical research is not as extensive as one would have thought and has not provided yet conclusive results on this hot topic. In a classic meta-analysis, Feldman (1987) reviews 29 papers published between 1950 and 1984 to conclude that “on average, there is a very small positive association between the two variables”. Hattie and Marsh (1996) also rely on a meta-analysis based on 58 contributions to obtain a more pessimistic conclusion, openly speaking of the relation between research and teaching as an “enduring myth”.

More recent research does not establish anything close to a stylized fact, as it also provides mixed evidence. Thus, Noser, Manakyan and Tanner (1996) differentiate the effect of research depending on the level of the courses taught: a positive but small relation is found for undergraduate level, whereas mixed results appear at the graduate level. Shin (2011) finds different signs depending on the indicator used for research productivity. More conclusive, Marsh and Hattie (2002) find a close-to-null relation between the two activities. The papers above show that a crucial point in this empirical literature comes from the way research productivity and teaching performance are measured. Regarding the case of teaching performance, student’s evaluations of teaching (SET) are in most cases the only indicator available. An obvious shortcoming is that it does not measure the improvement in the student’s knowledge, but, in the best of the cases, the student’s own perception of this improvement. This perception is affected by the teacher’s performance, but also by their own expectations about their outcome in the courses.³ If students are more sensible to

³ There is a huge literature analyzing the determinants of the SETs. In a classic survey, Marsh (1987) concludes favorably about the validity and reliability of student’s evaluations. However, later empirical research points out that SETs can be affected by both subjective and objective issues. Thus, Krautman and Sander (1999), Isely and Singh (2005) or Ewing (2012) find and impact of expected grade on SETs, whereas Bedard and Kuhn (2008) conclude that

expected grades than to their own learning, active researchers (which are supposed to be stricter) could be punished in SETs despite having a positive effect in student's learning.

For research productivity, most of the published empirical research uses the number of publications as the main indicator. However a few papers opt for qualitative indicators such as citation counts (Stack, 2003, among the most recent) or the quality of the journals (Noser, Manakyan and Tanner, 1996). Not surprisingly, this is not an innocuous option. In Feldman (1987), papers using quantitative indicators show a positive but weak relation with SETs, whereas in most cases, those based on citation counts do not get any. Neither do subsequent papers such as Gomez-Mejia and Balkin (1992); however, these authors report a positive correlation between teaching evaluations and publications in top-tier journals, and no correlation at all in the case of books. The opposite outcome can be found in Shin (2011). This author performs a covariance analysis to find that teaching is positively related to books and domestic journal publication and negatively related to international journal publications.⁴ Besides, this former association is found to be stronger in some academic disciplines.

The results above take to the forefront the heterogeneity existing in the research field. This heterogeneity has been analyzed in the literature at two levels: academic disciplines and individual researchers. First, two different papers focus on the existence of heterogeneity at academic discipline level, offering, once again, mixed evidence. Both Porter and Umbach (2001) and Marsh and Hattie (2002) rely on a multilevel model which allows to differentiate the relationship between teaching and research across departments. The results were hardly comparable, as the latter used SETs as the dependent variable, whereas the former tried to explain research productivity, being the teaching load one of the explanatory variables. The interesting point, however, is that academic discipline did matter for the former, whereas it did not for the latter.

A second aspect regarding heterogeneity in research productivity refers to the obvious fact that the number of publications is not evenly distributed across faculty. On the contrary, top researchers concentrate a high share of the total number of publications, which implies the appearance of skewness in the research productivity distribution. Stack (2003) reports a positive and significant

class size has a negative impact, although McPherson (2006) delimits this impact at the principles level. On the other hand, some recent papers question the link between SETs and teacher quality (Carrell and West, 2010) or effective learning by students (Beleche, Fairris and Marks, 2012).

⁴ This author interprets his result in terms of the narrowness associated to good researchers (who mostly publish in international journals). Therefore, according to him, this outcome is reflecting the role played by those abilities which characterize good and bad researchers/instructors.

link between research and teaching, by introducing non linearity in the estimation process against the linear relationship assumed by most empirical papers. This represents a promising line of research which will be followed in the present paper.

To summarize the literature reviewed in this section, we could conclude that, with very few exceptions, there is some evidence of a weak relation between teaching and research. Some recent papers have focused on asking why this weakness in the relationship. Answers have ranged from emphasizing methodological issues to pointing the relevance of context in explaining the variance in the relation. All these approaches raise interesting issues which are developed to different degrees in the analysis we present in the following sections.

III. Framework

A. Professors as instructors, researchers and managers

Most academic positions in universities imply three types of duties, namely teaching courses, conducting research in their respective fields, and performing administrative duties. There is some flexibility as to how professors allocate their time between the three tasks (some instructors also work outside the university, for example, as consultants). Let us summarize some statements regarding all three activities that will help to define the theoretical framework for our analysis and, subsequently, interpret our results.

With regard to the amount of teaching for a given position, it is somewhat fixed, except that it can be reduced by performing administrative duties that carry a statutory reduction in the teaching load. The amount of teaching is monitored by academic authorities.

The quality of teaching is usually not a crucial issue for promotion, so instructors do not have a strong incentive to deliver good lectures as they do have for publishing their research. Instead, they usually have to maintain a minimum standard in the student satisfaction surveys.⁵

In the case of research, instructors have incentives to perform and publish research because it has an important effect on their promotion. Unless they are already at the top of their academic careers, the salary is somewhat increased by the research output. However, it does not have the

⁵ At the university used in this research, teachers are reprimanded if they have low scores in the student reviews.

compulsory nature of teaching, so that many full time university teachers do little or no research.⁶ Even high quality research does not mean an automatic reduction in the teaching load.

The administrative load of each individual is variable and negotiable. It usually requires acquiescence in the part of the instructor. However, administrative duties usually carry a salary increase, a statutory teaching load reduction and, sometimes, an additional guaranteed sabbatical year and a higher status within the university community.

B. A model

Consider the following straightforward extension of Becker's (1975) professorial decision making model, adding to the original research-teaching setting a third type of output, university administration.

The model assumes that a professor i maximizes the utility function $U_i(Q_T, Q_R, Q_A, Q_C)$, whose arguments correspond to the four outputs Q_{ji} , (where $j = \text{Teaching, Research, Administration, Consumption}$) yielded by the professor's actions. The professor decides on the time spent on each one of the four activities. The four outputs monotonically increase in the time spent on each one of them, although potentially synergic activities like teaching and research may mutually benefit each other and the time spent on one of them may raise the output of the other. In fact, due to this interaction between potentially synergic activities, net complementarity or substitutability may be observed among them, despite the clear cut net substitution effect among the times dedicated to the generation of the four outputs as dictated by the time identity: $T_{R_i} + T_{T_i} + T_{A_i} + T_{C_i} = T_i$. Note that, originally, the model's comparative statics aimed at yielding testable hypotheses for empirical and normative analysis of different rewards to teaching and research. Nevertheless, the model's implications for technological shocks affecting the relative shadow prices of the outputs to an individual professor are immediately applicable to the heterogeneity usually observed among individuals or groups of professors, provided that their idiosyncratic and institutional differences are interpreted as between-professor changes in the parameters α_{jji} , α_{jki} of the expression $Q_{ji} = \alpha_{jji} \cdot T_{ji} + \alpha_{jki} \cdot T_{ki}$, specifying professor-specific direct and cross-activity impacts of the time dedicated, respectively, to activities j and k on the output of j . As noted by Becker (1975), the indeterminate effect of one activity on the other allows for multiple predictions concerning the

⁶ There are many reasons for not doing research, like lack of training, lack of infrastructure, lack of a research culture and group atmosphere, lack of interest, personal inability or lack of incentives (e.g. a high opportunity cost of time).

substitutability or complementarity among a professor's outputs, which here include administrative tasks. Thus, it is left open for empirical research to determine the signs of such cross-activity effects.

IV. The data

The dataset comprises yearly panel data which contain information on 69 academic variables for each of the 604 individual instructors of the University for a period of five years. The reference period is 2002-2006, during which the University has systematically compiled individual information on a large number of variables concerning each faculty member's teaching quality and quantity, together with information on their research performance and administrative duties. Since all the instructors of the university are required to undergo teaching evaluations, we avoid a possible self-selection bias. The first 23 variables correspond to indices of teaching. The next 14 are indices related to administrative duties, and the last 32 indices of research. The 604 professors are from 25 academic departments covering a variety of different disciplines including humanities, social sciences, economics, management, biomedical, natural sciences and engineering.

As we are interested in relating the quality of teaching with administrative duties, research, and other variables, we use the following synthetic indices to approximate the concepts of interest:⁷

Teaching effectiveness: Denoted by *Teachqual*, teaching effectiveness is assessed with students' evaluations of teaching based on a standard university survey evaluating on a scale between 0 and 9 a professor's performance.

Research outcome: Denoted by *Research1*, the research attainment of each academic in the university is available through an online application managed by each department. An independent panel assesses all outputs and applies a similar set of rules for all departments as to what counts as research productivity. The contributions are calculated using a quality adjusted index that weights more heavily research published in well ranked international peer-reviewed journals than other types of publications. A moving sum of the previous four years is calculated for each individual so that, for each relevant year of the panel, information accounts only for the more recent research. Several other variations of the research outcome measures were also considered.

⁷ Several other indices have also been constructed and tested. Thus, the abbreviations of the variables reported here correspond to a larger set of variables tested in alternative model specifications. The notation and numbering is aimed at distinguishing them from other variables (not reported here) measuring similar magnitudes.

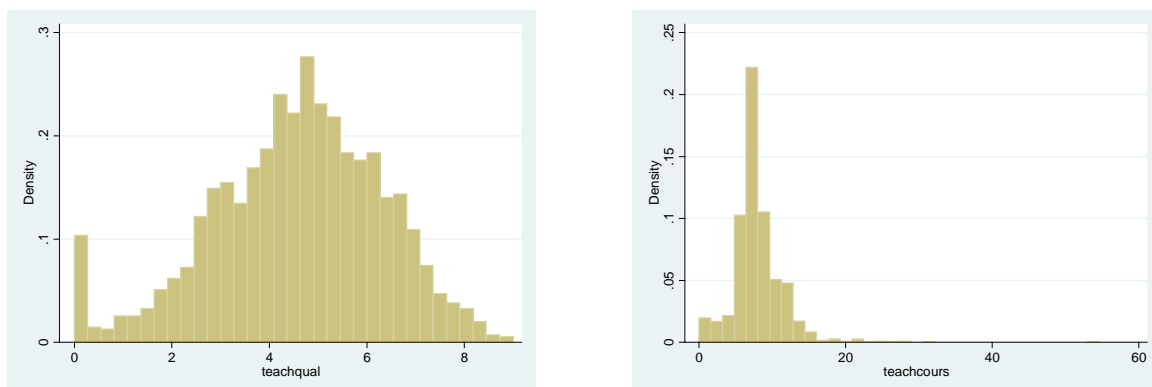
Quantity of teaching: Denoted by *Teachcours*, the teaching of each academic is the weighted sum of the number of undergraduate and postgraduate courses taught. Other measures that include, for instance, the number of courses taught abroad were also considered.

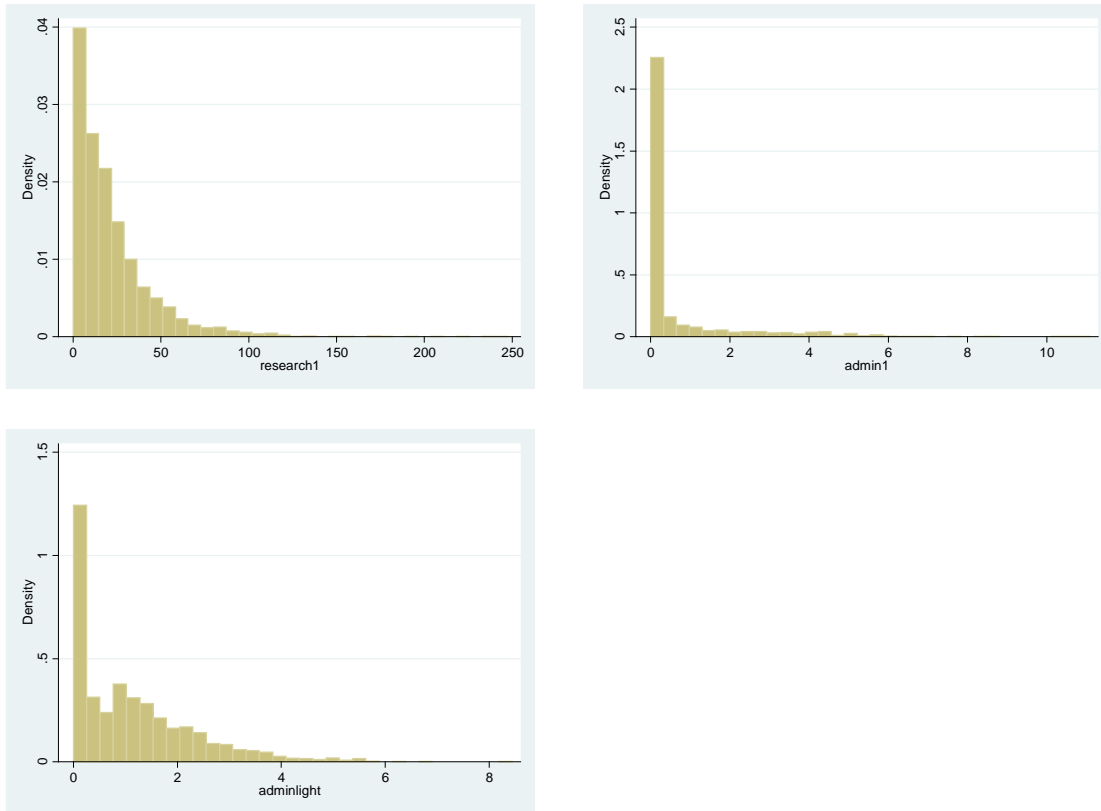
Administrative duties: The quantity of heavy administration duties, denoted by *Admin1* is measured by the sum of the number of years in the most demanding positions held by each academic within the university. These positions entail mandatory reductions in the quantity of teaching. We also consider other less demanding tasks denoted by *Adminlight* measured as the sum of the number of years served in less demanding administrative positions held by each academic within the university. These positions do not have a mandatory reduction in the teaching load.

With respect to other variables, we compute the *Participation in Academic Commissions* as the number of relevant Academic Commissions in which each academic participated in a given year. The *Publication of teaching materials* is the number of books and multimedia teaching materials with ISBN published by each instructor during the previous four years of each year in the sample. For each of the 25 departments of the university, the *Department* is coded using a binary variable, =1 if the instructor belongs to a given department and =0 otherwise. And finally, the *Gender* is coded using a binary variable (Male) taking the value 1 for females and 0 for males.

In Figure 1 we show the histograms of the indices mentioned.

Figure 1. Histograms of education indexes





The figures suggest that there is sufficient variation in the indices to make valid inferences. Several of them show accumulations at zero, reflecting that many professors do not do any research or administration. They also show that a substantial number of professors show a very low quality of teaching.

Figure 2. Teaching quality of researchers and non-researchers

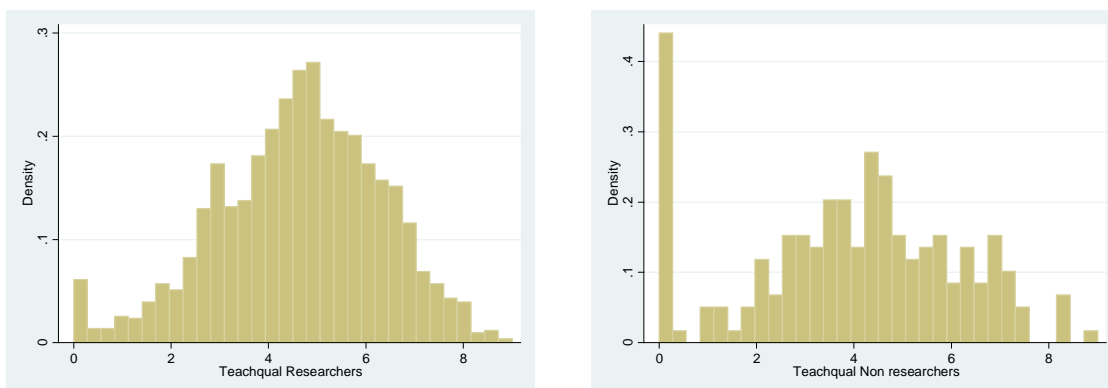


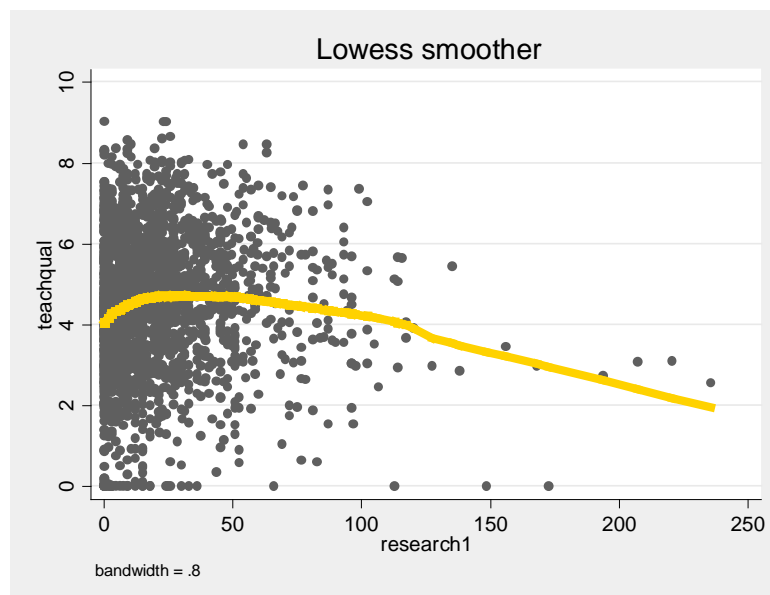
Figure 2 shows the histograms of teaching quality, divided into two categories. The left panel shows the scores for those professors who can obtain some research output ($Research1 > 0$). The

right hand side panel contains the scores of those professors with no research output. Notice that we find a higher accumulation of individuals with very low scores in the group of non-researchers.

Researchers have, on average, higher teaching loads than non researchers. The average teaching load for researchers with more than 5 teaching quality points is 8.35, while the average for non researchers is 6.87. Therefore, we find that, on average, researchers teach 21.5 % more than non researchers. This is possibly due to the existence of part time lecturers who, in general, do not perform research. However, we do not eliminate these lecturers from the sample, since we are interested in comparing the quality of their teaching with the rest of the professors in the sample.

It would be very interesting to discriminate between the quality of undergraduate and graduate courses. That would help us to test the hypothesis that research influences the quality of teaching at the graduate level, but less so at the undergraduate level. However, in our sample we only have aggregate student satisfaction surveys.

Figure 3: Teaching quality vs quality-adjusted research. Non parametric fit



In Figure 3 we observe that when we pool all observations and perform a non-parametric adjustment using the Lowess smoother, we observe a nonlinear relationship between teaching quality and published research, with a maximum of 4.91 (over 9) *Teachqual* points on the students' satisfaction scale at, approximately, *Research1* = 23 (quality-adjusted publication record), using a window of 5 observations. This suggests that the quality of teaching increases with the amount of published research up to a maximum level beyond which teaching quality

decreases slightly up to 50 quality-adjusted publication points, remaining in all cases above the teaching quality of low research performers. Only for the upper 10% of our sample's research performance the relation between teaching quality and research becomes negative. This exploratory analysis needs to be completed by considering other relevant variables in a multivariate model presented in the following section.

V. Econometric analysis and empirical results

In this section, we present an econometric model estimated from the yearly panel dataset using the indices shown in the previous section.

Some key questions that we would like to answer in this paper are: (1) How does research affect teaching quality? (2) How do administration duties affect teaching quality?, (3) How does the teaching load affect teaching quality? (4) How is the quality of teaching affected by complementary pedagogic activities, such as participating in committees, publishing teaching materials or taking teaching improvement courses? (5) Are there gender or department-specific differences?

We estimate the model shown in Table 4, reporting our favorite equation for teaching quality. The random effects panel data model estimated by generalized least squares allows for nonlinearity in Core published research, as suggested by Figure 3 and also nonlinearity in the Teaching load. The equation can be thought of as a multivariate generalization of Figure 3. The Hausman test (Hausman, 1978) for estimating the equation with panel data by covariance or using the generalized least squares estimates, takes a value of 11.93 with a p -value of 0.15. Therefore, we do not reject the null of no correlation between individual effects and regressors, so we can use the Generalized Least Squares estimates in Table 4.

Table 4. Teaching quality vs its determinants. Random effects panel data model estimated by Generalized Least Squares

Dependent variable: Teaching quality

<i>Explanatory variables</i>	<i>Estimate</i> (z-statistic)
Light administration duties	-0.064 (-2.47)
Core published research	0.0086 (2.58)
Core published research square	-0.000075 (-2.73)
Teaching load	0.479 (9.95)
Teaching load square	-0.031 (-9.98)
Participation in commissions	0.116 (2.77)
Books, multimedia teaching materials with ISBN	0.289 (3.60)
Education department	0.358 (3.51)
Female	0.276 (2.24)
Constant	2.977 (14.70)
Number of observations	
	2008
Estimation technique	
	Generalized Least Squares
Joint significance	
	$\chi^2(9)=177.25$ Prob> $\chi^2=0.0000$

NOTE: z-statistics are distributed as approximately normal $N(0,1)$ under the null. All estimates are significant at the 1% except for light administration duties and females, which are significant at the 1.4 and 1.7 %, respectively. STATA 12 is used for all calculations.

We find nonlinear relationships that are summarized by the positive coefficients of the linear terms in both equations, and the negative coefficients of the square terms. These estimates taken together suggest a maximum for *Teachqual* corresponding to a value of *Research1* of 57.55, which again suggests that research is positively associated with teaching quality for most of the observed levels of research performance. On the other hand, the maximum teaching quality is achieved at a teaching load of 7.63 units.

The nonlinearities found here agree with a common sense interpretation suggesting that, within a given range, research may be positive for teaching quality until a maximum is reached, beyond which higher levels of research may hinder the quality of teaching. Similarly, for some (low) values of teaching load, more teaching may improve teaching quality, while for higher levels, too much teaching reduces quality. This interpretation is natural and intuitively appealing. In fact, it may be unconceivable to think that research or teaching load will always be positively related to teaching quality for all possible values of research and teaching load. Searching for a linear or monotonic relationship, like has been done in the previous literature, (Marsh and Hattie, 2002) is methodologically incorrect and leads to essentially flawed results.

The main conclusions from Table 4 are the following:

In general, the quality of university teaching is positively affected by published research. Specifically, research output improves teaching quality for small amounts of research, below the median performance in our sample it reaches a maximum at 57.55 Core research contributions and decays slowly for higher values of research. That is, usual amounts of research are positive for teaching, while high research output may hinder teaching quality. In fact, two thirds of the professors could improve their teaching by doing more research. Both coefficients, 0.0086 with a z -statistic 2.58 and -0.000075 with a z -statistic of -2.73, are significant at the 1% and with the right signs.

The quantity of teaching has a positive effect on teaching quality for relatively low teaching loads, which is compatible with the significantly lower teaching evaluations obtained by part-time teachers. The quality peaks around 7.63 courses. In summary, moderate amounts of teaching are beneficial for the quality, while higher teaching loads may decrease teaching quality. Since the average teaching load is 8 in the period under study, the observed arrangement does not seem to be far from the optimum. Both coefficients, 0.469 with a z -statistic of 9.95 and -0.030 with a z -statistic of -9.98, are statistically significant at the 1% and with the right signs.

The quality of teaching is not affected by heavy administration duties. This variable is insignificant when entered in Table 4. This may be because these appointments entail mandatory reductions in the workload that compensate for administrative work. However, the quality of teaching is negatively affected by light administrative duties. This variable has a negative effect of -0.064 with a significant z -statistic of -2.47. This may arise because many of these duties do not entail a compensating reduction in teaching loads. The quality of teaching is positively related to

the participation in say, relevant committees, such as self evaluation commissions, creation/reform of degree curricula, etc. The estimate is 0.116, with a significant z -statistic of 2.77.

The quality of teaching is positively associated to the elaboration of books or multimedia teaching materials with ISBN. The point estimate is 0.289 with a significant z -statistic of 3.60.

Faculty members at the Education Department obtain better teaching results than the rest of the sample. They obtain, on average, a score that is 0.358 points higher than their colleagues of other departments.

As far as gender effects are concerned, we find that female professors obtain better teaching results than their male counterparts. The coefficient of Female is 0.276 with a significant z -statistic of 2.24, which means that female professors obtain, on average, a score that is 0.276 higher than their male counterparts.

Finally, courses of pedagogic enhancement of the Teaching Support Unit have no effect on teaching quality. This variable is insignificant when entered in the equation of Table 4.

C. Advantages of panel data

It is well known (Mundlak 1987, Hsiao 2003, Wooldridge 2010) that the use of individual panel data allows a more sophisticated analysis than cross section data.⁸ First, with panel data, the number of data available is typically higher than typical cross sections, allowing for more degrees of freedom and a more precise estimation of the parameters of interest. Second, and most important, they allow us to control for individual heterogeneity, such as individual ability, by using individual specific dummies. The use of individual specific dummies also allows us to diminish the bias due to omitted variables. For instance, omitted variables that do not change or evolve slowly over time, (such as the natural ability, the seniority, number of students per group, etc.) can be captured by the different individual constants.

D. Omitted variables tests

We have performed several tests to check for the possible omission of relevant terms such as a quadratic one for *Adminlight*, which turns out to be insignificant with a z value of 0.26, and both heavy and medium administrative duties with z values of 0.16, which are also insignificant. We have also tested for the omission of other candidate explanatory variables that measure activities

⁸ This type of data has been used in most of the previous literature.

related to the quality of teaching or directly designed to boost teaching quality. All of them are insignificant. The variables include courses of pedagogic enhancement, publishing teaching notes with ISBN, teaching articles published in journals or in the web, and the participation as instructor in experiences of European harmonization and advising.

E. Scenario

With respect to the recent law in Spain which imposes a mandatory increase in the teaching load of tenured non researchers our results predict a negative impact on teaching quality. Specifically, a typical non researcher whose teaching load increases by 8 ECTS, will see his teaching load increase by approximately one third. According to our model, this will imply a noticeable reduction in his teaching quality by 0.75 points. On the other hand, decreasing the teaching load of senior researchers by one third will have a positive effect on their teaching quality by approximately 0.17. Of course, for a given department, the aggregate effect will depend on the mix of non researchers and senior researchers. However, from rough calculations and the fact that the law forms part of the austerity measures aimed at drastic budget cuts, it is easy to predict that the non researchers affected by the increase in the teaching load will largely outnumber those few eligible for a decrease. Together with the relative sizes (0.75 against 0.17) of the negative and positive effects, this implies an unambiguous reduction in teaching quality. We summarize the scenario in Table 5.

Table 5. Effects of the Royal Decree of 2012, April 21st: Decreasing the teaching load of a senior researcher and increasing the teaching load of a typical non-researcher

	Teaching load before the law	Variation	Teaching load after the law	Change in teaching quality
Researcher (<i>researchI</i> ≥ 5)	10	-3	7	0.17
Non-researcher (<i>researchI</i> < 5)	10	+3	13	-0.73

VI. Conclusions

Academic jobs require simultaneous performance of a variety of duties, including teaching, research and administration captured in a large number of alternative performance measures available on each one of them. Assuming that teaching is a university's major contribution to the society, we have addressed the question whether teaching quality is affected by other academic duties. Our results show that the answer to this question is not as straightforward as assumed in previous studies restricting attention on linear or monotonic responses. We find non linear effects of the major determinants of teaching quality, among which are research performance and a professor's teaching load. These effects and the role of a number of other variables are identified due to our particularly rich dataset that allows for a more sophisticated statistical treatment of the relevant information. The dataset is a an annual panel of 604 professors from a wide variety of disciplines, types of professors and age ranges, at a medium size public Spanish University, for the period 2002-06.

The main conclusions of the paper are:

1. Non researchers are approximately five times more likely than researchers to be poor professors. Specifically, 14.8% of non researchers obtain ratings below 1, (on a 0-9 scale), while only 3.04% of researchers are rated below 1.
2. At the same time, on the whole, professors who conduct research are better professors than those who do not. The average satisfaction reported by students of non researchers is 3.96, while the average satisfaction reported by students of a typical researcher is 4.71, that is, 19% higher.
3. Researchers teach 21.5% more courses than non-researchers.
4. In general, the quality of university teaching is positively affected by published research. Specifically, research output improves teaching quality for small amounts of research, below the median performance in our sample, reaches a maximum and decays slowly for higher values of research output. That is, usual amounts of research are positive for teaching, while high research output may hinder teaching quality. In fact, two thirds of the professors could improve their teaching by doing more research.
5. The quantity of teaching has a positive effect on teaching quality for relatively low teaching loads. The quality peaks around the average teaching load observed in our sample. In summary, moderate amounts of teaching, are beneficial for the quality, while higher teaching

loads may decrease teaching quality. Since the average teaching load was 8 in the period under study, the observed arrangement does not seem to be far from the optimum.

6. The quality of teaching is not affected by heavy administration duties. This may be because these appointments entail mandatory reductions in the workload to compensate for administrative work.
7. The quality of teaching is negatively affected by light administrative duties. This may arise because many of these duties do not entail a compensating reduction in teaching loads.
8. The quality of teaching is positively related to the participation in say, relevant committees, such as self evaluation commissions, creation/reform of degree curricula, etc.
9. The quality of teaching is positively associated to the elaboration of books or multimedia teaching materials with ISBN.
10. Faculty members at the Education Dept. obtain better teaching results than the rest of the sample.
11. Female professors obtain better teaching results than their male counterparts.
12. Courses of pedagogic enhancement of the Teaching Support Unit have no effect on teaching quality.

Our results refute the myth that non-researchers are better professors than researchers. Since research generally increases the quality of teaching, there seems to be little ground for a stand-alone approach to the production of academic outputs. Furthermore, government and university authorities should take the interaction between different types of duties seriously into account because, as our results suggest, increasing the teaching load of one type of professor may reduce costs but it may have non-negligible negative effects on teaching quality.

The recent 2012 Royal Decree of April 21st in Spain modifies the teaching load of researchers and non researchers. Thus, the effects of increasing or decreasing one or the other professor type's teaching load on teaching quality need to be clarified in order for the qualitative consequences of the law to be assessed. In our model, the number of courses taught (*Teachcours*) by an instructor can be regarded as weakly exogenous. The number of courses taught is determined by the teaching needs and resources of the department, the type of contract (full time or a variety of part time positions), and the teaching reduction due to some administrative appointments. Our findings

indicate that modifying the teaching shares of researchers and non-researchers will have a negative impact on overall teaching quality in the University. In fact, the example constructed on the effects that the Decree is expected to have on teaching quality in Spain has yielded a clear-cut negative prediction.

The use of panel data allows us to control for different kinds of heterogeneity, such as the individual characteristics of the professors, their natural abilities, their department, and also to avoid the biases due to omitted variables such as the availability of teaching assistants, the size of the class and the seniority of the professor, among others (Hsiao, 2003).

There are no obvious reasons why these results would vary substantially in other universities or countries. However, it would be interesting to look for more universities which systematically collect similar data on their professors' teaching and research performance to make cross country comparisons.

VII. References

- Bedard, Kelly, and Peter J. Kuhn. 2008. "Where Class Size Really Matters: Class Size and Student Ratings of Instructor Effectiveness." *Economics of Education Review*. 27 (3): 253-265.
- Becker, William E. Jr. 1975. "The University Professor as a Utility Maximizer and Producer of Learning, Research, and Income." *The Journal of Human Resources*. 10(1): 107-115.
- Beleche, Trinidad, David Fairris, and Mindy Marks. 2012. "Do Course Evaluations Truly Reflect Student Learning? Evidence from an Objectively Graded post-test." *Economics of Education Review*. 31 (5): 709-719.
- Carrell, Scott E., and James E. West. 2010. "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors." *Journal of Political Economy*. 118 (3): 409-432.
- Ewing, Andrew M. 2012. "Estimating the Impact of Relative Expected Grade on Student Evaluations of Teachers." *Economics of Education Review*. 31 (1): 141-154.
- Feldman, Kenneth A. 1987. "Research Productivity and Scholarly Accomplishment of College Teachers as Related to their Instructional Effectiveness: A Review and Exploration." *Research in Higher Education*. 26 (3): 227-298.

- Friedrich, Robert J., and Stanley J. Michalak Jr. 1983. "Why Doesn't Research Improve Teaching?: Some Answers from a Small Liberal Arts College." *The Journal of Higher Education*. 54 (2): 145-163.
- Gomez-Mejia, Luis R., and David B. Balkin. 1992. "Determinants of Faculty Pay: An Agency Theory Perspective." *The Academy of Management Journal*. 35 (5): 921-955.
- Hattie, John, and Herbert W. Marsh. 1996. "The Relationship between Research and Teaching: A Meta-analysis." *Review of Educational Research*. 66 (4): 507-542.
- Hausman, Jerry A. 1978. "Specification Tests in Econometrics" *Econometrica*. 46 (6): 1251-1271.
- Hsiao, Cheng. 2003. *Analysis of Panel Data*, 2nd Edition, Econometric Society Monographs, Cambridge University Press.
- Isely, Paul, and Harinder Singh. 2005. "Do Higher Grades Lead to Favorable Student Evaluations?" *Journal of Economic Education*. 36 (1): 29-42.
- Krautman, Anthony C., and William Sander. 1999. "Grades and Student Evaluation of Teachers." *Economics of Education Review*. 18 (1): 59-63.
- Lindbeck, A., and D.J. Snower. 2003. "The Firm as a Pool of Factor Complementarities." IZA Discussion Paper 882.
- Marsh, Herbert W. 1987. "Student's Evaluation of University Teaching: Research Findings, Methodological Issues, and Directions for Future Research." *International Journal of Educational Research*. 11 (3): 263-253
- Marsh, Herbert W., and John Hattie. 2002. "The Relation Between Research Productivity and Teaching Effectiveness: Complementary, Antagonistic or Independent Constructs?" *The Journal of Higher Education*. 73 (5): 603-641.
- Martin, B.R. 2012. "Are Universities and University Research under Threat? Towards an Evolutionary Model of University Specialization." *Cambridge Journal of Economics*. 36 (3): 543-565.
- McPherson, Michael A. 2006. "Determinants of How Students Evaluate Teachers." *Journal of Economic Education*. 37 (1): 3-20.

- Mundlak, Yair. 1978. "On the Pooling of Time Series and Cross Section Data." *Econometrica*. 46 (1): 69-85.
- Noser, Thomas C., Herman Manakyan, and John R. Tanner. 1996. "Research Productivity and Perceived Teaching Effectiveness: A Survey of Economics Faculty." *Research in Higher Education*. 37 (3): 299-321.
- Porter, Stephen R., and Paul D. Umbach. 2001. "Analyzing Faculty Workload Data Using Multilevel Modeling." *Research in Higher Education*. 42 (2): 171-196.
- Shin, Jung C. 2011. "Teaching and Research Nexuses across Faculty Career Stage, Ability and Affiliated Discipline in a South Korean Research University." *Studies in Higher Education*. 36 (4): 485-503.
- Spiller, Pablo T., and Bennet A. Zelner. 1997. "Product Complementarities, Capabilities and Governance: A Dynamic Transaction Cost Perspective." *Industrial and Corporate Change*. 6 (3): 561-594.
- Stack, Steven. 2003. "Research Productivity and Student Evaluation of Teaching in Social Science Classes: A Research Note." *Research in Higher Education*. 44 (5): 539-556.
- Verburgh, Jan E., and Sari Lindblom-Ylänne. 2007. "Investigating the Myth of the Relationship between Teaching and Research in Higher Education: A Review of Empirical Research." *Studies in the Philosophy of Education*. 26 (5): 449-465.
- Vidal, Javier, and Miguel A. Quintanilla. 2000. "The Teaching and Research Relationship within an Institutional Evaluation." *Higher Education*. 40 (2): 217-229.
- Walstad, William, and Sam Allgood. 2005. "Views of Teaching and Research in Economics and Other Disciplines." *American Economic Review*. 95 (2): 177-183.
- Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd edition, The MIT Press Cambridge, Massachusetts, London, England.