Report to the UTK Faculty Senate Budget and Planning Committee on Analysis of Faculty Salary Data based upon Gender

Louis J. Gross, Faculty Senate President and Professor of Ecology and Evolutionary Biology and Mathematics
Bruce A. Johnson, Faculty Senate Graduate Assistant and Graduate Student in Electrical and Computer Engineering
April 28, 2007

This report is based upon activities included in a proposal submitted to the Senate Budget and Planning Committee, Provost Holub and Chancellor Crabtree in November 2006. The proposal arose though a prior suggestion submitted to Institutional Research by the Senate Budget and Planning Committee in October 2005. Discussions about the proposed effort were also held with the UTK Commission for Women who encouraged the Chancellor to support the study. With the support of the Chancellor and Provost, data to carry out the analysis were provided by Donald Cunningham, Interim Director of Institutional Research.

OBJECTIVES

1. Evaluate the hypothesis "Are female faculty salaries across UTK significantly different from those of males when one takes into account differences in units, longevity and rank".
2. Develop a methodology to quantitatively analyze the impact of any potential equity adjustment plan on salary distributions across campus.
3. Develop a methodology that is potentially extensible to analysis of other factors possibly affecting salary distributions across campus, including minority status.
4. Develop a methodology to analyze unit-level impacts on any gender-based salary differences across campus (e.g. provide a method to partition the contribution of different units to the distribution of salary differences between genders).

BACKGROUND

The methods applied in the various studies of salary differences carried out by the Office of Institutional Research at the request of the Commission for Women and the administration have been either a regression analysis, determining under the assumptions of the regression whether there is any significant contribution of gender to the salaries of UTK faculty, or a comparison of the differences in average salaries between certain units, with no statistical tests applied. The regression analysis, though providing some insight relative to the above objectives, is constrained in interpretation of its results. These results have noted over many years that on average across campus female faculty members have lower salary levels than males.

Constraints on the interpretation of the regression analyses arise due to the assumptions involved in the analysis (normality of variables, linear effects) as well as the manner in which the data are aggregated in the analysis of certain university units. While certain of the regression assumptions are not strict (e.g. a normal distribution for the variates), this method does not readily allow evaluation of the hypothesis given under Objective 1 above due in part to the small sample sizes when faculty members are disaggregated to department/college and rank. As these studies do not provide unit-by-unit comparisons, the reported differences could be discounted as arising from gender differences not in salaries directly, but in fields of scholarship.

Each year the Faculty Senate Budget and Planning Committee carries out a salary analysis, using data provided by Institutional Research. This study is designed to provide summary information by unit and rank, with a main objective being to allow comparisons to various peer institution group salary averages. It does not provide any information on any potential differences by faculty subgroup such as gender.

An additional analysis has been produced for the Chancellor by Institutional Research in which (according to discussions with the Chancellor - these reports have not been viewed by the authors of this report) a regression is carried out unit-by-unit with salaries regressed against years of service. This allows for particular outliers within a unit to be identified, focusing attention on whether there are reasons for these individuals to have a salary that is far from the regression line for that unit.

METHODS

Description of Resampling:

An approach that is more appropriate than regression to evaluate the hypothesis stated in Objective 1 is a resampling methodology. Resampling methods are applied when it is not clear what the underlying statistical distribution for an observed outcome of an experiment should be, and when there are limited data so that statistical approaches which rely on large sample-sizes cannot be used. These methods involve using a given set of a data again and again, re-assorting them in a random manner to calculate an underlying statistic many, many times. This allows the construction of the distribution for the underlying statistic (this is called "bootstrapping" the sample) and the observed value of this statistic can then be compared to the distribution. Confidence intervals (the "bootstrap confidence intervals") can be directly computed from the resampling distribution of the statistic. If the observed value falls within the chosen confidence level, then the alternative hypothesis is rejected. Resampling methods have become a standard statistical tool (Efron and Tibshirani, 1993; Good, 1999; Kaplan, 1999), which has been applied to analyze gender salary differences (de la Rica et al., 2006; Millimet and Wang, 2006) though we have not found any previous application of this to analysis of faculty salaries.

Application of Resampling (Method 1):

To apply resampling to evaluate the hypothesis in Objective 1, we need to specify a test statistic that is appropriate to measure potential gender-based salary differences. Though a variety of such measures may be used, we have chosen the simplest - the difference between male and female salaries within a unit (department, or college for those colleges without departments), and within rank. We focus here only on faculty at the Assistant Professor rank or above. The method essentially reassigns, randomly, independent of gender, salaries within each unit and rank, to the faculty, using exactly the salaries within that unit. Then the salary differential between males and females in that unit is used as one bootstrap sample, the process is repeated for all units and the average difference across the entire campus is then used as one bootstrap sample of the statistic of concern (average difference between male and female salaries across the campus, accounting for differences in units and ranks, D). This process is repeated many times, giving many values of D, and this generates a distribution for D, illustrated by a histogram. This is the distribution of average faculty salary differences across campus, under the null hypothesis that salaries are equally likely to be assigned to males or females within each rank and unit. Details of the method are included in the appendix.

This process, for each bootstrap sample, creates a "virtual UTK faculty salary allocation". It recreates departments with faculty distributed by rank and salary with exactly the same structure as current (e.g.

the same numbers of faculty at each rank in each unit, the same total salary per unit, and exactly the same individual salaries in the unit). It randomly assigns each salary as male or female, using the gender distribution within the unit and rank, so there are exactly the same numbers of males and females in each unit and rank. In this analysis, we ignore any units in which there are not both genders present at a rank - we count faculty "gender pairs". For each bootstrap sample, the difference in average salary (male minus female) in each unit and rank, is averaged across the entire campus, weighting each unit/rank value according to the number of faculty pairs in each average - e.g. Min(female,male). This weighting provides a scaling to ensure that units/ranks with the most balanced gender ratios are counted more heavily than those with the least balanced gender ratios. This provides a single statistic D for weighted mean salary difference for each "virtual UTK" created.

This distribution of D from the large number of "virtual UTKs" created is then used to determine the significance level of the test of the hypothesis in Objective 1. The hypothesis is evaluated by computing (in exactly the same manner as above) the actual value of D, D*, from the current salary data. If the value of D* is within the 95% confidence interval of the distribution of D from the resampling, then the alternative hypothesis (e.g. that there is a statistically significant gender difference in salaries at UTK) is rejected at the 5% level (two-tailed test).

Alternative Statistic using Faculty Gender-Pair Data (Method 2):

The above method uses the difference in mean salaries for males and females within units/ranks calculated from each bootstrap sample. An alternative to this is not to consider the mean difference over all faculty in a unit, but to calculate the mean difference in salaries within a unit based only on the actual number of gender pairs. Thus if a unit/rank has 3 females and 8 males, after the salaries are assigned randomly to gender, the statistic takes the difference between the average of the 3 assigned female salaries and the average of a randomly chosen set of 3 male salaries. This is then repeated across all units and ranks as in Method 1 to compute the campus-wide statistic E, and the same resampling method is applied as above to generate the distribution of E. Again, the contribution of each unit/rank to E is weighted by the number of gender-pairs in that unit/rank.

The objective of Method 2 is to evaluate whether the exact statistic used affects the resulting hypothesis test, though we expect as the number of samples taken increases, the distribution of E should approach that of D. This constitutes one verification that the method we have chosen is not sensitive to the details of the resampled statistic.

Accounting for Longevity:

There are several additional assumptions that underlie the above analyses. It may be that the length of service within units and ranks differs between genders, and salary may increase with length of service. To account for this effect, we do not include longevity pay in the above analysis - e.g. we utilize the base salary for all faculty, excluding longevity pay. Additionally, we use longevity pay as a proxy measure for longevity of service within rank (since we do not have data on time in rank). This is an additional analysis (Longevity analysis) that considers individual faculty not just by rank in a unit, but by whether their longevity pay is greater or less than the average longevity pay within that unit/rank. Then Methods 1 and 2 above are applied again, but with an additional subgrouping within unit/rank, that of higher than mean or smaller than mean longevity pay. We have only used this difference in longevity at the Associate and Full Professor levels - we do not account for it at the Assistant Professor level.

Note that longevity pay is not a perfect measure of length of time in rank since recently hired individuals may have had long service at that rank at another institution, but have low longevity pay here. Note as well that accounting for longevity in this way reduces the number of gender-pairs as it adds an additional subgroup to each unit/rank. Using a finer division of longevity would lead to further reductions in the number of gender pairs, further reducing the sample sizes usable from the data. Our objective in this analysis is to establish whether or not a simple, basic method to account for some aspects of length of service generates different conclusions from the methods which do not account for longevity.

Productivity and Merit:

As Chancellor Crabtree has noted many times, UTK is a meritocracy, in which merit should be rewarded with higher compensation. A basic assumption in all the above methods is that there are no gender-based differences in productivity which could lead to differential salaries due to merit pay, Chairs of Excellence, named professorships, or other compensation tied to merit. We have no evidence either here or from the studies we are aware of elsewhere that there are gender differences in productivity among faculty. We did consider the possibility of utilizing the faculty evaluations conducted each year (e.g. exceeds expectations, meets expectations, etc.) to potentially account for productivity differences. However an analysis of the distribution of these rankings over several years in the College of Arts and Sciences shows tremendous variance year to year, and between departments, for the number of faculty ranked as exceeding expectations. Given this, we have little justification for using such evaluation data to account for productivity differences between faculty. Thus we make no claim here that the analysis accounts for productivity differences which may occur between genders.

Data utilized:

Donald Cunningham, Interim Director of the Institutional Research, provided the data on faculty salaries as an Excel spreadsheet that included for each faculty member the following:

Rank, gender, unit identifier, base salary, longevity pay

For confidentiality and to reduce chances of bias in the analysis, no faculty identifiers such as name or personnel number were used, and the unit identifiers were numbers assigned to each unit by IR. We were not provided a master list of which units corresponded to which unit identifier. Note that the salary data provided were as of the Fall 2006 payroll, so these did not include merit and other raises that were given to faculty in early 2007. The salaries used are nine-month, do not include longevity pay and include at least one unit (Library) that is not included in some other studies. Thus the salary figures should not be expected to be exactly the same as those given in either the annual Senate Faculty salary analysis (based upon October 2006 data) or the gender salary analysis produced by IR - posted at
http://oira.tennessee.edu/stats/facsal/facsal-06.pdf
that is based on 2005-6 salary data.

Implementation:

The above analyses were coded using the scientific software package Matlab Release 7.2, and carried out only after the methods noted above were developed (e.g. the process of development of the methods and codes was carried out before any analysis of the salary data was done). The shuffling method implemented in the resampling is the one applied in shuffle.m from Kaplan (1999). A test data set was utilized to verify that the calculations in the code were correct.

RESULTS

We include in Table 1 basic descriptive statistics for the salary data utilized in this study. While not identical to the values in the other studies on faculty salaries noted above, the results are quite similar in that there are differences in average salaries across ranks for males and females, and median salaries are smaller than mean salaries. Note that the salaries used in this analysis include a unit (Library) which is not included in some of the other studies. The salary values in Table 1 are for nine-months and exclude longevity pay.

Table 1. Summary of Faculty Salary Data

Rank	Number in Rank	Mean Salary Males	Mean Salary Females	Median Salary Males	Median Salary Females
Assistant	358	60867	54269	57519	53027
Associate	352	73500	66437	72377	62948
Professor	492	97048	84532	89480	82108

Analysis of Objective 1 can be obtained by considering the below four figures. These graph the histograms for the distributions of the resampled statistics D and E for both the case in which longevity status is included and the case in which longevity status is ignored, using 4000 bootstrap samples in each case. The calculated value for the statistic using the actual faculty salaries, D*, lies well outside the 95% confidence intervals (indeed outside essentially the entire range of resamples generated) in all cases. Thus the null hypothesis is strongly rejected. ***The evidence is very strong that the differences in salary between males and females across UTK do not arise from chance assignments of salaries, nor are they explained by differences in gender distributions across units, ranks or longevity status.***

Figure 1: Distribution of 4000 resamples of statistic D using Method 1 and accounting for longevity. The D* value of 3316 is calculating using the actual faculty salaries.

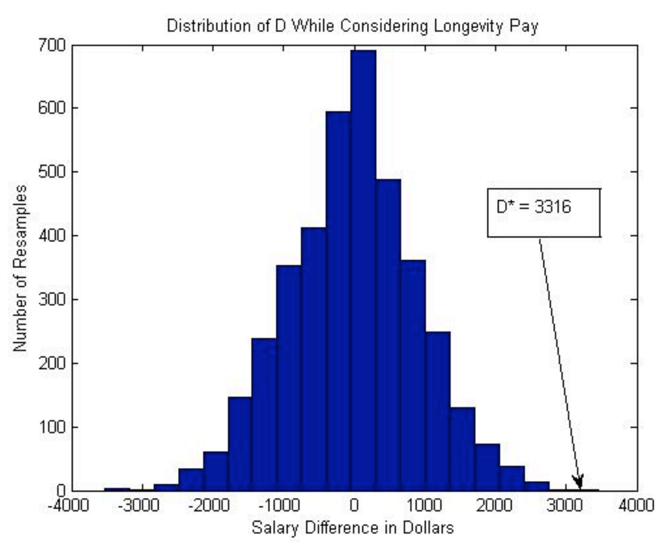


Figure 2: Distribution of 4000 resamples of statistic D using Method 1 without accounting for longevity. The D* value of 3670 is calculating using the actual faculty salaries.

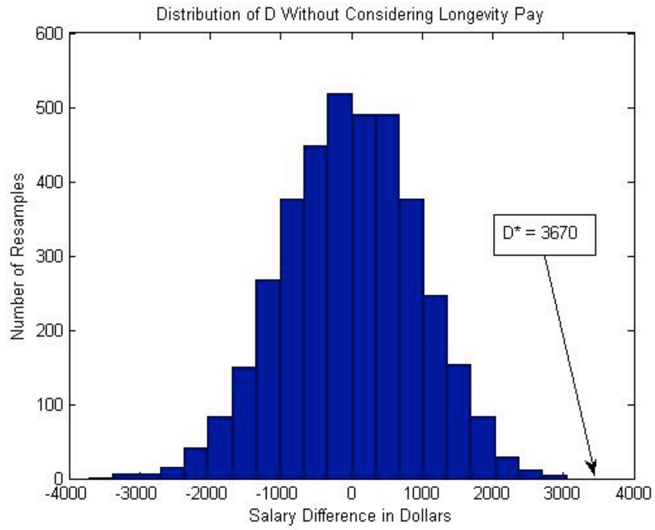


Figure 3: Distribution of 4000 resamples of statistic E using Method 2 and accounting for longevity. The appropriate D* value of 3316 to compare with is off the scale of this graph.

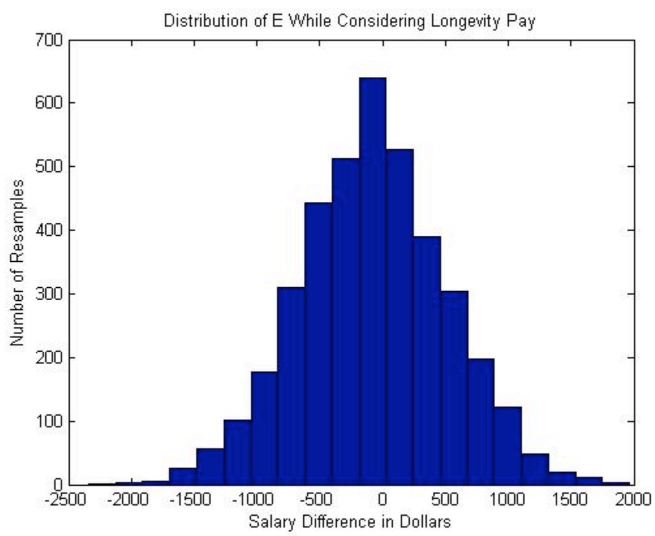
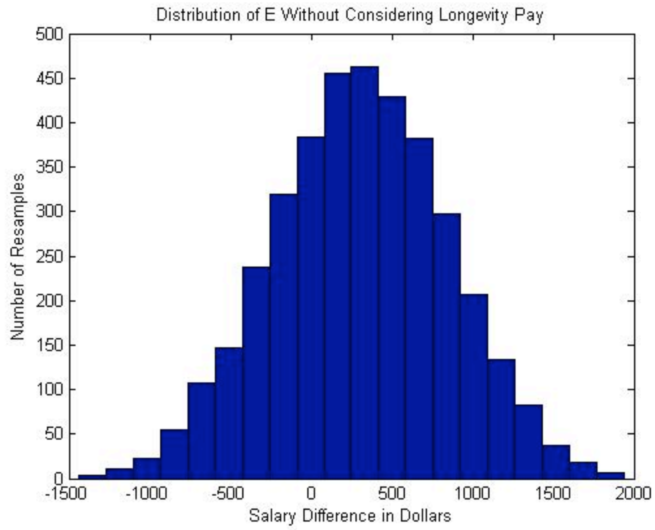


Figure 4: Distribution of 4000 resamples of statistic E using Method 2 without accounting for longevity. The appropriate D* value of 3670 to compare with is off the scale of this graph.



This resulting conclusion arises in both the resampling methods applied – providing evidence that the details of the resampling method do not play a role in the conclusion. As an additional evaluation of this, however, it would be useful to compare the distribution of E*, the statistic which would arise by carrying out the Method 2 resampling, keeping the existing salaries for each gender within each unit/rank fixed at the observed values but shuffling them within gender in the manner used to calculate E. The distribution for E* could then be compared to the distribution of E for an additional evaluation of the impact of resampling. Since the mean of the E* distribution will be D*, for large numbers of resamples, we do not expect any different conclusions to arise from this additional analysis.

Regarding Objective 2 of this study, the impact of any modification of faculty salaries can be readily analyzed by carrying out the same calculations as incorporated here using the modified salaries. For example, the data indicate that average female faculty salaries would need to be increased by 12%, 11% and 15% at the Assistant, Associate and Full Professor levels respectively, to match the average salaries of males. However carrying out such a pattern of increases across the board would not necessarily cause a different conclusion than that derived here for the null hypothesis of equivalent salaries across gender. The distribution of gender across ranks and units could still play a role.

Regarding Objective 3, there is no inherent reason why the methods developed here could not be as readily applied to any other categorization of faculty and a similar salary analysis be applied. Difficulties could arise in making inferences if the categorization leads there to be small numbers of faculty within the subgroups, but the resampling methodology would still apply.

Regarding Objective 4, Table 2 presents a breakdown by unit as to the contribution of each unit to the calculated value of D* - it utilizes the values C(u,r,l) in (1) below, normalized by the sum of all gender pairs across the campus. The order in each table is based upon the magnitude of departmental contributions to D*, from most negative to most positive. The values shown in Table 2 are for the case in which longevity effects within ranks are taken into consideration, while those in Table 3 are the case in which longevity status is not considered. The sum of all departmental contributions in these Tables gives the appropriate D* (for the cases in which longevity is taken into account and that in which it is not). The departmental contributions are negative if average female salary is higher than average male salary (weighted by number of gender pairs) and positive if the reverse is true. The departmental contributions indicate that there are some units that much more greatly contribute to the observed differences in salary across gender than other units.

Also shown in Tables 2 and 3 are the number of gender pairs used in the analysis in each department. The number of gender pairs is lower in a few cases when longevity is taken into account, due to the additional subgrouping when longevity is included. As a method to account for the differences in numbers of gender pairs across departments, Tables 2 and 3 provide the contribution to D* per gender pair in the unit. While the contributions of each department to D* vary somewhat when longevity status is taken into account versus when it is not, there is a very consistent grouping of units at both the very positive and the very negative end of the range. This indicates that longevity status does not greatly modify which departments contribute the most to the unequal salary distributions across gender at UTK.

One possible response to the salary differences reported here and elsewhere would be to focus attention on those units which contribute the greatest to D*. We do not provide here information on which units these are, as it is an administrative matter to carry out a more detailed analysis of the situation within the units which most contribute to the unequal gender distribution in faculty salaries. Such a more detailed analysis would take into account the number of faculty pairs in the unit and the contribution per pair shown in the Tables, determining for example whether the large contribution of a unit is due only to a large number of faculty gender pairs in that unit (and therefore larger contribution to D*). The differences arising in units with large contributions to D* could come about due to particular circumstances in that unit, such as the history of hires of high-salaried Chairs for example. The data supplied here provides a basis for determining whether some corrective action is appropriate, potentially prioritizing this across the large number of departmental units on campus. Thus this Table can focus administrative attention on particular units from which perhaps a reassessment of salaries could have a large impact on the conclusions of this report regarding gender differences in salary.

Table 2: Contributions to the statistic D* arising from each Department, using Method 1 and taking longevity into account. The Department Contributions sum to D* = 3316, the Number of Pairs is the number of gender pairs for each Department, and the Contribution per Gender Pair is the Department Contribution divided by the Number of Pairs.

Department Number	Department Contribution	Number of Pairs	Contribution per Gender Pair
56	-102	7	-14.5
58	-100	3	-33.2
21	-96	2	-47.8
49	-61	3	-20.1
65	-60	9	-6.6
53	-52	7	-7.3
15	-44	2	-21.7
61	-40	3	-13.2
23	-39	1	-38.1
35	-33	3	-10.7

52	-25	2	-12.5
43	-19	2	-9
25	-18	2	-8.5
11	-17	1	-16.3
55	-14	1	-13
51	-6	2	-2.6
3	0	0	0
5	0	0	0
14	0	0	0
18	0	0	0
19	0	0	0
20	0	0	0
27	0	0	0
38	0	0	0
44	0	0	0
47	0	0	0
62	0	0	0
40	0	3	0.2
37	2	9	0.2
42	2	1	2.8
6	4	1	4.3
1	9	1	9.8
31	16	2	8.3
7	24	1	25
24	26	8	3.3
28	34	6	5.8
30	37	2	18.8
29	42	1	42.6
32	45	3	15.2
34	46	1	46
57	52	9	5.8
63	65	1	65.7
41	69	3	23.1
9	72	7	10.4
50	83	3	28
48	87	3	29
45	91	3	30.3
10	97	2	48.8
17	102	1	103
36	103	4	25.9
12	103	3	34.5
13	104	3	34.9
22	105	3	35.1
46	113	10	11.4
39	114	13	8.8
26	115	12	9.6
8	129	5	26
4	135	1	135.3
54	150	2	75.3
59	163	12	13.6
16	163	2	81.6

60	187	10	18.8
2	200	4	50.2
64	235	3	78.5
33	352	10	35.2
66	535	30	17.8

Table 3: Contributions to the statistic D* arising from each Department, using Method 1 without taking longevity into account. The Department Contributions sum to D* = 3670, the Number of Pairs is the number of gender pairs for each Department, and the Contribution per Gender Pair is the Department Contribution divided by the Number of Pairs.

Department Number	Department Contribution	Number of Pairs	Contribution per Gender Pair
29	-186	2	-92.8
58	-89	4	-22
56	-82	7	-11.7
65	-61	9	-6.8
61	-55	4	-13.6
53	-50	7	-7
19	-42	1	-41.9
21	-42	2	-20.9
15	-41	2	-20.2
32	-38	4	-9.4
35	-31	3	-10
43	-21	2	-10.2
49	-19	3	-6.2
25	-19	2	-9.1
52	-7	2	-3.4
3	0	0	0
5	0	0	0
14	0	0	0
18	0	0	0
20	0	0	0
27	0	0	0
38	0	0	0
44	0	0	0
47	0	0	0
62	0	0	0
51	0	2	0.2
42	2	1	2.6
23	5	2	2.8
40	7	3	2.5
11	8	1	8.4

1	9	1	9.1
63	10	3	3.5
17	15	2	8
34	24	1	24.8
31	25	2	12.7
28	26	6	4.5
7	34	1	34.1
6	43	2	21.6
22	53	3	17.8
30	53	3	17.9
54	55	3	18.4
37	61	9	6.8
24	65	9	7.3
57	74	9	8.3
10	78	2	39.3
39	80	13	6.2
50	84	3	28.3
4	93	1	93.8
8	102	5	20.6
41	108	3	36.3
26	115	12	9.6
9	118	7	16.9
36	128	5	25.6
55	135	2	67.6
12	135	3	45.3
16	139	2	69.8
46	154	11	14.1
13	155	3	51.8
48	175	4	44
2	187	4	46.8
60	219	10	21.9
64	219	3	73.1
59	231	12	19.3
33	344	10	34.4
45	400	4	100.1
66	456	30	15.2

REFERENCES

de la Rica, S. et al. 2006. Ceilings or Floors? Gender Wage Gaps by Education in Spain. http://www.ehu.es/SaradelaRica/docs/ceilingflrevnov06.pdf

Good, Phillip I. 1999. Resampling Methods: *A Practical Guide to Data Analysis*. Birkhauser.

Efron, Bradley and R.J. Tibshirani. 2003. *An Introduction to the Bootstrap*. Chapman and Hall.

Kaplan, Daniel 1999. *Resampling Stats in Matlab*. Resampling Stats., Inc.

Millimet, D. L. and L. Wang. 2006. A Distributional Analysis of the Gender Earnings Gap in Urban China. Contributions to Economic Analysis and Policy. Berkeley Electronic Press. http://www.bepress.com/cgi/viewcontent.cgi?article=1461&context=bejeap

Appendix: Details of Methods

Note that the below details are for the cases in which longevity status is included in the analysis. For the cases in which longevity is not included, identical methods are employed but ignoring longevity status in each calculation. Including longevity status reduces in general the number of unit/rank groups which are included in the analysis as it is more likely that a particular unit/rank would not have at least one gender pair in the same longevity status.

A. Calculation of D*

This is the base statistic computed from the actual salary figures for faculty, used to compare to the distribution of the statistic D, which is calculated in the same manner as D* from resampled salary data.

Definitions:

u represents the unit identifier, u = 1,…,U where U is the number of units (in our case there are 66 units)

r represents the rank identifier, r = 1, 2, 3 for Assistant, Associate and Full Professor respectively

l represents the longevity status identifier, l=1,2 where l=1 corresponds to longevity pay below the average for that u, r and l=2 corresponds to longevity pay above or equal to the average for that u, r

g represents gender identifier, g=1,2 with g=1 denoting male and g=2 denoting female

F(u,r,l) = Number of females in unit u, at rank r, with longevity status l

M(u,r,l) = Number of males in unit u, at rank r, with longevity status l

T(u,r,l) = Total number of faculty in unit u, at rank r, with longevity status l
= F(u,r,l) + M(u,r,l)

N(u,r,l) = Min (F(u,r,l), M(u,r,l)) – this is the number of gender pairs in unit u, rank r and longevity status l

S_i(u,r,l,g) = Salary (Base, not including longevity pay) of the i^{th} faculty member in unit u, rank r, longevity status l and gender g. Here i = 1,2, …, M(u,r,l) if g=1 and i = 1,2, …, F(u,r,l) if g=2.

The contribution of unit u, rank r, and longevity l to D* is given by

$$(1) \quad C(u,r,l) = \left(\bar{S}(u,r,l,1) - \bar{S}(u,r,l,2)\right) N(u,r,l)$$

where

$$\bar{S}(u,r,l,1) = \frac{\sum_{i=1}^{M(u,r,l)} S_i(u,r,l,1)}{M(u,r,l)}$$

and

$$\bar{S}(u,r,l,2) = \frac{\sum_{i=1}^{F(u,r,l)} S_i(u,r,l,2)}{F(u,r,l)}$$

Then D* is obtained by averaging the C(u,r,l) values over all units, ranks and longevities

$$(2) \quad D^* = \frac{\sum_u \sum_r \sum_l C(u,r,l)}{\sum_u \sum_r \sum_l N(u,r,l)}$$

B. Resampling Method 1

To create the resampled distribution for D, we utilize random shuffling of the salaries within each unit, rank and longevity status. For each u, r, l, the T(u,r,l) salaries are shuffled to create a new set of salaries

$$\left\{S_i^j(u,r,l,g)\right\} \quad for\ j = 1,....,R$$

where R is the number of bootstrap samples to be used. This process randomly assigns salaries within each unit, rank and longevity status group independent of gender, maintaining the same salaries within this group, as well as the number of each gender. The same process as described above is used to calculate the right hand side of (2) for this bootstrap sample, giving D^j for the j^{th} bootstrap sample of the statistic D.

For the R bootstrap samples of D, calculate the 95% confidence interval by sorting the R values of D^j into order, and determining the range of values which contain the middle 95% of the R values (e.g. if R=1000, then the 95% confidence interval is the range from the 25^{th} value in the list to the 975^{th} value). The number R of bootstrap samples utilized was chosen so that the standard deviation of the distribution of the bootstrap samples changed by less than 1% when the number of bootstraps was doubled, for the case in which longevity was not included. This led to R=4000, starting with R=500 and doing successive doublings.

The null hypothesis (there is no gender difference in salary) is rejected at the 95% level if the D* value falls outside the 95% confidence interval.

Note that the shuffling method used is the one implemented in shuffle.m from Kaplan (1999) and the randomization utilizes Matlab's internal pseudo-random number generator, rand.

C. Resampling Method 2

Proceed as in Method 1 by random shuffling of the salaries within each unit, rank and longevity status to create

$$\left\{S_i^j(u,r,l,g)\right\} \quad for\ j = 1,....,R$$

If M(u,r,l) ≥ F(u,r,l) then randomly sample N(u,r,l) values from the set

$$\left\{S_i^j(u,r,l,1)\right\}$$

and compute the sum of this subgroup of the assigned male salaries as well as the sum of all the assigned female salaries

$$(3) \quad \bar{S}^j(u,r,l,g) = \sum_{i=1}^{N(u,r,l)} S_i^j(u,r,l,g) \quad for\ g = 1,2$$

The contribution to the resampled statistic E^j from this unit/rank/longevity is then

$$(4) \quad C^j(u,r,l) = \bar{S}^j(u,r,l,1) - \bar{S}^j(u,r,l,2)$$

and the j^{th} bootstrap estimate of E is

$$(5) \quad E^j = \frac{\sum_u \sum_r \sum_l C^j(u,r,l)}{\sum_u \sum_r \sum_l N(u,r,l)}$$

Continue as in Method 1 to find the 95% confidence interval and test the null hypothesis.