1-1998

Consistent Estimation of Faculty Rank Effects in Academic Salary Models

James G. Strathman
Portland State University

Consistent Estimation of Faculty Rank Effects in Academic Salary Models

James G. Strathman*

January 1998

DP 98-1

Center for Urban Studies
College of Urban and Public Affairs
Portland State University
Portland, OR 97207

* The author is a professor of urban studies and planning, and interim director of the Center for Urban Studies at Portland State University. The comments and assistance of Kathi Ketcheson, Tim O'Brien, Matt Ransom, Ellery Reed, Anthony Rufolo, Tim Schindler and Kay Woodford are gratefully acknowledged.

ABSTRACT

Faculty rank is often included as an explanatory variable in academic salary models. Because there is reason to believe that this results in specification bias, rank effects should be estimated endogenously in salary models. A salary model in which faculty rank is endogenous is estimated in this paper and the results are compared with those obtained from a conventionally specified alternative.

INTRODUCTION

Academic salary models often include faculty rank as an explanatory variable (e.g., Bellas, 1993; Braskamp and Johnson, 1978; Gordon et al., 1974; Moore, 1993). Rank serves as an important proxy for performance because it can reflect, in part, aspects of a faculty member's work that are very difficult to quantify. These aspects, which are commonly identified in an institution's promotion, tenure and merit guidelines, include instructional performance and service to one's profession, institution and community. When combined with other more easy to quantify performance attributes focusing mainly on research and publication activity, rank helps to provide a more complete and representative picture of a faculty member's contribution to his or her institution.

As has been widely discussed, the problem with including faculty rank in salary models is that this variable may be tainted by the inequities that salary models are frequently designed to address. Gender-related inequities have received the greatest attention. To the extent that the promotion process is not gender-neutral, the inclusion of faculty rank will mask underlying gender differences in salaries (Barrett and Sansonetti, 1988; Boudreau et al., 1997; Gunderson, 1989; Moore, 1993; Ramsay, 1979).

Given that salary models have been used to provide evidence of employment discrimination, it has been a responsibility of the courts to decide on the appropriate treatment of rank. According to Barrett and Sansonetti (1988: 511), the courts "... will criticize regressions that include rank when it has already been found that discrimination in promotion or rank was proven." Thus the inclusion of rank in academic salary models has been allowed when separate analysis has found no significant gender effect in promotion.

Although evidence from promotion models provide a safeguard against overlooking indirect sources of discrimination in models of academic salaries, there should be a more basic concern about the use of faculty rank variables. Both salary and faculty rank models include similar explanatory variables. It should thus be expected that rank will be correlated with the error term when it is included in salary models, and that this will contribute to biased parameter estimates. As a result, rank should be treated as an endogenous variable in salary models in order to eliminate the potential for estimation bias (Ramsay, 1979).

This issue is addressed in the remainder of the paper. In the next section the nature of the problem and a correction for it are presented. This is followed by the development of an empirical salary model for a sample of university faculty. The findings from this model are then presented and compared to those obtained from a model in which rank is treated as exogenous. The paper concludes with a discussion of the implications of these findings.

ACADEMIC RANK AND ESTIMATION BIAS

The determinants of academic salaries and rank can be stated in the following general forms:

$$\text{Salary} = f(\text{Rank, Scholarly Activity, Gender, Discipline}) \quad 1)$$

$$\text{Rank} = f(\text{Scholarly Activity, Gender, Experience}) \quad 2)$$

Salaries are defined to be a function of faculty rank, prior scholarly activity (often represented by the number of publications and proxies for quality, such as citations) and disciplinary differences (to account for supply and demand conditions in the academic labor market). A dichotomous sex term is also included to test whether gender-based salary differences remain after accounting for other salary determinants. Salary studies by Bellas (1993), Braskamp and Johnson (1978) and Gordon et al. (1974) are examples that follow this form.

Alternatively, faculty rank is defined to be a function of prior scholarly activity and experience (typically represented by the number of years since receipt of one's terminal degree and sometimes supplemented by information on length of service at an institution). Again, a dichotomous sex term is included to test whether gender-based differences remain after accounting for the other determinants of faculty rank. Studies by Hoffman (1977), Ramsay (1979) and Raymond et al. (1993) are examples employing this form.

With the addition of a stochastic error term equations 1 and 2 can be estimated empirically. The error terms represent the collective unobserved effects of numerous omitted variables, as well as any errors in measuring the dependent variable. To ensure that parameter estimates are consistent the explanatory variables cannot be correlated with the error term. However, as Ramsay (1979) has noted, faculty rank is likely to be correlated with the error term in equation 1, given its correlation with the error term in equation 2. The consequence of this correlation is that the estimated effect of rank on salary will be biased. To mitigate this problem, salary models should treat rank as an endogenous variable.[1]

The appropriate remedy in this case depends on whether the faculty rank variable is also subject to simultaneous equations bias. This would be the case if salary levels were determinants of faculty rank. This seems unlikely, which points to the choice of an Instrumental Variable (IV) technique rather than a simultaneous equations estimator (e.g., two or three-stage least squares).

In the IV approach a first stage regression is estimated for rank from variables that are correlated with it but not with the error term. Then, in the salary regression, the observed values of rank are replaced by their predicted values. In the present case, a number of variables explain both salary and rank, suggesting that the IV approach will produce both consistent and efficient parameter estimates.

Ramsay (1979) has also addressed the problem of estimated bias involving the treatment of faculty rank in salary models. His solution involved estimating a recursive model comprised of a first stage rank equation, specified like equation 2 above, and a second stage salary equation containing predicted rank, disciplinary terms, a gender term and a term to capture price trends from the time of appointment. The gender term is the only variable that is common to Ramsay's two equations. This would imply, for example, that the returns to publications are captured entirely through the process determining one's rank. This seems to be overly restrictive given inclusion of publication activity in many salary models.

The IV approach employed in this paper estimates equation 2 in the first stage and equation 1 in the second stage. In the second stage, observed rank is replaced by its predicted value from the first stage. Terms related to gender, publications, citations, and discipline are included in both equations, which will provide a means of determining their direct and indirect (i.e., through promotion) effects on academic salaries.

MODEL SPECIFICATION

As stated above, the determinants of faculty salaries include factors representing rank, academic discipline, and scholarly productivity and recognition. The model to be estimated here is specified as follows:

$$\text{Salary} = f(\text{Professor, Male, ENG, SBA, Cites, Cites}^2\text{, Articles, Articles}^2), \text{ where} \qquad 3)$$

Salary = Faculty member's salary for the 1994-95 academic year;

Professor = The predicted probability that the faculty member holds the rank of professor;

Male = A dummy variable equaling 1 if the faculty member is a man, and 0 if the faculty member is a woman;

ENG = A dummy variable equaling 1 if the faculty member's appointment is in the School of Engineering, and 0 otherwise;

SBA = A dummy variable equaling 1 if the faculty member's appointment is in the School of Business Administration, and 0 otherwise;

Cites = The number of citations the faculty member's published works received between 1988 and 1993;

$Cites^2$ = The number of citations squared;

Articles = The number of articles published by the faculty member between 1988 and 1993;

$Articles^2$ = The number of articles squared;

Dummy variables are specified for appointments in the schools of business and engineering in order to estimate the salary premiums that faculty in those schools command. Gordon et al. (1974) found that, holding other factors constant, faculty in the areas of health, engineering and science received substantial premiums over faculty in other fields.

Citations and articles are both specified in linear and quadratic forms. This reflects the expectation of diminishing returns, as estimated by Diamond (1986). If so, this would mean that the value of an additional citation or article would be greater for someone with, say, five than for someone with ten citations or articles. Diminishing returns will be exhibited if the parameter estimates on the linear and quadratic citation/publication terms are positive and negative, respectively.[2]

In the first stage the following equation is specified to estimate the instrumental variable for rank:

Professor = (Experience, Male, Articles, Articles2, Cites, Cites2, ENG, SBA) 4)

In this instance, the variable experience (the number of years since receipt of terminal degree) was added to identify the equation. To simplify the analysis, rank is represented as a dichotomous variable (whether or not the faculty member is a professor). This simplification should still provide meaningful insights because the differentials of most interest relating to salary and promotion practices are greatest at this level.

Equation 4 is estimated as a linear probability model, given Heckman's (1978) observation that its simplicity in comparison with probit or logit specifications does not necessarily suffer a loss of estimating efficiency. The estimated probabilities are then substituted for the rank variable in the academic salary model.

DATA

The academic salary model is estimated is estimated from data on 351 Portland State University faculty with tenured or tenurable instructional appointments during the 1994-95 academic year. Excluded were faculty with administrative, fixed term, research, or adjunct appointments. The data employed in the analysis is described below.[3]

Biographical Data. The *1994-1995 Portland State University Bulletin* provided information on rank, experience (years since the receipt of terminal degree), college or school appointment, and gender.

Publication Data. Data on published articles was obtained through a search of the UNCOVER database. This database consists of titles and abstracts from periodicals maintained by subscribers to the Colorado Alliance of Research Libraries (CARL). It covers over 13,000 periodicals and extends from 1988 to the present.

Citation Data. Data on citations from 1988 to 1993 were collected from the *Social Science Citation Index* and the *Science Citation Index.*

Salary Data. The list of instructional faculty and salary data was obtained from PSU's Office of Institutional Research and Planning.

RESULTS

Parameter estimates for the salary model are presented in Table 1. For comparison, results are presented for both ordinary least squares (OLS) and IV-estimated versions. Rank is treated as exogenous in the OLS version of the model, while it is estimated endogenously in the IV version.

The general performance of both models is quite good, with each explaining about 75 percent of the variance in faculty salaries, and nearly all of the variables being statistically significant and having the expected effect. Both models estimate that publications have a positive effect, with each published article contributing over seven hundred dollars in additional salary. The hypothesized diminishing returns to publication are not supported, however, as neither of the quadratic terms is significant. Alternatively, the returns to citations are positive and diminishing in both models. The IV model, for example, estimates the value of a faculty member's first citation to be $246 and the value of his or her tenth citation to be $201. These values lie within the range of those estimated and reported by Diamond (1986). Both models also estimate similar salary premiums for faculty with appointments in engineering and business. These faculty are estimated to receive approximately twelve and twenty thousand dollars more, respectively, than faculty appointed in the university's other schools and colleges.

(Table 1 about here)

As expected, the most noticeable difference between the OLS and IV model estimates are associated with the rank and gender variables. The OLS model estimates a salary premium of nearly twelve thousand dollars for professors, while the fourteen thousand dollar premium estimated in the IV model is nearly twenty percent larger. The OLS model also estimates that men earn about fifteen hundred dollars more than women, while the gender gap estimated in the IV model is nine hundred dollars. Also, the gender difference in salaries is significant in the OLS model, but not in the IV model.

The estimates associated with gender and rank in the OLS and IV models are consistent with the contention that the most likely source of discrimination among faculty is in the promotion process. In the present example, the OLS model treats rank as exogenous and, holding it constant, estimates a significant gender gap in salaries. The IV model treats rank as endogenous and estimates a much smaller gender gap. An important question associated with the IV model results, then, is whether a significant gender difference is associated with the likelihood of holding the rank of professor.

Table 2 reports the first stage linear probability estimates from the IV model. The likelihood of holding the rank of professor is estimated to increase significantly with experience and with publications. Men are also estimated to be nearly 11 percent more likely than women to hold the rank of professor (.47 versus .42, estimated at the means of the right hand side variables), consistent with the contention stated above, but the difference is not significant.

(Table 2 about here)

CONCLUSIONS

This paper has focused on the specification of faculty rank in academic salary models. The main contention, that correlation between faculty rank and the error term lead to inconsistent estimates of the effect of rank on salary, was supported by findings from alternative models in which rank was treated as an endogenous and an exogenous variable. Also, others have noted that rank and gender effects can be confounded, providing another reason to treat rank as an endogenous variable. The present study's findings are consistent with this latter issue, but they lacked statistical significance.

More generally, the IV approach developed in this paper provides a consistent framework for estimating the effects of rank and evaluating gender differences in two critical related areas, salaries and promotion. The first stage of the IV model can identify gender differentials in the promotion process, an underlying potential source of salary differences. The second stage addresses the salary distinction directly, with any underlying gender differential in promotion accounted for. Thus the model consistently addresses two potential sources of wage discrimination in the academic labor market. Because these two sources are sometimes confounded, the IV model also addresses the question of whether or not to remedy gender differences through salary or promotion processes. In this case the response could be the promotion process (if the gender effect is significant there and not in the salary equation), the salary process (if the gender effect is significant there and not in the promotion equation), or both (if the gender effect is significant in both equations).

FOOTNOTES

1. One potential remedy would be to drop the rank variable from the salary model. As Boudreau et al. (1997) and Ramsay (1979) point out, however, eliminating this variable may result in specification error given its strong predictive effect on salaries.
2. Information on the number of books in print authored by PSU faculty was also collected. The number of faculty with books in print was fairly small, however, and initial analysis showed no effect of book publishing on salary. It was therefor dropped from the model.
3. The data is available from the author upon request.

REFERENCES

Barrett, G.V. and Sansonetti, D. M. (1988). Issues concerning the use of regression analysis in salary discrimination cases. *Personnel Psychology* 41(3): 503-516.

Bellas, M. (1993). Faculty salaries: Still a cost of being female? *Social Science Quarterly* 74(1): 62-75.

Boudreau, N., Sullivan, J., Balzer, W., Ryan, A., Yonker, W., Thorsteinson, T., and Hutchinson, P. (1997). Should faculty rank be included as a predictor variable in studies of gender equity in university faculty salaries? *Research in Higher Education* 38(3): 297-312.

Braskamp, L.A., and Johnson, D.R. (1978). The use of a parity-equity model to evaluate faculty salary policies. *Research in Higher Education* 8(1): 57-66.

Diamond, Arthur M. Jr. (1986). What is a citation worth? *Journal of Human Resources* 21(2): 200-215.

Gordon, N.M., Morton, T.E. and I.C. Braden. (1974). Faculty salaries: Is there discrimination by sex, race, and discipline? *American Economic Review* 64(3): 419-427.

Gunderson, M. (1989). Male-female wage differentials and policy responses. *Journal of Economic Literature* 27(1): 46-72.

Heckman, J.J. (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica* 46(6): 931-959.

Hoffman, E.P. (1977). Determinants of faculty rank: The discriminant analysis approach. *Quarterly Review of Economics and Business* 17(2) 79-88.

Moore, N. (1993). Faculty salary equity: Issues in regression model selection. *Research in Higher Education* 34(1): 107-126.

Ramsay, G. (1979). A generalized multiple regression model for predicting college faculty salaries and estimating sex bias. In T.R. Pezzullo and B.E. Brittingham (eds.), *Salary Equity*, pp. 37-53. Lexington MA: Lexington Books.

Raymond, R.D., Sesnowitz, M.L. and Williams, D.R. (1993). Further evidence on gender and academic rank. *The Quarterly Review of Economics and Finance* 33(2): 197-215.

Table 1

Parameter Estimates for the OLS and IV Salary Models
(Dependent Variable = 1994-95 Salary)

Variable	Means[1]	OLS Model[2]	IV Model[2]
Professor	.46 (.50)	11957 (18.10)*	14237 (13.35)*
Male	.69 (.46)	1522.4 (2.13)*	906.7 (1.21)
Articles	1.33 (2.23)	738.2 (2.92)*	716.1 (2.82)*
$Articles^2$	6.73 (35.0)	-10.1 (-0.63)	-11.9 (-0.73)
Citations	7.24 (16.02)	264.7 (6.09)*	249.0 (5.66)*
$Citations^2$	308.27 (1387.1)	-2.59 (-5.22)*	-2.54 (-5.08)*
Engineering Faculty	.10 (.30)	12420 (11.62)*	12789 (11.82)*
Business Faculty	.11 (.31)	19968 (19.29)*	20020 (19.26)*
Intercept		36199 (54.96)*	35657 (51.64)*
R^2		.75	.74[3]
SEE		5821	5845
n		351	351

1. Standard deviations are in parentheses in this column.
2. The values in parentheses in these two columns are t-ratios (OLS Model) and asymptotic t-ratios (IV Model). Values that are significant at the .05 level are denoted with an asterisk.
3. In the IV model the reported R^2 statistic is the squared correlation between observed and predicted salaries.

Table 2

Parameter Estimates for the First Stage Linear Probability Equation
(Dependent Variable = Professor)

Variable	Coefficient[1]
Years Experience[2]	.033 (14.68)*
Male	.046 (0.99)
Articles	.037 (2.25)*
$Articles^2$	-.001 (-0.73)
Citations	.004 (1.49)
$Citations^2$	-.0002 (-0.47)
Engineering Faculty	.029 (0.41)
Business Faculty	.066 (1.00)
Intercept	-.213 (-4.15)*
R^2	.45
SEE	.37
n	351

1. Values in parentheses are t-ratios. An asterisk indicates significance at the .05 level.
2. The mean and standard deviation for experience are 16.9 and 9.5 years, respectively.