Analyzing Faculty Workload Data Using Multilevel Modeling

Presented at the 40th Annual Association of Institutional Research Forum,
Cincinnati, Ohio, May 21-24, 2000

Stephen R. Porter, Ph.D.
School of Nursing, University of Maryland, Baltimore
655 West Lombard Street
Baltimore, MD 21201
Telephone: 410.706.3618
Fax: 410.706.0719
Email: porter@parsons.umaryland.edu

Paul Umbach
Office of Institutional Studies, University of Maryland, College Park
College Park, MD 20910
Telephone: 301.405.5335
Fax: 301.314.9443
Email: umbach@wam.umd.edu

Analyzing Faculty Workload Data Using Multilevel Modeling

Abstract

Research on faculty productivity fails to account for the hierarchical nature of the data. Faculty within an academic discipline more closely resemble one another than faculty in other disciplines, resulting in dependent observations and thus inaccurate statistical results. Unlike ordinary least squares, multilevel modeling takes into account this grouping effect. The paper analyzes the research productivity of 1,104 tenured/tenure-track faculty from the 1993 NSOPF survey to compare traditional regression models with a random coefficients model. The results indicate a large grouping effect on research productivity, and the statistical as well as the substantive results of the random coefficients model differ significantly from the regression approach.

Analyzing Faculty Workload Data Using Multilevel Modeling

Introduction

Because institutional researchers often advise high level administrators on policies that may significantly impact their campuses, accurate analyses of institutional data are essential. Yet in their analyses institutional researchers routinely overlook the effects caused by group membership in academic disciplines. Generally one of two approaches are taken. Either researchers build statistical models that examine data at the group level or organizational level and neglect differences in individuals, or they examine data at the individual level and ignore the impact of group membership. Both approaches can result in inaccurate parameter estimates, and thus lead to poor or even misleading policy analyses. Multi-level modeling techniques allow researchers to appropriately handle the complex organizational effects of colleges and universities and provide the tools necessary to arrive at more accurate results.

One example of policy that must be analyzed using multi-level modeling techniques is faculty productivity. In the last decade, one of the most salient policy issues in higher education has been the regulation of faculty work. As the demands on state revenues have grown, state legislators have begun focussing attention on increasing the productivity of faculty at state-supported universities as an alternative to increasing state spending (Layzell, 1996). Some legislators believe that significant cost savings would result if faculty, especially faculty at research universities, were required to do more teaching. In fact, one study by the (Maryland Higher Education System, 1994) argued that the University of Maryland, College Park, could save $20 million annually if all full-time, tenured and tenure-track faculty were required to teach five courses per year

While many researchers have examined the processes by which faculty workload is measured, few have studied the effects these mandates have on overall faculty research output (Middaugh, 1998). Such an analysis is vital given the substantial revenues generated by faculty research. A mandated increase in faculty teaching could decrease instructional costs, but these savings might by offset by a concomitant loss in research revenues. Because grant dollars are an important part of an institution's budget, efforts to save money by forcing faculty to teach more could paradoxically cause a loss in revenue.

While an understanding of faculty productivity is critical, little research has been done on the contextual effect of academic discipline on outputs. Faculty productivity across disciplines is a very complex empirical issue and previous researchers have ignored one of the most important effects, clustering. Understanding the effect of discipline is essential in analyses of faculty productivity. Using the 1993 National Study of Postsecondary Faculty, we compare the statistical and substantive results of multilevel or hierarchical linear modeling to the traditional regression approach. Our results indicate large differences between the two statistical techniques.

Previous research

While faculty productivity has been the focus of numerous analyses, previous researchers have largely ignored the multilevel nature of their data, not only in their choice of statistical methodology, but also in their samples and variable construction. Because the factors affecting faculty research productivity vary so much between institution types, academic fields and even modes of research productivity, only an approach that takes this variation into account can hope to accurately shed light on this important facet of higher education. We believe research in this area must meet four criteria:

- Use of the appropriate unit of analysis: the individual.
- Analysis is based on a sample of homogenous institutions.
- Use of a dependent variable(s) that distinguish between heterogeneous modes of research.
- Use of a statistical technique that takes into account the clustered nature of the data.

We address each of these criteria in turn.

Some researchers have analyzed faculty productivity with data collected at the institutional, departmental or discipline level rather than the individual level (Baird, 1991; Bentley & Blackburn, 1990; Dundar & Lewis, 1998; Gander, 1999; Olson, 1994; Perry, Clifton, Menec, Struthers, & Menges, 2000). In essence this approach takes individual data and uses unit means as explanatory variables, usually combined with unit-level data such as budgetary allocations. Such an approach allows an analysis of numerous departments or disciplines in institutions across the country, resulting in very generalizable results.

Unfortunately analyses of this type are prone to what is known as an "ecological fallacy" (Robinson, 1950; King, 1999; Kreft & De Leeuw, 1998), in which aggregate-level results may substantially differ or even be the reverse of individual-level results. For example, an analysis based on departments might reveal that departments with higher proportions of female faculty are less productive, while an analysis of the same departments at the individual level might reveal that female faculty are more productive than male faculty. While aggregate department or discipline-level data about faculty research productivity for the nation is more readily available than individual-level data, the severe biases introduced by aggregation renders the use of these data for multivariate analyses almost impossible (but see the solution proposed by King, 1999). To understand individual-level behavior, we must use individual-level data.

In addition to using samples of disciplines rather than samples of individuals, some researchers have analyzed samples combining a wide variety of post-secondary institutions (Bellas & Toutkoushian, 1999; Fox, 1992; Gander, 1999; Wanner et al., 1981). Bellas and Toutkoushian (1999), for example, combine two-year institutions with four-year institutions from several different Carnegie classifications (Carnegie Foundation for the Advancement of Teaching, 1994). Given that the strengths of the relationships between factors affecting research productivity and the dependent variable are likely to vary in strength between institutions, lumping heterogeneous institutions into one sample can obscure these relationships. While such an approach again assures generalizability to all major institutions in the country, such generalization comes at a price. A more profitable approach would focus on a smaller sample of more homogenous institutions, or the use of a methodology that takes this structure of the data into account.

Besides sample selection, the construction of the dependent variable has also been problematic in some studies, with researchers combining various types of research outputs into a single measure or index (Bellas & Toutkoushian, 1999; Buchmueller et al., 1999; Olson, 1994; Olsen & Simmons, 1996; Sullivan, 1996; Wanner et al., 1981). The concept of research productivity embraces many different modes of research, from presentations to journal publications and books to amount of grant dollars generated. Some require more effort to achieve than others do, and the amount of effort within a category can also vary; for example, in general an article appearing in a non-refereed journal probably took less time and effort to produce than an article appearing in a refereed journal. Combining research modes that vary in the amount of effort required to produce an outcome can only obscure the substantive results. Again, the trade-off is one between generalizability and clear results. Combining multiple modes into one index allows one to address the majority of faculty research output while simplifying the results, but at the cost of possibly obscuring interesting substantive results. A second drawback to such indices is that concrete policy recommendations can be difficult to make, because the substantive impact of a change in an independent variable is not always clear.

Our final criterion involves the appropriate statistical technique to be used when analyzing faculty productivity. All previous researchers have used simple cross-tabulations, correlations or regression analysis. Such techniques ignore the clustered nature of the data, because faculty members in an academic field often more closely resemble one another than faculty in other fields (Austin, 1996). Clustering of the data can radically affect the substantive results of any analysis, as will be explained below.

Our review of the literature has focused on one theme: data on faculty research productivity can be quite heterogeneous, and this heterogeneity must be taken into account in any analysis. Only by constructing samples and variables as homogenous as possible, and explicitly modeling (at least partially) the remaining heterogeneous structure of the data can we begin to truly understand why some faculty produce more than others.

Methodology and data

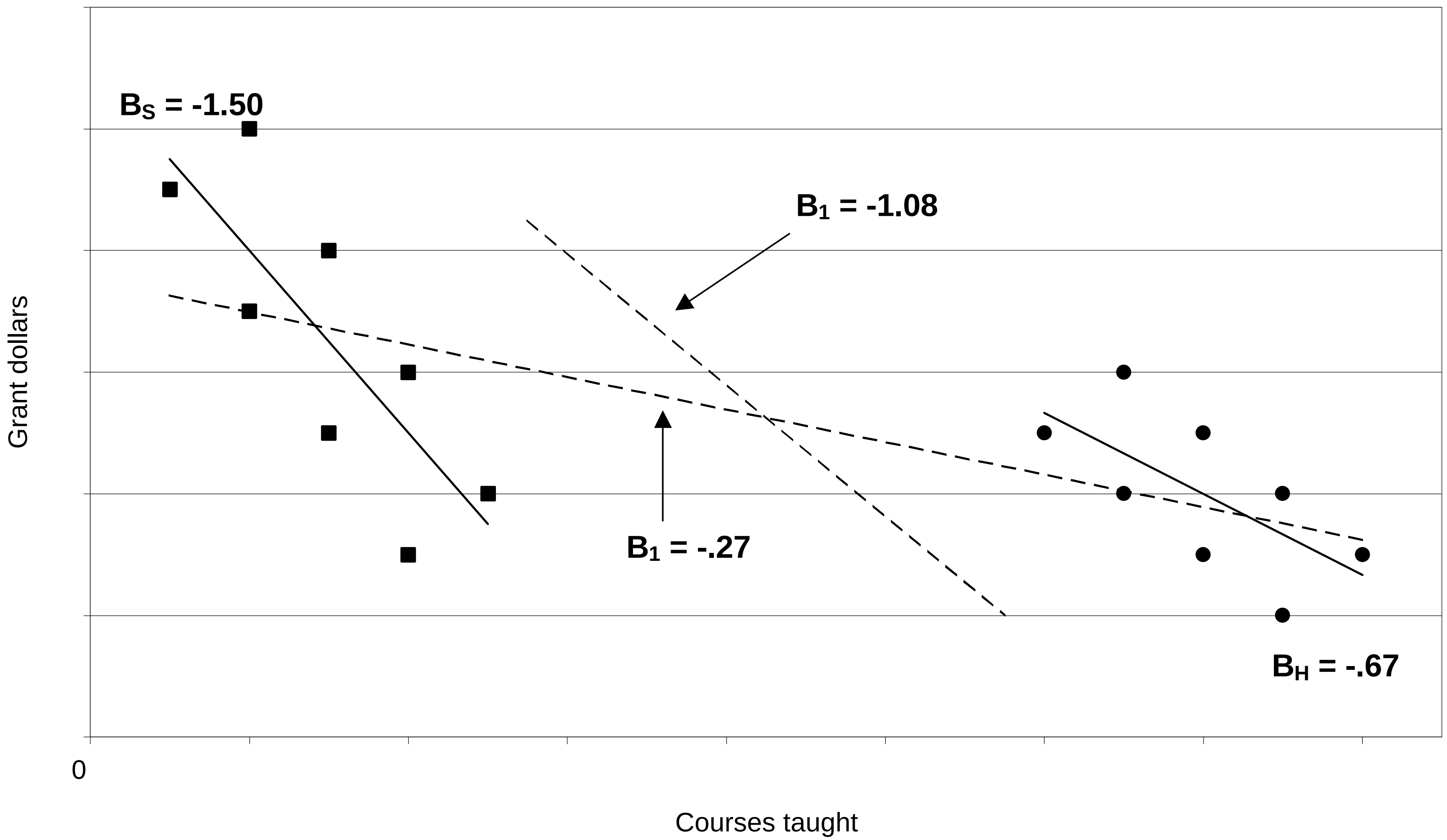
Clustering and estimation strategies

The impact of clustering on ordinary least squares (OLS) regression analysis can be seen in Figure 1, which graphs faculty in two hypothetical academic departments. Faculty members in the first department, in the upper left-hand corner of the figure, tend to teach few courses while generating substantial grant dollars, while faculty in the second department tend to teach more courses while bringing in fewer grant dollars. The first department can be considered a stylized hard sciences department, while the second can be viewed as a department in the humanities. The lines through each group of faculty, labeled B_S and B_H, represent the relationship between courses taught and grant dollars generated for the faculty in that department.

If we estimated a simple OLS equation of the following form

$$y_i = a + bx_i + e_i \qquad (1)$$

Figure 1. Estimates of the Relationship Between Grant Dollars and Number of Courses Taught



where y_i represents grant dollars generated by faculty member i, a represents the intercept, b the coefficient describing the relationship between courses taught and grant dollars, x_i the number of courses taught by the faculty member i, and e_i a random error ~N(0, s^2), the result would be the long dotted line with the slope coefficient B_1= -.27. Ignoring the clustered nature of the data, OLS fits a line that best describes the data, with the result that the relationship between courses taught and grant dollars (either -.67 or -1.50, depending on the department) is underestimated to be -.27.

One solution to this problem would be to include a dummy variable for one of the two departments. This allows the intercepts to differ for each department, while constraining the slope coefficient for courses taught (shown in the figure by the short dotted line B_1= -1.08) to be the same for each department. This estimate is much closer to the individual slopes for the two departments, B_S and B_H. The full solution under OLS would be to include an interaction term between the department dummy variable and courses taught variable to allow B_1 to vary between the two departments, so that the correct estimates of B_S and B_H are derived.

Fitting multiple group-based dummy variables and interaction terms is only practical when the number of groups or clusters is quite small. Such a strategy cannot be used with data on a university with numerous departments, or a sample such as the one used here where faculty are distributed among 100+ academic fields. Two problems occur. First, multicollinearity among the dummy variables and interaction terms yield traditional hypothesis tests useless, and may also affect the standard errors of other variables in the model if they vary by group. Second, this procedure can provide inaccurate estimates, create difficulty interpreting the results and significantly reduce the degrees of freedom (Stapleton & Lissitz, 1999). These models assume that data for each faculty member is independent from other observations; however, the nesting effect of disciplines at a research university causes a significant grouping effect and therefore results in dependent observations.

Multilevel or hierarchical linear statistical techniques were developed to estimate data clustered by groups. While OLS only accounts for variance at the individual level, multilevel techniques take into account variance at both the individual and group levels, thus allowing intercepts and slope coefficients for selected variables to vary across groups. These models permit the researcher to examine the contextual effect of belonging to a particular group without compromising degrees of freedom and accuracy of estimates. They are most commonly used to study group effects on individual-level behavior. The classic example from the literature is the effect of a student's socioeconomic background on math achievement, while also taking into account the socioeconomic status (SES) of other students in the school by including school-level means of student SES. Multilevel models are estimated with a variety of maximum likelihood techniques (see the discussion in (Kreft & De Leeuw, 1998) and (Bryk & Raudenbush, 1992))[1].

One specific subset of multilevel models is termed random coefficient models (Kreft & De Leeuw, 1998; Greene, 1997). These models do not contain any variables measured at the group level such as mean SES. Instead, the intercept is allowed to vary by groups, and sometimes the coefficients of select independent variables. Using the grant dollars example

[1] We estimated the multilevel models using PROC MIXED in SAS. See (Singer, 1999) for details.

above, the intercepts for each academic discipline are allowed to vary, and the coefficients for courses taught are also allowed to vary by discipline, but no discipline specific means are included as explanatory variables.

This can be seen more clearly using equations:

$$y_{ij} = a_j + b_j x_{ij} + e_{ij} \qquad (2)$$

$$a_j = g_{00} + u_{oj} \qquad (3)$$

$$b_j = g_{10} + u_{1j} \qquad (4)$$

Equation (2) is the same as (1), with one additional subscript j signifying academic discipline membership. Thus (2) is the amount of grant dollars generated by individual i in discipline j, with this amount the function of a discipline-specific grant dollar amount (a_j) plus the amount of courses taught by individual i in discipline j (x_{ij}) conditional on the relationship between courses taught and grant dollars in discipline j (the coefficient b_j), plus a random error term for every individual i in each j disciplines.

Equations (3) and (4) show how both the intercept and regression coefficient for courses taught varies. The intercept or level of grant dollars (a_j) for discipline j is the function of g_{00}, the average amount of grant dollars for all disciplines, plus an amount or deviation u_{oj} which varies by discipline. Similarly, the coefficient b_j is a function of g_{10}, the average effect of courses on grant dollars generated, plus a deviation that varies by discipline.

Although the estimation techniques differ, in terms of substantive theory this model is the same as a regression model with dummy variables and interaction terms with courses taught for all j-1 disciplines. Unlike OLS, however, the degree of freedom and other problems do not occur. In addition, the interpretation of the results is clearer. Under OLS with multiple dummy variables and interaction terms, the interpretation of the courses taught variables is the impact of courses taught for the excluded academic discipline, or reference group. (Recall that with j groups, only j-1 dummy variables and interaction terms can be entered into the regression equation, with the intercept and independent variable on which the interaction terms are based representing the effects for the one group that was excluded from the equation.) In a random coefficients model l_{10} becomes the average impact of courses taught on discipline.

In many applications in institutional research we are less concerned with discipline-specific results and more concerned with being able to make global policy conclusions; for example, telling the President or Provost that increasing course loads by one course would cause the average amount of grant dollars generated to drop by X dollars. Random coefficient models allow us to draw such substantive conclusions from our results. While OLS cannot handle the large number of academic disciplines that are present in most institutions of higher education, random coefficient models provide better estimation of more theoretically relevant models.

Sample data

We use the 1993 National Study of Postsecondary Faculty data (NSOPF) (U.S. Department of Education, 1998) collected by the National Center for Education Statistics to obtain a "general model" for postsecondary institutions in the United States, avoiding the generalization problem posed by using data from a single institution. The weighted sample is all full-time tenured or tenure-track faculty in Fall 1992 (excluding chairs) holding the rank of assistant, associate or full professor at public Research I and II institutions. By only using data from Research I and II institutions we reduce the heterogeneity between institutions. We exclude chairs from the analysis given their unique administrative burden and its likely impact on research productivity. (One strategy would be to include a dummy variable for chairs in our model. But such a solution cannot take into account the likely possibility that the slope coefficients for some variables differ for chairs. Given their small numbers exclusion seemed the best approach.)

The NSOPF survey asked faculty to choose their principle fields of teaching and research from a detailed list of academic disciplines and major fields of study (see Appendix A). Principal field of teaching is the variable used to cluster faculty into 104 homogenous groups. We believe principle field of teaching more closely corresponds with a faculty member's academic background and department than field of research. If the faculty respondent did not list a teaching field we used their principal research field instead. Respondents failing to list either a teaching or research field, or whose field was in the academic discipline of vocational training or health sciences were excluded from the analysis.

Model

Dependent variables

We use two dependent variables to measure faculty research productivity: publications over a two-year period and the dollar amount of external research funding. The means and standard deviations for the dependent (and independent) variables are displayed in Table 1. Respondents to the NSOPF survey were asked about the number of publications in a variety of categories in the two years previous to the survey. We summed the number of articles published in refereed professional or trade journals, creative works published in juried media and chapters in edited volumes into a single measure of refereed publications.

The second dependent variable is the total external grant dollars for the 1992-1993 academic year on which the faculty member was a principal or co-principal investigator. Note that this formulation excludes funds from the faculty member’s institution, as well as grants on which the faculty member worked as a staff member.

Because of the nonlinear relationship between our dependent and independent variables, the dependent variables are logged to ensure normality (Olson, 1994); see also the discussion in (Tufte, 1974), pp. 108-134). Comparisons of regression equations using the logged and unlogged variables show much better model fit using the logged version.

Table 1. Means and Standard Deviations of Variables

	Mean	SD
Grant dollars	154274.984	857215.465
Grant dollars (logged)	4.792	5.631
Publications	5.059	4.858
Publications (logged)	1.236	0.896
Research assistantship	0.489	0.500
Scholarship	0.584	0.493
Ph.D.	0.869	0.337
Professional	0.057	0.232
Opportunity-research	2.233	0.757
Opportunity-teach	1.710	0.750
Full professor	0.421	0.494
Associate professor	0.319	0.466
Years in rank	8.631	7.341
Courses-undergraduate	1.101	1.087
Courses-graduate	0.531	0.723
Age	48.061	9.576
Non-white	0.176	0.381
Married	0.818	0.386
Female	0.241	0.428
Female*Married	0.163	0.370
Female*Children	0.196	0.619
Children	1.565	1.366

Independent variables

We use five groups of variables to model faculty research productivity. In our model research productivity is a function of human capital, personal tastes, career status, teaching workload, demographics and academic discipline.

Human capital and tastes will certainly affect research productivity. Those most able and those who like to perform research will, all things being equal, be more productive. Human capital is proxied with four dummy variables. The first measures whether a faculty member had received a fellowship, scholarship or grant in graduate school. Because these financial supports are usually awarded based on merit, this variable proxies raw ability. In addition to ability, research training will also affect productivity (Buchmueller et al., 1999; Wanner et al., 1981). The second dummy variable measure whether the faculty member had a research assistantship in graduate school, while two other variables indicate the highest degree earned, either a Ph.D. or a professional degree.

Research productivity is also a matter of personal taste, so we include two variables measuring attitudes towards doing research. Respondents were asked to rate the importance on a scale of one to three (three being “very important”) of several factors in a hypothetical decision to leave their current position. Two of these factors are “greater opportunity to teach” and “greater opportunity to do research.” We hypothesize that faculty who rate opportunities to do research as very important will be more productive, while faculty rating opportunities to teach will be less productive. This is due to individual tastes: faculty who would prefer to spend their time doing research are likely to place emphasis on research in their hypothetical decision to leave, while faculty preferring teaching are likely to emphasize the opportunity to teach. One can argue that these variables also measure institutional climate, in that faculty facing administrative pressure to publish or teach may rate these factors as very important. But given the sample, Research I and II institutions, the pressure to publish is present at all the institutions in our sample and probably does not vary widely.

Faculty career status should also affect research productivity. We expect productivity to vary by faculty rank, but in the opposite direction that some might hypothesize. Because of the need to achieve tenured status, one could argue that junior faculty should be producing more research than senior faculty. Disregarding the controversy over this hypothesis, this formulation ignores the cross-sectional nature of our data. There is a substantial selection effect that is not capture by our sample, in that faculty who achieve tenured status have (in theory) proven themselves able to perform substantial research, while those less able are filtered out by the tenure and promotion process. Thus these variables may actually capture human capital effects more than career effects, but in either case they should be included as controls.

Years in current rank is also an important variable when predicting faculty productivity (Stapleton & Lissitz, 1999). Including only faculty rank does not provide the full picture of faculty experience. The productivity of full and associate professors who have held that rank for several years is quite different than those recently promoted. Accounting only for three faculty ranks would overlook the differences of people within those ranks.

Besides research, faculty are also expected to teach, often at both the undergraduate and graduate levels. Given the time demands from teaching, faculty teaching numerous courses will, all things being equal, be less productive than faculty with lighter teaching loads (Fox, 1992; Neumann, 1996; but see also Braxton, 1996).

We include two variables to measure teaching load: the number of undergraduate courses taught in the Fall 1992 semester and the number of graduate courses taught in the Fall 1992 semester. Given disparities between institutions as to what comprises a credit hour, and the fact that courses involving internships and field research can generate large numbers of credit hours for one course, we believe that a simple count of courses provides a more accurate measure of teaching load than the sum of course credit hours.

We separated teaching load into undergraduate and graduate courses because the course demands for undergraduate courses can be quite different from graduate courses. While it is difficult to control for the variability in the demands that each course has on a faculty member, separating the undergraduate and graduate courses does provides some control.

We also include a set of demographic variables to control for differences between faculty: age in years, a dummy variable if the faculty member is a minority, a dummy variable indicating the faculty member was married, a dummy variable for gender, and number of children. Because of the cross-sectional nature of the data, the interpretation of the age variable is problematic because we cannot distinguish between aging and cohort effects (Lawrence & Blackburn, 1988; Levin & Stephan, 1989); however, age still remains an important control variable. We also include two interaction terms to control for the possible differential effect of being married and having children on female faculty.

Finally, our models are organized around the 104 academic disciplines as designated by NCES (See Appendix A). For the OLS models 103 dummy variables (with “Other Social Sciences” as the excluded or reference discipline) were added to the model to control for discipline. In the random coefficients models, academic discipline was made the intercept.

Results

We estimated two models for each dependent variable. The first model uses OLS with 103 dummy variables for academic discipline. The second was a random coefficients model (RCM) in which the independent variables are fixed, similar to OLS, with two exceptions. First, the intercept is allowed to vary by discipline. Second, the coefficients for undergraduate and graduate courses taught are also allowed to vary by discipline.

While these choices are based on theory, they can also be tested like any hypothesis. In the RCM approach both a fixed effect for the intercept and courses variables, and the variances of the discipline-specific deviations from the intercept and deviations from the main course effects, can be estimated and tested for statistical significance. If the intercept or the course coefficients do not vary by discipline, the significance test for their variances will be negative.

Grant dollars

Before estimating the full models, we calculated the impact of academic discipline on publications. The amount of variance that discipline explains in publications, or the inter-class correlation (ICC), is .32. An ICC of this magnitude would indicate a grouping effect as a result of academic discipline. This effect must be controlled for in the models. If such a grouping effect was not present, a random coefficients approach would not yield dissimilar results from OLS.

The results for the both the OLS and RCM models are presented in Table 2. The results from both the OLS and RCM results reveal significant relationships between several of the independent variables and publications.

Faculty members that were research assistants while in graduate school on average generated more grant dollars. Education and rank have no significant impact on the ability to earn grant money. Yet the number of years in rank has a negative relationship with grant dollars. Those that have been in their current rank longer had less grant dollars on average than those who have less years in their current appointment.

Table 2. Grant Dollars and Publications Model Estimates

	Grant Dollars		Publications	
	OLS[a]	RCM	OLS[a]	RCM
Slope coefficients				
Intercept	4.914 *	5.652 **	1.682 **	1.693 **
	(2.037)	(1.405)	(0.355)	(0.234)
Research assistantship	1.922 **	2.092 **	0.135 *	0.195 **
	(0.326)	(0.311)	(0.057)	(0.052)
Scholarship	0.084	0.007	0.048	0.023
	(0.303)	(0.295)	(0.053)	(0.050)
Ph.D.	-0.251	0.233	0.386 **	0.499 **
	(0.603)	(0.559)	(0.105)	(0.091)
Professional	0.003	0.159	0.320 *	0.337 **
	(0.821)	(0.781)	(0.143)	(0.130)
Opportunity-research	0.404 *	0.398 *	0.129 **	0.135 **
	(0.200)	(0.194)	(0.035)	(0.033)
Opportunity-teach	-0.692 **	-0.749 **	-0.131 **	-0.144 **
	(0.202)	(0.197)	(0.035)	(0.034)
Full professor	0.922 +	0.837	0.469 **	0.427 **
	(0.526)	(0.510)	(0.092)	(0.087)
Associate professor	0.315	0.212	0.237 **	0.178 *
	(0.444)	(0.432)	(0.077)	(0.074)
Years in rank	-0.087 **	-0.082 **	0.004	0.002
	(0.029)	(0.028)	(0.005)	(0.005)
Courses-undergraduate	-0.607 **	-0.729 **	-0.038	-0.077 *
	(0.156)	(0.179)	(0.027)	(0.031)
Courses-graduate	-0.018	-0.138	-0.011	-0.025
	(0.224)	(0.232)	(0.039)	(0.045)
Age	-0.017	-0.021	-0.028 **	-0.025 **
	(0.028)	(0.027)	(0.005)	(0.005)
Non-white	-0.610	-0.439	0.123	0.094
	(0.488)	(0.472)	(0.085)	(0.080)
Married	0.449	0.576	0.123	0.095
	(0.497)	(0.488)	(0.087)	(0.083)
Female	-1.071	-1.399 +	-0.064	-0.145
	(0.767)	(0.742)	(0.134)	(0.126)
Female*Married	1.467	1.606 +	-0.051	0.030
	(0.894)	(0.868)	(0.156)	(0.147)
Female*Children	-0.268	-0.321	-0.015	-0.004
	(0.356)	(0.345)	(0.062)	(0.059)
Children	-0.053	-0.004	-0.034	-0.025
	(0.134)	(0.131)	(0.023)	(0.022)

(continued)

Slope variances				
Intercepts	-	9.130 **	-	0.114 *
		(2.533)		(0.051)
Courses-undergraduate	-	0.607 *	-	0.024 *
		(0.363)		(0.011)
Courses-graduate	-	0.391	-	0.041 +
		(0.568)		(0.027)
Residual	-	19.106 **	-	0.557 **
		(0.891)		(0.027)
Model statistics				
Adjusted R-square	0.398	-	0.267	-
SEE	4.447	-	0.775	-
F-test	6.97 **	-	4.29 **	-
-2 Log likelihood	-	6785.1	-	2917.1
N	1,104	1,104	1,104	1,104

Note: ** p<.01, * p<.05, + p<.10.
[c]oefficients and standard errors for 104 academic field dummy variables are not shown.

Preferences were also significantly related to grant dollars. On average, those that prefer research over teaching secured more grant dollars than those that prefer teaching over research.

In both OLS and RCM, the more undergraduate faculty taught, the less grant dollars they earned. Examining the slope variances in the RCM, it appears that the impact of the number of undergraduate courses taught is different for different disciplines

Publications

As with grant dollars we first calculated an ICC. The amount of variance that discipline explains in publications is .18, again a sizable ICC indicating a grouping effect as a result of academic discipline.

Human capital appears to be strongly related to the number of publications by a faculty member in the last two years. Faculty that had research assistantships while in graduate school produced more publications. Faculty holding a Ph.D. or professional degree published more than those that did not, and full professors and associate professors had more publications than associate professors.

Personal preferences also had a statistically significant relationship with publications in both the OLS and RCF models. Faculty that prefer to teach rather than do research publish less; and those faculty that prefer to do research publish more. The only demographic variable that appears to relate to publications is age. On average, the older the faculty member the fewer publications he or she produces, all else being equal.

In OLS, the number of undergraduate courses taught appears to have no significant relationship with the number of publications. However, in the RCM, undergraduate courses is statistically significant, and negatively related with publications. As the number of courses taught increases, publication output decreases.

Of particular interest when examining results of RCM are the intercept and slope variances and their significance. When academic discipline is set as the intercept, its relationship with the dependent variable is significant. This would support the notion that discipline is an important variable when examining publication productivity. Additionally, when the number of undergraduate and graduate courses were allowed to vary across disciplines, their relationship was significant with publications. In other words, the number of undergraduate and graduate courses taught has a differential impact on publications across disciplines.

Assessing the impact of the independent variables

While RCM and OLS yield similar results when examining statistical significance, their parameter coefficients are quite different. Examining the impact of one unit change in the statistically significant independent variables has on publications and grant dollars illustrates these differences (See Table 3).

When the dependent variable has been logged, the traditional interpretation of the coefficient for an independent variables changes. In this situation a one-unit change in the independent variable results in a percentage change in the dependent variable, at the mean level of the dependent variable.

For example, in the OLS model, a one-unit increase in the number of courses taught would on average yield a 45.5 percent decrease in grant dollars. With RCM, a one-unit increase on average results in a 51.8 percent decrease in grant dollars. By using OLS, the estimation of the impact of a one-unit change in courses would cause a 6.3 percentage point underestimation of the amount of grant dollars earned on average. When universities are bringing in millions of grant dollars a year, this underestimation is quite dramatic.

When estimating the impact of a one-unit increase in the number of courses taught on publications, a similar, yet less dramatic difference exists. Using OLS, a one-unit increase in courses would result in a 3.7 percent decrease in publications versus the 7.4 percent decrease found in RCM. Relying only on OLS would result in recognizing only half of the actual impact on publications caused by a one-unit change in the number of undergraduate courses.

In general, the impact of one unit change in the statistically significant independent variables is quite dramatic. Examining the impact using RCM reveals that on average, a faculty member that had a research assistantship earns approximately 710.1 percent more grant dollars than a faculty member that did not. With the mean level of grant dollars equal to $154,275, this means that all things being equal, a faculty member who had a research assistantship in graduate school would bring in about $1,095,500 versus $154,275 for one who did not.

Preferences appear to have a large impact on productivity, as well. A faculty member who has a one-unit increase in the preference to do research will on average produce 48.9

percent more grant dollars and 14.5 percent more publications. Note, however, that for these two variables the differences between OLS and RCM are minimal.

Table 3. Impact of One-Unit Change on Publications and Grant Dollars[a]

	Grant Dollars			Publications		
	OLS	RCM	**Difference**	OLS	RCM	**Difference**
Research assistantship	583.3%	710.1%	**-126.8%**	14.5%	21.5%	**-7.1%**
Ph.D.	-	-	**-**	47.1%	64.7%	**-17.6%**
Professional degree	-	-	**-**	37.7%	40.1%	**-2.4%**
Opportunity-research	49.8%	48.9%	**0.9%**	13.8%	14.5%	**-0.7%**
Opportunity-teach	-49.9%	-52.7%	**2.8%**	-12.3%	-13.4%	**1.1%**
Full professor	-	-	**-**	59.8%	53.3%	**6.6%**
Associate professor	-	-	**-**	26.7%	19.5%	**7.3%**
Years in rank	-8.3%	-7.9%	**-0.5%**	-	-	**-**
Courses-undergraduate	-45.5%	-51.8%	**6.3%**	-3.7%	-7.4%	**3.7%**
Age	-	-	**-**	-2.8%	-2.5%	**-0.3%**

[a]Because dependent variables are logged, impact is expressed as a percentage change in the dependent variable.

Conclusion

Institutional research professionals are increasingly called upon to employ statistical techniques to provide support institutional decision-making. Decisions about what modeling technique is appropriate are extremely important. The manner in which universities are organized (with faculty nested in departments, departments in nested colleges and colleges nested within universities) would suggest that multi-level modeling techniques are most appropriate. If these group effects are not accounted for in the type of modeling procedure employed, inaccurate coefficients and subsequently poor analyses are likely to be the result.

Our study of faculty workload shows the importance of using multi-level modeling when studying faculty productivity. Previous research on faculty productivity is flawed for several reasons: the use of inappropriate units of analysis, analysis based on heterogeneous institutions, the use of dependent variables that combine different modes of research productivity, and the use of analytical techniques that overlook the clustered nature of faculty as a result of field of study. We attempt to overcome the shortcomings of previous research through careful data selection and variable construction while also employing multi-level modeling techniques.

Our study highlights several important aspects of the study of faculty work that cannot be overlooked by institutional researchers. First, the group effect of academic field of study should be accounted for when modeling faculty productivity. Simply controlling for discipline with a series of dummy-coded variables can lead to inaccurate results. Employing a multi-level modeling technique like RCM will yield the most accurate coefficients and standard errors without compromising degrees of freedom. Only when a multi-level technique is employed can the impact of other factors on individuals be adequately assessed.

Second, faculty work is extremely complex and cannot be explained using single measures for research productivity. The ability to raise grant money and the ability to publish require different sets of skills. Variables that have a large impact on grant dollars earned do not have a similar impact on publications. Assuming that only one measure can assess research productivity or combining several measures into one does not address differences among faculty.

Besides questions of appropriate estimation methods, our results also have important implications for policy. Traditional OLS methods underestimate the impact of the several important predictors of faculty productivity. The random coefficient models show that the impact of having a research assistantship in graduate school, possessing a Ph.D., and undergraduate course loads are larger than the OLS models would lead us to believe. The magnitude of the variable effects have significant policy implications.

Increasing the diversity of the upper ranks of faculty is an important goal for most universities. But for minorities and women to achieve senior ranks they must produce research. Given our results that indicate the large impact of research training in graduate school on productivity, the role of graduate training and the pipeline for minority populations deserves further analysis.

Policy makers that attempt to boost teaching loads by mandating course minimums should be aware of the consequences. Adding only one course to faculty workloads will have significant ramifications for the amount of grant dollars brought in by a university and the number of publications produced. Further complicating policy decisions is the fact that an increase in undergraduate teaching loads affects disciplines differently. A policy that requires all faculty to teach a minimum number of courses will have a different impact on those teaching in engineering than those teaching in the arts.

Multiple academic disciplines are ubiquitous in institutions of higher education. Institutional researchers must take this into account. In the past, researchers have overlooked the effect of groups, which in turn has probably caused inaccurate empirical results. With the advent of new multilevel modeling techniques, we can now provide more accurate information to policy makers on college and university campuses.

Appendix - Distribution of Faculty by Field of Study and Academic Discipline

Code	Field of study and academic discipline	N	%
	AGRICULTURE		
101	Agribusiness & Agricultural Production	8	0.7
102	Agricultural, Animal, Food, & Plant Sciences	40	3.6
103	Renewable Natural Resources, including Conservation, Fishing, & Forestry	11	1.0
110	Other Agriculture	13	1.2
	ARCHITECTURE & ENVIRONMENTAL DESIGN		
121	Architecture & Environmental Design	10	0.9
122	City, Community, & Regional Planning	2	0.2
123	Interior Design	3	0.3
130	Other Arch. & Environmental Design	4	0.4
	ART		
141	Art History & Appreciation	5	0.5
143	Dance	2	0.2
144	Design (other than Arch. Or Interior)	4	0.4
145	Dramatic arts	11	1.0
146	Film Arts	2	0.2
147	Fine Arts	11	1.0
148	Music	26	2.4
149	Music History & Appreciation	2	0.2
150	Other Visual & Performing Arts	5	0.5
	BUSINESS		
161	Accounting	14	1.3
162	Banking & Finance	13	1.2
163	Business Administration & Management	4	0.4
164	Business administrative support (e.g. Bookkeeping, Office Management, Secretarial)	3	0.3
165	Human Resources Development	3	0.3
166	Organizational Behavior	3	0.3
167	Marketing & Distribution	6	0.5
170	Other Business	5	0.5
	COMMUNICATIONS		
182	Broadcasting & Journalism	15	1.4
183	Communications Research	7	0.6
190	Other Communications	8	0.7
	COMPUTER SCIENCE		
201	Computer & Information Sciences	19	1.7
204	Systems Analysis	1	0.1
210	Other Computer Science	2	0.2
	EDUCATION		
221	Education, General	1	0.1
223	Bilingual/Cross-cultural Education	3	0.3
224	Curriculum & Instruction	7	0.6
225	Education Administration	11	1.0
226	Education Evaluation & Research	4	0.4
227	Educational Psychology	3	0.3
228	Special Education	4	0.4
229	Student Counseling & Personnel Svcs.	3	0.3
230	Other Education	9	0.8
	TEACHER EDUCATION		
241	Pre-Elementary	4	0.4
242	Elementary	4	0.4
243	Secondary	4	0.4
244	Adult & Continuing	1	0.1

245	Other Teacher Ed. Programs	1	0.1
250	Teacher Education in Specific Subjects	8	0.7
	ENGINEERING		
261	Engineering, General	2	0.2
262	Civil Engineering	21	1.9
263	Electrical, Electronics, & Communication Engineering	36	3.3
264	Mechanical Engineering	19	1.7
265	Chemical Engineering	15	1.4
270	Other Engineering	21	1.9
280	Engineering-Related Technologies	4	0.4
	ENGLISH AND LITERATURE		
291	English, General	7	0.6
292	Composition & Creative Writing	9	0.8
293	American Literature	9	0.8
294	English Literature	29	2.6
295	Linguistics	11	1.0
296	Speech, Debate & Forensics	3	0.3
300	English, Other	6	0.5
	FOREIGN LANGUAGES		
311	Chinese (Mandarin, Cantonese, or Other Chinese)	3	0.3
312	French	10	0.9
313	German	9	0.8
314	Italian	2	0.2
315	Latin	3	0.3
316	Japanese	2	0.2
318	Russian or Other Slavic	6	0.5
319	Spanish	16	1.5
320	Other Foreign Languages	5	0.5
350	**HOME ECONOMICS**	21	1.9
360	**INDUSTRIAL ARTS**	1	0.1
370	**LAW**	23	2.1
380	**LIBRARY & ARCHIVAL SCIENCES**	9	0.8
	NATURAL SCIENCES: BIOLOGICAL SCIENCES		
391	Biochemistry	18	1.6
392	Biology	14	1.3
393	Botany	8	0.7
394	Genetics	7	0.6
395	Immunology	2	0.2
396	Microbiology	11	1.0
397	Physiology	7	0.6
398	Zoology	8	0.7
400	Biological Sciences, Other	24	2.2
	NATURAL SCIENCES: PHYSICAL SCIENCES		
411	Astronomy	2	0.2
412	Chemistry	27	2.5
413	Physics	27	2.5
414	Earth, Atmosphere, and Oceanographic (Geological Sciences)	19	1.7
420	Physical Sciences, Other	3	0.3
430	**MATHEMATICS**	52	4.7
440	**STATISTICS**	7	0.6
470	**PARKS & RECREATION**	4	0.4
480	**PHILOSOPHY AND RELIGION**	20	1.8
510	**PSYCHOLOGY**	32	2.9
520	**PUBLIC AFFAIRS (e.g. Community Services, Public Administration, Public Works, Social Work)**	10	0.9
530	**SCIENCE TECHNOLOGIES**	1	0.1

	SOCIAL SCIENCES AND HISTORY		
541	Social Sciences, General	4	0.4
542	Anthropology	12	1.1
543	Archeology	4	0.4
544	Area & Ethnic Studies	1	0.1
546	Economics	21	1.9
547	Geography	7	0.6
548	History	55	5.0
549	International Relations	1	0.1
550	Political Science & Government	26	2.4
551	Sociology	30	2.7
560	Other Social Sciences	8	0.1
	TOTAL	1,104	100.0

References

Austin, A. E. (1996). Institutional and departmental cultures: the relationship between teaching and research. New Directions for Institutional Research, 90, 57-66.

Baird, L. L. (1991). Publication productivity in doctoral research departments: interdisciplinary and intradisciplinary factors. Research in Higher Education, 32(3), 303-318.

Bellas, M. L., & Toutkoushian, R. K. (1999). Faculty time allocations and research productivity: gender, race, and family effects. Review of Higher Education, 22(4), 367-390.

Bentley, R., & Blackburn, R. (1990). Changes in academic research performance over time: a study of institutional accumulative advantage. Research in Higher Education, 31(4), 327-353.

Braxton, J. M. (1996). Contrasting perspectives on the relationship between teaching and research. 90, 5-14.

Bryk, A. S., & Raudenbush, S. W. (1992). Hierarchical Linear Models. Newbury Park, CA: Sage Publications.

Buchmueller, T. C., Dominitz, J., & Hansen, W. L. (1999). Graduate training and the early career productivity of Ph.D. economists. Economics of Education Review, 14, 65-77.

Carnegie Foundation for the Advancement of Teaching. (1994). A classification of institutions of higher education. Princeton, NJ: Carnegie Council for the Advancement of Teaching.

Dundar, H., & Lewis, D. R. (1998). Determinants of research productivity in higher education. Research in Higher Education, 39(6), 607-631.

Fox, M. F. (1992). Research, teaching and publication productivity: mutuality versus competition in academia. Sociology of Education, 65(4), 293-305.

Gander, J. P. (1999). Faculty gender effects on academic research and teaching. Research in Higher Education, 40(2), 171-184.

Greene, W. H. (1997). Econometric analysis. Upper Saddle River, NJ: Prentice Hall.

King, G. (1999). A solution to the ecological inference problem : reconstructing individual behavior from aggregate data. Princeton, N.J.: Princeton University Press.

Kreft, I., & De Leeuw, J. (1998). Introducing multilevel modeling. Thousand Oaks, California: Sage Publications.

Lawrence, J. L., & Blackburn, R. T. (1988). Age as a predictor of faculty productivity. Journal of Higher Education, 59(1), 22-38.

Layzell, D. T. (1996). Faculty workload and productivity: recurrent issues with new imperatives.

Review of Higher Education, 19(3), 267-281.

Levin, S. G., & Stephan, P. E. (1989). Age and research productivity of academic scientists. Research in Higher Education, 30(5), 531-549.

Maryland Higher Education System. (1994). Faculty teaching load analysis. Adelphi, MD: Maryland Higher Education System.

Middaugh, M. F. (1998). How much do faculty really teach? Planning for Higher Education, 27, 1-11.

Neumann, R. (1996). Researching the teaching-research nexus: a critical review. Australian Journal of Education, 40(1), 5-18.

Olsen, D., & Simmons, A. (1996). The research versus teaching debate: untangling the relationships. New Directions for Institutional Research, 90, 31-39.

Olson, J. E. (1994). Institutional and technical constraints on faculty gross productivity in American doctoral universities. Research in Higher Education, 35(5), 549-567.

Perry, R. P., Clifton, R. A., Menec, V. A., Struthers, C. W., & Menges, R. J. (2000). Faculty in transition: a longitudinal analysis of perceived control and type of institution in the research productivity of newly hired faculty. Research in Higher Education, 41(2), 165-194.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. Sociological Review, 15, 351-357.

Singer, J. D. (1999). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. Journal of Educational and Behavioral Statistics.

Stapleton, L. M., & Lissitz, R. W. (1999). Evaluating faculty salary equity using hierarchical linear modeling. Paper Presented at the 1999 American Educational Research Meeting, Montreal, Canada.

Sullivan, A. V. S. (1996). Teaching norms and publication productivity. New Directions for Institutional Research, 90, 15-21.

Tufte, E. R. (1974). Data analysis for politics and policy. Englewood Cliffs, N.J.: Prentice-Hall.

U.S. Department of Education. (1998). 1993 national study of postsecondary faculty user's manual public-use faculty and institution data. Washington, DC: National Center for Education Statistics.

Wanner, R. A., Lewis, L. S., & Gregorio, D. I. (1981). Research productivity in academia: a comparative study of the sciences, social sciences and humanities. Sociology of Education, 54(4), 238-253.