Faculty Equity Regression Study – 2022-23

April 18, 2023

A. Edwards and T. Lu

Introduction

Multiple regression analysis is a statistical technique that determines which independent variables appear to have a significant effect on a single dependent variable. The University of Illinois Urbana-Champaign began using multiple regression analysis in the early 1990s to examine the factors that might contribute to faculty salaries; this report describes the results of the 2022-23 study.

The study is divided into two parts. The first can be considered "diagnostic"; it attempts to determine whether there is a systematic, campus-wide bias in the setting of salaries based on inappropriate factors such as sex or race/ethnicity. If the regression coefficients for the sex and race/ethnicity terms are significantly different from zero, then these factors may be affecting salaries. We build regression models separately for each rank (full, associate, and assistant professors) and for all ranks combined to examine this question. In addition, we examine new assistant professors (in their first three years of assistant professor position) in a separate regression to see if there are any biases at this early, critical stage of salary determination.

The second part of this study aims to identify individual faculty members whose salaries are lower than would be expected given their rank, discipline, time in the workforce, and other "appropriate" factors; the inappropriate factors of sex and race/ethnicity are omitted. Each faculty member's factors are substituted into a regression equation to compute a "predicted" salary. Because our model lacks good measures of quality and productivity, it cannot predict salaries perfectly; we expect salaries to vary from the predictions due to quality and productivity. Nevertheless, the predictions give the campus and deans a place to begin discussions of whether individual salaries are set appropriately.

Summary of current results

Diagnostic models: Five regression models (professors, associate professors, all assistant professors, new assistant professors, and all ranks combined) were constructed to examine whether there were any systematic biases in setting of salaries based on sex or race/ethnicity. At the 5% significance level, none of the models showed a noticeable sex bias.

At the 5% significance level, two models showed bias on race/ethnicity groups. The Hispanic group was paid $6,509 higher than the White group in the All Ranks combined model. In the Full professor model, the Asian group was paid $5,943 higher than the White group and the Hispanic group was paid $15,591 higher than the White group, but the Other Non-White group was paid $15,983 less than the White group. It is possible that the interactive effects of race/ethnicity and other variables may explain some of the difference.

All results are summarized in Table 1, with additional details shown in Appendix A. Complete regression printouts are available at

http://www.dmi.illinois.edu/docs/reg/

Table 1. Summary of Significant Effects (p<.0500) found in diagnostic models

Model	Sex effects	Race/ethnicity effects
All faculty ranks combined	not significant	Hispanic was paid $6,509 more than White (p=0.0031)
Full professors	not significant	Asian was paid $5,943 more than White (p=0.0416); Hispanic was paid $15,591 more than White (p=0.0008); and Other Non-White was paid $15,983 less than White (p=0.0233)
Associate professors	not significant	not significant
All Assistant professors	not significant	not significant
New assistant professors (tenure codes 1,2,3) (also included in "All Assistant professors")	not significant	not significant

Identification of potentially underpaid faculty: To analyze individual salaries, a regression model was built omitting the sex and race/ethnicity terms. The "all-ranks-combined" regression cannot include some "quality" indicators such as years to reach full professor; the only "quality" indicator among the independent variables is whether the faculty member was hired in as an assistant professor or at a higher rank. Thus, the predicted salaries are based on factors that largely ignore quality and productivity.

The coefficients from this regression were then used to predict salaries of individual faculty members. The salaries predicted for each individual using this model represent the best estimate of salary from available and measurable faculty characteristics. Any deviation of a faculty member's actual salary from the predicted salary should be due entirely to characteristics we have not attempted to measure, notably quality and productivity.

The distribution of differences between actual and predicted salary, expressed as a percent of the predicted salary, is shown in Table 2. Women faculty members are 31% of the group with actual salaries *15% or more below predicted salaries; they are 9% of the overall women faculty population.*

Table 2. Faculty whose salaries vary from predicted salary

Range	Number and Percent of Men & Women by Salary Deviation						
	Women			Men			All
	Number	Row %	Col %	Number	Row %	Col %	
15% or more below prediction *	61	31%	9%	134	69%	12%	195
10-15% below	60	35%	8%	110	65%	10%	170
7-10% below	77	42%	11%	108	58%	9%	185
0- 7% below	177	40%	25%	271	60%	24%	448
0- 7% above	153	40%	21%	234	60%	20%	387
7-10% above	43	41%	6%	62	59%	5%	105
10-15% above	55	43%	8%	73	57%	6%	128
15% or more above prediction	91	37%	13%	156	63%	14%	247
All	**717**	**38%**	**100%**	**1148**	**62%**	**100%**	**1865**

* The percentages in Table 2 are not significantly different from those expected except for **15% or more below prediction**, where men are more represented, given the proportion of men and women on the faculty of the corresponding rows.

Next Steps

The salaries and predicted salaries of all faculty members will be examined by campus administrators, deans, and department heads to identify any inappropriate salaries and, if warranted, salary adjustments may be made.

More Details: This report is a management overview and omits much of the detail that would be presented in a published paper. Complete appendices and regression diagnostics are available on the web at http://www.dmi.illinois.edu/docs/reg/

Appendix A. Regression Results
Model used: Department dummy variables instead of peer salaries
Estimate of Coefficients for Each Independent Variable

Notes: The coefficients for each of the 80 departmental dummy variables are not included here but can be found on the web site http://www.dmi.illinois.edu/docs/reg

n/s = Coefficients are not significantly different from zero at the 5% level (Student's T test)

FY23 Prob |T| > 0: Using a two-tailed T-test, the probability that a parameter estimate for FY23 data is different from 0.0500 (5%) was used as the cutoff for significance in this study.

Starting year 2016-17, we report Faculty Salary Equity Regressions every three years.

A1. All Faculty Combined	FY14	FY15	FY16	FY17	FY20	FY23	FY23 Prob > |T|
Full Professor=Y	35,913	37,425	36,137	36,275	37,727	40,962	<.0001
Associate Prof=Y	6,523	6,662	5,966	7,294	n/s	6,594	<.0001
Administrator=Y	21,786	17,191	18,011	18,799	15,033	17,590	<.0001
Number of depts.	7,436	10,752	8,609	8,847	11,916	11,015	<.0001
First hired as an asst prof=Y	-12,985	-13,052	-13,270	-13,252	-15,010	-12,781	<.0001
Doctorate=Y	n/s	n/s	n/s	n/s	n/s	n/s	0.0985
Years from degree	473	536	608	633	725	701	<.0001
Sex=male	n/s	n/s	n/s	n/s	n/s	n/s	0.6433
Race=Native American	n/s	n/s	n/s	n/s	n/s	n/s	0.5988
Race=African American	n/s	n/s	n/s	n/s	n/s	n/s	0.5992
Race=Hispanic	n/s	5,355	n/s	n/s	n/s	6,509	0.0031
Race=Asian	n/s	n/s	n/s	n/s	n/s	n/s	0.3383
Race=Other	n/s	-4,995	n/s	n/s	n/s	n/s	0.9246
Y-axis intercept (b_0)	81,310	88,469	90,087	91,414	97,925	113,968	<.0001

A2. Full Professors	FY14	FY15	FY16	FY17	FY20	FY23	FY23 Prob > |T|
Administrator=Y	27,480	22,137	21,659	22,624	19,827	21,921	<.0001
Number of depts.	10,138	14,141	12,532	10,265	11,560	10,559	<.0001
First hired as an asst prof=Y	7,402	9,843	10,822	11,242	n/s	14,116	0.0001
Doctorate=Y	13,067	n/s	n/s	n/s	n/s	15,841	0.0046
Years from degree	951	1,050	1,087	1,123	1,274	1,304	<.0001
Sex=male	n/s	n/s	n/s	n/s	n/s	n/s	0.2548
Race=Native American	n/s	n/s	n/s	n/s	n/s	n/s	0.6638
Race=African American	n/s	n/s	n/s	n/s	n/s	n/s	0.1536
Race=Hispanic	n/s	12,935	n/s	11,042	n/s	15,591	0.0008
Race=Asian	n/s	n/s	n/s	n/s	n/s	5,943	0.0416
Race=Other	n/s	n/s	n/s	n/s	n/s	-15,983	0.0233
Years to reach full prof	-2,236	-2,607	-2,764	-2,708	-3,045	-2,899	<.0001
Y-axis intercept (b_0)	96,755	107,778	109,945	121,606	124,612	147,356	<.0001

A3. Associate Professors	FY14	FY15	FY16	FY17	FY20	FY23	FY23 Prob > |T|
Administrator=Y	8,903	7,678	9,931	13,429	8,779	9,628	<.0001
Tenured=Y	n/s	n/s	n/s	n/s	n/s	n/s	0.1554
Number of depts.	n/s	n/s	n/s	4,224	n/s	9,377	<.0001
First hired as an asst prof=Y	n/s	n/s	n/s	n/s	n/s	n/s	0.4278
Doctorate=Y	n/s	n/s	n/s	n/s	n/s	n/s	0.5076
Years from degree	-308	-279	-205	-175	n/s	-288	0.0020
Sex=male	n/s	n/s	n/s	n/s	n/s	n/s	0.2785
Race=Native American	n/s	n/s	n/s	n/s	n/s	n/s	0.3944
Race=African American	n/s	n/s	n/s	n/s	n/s	n/s	0.5664
Race=Hispanic	n/s	n/s	n/s	n/s	n/s	n/s	0.4227
Race=Asian	n/s	n/s	n/s	n/s	n/s	n/s	0.5924
Race=Other	n/s	n/s	n/s	n/s	n/s	n/s	0.3066
Years to reach assoc prof	-856	n/s	n/s	n/s	n/s	n/s	0.1270
Y-axis intercept (b_0)	109,970	113,241	111,086	106,703	112,696	125,477	<.0001

A4. All Assistant Professors	FY14	FY15	FY16	FY17	FY20	FY23	FY23 Prob > |T|
Number of depts	4,267	5,531	6,278	5,120	4,049	8,325	<.0001
Doctorate=Y	n/s	n/s	n/s	n/s	n/s	n/s	0.8548
Years from degree	245	421	287	226	355	n/s	0.5314
Sex=male	n/s	n/s	n/s	n/s	n/s	n/s	0.1983
Race=Native American	n/s	n/s	n/s	n/s	n/s	n/s	0.6661
Race=African American	n/s	n/s	n/s	n/s	n/s	n/s	0.1415
Race=Hispanic	n/s	n/s	n/s	n/s	n/s	n/s	0.4083
Race=Asian	n/s	n/s	n/s	n/s	n/s	n/s	0.5278
Race=Other	-2,356	n/s	-2,085	n/s	n/s	n/s	0.4630
Y-axis intercept (b_0)	90,121	91,145	91,194	94,601	105,017	109,589	<.0001

A5. New Assistant Professors*	FY14	FY15	FY16	FY17	FY20	FY23	FY23 Prob > |T|
Number of depts	10,369	6,538	7,301	4,418	n/s	8,969	0.0002
Doctorate=Y	n/s	n/s	3,769	n/s	n/s	n/s	0.9564
Years from degree	n/s	332	351	n/s	n/s	n/s	0.1206
Sex=male	n/s	n/s	n/s	n/s	n/s	n/s	0.9309
Race=Native American	n/a	n/s	n/s	n/s	n/a	n/a	n/a
Race=African American	n/s	n/s	n/s	n/s	n/s	n/s	0.0892
Race=Hispanic	n/s	n/s	n/s	n/s	n/s	n/s	0.7324
Race=Asian	n/s	n/s	n/s	n/s	n/s	n/s	0.3472
Race=Other	n/s	n/s	n/s	n/s	n/s	n/s	0.8436
Y-axis intercept (b_0)	76,582	89,362	92,041	100,066	102,026	106,845	<.0001

* New assistant professors are reported separately here and also in the regression for all assistant professors.

Appendix B -- Demographic Profile of Faculty Selected
B1. Men and Women Combined

		All Faculty	Full Professors	Associate Professors	Assistant Professors
Number		1865	829	534	502
Percent with an administrative appointment		20.1%	32.4%	18.9%	1.0%
Sex	Women	717	228	243	246
	Men	1148	601	291	256
Race/Ethnic Group	Am. Ind./Alaska Nat.	4	1	1	2
	Asian	351	147	109	95
	African-American	88	31	28	29
	Nat. Hawaiian/P. I.	0	0	0	0
	Hispanic	141	52	37	52
	White	1148	580	333	235
	Other Non-White	133	18	26	89
Faculty Type	Regular	1787	818	492	477
	Library	78	11	42	25
Tenure status	Tenure Track	506	0	4	502
	Indefinite Tenure	1359	829	530	0
First rank Hired In	Associate or full professor	379	309	70	0
	Assistant Professor	1486	520	464	502
Highest Degree	Not doctoral level	182	66	75	41
	Doctoral level	1683	763	459	461
Years since degree	Mean	19.3	28.2	17.0	7.1
	High	63.7	63.7	51.0	21.7
Age	Mean	50.0	58.0	48.4	38.4
	High	87.6	87.6	77.3	53.9
	Low	28.0	36.2	32.8	28.0
9-month, 100% salary	Mean	137,366	171,103	114,247	106,246
	High	436,090	436,090	299,401	262,623
	Low	52,683	86,447	61,481	52,683
Years at UIUC	Mean	13.2	19.9	11.9	3.4
	High	53.3	53.3	47.4	9.5
Mean Years from hire	To Associate professor	4.9	4.4	5.4	-
	To Full professor	8.8	8.8	-	-

Appendix B -- Demographic Profile of Faculty Selected
B2. Women only

		All Faculty	Full Professors	Associate Professors	Assistant Professors
Number		717	228	243	246
Percent with an administrative appointment		19.1%	37.7%	19.3%	1.6%
Race/Ethnic Group	Am. Ind./Alaska Nat.	3	1	1	1
	Asian	131	29	53	49
	African-American	48	13	15	20
	Nat. Hawaiian/P. I.	0	0	0	0
	Hispanic	60	21	17	22
	White	430	162	147	121
	Other Non-White	45	2	10	33
Faculty Type	Regular	660	220	214	226
	Library	57	8	29	20
Tenure status	Tenure Track	248	0	2	246
	Indefinite Tenure	469	228	241	0
First rank Hired In	Associate or full professor	118	84	34	0
	Assistant Professor	599	144	209	246
Highest Degree	Not doctoral level	91	25	36	30
	Doctoral level	626	203	207	216
Years since degree	Mean	16.9	27.1	17.0	7.2
	High	63.7	63.7	37.7	21.7
Age	Mean	48.1	57.5	48.8	38.7
	High	87.6	87.6	70.8	53.9
	Low	28.0	40.4	34.0	28.0
Years at UIUC	Mean	11.3	19.2	12.0	3.3
	High	44.3	44.3	34.4	9.4
Mean Years from hire	To Associate professor	5.0	4.5	5.4	-
	To Full professor	9.6	9.6	-	-

Appendix B -- Demographic Profile of Faculty Selected
B3. Men only

		All Faculty	Full Professors	Associate Professors	Assistant Professors
Number		1148	601	291	256
Percent with an administrative appointment		20.7%	30.4%	18.6%	0.4%
Race/Ethnic Group	Am. Ind./Alaska Nat.	1	0	0	1
	Asian	220	118	56	46
	African-American	40	18	13	9
	Nat. Hawaiian/P. I.	0	0	0	0
	Hispanic	81	31	20	30
	White	718	418	186	114
	Other Non-White	88	16	16	56
Faculty Type	Regular	1127	598	278	251
	Library	21	3	13	5
Tenure status	Tenure Track	258	0	2	256
	Indefinite Tenure	890	601	289	0
First rank Hired In	Associate or full professor	261	225	36	0
	Assistant Professor	887	376	255	256
Highest Degree	Not doctoral level	91	41	39	11
	Doctoral level	1057	560	252	245
Years since degree	Mean	20.9	28.7	16.9	7.0
	High	63.7	63.7	51.0	19.7
Age	Mean	51.2	58.2	48.2	38.1
	High	83.9	83.9	77.3	52.7
	Low	29.4	36.2	32.8	29.4
Years at UIUC	Mean	14.3	20.2	11.8	3.5
	High	53.3	53.3	47.4	9.5
Mean Years from hire	To Associate professor	4.8	4.4	5.4	-
	To Full professor	8.5	8.5	-	-

Appendix C. Methodology

General approach
This model assumes that the salary paid to a faculty member (the "dependent variable") is a linear function of a set of "independent variables", x_1 to x_n:

$$predicted\ salary = b_0 + b_1x_1 + b_2x_2 + \ldots + b_nx_n$$

The symbols $x_1 \ldots x_n$ are the values of the independent variables, e.g. age. The symbols $b_0 \ldots b_n$ are constant coefficients; the regression model attempts to estimate these coefficients and determine which, if any, are significantly different from 0. If reliable estimates of the regression coefficients can be obtained, we may predict what the salary should be for any faculty member for whom we have the values of the independent variables. The actual salary of a faculty member may differ from the predicted salary because of:

- Error in the specification of the model. The terms may not be linear, for example.
- Critical factors may have been omitted which cause changes in salary. Certainly, the quality of a faculty member's work is one independent variable which is difficult to quantify and include.
- Error in measurement of one of the variables. For example, the dependent variable salary can be calculated in several equally valid ways.

Faculty members were identified and relevant data for each faculty member were pulled from the administrative computer databases. The data were entered into the computer databases for statistical analysis. A total of 1865 faculty members were identified; demographic characteristics are in Appendix B.

Initial selection of faculty*:* Faculty were defined as any person who holds a currently active tenured or tenure-track job on the Urbana campus, which includes campus and central administration employees located on this campus, whose employment status was "active" on October 15 and at least one appointment extending past May 15. We eliminated all faculty with a "T" contract (terminated) and faculty who were retiring during the year.

Dependent variable: 9 month, 100% Time Salary
Calculation of a meaningful salary for each faculty member was a challenge because of the many ways employees are coded on the payroll. For the purpose of this study, we included all appointments which appeared to be continuing past the academic year, including zero percent administrative stipends. Short term or insignificant appointments (under 60 days and under $350) or lump sum payments were excluded. Appointments active on October 15 were used unless an individual's appointments changed during the year; in these cases, the Mid-year salary (March 15) or the salary at the end of the academic appointment year (August 15) was used.

All salaries were adjusted to represent payment for a nine-month period at 100% time.

Independent variables
Data for the following independent variables were collected. Derivation of each item is described below.

Current faculty rank
Highest degree earned
Years since the highest degree was awarded
Rank into which faculty member was first hired as tenure-system faculty
Years from first hire as tenure-system faculty to reach associate professor
Years from first hire as tenure-system faculty to reach full professor
Number of departments in which a continuing appointment is held
Starting rank at first hiring
Whether the faculty member holds any administrative appointments
Sex
Race and Ethnicity (Hispanic or Not Hispanic): as reported to IPEDS
Percent faculty appointment
Type of faculty appointment (regular or library)

Data pulled from Enterprise Data Warehouse (EDW) database

For each faculty member, the following demographic data was pulled from the EDW:

- Name
- UIN
- Date of first employment as tenure-system faculty at UIUC
- Race/ethnicity code
- Sex
- Tenure appointment college and department code
- Leave codes (to identify those on sabbatical leave, disability leave, leave without pay, etc.)
- Highest degree, degree level, and degree date, when available

Each faculty member may have many different jobs. All jobs not paid on an hourly basis for these faculty members were selected and the following appointment information was downloaded:

- Job department
- Job E-class (to determine if the annual salary was paid out 9/12, 10/12 or 12/12)
- Start and end dates
- Percent time
- Annual salary
- Monthly salary
- Position class code

Data pulled from faculty vitas on the web, from department records, and from the Grey Book (supplement to the BOT minutes with all academic salaries and ranks)

- Highest degree, degree level (whether it was a doctoral, terminal, master, or bachelor degree) and degree date (When in doubt, departments were called to verify the degree level. JD degrees were classed as doctoral level, MFA and MARCH degrees were classed as terminal)
- Date highest degree was awarded (in some cases, we had to call departments for this information when the degree was noted as "expected" on the application form). For faculty members with no degree at all, we used year from age 25 to estimate the years the person had been in the workforce.
- Rank into which faculty member was first hired
- Date of promotion to associate professor (if any)
- Date of promotion to full professor (if any)

Derived data elements

From the downloaded and manually collected data, the following were calculated:

- Highest faculty rank: all administrative and academic professional ranks were ignored. Faculty holding library or extension faculty appointments in addition to appointments with regular faculty rank were classed as regular faculty, regardless of which appointment had a greater percent.
- Highest tenure code:
  - If any tenured appointment was found, code is A
  - If no tenured appointment is found, this code is 1-7 or Q.
- Years since degree to January 1 in the academic year under study.
- Number of different departments in which a continuing appointment is held
  - Includes any department where the faculty member held a zero percent appointment or more that was active on Oct. 15.
- Years from first hire at UIUC to January 1 in the academic year under study.
- Years from first hire to promotion to associate professor & to full professor
  - These data elements will be 0 for those hired in at the associate or full professor level. For faculty who left campus at one rank and returned at a higher rank, an estimate of reasonable promotion dates was made.
- Tenure department
  - This was needed to set a dummy variable for the department. When a faculty member had tenured appointments in multiple departments, the department with the highest percent appointment was used. If all tenured appointments had identical percents, the department with the same home department code was used.

Administrator flag

Administrators were defined as:

All top executives

All department head/chairs that could be identified from appointments

Faculty whose administrative appointment percent was larger than their faculty percent

"Administrative" appointments were defined as academic appointments with tenure code=N and a rank/class code not in the faculty range.

Faculty members with a 0% administrative appointment with pay at least 5% or more of total salary.

Executive flag

The president, vice president for academic affairs, chancellor, vice chancellors, Provost, Vice Provosts, and deans were marked as top executives and excluded from the analyses. Former holders of any of these offices may also be flagged and excluded.

Percent time

Total percent on all appointments active October 15 (or August for those with midyear changes) was calculated.

9-month, 100% equivalent of salary on all continuing appointments

All faculty whose appointments changed after Oct. 15 (change in percent, change in salary, or new appointments beginning after that date.) were identified. For employees with no such midyear changes, only appointments active on Oct. 15 were totaled. For employees with a midyear change, appointments active on August 15 at the end of the appointment year were totaled.

Temporary appointments were eliminated. All other on-going appointments were included.

All salaries were adjusted to be 9-month, 100% equivalents. If the job had an employee class code indicating the period of service was 10 months, the annual salary was multiplied by 9/10. If the appointment was for 11 months service, the annual salary was multiplied by 9/11. For all other appointments, the annual salary was used without adjustment. This yields the salary rate for a 9-month period of service. The nine-month equivalent salary and the percent (unadjusted) for all appointments active on Oct. 15 (or Aug 15 if a mid-year change took place) were totaled for an individual to derive the person's actual current 9-month salary rate. If an individual's total percent time was less than 100%, the calculated salary was adjusted to a 100% equivalent by multiplying it times 100/(total percent time).

Dummy variables for each department

A dummy variable (1/0) was created for each department but one. The coefficient for this variable represents the disciplinary difference in salaries between a department and the department left out (in this case, Agricultural & Consumer Economics).

Dummy variables for race/ethnicity

1/0 for Native American, Asian, African American, Hispanic, Other.

Refining the model

As in the previous study, we eliminated "top executives" (dean level and higher) from the regression analyses. Once the set of independent variables was created and verified, multivariate linear least-squares regression models were built using SAS. Regressions with all faculty members combined and separate regressions by rank were run and the results tabulated. Several other specialized regressions were run as described in the Appendix E.

Determining if an independent variable is a significant factor in determining salary levels

If the coefficient for an independent variable is significantly different from zero, then that variable appears to have a significant effect on salary. To determine if a coefficient was significantly different from zero, we used a Student's T test to estimate the probability that the regression coefficient for that factor was zero. If the probability was 5% or less, we assumed the factor was a significant contributor to salaries. It is important to note that this 5% level is somewhat arbitrary; a similar study performed at the University of Wisconsin (Madison) used a 10% level for significance.

By looking at the estimate of the coefficient for each of the independent variables, we can see the magnitude and direction of the effect each has on salary. If the coefficient for the dummy variable for males is $1000, for example, and if that coefficient is significantly different from 0, we would conclude that being male generally is associated with a salary increase of $1000, all other factors being equal.

Appendix D. Regression Statistics

Overall Statistics for Each Model

Who was included in the model	Coefficient of determination (R-squared)*	Model degrees of freedom	F-value statistic for model	Probability that model is significant
All Faculty	0.8360	92	98.18	<0.0001
Full Professors	0.7473	89	24.56	<0.0001
Associate Professors	0.8912	91	39.80	<0.0001
Assistant Professors	0.9832	86	282.58	<0.0001
New Assistant Professors	0.9863	76	154.80	<0.0001

*This is the fraction of variance of salary "explained" by the regression model

More complete regression diagnostics are available at http://www.dmi.illinois.edu/docs/reg/

Appendix E. Other models examined

Two variants on the regression model were examined. The regression output for each of these is posted at http://www.dmi.illinois.edu/docs/reg/

Using peer salaries instead of dummy variables for each department
Through the 1999-2000 study, we had used an average assistant professor salary for each Illinois department and its peers as a proxy for the starting salary in the discipline. Because this factor has always been the most significant factor in each analysis and because in previous models, it was one of the more difficult measures to derive, the Committee on the Status of Women suggested we replace it with a dummy variable for each department. For several years, we continued running this regression in addition to the regressions with dummy variables. Due to time constraints, we have not repeated this analysis since then.

Replacing the dependent variable (actual salary) with log(actual salary)
This model is frequently used for salary analyses because raises tend to be granted as percentage increases, not as flat dollar amounts. In fact, in the original study in FY94, we tried using log(salary) instead of salary as the dependent variable. At that time, we elected to use salary as a dependent variable because

(1) while log(salary) shows a small increase in the goodness of fit, the two models did not differ greatly in overall significance; and
(2) using log(salary) as a dependent variable makes the coefficients for the independent variables harder to explain to a general audience.

We tried a log(salary) model again with each subsequent year's processing. As expected, there was a slight increase in the goodness of fit (R^2=0.87 as opposed to 0.84 with the linear model). The independent variables that were significant contributors to the salary are similar to those found significant in the linear model; however, no significant difference is found for women using this model.

Examining the interaction of sex with other independent variables in the regression

The Committee on the Status of Women suggested that we should also examine the interaction of sex with other variables, such as years from degree or years from first hire to promotion. To test the significance of these interactions, we examined regressions where we added an interaction term to the model:

$$predicted\ salary = b_0 + b_1x_1 + b_2x_2 + \ldots + b_nx + b_{1*2}(x_1\ x_2)$$

To evaluate the importance of these interactive terms, we look at the significance of the coefficient for the interactive term (b_{1*2} above), the significance of the improvement in the overall predictive accuracy of the model, and the proportion of the variance of the model due to the interactive term ("eta squared"). A summary of results is shown in the table below, and complete diagnostics are available at http://www.dmi.illinois.edu/docs/reg/

Summary of Results Testing Interactive Terms

Interactive term	Interactive Term Coefficient is significant (5% level)?		Overall model improvement	
	All Faculty	Full Professors	All Faculty	Full Professors
Sex x Associate professor flag	No	-	0.00% (n/s)	-
Sex x Full professor flag	No	-	0.00% (n/s)	-
Sex x Years from degree	No	No	0.01% (n/s)	0.00% (n/s)
Sex x Has administrative appointments	No	No	0.00% (n/s)	0.13% (n/s)
Sex x Number of departments	No	No	0.01% (n/s)	0.02% (n/s)
Sex x First Rank=assistant professor	Yes	No	0.04%	0.07% (n/s)
Sex x Years to reach full professor	-	No	-	0.11% (n/s)

All faculty regression: Interactive terms of sex with associate professorship, full professorship, years from degree, having administrative appointments, and number of departments were not significant; but interactive terms of sex with first hired as assistant professor was significant at the 5% level. The proportion of the variance of the model from each of the interactive terms was very small -- the contribution to the overall variance is no more than 0.04% for each of the interactive terms. We conclude that the interaction of sex with each of these variables is small even in the case with significant interactive term.

Full professor regression: None of the interactive terms of sex with years from degree, having administrative appointments, number of departments, first hired as assistant professor, and years to reach full professor were significant at the 5% level. The proportion of the variance of the model from each of the interactive terms was very small -- the contribution to the overall variance is no more than 0.13%.