DOCUMENT RESUME

ED 447 149 TM 031 963

| | |
|---|---|
| AUTHOR | Stapleton, Laura M.; Lissitz, Robert W. |
| TITLE | Evaluating Faculty Salary Equity Using Hierarchical Linear Modeling. |
| PUB DATE | 1999-04-19 |
| NOTE | 29p.; Paper presented at the Annual Meeting of the American Educational Research Association (Montreal, Quebec, Canada, April 19-23, 1999). |
| PUB TYPE | Reports - Evaluative (142) -- Speeches/Meeting Papers (150) |
| EDRS PRICE | MF01/PC02 Plus Postage. |
| DESCRIPTORS | *College Faculty; *Equal Opportunities (Jobs); Higher Education; *Regression (Statistics); Simulation; *Teacher Salaries |
| IDENTIFIERS | *Hierarchical Linear Modeling |

ABSTRACT

This paper presents results from a comparison of the multiple regression (MR) approach to examining faculty salary equity (with clusters for the various disciplines) and hierarchical linear modeling (HLM) for the same problem. The comparison was done in two steps. First, a practical example of applying both techniques, using empirical data, is provided, and then a simulation study examined varying data conditions to show the differences resulting from these two statistical approaches. For the empirical example, data were used for 1,216 tenured and tenure-track faculty grouped into 7 discipline clusters. Variables were: (1) discipline; (2) rank; (3) years in rank; (4) refereed articles; (5) presentations; (6) sponsored research; and (7) gender. For the simulation, 5 discipline clusters, each of 10 departments, were generated for 20 faculty observations for each, a total of 1,000 observations. The empirical study showed that both approaches provided similar results, but that the HLM model could provide more information, although with a more complicated picture of the salary process. Simulation results confirm that if departments within discipline clusters are not homogenous with regard to the percentage of males and salary levels (controlling for all other variables), the analysis may be biased when using MR. An appendix shows cluster averages and the range of department averages. (Contains 14 tables, 4 figures, and 18 references.) (SLD)

# Evaluating Faculty Salary Equity Using Hierarchical Linear Modeling

Laura M. Stapleton and Robert W. Lissitz
University of Maryland, College Park

Paper presented at the Annual Meeting of the American Educational Research Association
Montreal, Canada
April 19, 1999

Correspondence should be directed to the first author at:

2119 Main Administration Building
University of Maryland
College Park, MD 20742

e-mail: lstaplet@accmail.umd.edu
phone: 301/405-5592
fax: 301/314-9443

With increasing emphasis on diversity issues in higher education in general, and specifically in faculty employment, it is foreseeable that equitable pay for faculty across gender and racial groups will continue to be an important policy issue for institutional administration. It is therefore imperative that colleges and universities use valid methods when examining salary equity. Universities and colleges, as well as systems of higher education, often undertake studies of whether there exist pay inequities between groups of faculty on campuses. Examples of such equity studies and discussions of methodological concerns have been plentiful in the literature over the past decade, however, no single method of undertaking such studies has been embraced by the research community. Two higher education unions, American Association of University Professors and United University Professions, have developed "kits" for researchers to use as a guide in undertaking salary equity studies (Scott, 1977; Haignere et. al., 1996), but over time some of their guidelines have been questioned.

In general, the objective of most faculty salary equity studies is to check for systematic gender and/or race bias at an aggregate level (Haignere et. al., 1996). Once the decision is made to undertake a study at an institutional level (as opposed to within one academic unit only), questions arise as to the statistical method to be used and the variables to be included in the analysis. In their reviews of several case studies, Balzer et. al. (1996) and Moore (1993) indicate that the vast majority of studies utilize a multiple regression approach. Variations used within multiple regression approaches include "direct" (or classic) regression, reverse regression (regressing a merit measure on salary and sex or race to identify discrimination in the assignment of merit), two-step regression (entering the variables of interest, such as gender or race in a second step, after the effects of predictor variables have been controlled for), and even step-wise regression (including only those variables based on ordinary least squares optimization) (Moore, 1993; Haignere et. al., 1996).

Equally important to the choice of statistical method are the predictor variables to be included in the model. Consensus seems to have been reached on a limited number of variables, such as years of experience and some measure of discipline or market value (Haignere et. al., 1996). Other variables that have been suggested include academic rank, initial salary, and productivity measures (Balzer et. al., 1996; Snyder et. al., 1994).

## *Concerns with Current Methods*

There have been many concerns voiced about the various methods employed in faculty salary equity studies over the past two decades. These concerns fall into three general categories: choice of variables and the error associated with those variables, interpretation, and statistical technique. While the focus of the current paper is on the statistical technique, a brief review of the other concerns is warranted.

### *Choice of Variables*

The issue of omitted variables is not a minor one, as Boudreau et. al. (1997) demonstrated that "conclusions regarding the presence or absence of gender discrimination do differ ...depending on the particular variables or factors that are included in the model to predict salary" (p. 298). In particular, they argue that exclusion of a faculty member's academic rank can lead to inappropriate conclusions. The use of rank, and initial salary as well, has been hotly contested due to those variables' tendency to mask earlier salary and promotion discrimination in a faculty member's career (Scott, 1977; Boudreau et. al., 1997).

Another common concern regarding the variables from which to predict salary that are used in most studies include the lack of some measure of merit or productivity (Moore, 1993). Most researchers indicate that measures of productivity are important to include but lament that those types of data are not readily available. In fact, the most recent salary equity study "kit" does not include a mention of such measures (Haignere et. al., 1996).

Additionally, researchers have complained about the inclusion of variables which are known to contain measurement error, in particular, measures of merit or productivity (Birnbaum, 1979; Millsap & Meredith, 1994). This complaint has led to the use of reverse regression to produce predictors which were supposedly free of measurement error. However, McFatter (1987) and Everett (1990) have countered with models which offer latent measures to address the problem.

*Interpretation*

Moore (1993) has indicated that the emphasis placed on statistical significance of the multiple regression results in salary equity studies is unwarranted. She claims that, because these studies are typically performed on a population, not a sample, issues of inference are moot. This opinion, however, is not widely shared (see Haignere et. al., 1996, for a review of the discussion.)

*Statistical Technique*

Although they provide neither solutions nor suggested practices, Hengstler and McLaughlin (1985) summarize the common concerns associated with the use of multiple regression in salary equity studies. Most important to the present study is the concern that multiple regression relies on assumptions that may not be appropriate for the data in question. For example, warnings about possible heteroscedasticity abound in the literature, however, few salary studies ever examine their data for this condition (Balzer et. al., 1996; Millsap & Meredith, 1994).

In multiple regression, there are a few basic assumptions that should be met, or at least examined prior to undertaking an analysis. These assumptions include normally-distributed residuals at each value of X, equal variances of residuals across values of X, and independent (or random) residuals across observations (Cohen & Cohen, 1983). It is this final assumption that is of utmost interest in this paper.

*Independence Assumption*

One of the fundamental assumptions of traditional statistical techniques such as ANOVA and multiple regression analysis is that data are obtained from independent observations, thus resulting in error terms that are independent. With data that are hierarchically clustered, this assumption is likely being violated: specific to the issue of faculty pay equity, faculty who are in a given department, such as English, are more likely to be like each other than like faculty in another department, such as Physics. They are likely to have a shared concept of the mission of a department, a shared expectation of research productivity, and so on. Therefore, the resulting clusters in these types of studies will be characterized by some homogeneity.

Why is independence of observations necessary, statistically speaking? The assumption of independent observations, while not absolutely necessary for the estimation of parameters (such as regression coefficients), is crucial for the estimation of variance and covariance, and therefore standard errors of estimated parameters (Lee et al, 1989). Kish and Frankel (1974), in empirical studies of large samples found that parameter estimates were robust to violations of the assumption of independent observations with group sizes that were not wildly varying, but their classic article delineates the (sometimes drastic) underestimation of sample variance of the parameters that can occur using traditional analysis methods.

Often researchers will assume that their data are independent, even when obvious clustering has occurred, if only out of unfamiliarity with the issues involved. The traditional formulas for standard errors in statistics textbooks and incorporated into most statistical computer programs are based on the simple random sampling with replacement design (Lee et al, 1989). Because these formulas assume that the correlation of the error terms is zero, when analyzing clustered data, a researcher will underestimate the sample variance of the parameter. This underestimate will provide narrower confidence intervals around parameter estimates and will result in the researcher rejecting the null more often than appropriate. In other words, the chance of making a Type I error increases. Scariano and Davenport (1987) reported on a simulation study which estimated the true Type I error rates under conditions of dependent clustering in ANOVA. For example, with only modest levels of dependency and two means, the true Type I error was .57 for group sizes of 100, far from the assumed .05 Type I error rate.

There are two general categories of approaches to properly analyze data resulting from a cluster sampling design. The first of these uses traditional statistical techniques (such as OLS regression), but employs special procedures, such as Balanced Repeated Replication and Jackknife Repeated Replication, to estimate the standard errors of the parameter estimates. The second approach is to model the data in a multi-level fashion, mirroring the hierarchical structure of the data. These approaches have been termed design-based and model-based respectively in the literature (Kalton, 1983). In this paper, a model-based approach will be examined.

*Dependent Observations in Salary Equity Studies*

Data used in salary studies are problematic. A faculty member's salary should be a function of individual characteristics such as years of experience and productivity. However, to include only individual-level variables would constitute an "individualistic" or "psychologistic" fallacy -- there are also contextual variables which interact to affect salary. (For a more in-depth discussion of individualistic fallacy, see Diez-Roux, 1998). Given the same individual-level characteristics, we would not expect an English faculty member to receive the same salary as a Physics faculty member. There is something about the discipline context which affects salary, such as competition, and societal or market value.

Previous research in salary equity study methodology has recognized this problem of contextual effect. Statistical models typically include some measure of discipline -- usually a set of dummy variables to indicate broad discipline categories (Snyder, et. al., 1994, Haignere, et. al., 1996, Moore, 1993). There are two different procedures that researchers have followed with dummy variables. The more popular procedure is to use dummy variables to reflect broad groups, or clusters, of departments. This procedure would result in dummy variables such as "Social Science" and "Humanities". Faculty within departments that housed social science programs, such as Psychology and Economics would be assigned to the "Social Science" cluster, while faculty within departments such as Art and English would fall into the "Humanities" cluster. Using this procedure, usually about five to ten dummy variables are constructed. A second procedure would be to dummy code each department as a separate entity. Published studies exist where up to 87 dummy variables were constructed to represent departments (Ransom, 1993). With either procedure, the opportunity to create interactions, or cross-products, is available, however rarely used.

It is argued in this paper that the use of these procedures, using broad discipline indicators, can provide inaccurate estimates of gender effects in salary. Because most salaries are basically set at a department level, the contextual unit should be the department. Most universities, however, have dozens of departments and dummy coding each one becomes cumbersome and is a drain on degrees of freedom.

Salary studies within disciplines outside higher education have been criticized in the same manner; The "methodology typically used is a single-level regression analysis, which describes individuals but neglects context or industry" (Kreft & de Leeuw, 1994, p.321). Kreft and de Leeuw provide an example looking at a comparison of multiple regression techniques and random coefficient modeling across twelve industries ranging from retail to manufacturing and the military. They suggest that the choice of analytic method should reflect the context of the data, as well as the data collection scheme.

The use of multiple regression to study nested data is neither a new nor unique problem. Bryk and Raudenbush (1986) indicate that "despite forceful warnings, single-level linear-model analyses of school effects abound...In the past, analysts clung to single-level models not out of conviction but because of the absence of viable alternatives" (p. 1). Accessible, viable alternatives now exist. Statistical software packages which provide for multilevel regression modeling, such as Hierarchical Linear Modeling (HLM), MLwiN, SAS PROC MIXED, and VARCL, are available for personal computers and at fairly inexpensive prices. Procedures for using these packages are outlined in Kreft & de Leeuw (1998), Singer (1999), and Hox (1994).

*Hierarchical Linear Modeling*

Multilevel regression models are a category of regression-based models, including hierarchical linear models, random coefficient models, and variance component models. Conceptually, the multilevel regression model can be viewed as a hierarchical system of regression equations. The discussion that follows will be based on experience using the HLM software for hierarchical linear modeling and therefore, in the remainder of the paper, we will use "HLM" simultaneously to refer to the statistical technique as well as the software.

HLM estimates linear equations to explain outcomes for individuals within groups as a function of the characteristics of the groups as well as the individuals (Arnold, 1992). There are two advantages of applying HLM to the study of salary equity: (1) the technique can model the effects on salary of faculty characteristics (within-group variables), such as years of experience and gender, while moderating these effects by considering departmental differences (between-group variables), such as differential salary structures; and (2) it can examine these phenomena while explicitly modeling the within-group dependencies. Faculty within a unit share rewards, stresses, and expectations in common, therefore they share variance introduced as a dependency.

Before undertaking a multilevel regression analysis, it is important to first determine whether the data exhibit clustering effects or perhaps whether theory indicates that the data are clustered. While the theory is a matter of in-depth understanding of the issues surrounding your data, the clustering effect can be statistically estimated. A measure of the variance in a dependent variable which is accounted for solely by the grouping variable is called the intra-class correlation (ICC) (Kenny & Judd, 1986). This measure can be calculated using components from a simple ANOVA using the following formula:

$ICC = (MS_B - MS_W) / (MS_B + (c-1)MS_W)$

where $MS_B$ = means square between groups,
$MS_W$ = means square within groups, and
c = the common group size in the balanced case, or the average group size if groups are unbalanced.

The ICC can range from -1/(c-1) to +1. Previous research indicates that with geographically-determined clusters such as households, the intraclass correlation is relatively low on demographic variables (such as age and gender) and higher for socioeconomic variables and attitudes (Kalton, 1977). In educational studies, the ICCs have been found to be rather high: for example, between .3 and .4 due to classroom components when examining mathematics achievement for U.S. eighth graders (Muthen, 1996).

To show these group-level relationships visually, suppose for simplicity that one has a theory that salary is a function of number of advisees. The familiar multiple regression model would be:

$Y_i = \beta_0 + \beta_1 X_i + e_i$

where $Y_i$ = an individual faculty member's salary,
$X_i$ = an individual faculty member's number of advisees,
$\beta_0$ = the intercept for salary across all faculty members,
$\beta_1$ = the beta coefficient, or slope, for number of advisees, and
$e_i$ = the residual for each individual faculty member.

Therefore, for a given individual, i, his salary is the sum of the intercept, some linear function of his number of advisees ($X_i$), and some residual, as depicted in Figure 1.

------------------------------------------

Insert Figure 1 about here

------------------------------------------

However, if one hypothesizes that the relationship between salary and number of advisees is dependent on the context of a department, a hierarchical model might be:

$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + r_{ij}$
$\beta_{0j} = \gamma_{00} + u_{0j}$
$\beta_{1j} = \gamma_{10} + u_{1j}$

where $Y_{ij}$ = the salary of an individual faculty member i in department j,
$X_{ij}$ = the number of advisees of an individual faculty member i in department j,
$\beta_{0j}$ = the intercept of salary for all faculty members in department j,
$\beta_{1j}$ = the beta coefficient, or slope, for number of advisees in department j, and
$r_{ij}$ = the residual for each individual faculty member in department j.

HLM estimates each of these parameters simultaneously using a restricted maximum likelihood estimation method. Note that the intercept and slope coefficients are no longer fixed effects for all individuals. These coefficients are now random, dependent on the department in which the faculty member is employed. These coefficients are each comprised of a fixed component ($\gamma_{00}$ and $\gamma_{10}$) which represents the average intercept and slope across the departments, and a random component ($u_{0j}$ and $u_{1j}$) which represents the residual at the group level. The assumptions for this two-level model are:

1) each $r_{ij}$ is independent and normally distributed with a mean of 0 and a variance $\sigma^2$ for every level 1 unit i within each level-2 unit j,
2) the level 1 predictor ($X_{ij}$) is independent of $r_{ij}$ [$Cov(X_{ij},r_{ij})=0$], and
3) the errors at level 1 and level 2 are independent [$Cov(r_{ij}, u_{0j})=0$].

Let's take an example. Suppose, after running HLM using this model, we found the following coefficients, as depicted in Figure 2:

$$Salary = \beta_{0j} + \beta_{1j} * Years + e_{ij}$$
$$\beta_{0j} = 52{,}000 + u_{0j}$$
$$\beta_{1j} = 0 + u_{1j}$$

-----------------------------------------

Insert Figure 2 about here

-----------------------------------------

The intercept for each faculty member depends on his department. The (weighted) average intercept is \$52,000, but some departments may have \$62,000 ($u_{0j}$ = \$10,000) and some may have \$45,000 ($u_{0j}$ = -\$7,000). Similarly, for the slope on number of advisees, the average across departments is \$0, however, the slope may vary depending on department. HLM provides estimates of the amount of variance of the within-group residual ($r_{ij}$) as well as each of the between-group residuals ($u_{0j}$ and $u_{1j}$). In this example, there is relatively little variance in the residual for the slope ( $u_{1j}$). Significance tests of these between-group residuals are available to determine whether the parameters truly vary across groups.

The traditional technique used in equity studies, of accounting for differences in discipline by adding discipline clusters as dummy variables, is an attractive alternative in modeling faculty salaries, however, it does not necessarily fix all potential problems. If the four groups of faculty who are displayed in Figure 2 had each been from a different discipline cluster, then the traditional method of dummy variable coding would provide estimates that were similar to HLM estimates. However, if these four groups of faculty represented four departments that were lumped together to form a discipline cluster, the MR results would indicate that there was a relationship between number of advisees and salary. In fact, there is no relationship between number of advisees and salary. But the grouping of departments into large clusters masks this lack of relationship. Because the departments with the higher average salaries also have higher average advisees, the relationship is attributed to the individual level. Diez-Roux (1998) and others in sociology have long labeled this ascription of group level properties to the individual an "ecological fallacy."

*A Comparison of Using MR and HLM in Faculty Salary Equity Studies*

This paper presents results from a comparison of the multiple regression approach to examining salary equity (with discipline clusters) and the approach of hierarchical linear modeling to the same problem. The comparison is done in two steps. First, a practical example of applying multiple regression and hierarchical linear modeling techniques, using empirical data, are provided. Next, results from a simulation study which examined varying data conditions are described to show differences in the results from these two types of statistical procedures.

*Empirical Example*

*Method*

The differences between results from applying MR and HLM techniques are highlighted using faculty data from a large public research institution in the Mid-Atlantic region. Of interest in this study is whether the two methods provide similar results and whether those results are sensible and interpretable for administrators. There were 1,216 tenured and tenure-track faculty on the Fall 1997 university employee census who were part of the analysis. Faculty with administrative duties, such as department chairs, were excluded from the analysis. Additionally, only full-time faculty who had appointments in instructional units were included. The variables used in the analysis included the following:

*Discipline:* As typical in most multiple regression faculty salary equity studies, several groups of disciplines were created and represented by dummy variables (coded 0/1). Seven clusters were created and included AGLIFE (Agricultural, Natural Resources, and Life Sciences), PHYSENGN (Physical Sciences and Engineering), SOCSCI (Social Sciences), HUMAN (Humanities and Arts), EDUCHLTH (Education, Kinesiology, and Health Education), BMGT (Business and Management), and PROFCOLL (Professional programs of Journalism, Public Affairs, Architecture, and Library Sciences).

*Rank:* Four ranks were created and were represented by three dummy variables (coded 0/1). These dummy variables included PROF (rank of professor), PERMASSC (rank of "permanent" associate professor), and STRVASSC (rank of "striving" associate professor). Assistant professors were the reference group and were identified by observations with zero values for each of the these three dummy variables. It was recognized that for the rank of associate professor, years in rank is often related to salary in a curvilinear fashion, with a relatively steep positive relationship in the early years and a less steep, perhaps even negative, slope for associate professors with long tenure at the institution. Therefore, associate professors were divided into "striving" associates and "permanent" associates. By visual inspection of the relationship between salary and years in rank for associate professors, it was determined that the change in slope occurred at about the ten year mark. Therefore, faculty with less than ten years in the associate professor rank were termed "striving associate professors" while those with ten or more years were considered "permanent associate professors."

*Years in rank:* The number of years (YRSRANK) the faculty member had been at their current rank at the institution.

Productivity variables were taken from the institution's annual accountability report of instructional faculty workload. This report is mandated by the state and reported with the individual faculty as the unit of analysis. For this reporting process, faculty are asked to fill out a survey on non-instructional activity, including information on publications, research, and awards.

These data were averaged for each faculty member for the academic years 1996-97 and 1997-98 to arrive at a more stable estimate of faculty productivity. For those faculty who did not provide information in one of the years, a one-year figure served as the estimate. While there was some hesitation to use self-report measures in this analysis, we believed that the measures had some degree of validity because the department chairs reviewed each faculty member's responses to the productivity questionnaire, and in some cases, this same questionnaire was used in the promotion and tenure process.

*Refereed articles:* The average yearly number of refereed articles (REF) published.

*Presentations:* The average yearly number of presentations (PRES) given.

*Sponsored research:* The average yearly expended dollar amount in sponsored grants and contracts (GRANT) associated with the faculty member. These data were obtained from the Office of the Comptroller.

Two additional variables, books published and creative activities, were initially used in the analyses, but were never found to have a significant relationship with salary and were therefore dropped from this exploratory study.

*Gender:* Coded 1 for males, 0 for females.

*Descriptive Statistics*

As described above, there were 1,216 faculty in 62 departments which were grouped into seven discipline clusters. Table 1 contains the descriptive statistics for the 1,216 faculty as a whole.

Table 1
Descriptive Statistics for Sample

| Variable Name | Mean | St. Dev. |
|---|---|---|
| SALARY | 66,271.29 | 20,049.65 |
| PROF | 0.48 | 0.50 |
| PERMASSC | 0.12 | 0.32 |
| STRVASSC | 0.23 | 0.42 |
| YRSRANK | 9.32 | 8.44 |
| GRANT | 76,728.07 | 252,883.00 |
| REF | 2.58 | 4.53 |
| PRES | 2.91 | 3.52 |
| GENDER | 0.78 | 0.41 |

Overall, females at the institution are paid less than males, and this relationship holds true within each of the seven discipline clusters (see Table 2).

Table 2
Average Salaries, by Discipline Cluster and Gender

| Cluster | Females | Males |
|---|---|---|
| Total | $58,920 | $68,290 |
| AGLIFE | $59,646 | $61,156 |
| PHYSENGN | $61,631 | $73,026 |
| PROFCOLL | $60,846 | $77,986 |
| EDUCHLTH | $55,762 | $59,874 |
| SOCSCI | $61,660 | $73,311 |
| BMGT | $84,743 | $91,109 |
| HUMAN | $54,891 | $57,411 |

The Appendix provides descriptive statistics for all variables within each cluster. Additionally, the number of departments, and the range of the department averages within each of the clusters is displayed.

The first-order correlations for the variables of interest are provided in Table 3.

Table 3
Total Sample Correlation Coefficients for Individual Level Data

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 1. SALARY | 1.00 | | | | | | | | |
| 2. YRSRANK | .19 | 1.00 | | | | | | | |
| 3. PROF | .65 | .24 | 1.00 | | | | | | |
| 4. PERMASSC | -.24 | .47 | -.35 | 1.00 | | | | | |
| 5. STRVASSC | -.24 | -.35 | -.53 | -.20 | 1.00 | | | | |
| 6. GRANT | .21 | ns | .09 | -.08 | ns | 1.00 | | | |
| 7. REF | .20 | -.12 | .12 | -.16 | ns | .20 | 1.00 | | |
| 8. PRES | .20 | -.17 | .08 | -.16 | *.06* | .27 | .39 | 1.00 | |
| 9. GENDER | .19 | .25 | .17 | .08 | -.12 | .08 | .06 | ns | 1.00 |

all correlations listed are significant with p<.01, except the correlation between PRES and STRVASSC (italicized) where p<.05

*Multiple Regression Procedure*

In an attempt to find a small number of variables to be modeled in using the MR and HLM techniques, all analyses were considered to be exploratory and therefore several models were examined before arriving at a final MR model. All data, except the dummy variables, were transformed to be in z-score form, to allow for easier estimation when running the HLM model.

Initially, the arts and humanities (HUMAN) discipline cluster was used as the reference group, however, no significant differences from the reference group were found for the education and health cluster (EDUCHLTH) and therefore the two clusters were combined as the reference category.

A full model was developed and run, and after this step, GENDER was added to the model. There was not a significant change in $R^2$, indicating that gender is not significantly related to salary when other variables are held constant. The final model was:

$$\begin{aligned}\text{SALARY}_i = \beta_0 + \\ &\beta_1\text{AGLIFE}_i + \beta_2\text{PHYSENGN}_i + \beta_3\text{PROFCOLL}_i + \beta_4\text{SOCSCI}_i + \beta_5\text{BMGT}_i + \\ &\beta_6\text{PROF}_i + \beta_7\text{PERMASSC}_i + \beta_8\text{STRVASSC}_i + \\ &\beta_9\text{YRSRANK}_i + \beta_{10}\text{GRANT}_i + \beta_{11}\text{REF}_i + \beta_{12}\text{PRES}_i + \\ &\beta_{13}\text{GENDER}_i + e_i\end{aligned}$$

*Multiple Regression Results*

The $R^2$ in the final model was .610, indicating that over 60 percent of the variance in salary could be accounted for by the variables in the model. It was interesting to note that the addition of the productivity variables, GRANT, REF and PRES, increased the $R^2$ by about 19 percentage points -- not a trivial amount. These results support Williford's findings with regard to the importance of using productivity measures (1998). Although productivity measures are often not used because of unavailability, researchers should try to include such information when undertaking salary equity studies. Table 4 contains the multiple regression estimates from the final model.

Table 4
Final MR Model Estimates

| Variable | Estimated Beta | Std. Err. | T-ratio | p-value |
|---|---|---|---|---|
| INTERCEPT | -1.088 | 0.061 | -17.894 | <0.001 |
| Set of Discipline Indicators | | | | |
| AGLIFE | 0.161 | 0.055 | 2.933 | 0.003 |
| PHYSENGN | 0.401 | 0.050 | 7.999 | <0.001 |
| PROFCOLL | 0.569 | 0.088 | 6.432 | <0.001 |
| SOCSCI | 0.486 | 0.061 | 7.997 | <0.001 |
| BMGT | 1.592 | 0.086 | 18.611 | <0.001 |
| Set of Rank Indicators | | | | |
| PROF | 1.338 | 0.057 | 23.608 | <0.001 |
| PERMASSC | 0.165 | 0.084 | 1.964 | 0.050 |
| STRVASSC | 0.324 | 0.058 | 5.586 | <0.001 |
| YRSRANK | 0.074 | 0.024 | 3.078 | 0.002 |
| GRANT | 0.098 | 0.019 | 5.100 | <0.001 |
| REF | 0.054 | 0.020 | 2.689 | 0.007 |
| PRES | 0.112 | 0.020 | 5.517 | <0.001 |
| GENDER | 0.042 | 0.047 | 0.890 | 0.374 |

Because all data, except for those variables that were dummy coded, were in z-score form, the results can be interpreted in terms of standard deviations -- one standard deviation change in salary is approximately $20,000. Faculty in the BMGT cluster receive about 1.6 standard deviations more salary than faculty in the Arts and Humanities, Education and Health fields. Faculty who are full professors receive 1.3 standard deviations more salary than assistant professors. For each additional standard deviation of years in rank (about 8 years), a faculty member receives an additional .07 standard deviation in salary. And for every standard deviation of grant dollars awarded ($250,000), refereed articles published (4.5), and presentations given (3.5), a faculty member's salary increases by .10, .05, and .11 standard deviations respectively about $2,000, $1,000, and $2,200).

The overall interpretation meets with general expectation -- the highest paid faculty are full professors who have been at that rank for some time, obtain grant funding, publish articles, present papers, and reside in the Business school. Faculty receiving the lowest salary would be new assistant professors who do not obtain grants, publish articles, or present papers, and are in the Humanities or Education fields.

*Hierarchical Linear Modeling Procedure*

The first step in multilevel modeling is to estimate the amount of variance in the dependent variable that can be accounted for solely by the grouping variable. There are a few ways to accomplish this. In this case, we ran a simple ANOVA with salary as a dependent variable and the 62 departments as the class variable.

$ICC = (MS_B - MS_W) / (MS_B + (c-1)MS_W)$

where $c = n/j = 1216/62 = 19.6$ (average group size)

$ICC = (5.802 - .746) / (5.802 + (19.6-1)*.746)$

$= .257$

This indicates that nearly 26 percent of the variance in salary can be accounted for by department groupings. This seems to be reasonable in the context of higher education. The market value for Electrical Engineering faculty may be quite different from the market value for History faculty, holding constant the productivity and tenure of the faculty. It is conceivable that there would be a within-group dependency in higher education units. A department chair and the dean work together to assign salary within each department, therefore, salaries of two faculty members should not be considered independent within a given department. With an intra-class correlation as high as 26 percent, a multilevel analysis is warranted.

First, a "null" model was run in HLM. This null (or random effects) model provides an additional estimate of intra-class correlation and serves as a base model to determine the variance accounted for with future models. The results of estimating this model are displayed in Tables 5 and 6.

$SALARY = \beta_{0j} + r_{ij}$

$\beta_{0j} = \gamma_{00} + u_{0j}$

where $\gamma_{00}$ = the grand department mean salary (in z-score form)

$u_{0j}$ = the group mean deviation from the grand mean

Table 5
Null Model: Fixed Effects

| | Coefficient | Std Error | T-ratio | p-value |
|---|---|---|---|---|
| Intercept | | | | |
| $\gamma_{00}$ | -0.120 | 0.066 | -1.821 | 0.068 |

Table 6
Null Model: Random Effects

| | Std. Dev. | Variance Component | df | Chi-square | p-value |
|---|---|---|---|---|---|
| Intercept | | | | | |
| $u_{0j}$ | 0.462 | 0.213 | 61 | 499.058 | <0.001 |
| Level-1 residual | | | | | |
| $r_{ij}$ | 0.863 | 0.744 | | | |

Another estimate of the intra-class correlation can be calculated from these results (Bryk & Raudenbush, 1992).

$$\begin{aligned} ICC &= var(u_{0j}) / [var(u_{0j}) + var(r_{ij})] \\ &= .213/(.213+.744) \\ &= .222 \end{aligned}$$

This estimate is somewhat smaller than that provided by the ANOVA results (22% versus 26%). This is perhaps due to group sizes that are unbalanced -- the groups range from 5 to 68 faculty in a department. The ANOVA-based calculation is based on a common group size (c), however the average group size is often used when group sizes are unequal. While this substitution has been shown to be somewhat robust with unbalanced data, the variance group sizes in this data set are quite large. Regardless, the null model estimates indicate that, indeed, the between-group variance in salary is statistically significantly greater than zero ($\chi^2$=499.053, p<.001) and therefore multilevel modeling is warranted.

The HLM model was set up to mirror the final MR model as closely as possible, with the exclusion of the cluster dummy variables. Once the basic model was created, some additional exploratory modeling was undertaken. An additional complication, however, was that the most reasonable interpretations would come from a model that was group-mean centered, as opposed to grand-mean centered or uncentered. Kreft and de Leeuw (1998) and Bryk and Raudenbush (1992) provide in-depth discussions of the various centering options. With group-mean centering, the group level averages for all variables that are group-mean centered should be entered into the intercept at level 2 ($\beta_{0j}$). This addition makes for a model which appears much more complicated.

The level 1 model should look familiar (however note that the variables in italics are centered around their group mean, ie. $\mathit{YRSRANK}_{ij}$ =[$\mathrm{YRSRANK}_{ij}$ - $\mathrm{YRSRANK}_{.j}$]):

$$\begin{aligned} \mathrm{SALARY} = \beta_{0j} + \beta_{1j}\mathrm{PROF}_{ij} + \beta_{2j}\mathrm{PERMASSC}_{ij} + \beta_{3j}\mathrm{STRVASSC}_{ij} + \beta_{4j}\mathit{YRSRANK}_{ij} \\ + \beta_{5j}\mathit{GRANT}_{ij} + \beta_{6j}\mathit{REF}_{ij} + \beta_{7j}\mathit{PRES}_{ij} + \beta_{8j}\mathrm{GENDER}_{ij} + r_{ij} \end{aligned}$$

Salary is hypothesized to be a function of an intercept, the rank of the faculty member, their years in that rank, their productivity in terms of grants, refereed articles and presentations, and possibly, their gender, all within the context of their department.

At level 2, the intercept for a given department was assumed to be a function of the average years in rank for the department faculty, the average grants, average refereed articles, average presentations, and the percent of faculty in the department who are male, and some residual that is not explained by these group averages.

$$\begin{aligned} \beta_{0j} = \gamma_{00} + \gamma_{01}\mathrm{AVEYRS}_j + \gamma_{02}\mathrm{AVEGRANT}_j + \gamma_{03}\mathrm{AVEREF}_j + \\ \gamma_{04}\mathrm{AVEPRES}_j + \gamma_{05}\mathrm{PCTMALE}_j + u_{0j} \end{aligned}$$

The effect of being a full professor (PROF) was assumed not to vary across departments and was therefore fixed to a constant, $\gamma_{10}$.

$$\beta_{1j} = \gamma_{10}$$

Likewise, the effect of being a permanent associate professor (PERMASSC) was assumed not to vary across departments.

$$\beta_{2j} = \gamma_{20}$$

And the effect of being a striving associate professor (STRVASSC) was assumed not to vary across departments.

$$\beta_{3j} = \gamma_{30}$$

The effect of years in rank (*YRSRANK*) was assumed not to vary *randomly* across departments, but it was assumed that the effect would be moderated by the average number of years in rank of the faculty in the department.

$$\beta_{4j} = \gamma_{40} + \gamma_{41}AVEYRS_j$$

The effect of grant dollars expended (*GRANT*) was assumed not to vary *randomly* across departments, but it was assumed that the effect would be moderated by the average number of grant dollars obtained by the faculty in the department.

$$\beta_{5j} = \gamma_{50} + \gamma_{51}AVEGRANT_j$$

The effect of refereed articles published (REF) was assumed to *randomly* vary across departments.

$$\beta_{6j} = \gamma_{60} + u_{6j}$$

The effect of paper presentations (PRES) was assumed to be fixed across all departments.

$$\beta_{7j} = \gamma_{70}$$

The effect of gender (GENDER) was assumed to be fixed across all departments.

$$\beta_{8j} = \gamma_{80}$$

*HLM Model Results*

The results from the final (exploratory) model are shown in Tables 7 and 8.

Table 7
Final Model: Fixed Effects

| | Coefficient | Std Error | T-ratio | p-value |
|---|---|---|---|---|
| Intercept | | | | |
| $\gamma_{00}$ | -0.088 | 0.049 | -1.801 | 0.077 |
| $\gamma_{01}$ (AVEYRS) | -0.177 | 0.167 | -1.060 | 0.294 |
| $\gamma_{02}$ (AVEGRANT) | 0.035 | 0.121 | 0.288 | 0.774 |
| $\gamma_{03}$ (AVEREF) | 0.240 | 0.177 | 1.354 | 0.181 |
| $\gamma_{04}$ (AVEPRES) | 0.119 | 0.152 | 0.780 | 0.439 |
| $\gamma_{05}$ (PCTMALE) | 0.581 | 0.272 | 2.134 | 0.037 |
| PROF | | | | |
| $\gamma_{10}$ | 1.311 | 0.054 | 24.194 | <0.001 |
| PERMASSC | | | | |
| $\gamma_{20}$ | 0.162 | 0.079 | 2.048 | 0.045 |
| STRVASSC | | | | |
| $\gamma_{30}$ | 0.322 | 0.055 | 5.885 | <0.001 |
| *YRSRANK* | | | | |
| $\gamma_{40}$ | 0.123 | 0.026 | 4.793 | <0.001 |
| $\gamma_{41}$ (AVEYRS) | -0.144 | 0.063 | -2.278 | 0.027 |
| *GRANT* | | | | |
| $\gamma_{50}$ | 0.150 | 0.029 | 5.243 | <0.001 |
| $\gamma_{51}$ (AVEGRANT) | -0.044 | 0.019 | -2.338 | 0.023 |
| *REF* | | | | |
| $\gamma_{60}$ | 0.131 | 0.045 | 2.931 | 0.005 |
| *PRES* | | | | |
| $\gamma_{70}$ | 0.055 | 0.021 | 2.636 | 0.011 |
| GENDER | | | | |
| $\gamma_{80}$ | -0.002 | 0.046 | -0.050 | 0.961 |

Table 8
Final Model: Random Effects

| | Std. Dev. | Variance Component | df | Chi-square | p-value |
|---|---|---|---|---|---|
| Intercept | | | | | |
| $u_0$ | 0.349 | 0.122 | 56 | 679.912 | <0.001 |
| REF | | | | | |
| $u_6$ | 0.177 | 0.031 | 61 | 119.565 | <0.001 |
| Level-1 residual | | | | | |
| $r_{ij}$ | 0.571 | 0.326 | | | |

To interpret the results, it may help to view the model in one large equation, plugging the level-2 fixed and random components into their respective level 1 place markers:

$$\begin{aligned} \text{SALARY} = \ & \gamma_{00} + \gamma_{01}\text{AVEYRS}_j + \gamma_{02}\text{AVEGRANT}_j + \gamma_{03}\text{AVEREF}_j + \\ & \gamma_{04}\text{AVEPRES}_j + \gamma_{05}\text{PCTMALE}_j + \\ & \gamma_{10}\text{PROF}_{ij} + \gamma_{20}\text{PERMASSC}_{ij} + \gamma_{30}\text{STRVASSC}_{ij} + \\ & [\gamma_{40} + \gamma_{41}\text{AVEYRS}_j]\ \mathit{YRSRANK}_{ij} + \end{aligned}$$

$[\gamma_{50} + \gamma_{51}\text{AVEGRANT}_j]\ \textit{GRANT}_{ij} +$
$[\gamma_{60} + u_{6j}]\ \textit{REF}_{ij} +$
$\gamma_{70}\ \textit{PRES}_{ij} +$
$\gamma_{80}\ \text{GENDER}_{ij} +$
$+ u_{0j} + r_{ij}$

And with the estimated $\gamma$ coefficients (those in bold are significant at $\alpha = .05$):

SALARY = $-.088 + -.177\text{AVEYRS}_j + .035\text{AVEGRANT}_j + .240\text{AVEREF}_j +$
$.119\text{AVEPRES}_j + \mathbf{.581}\text{PCTMALE}_j +$

$\mathbf{1.311}\text{PROF}_{ij} + \mathbf{.162}\text{PERMASSC}_{ij} + \mathbf{.322}\text{STRVASSC}_{ij} +$

$[\mathbf{.123} + \mathbf{-.144}\text{AVEYRS}_j]\ \textit{YRSRANK}_{ij} +$
$[\mathbf{.150} + \mathbf{-.044}\text{AVEGRANT}_j]\ \textit{GRANT}_{ij} +$
$[\mathbf{.131} + u_{6j}]\ \textit{REF}_{ij} +$
$\mathbf{.055}\textit{PRES}_{ij} +$
$-.002\text{GENDER}_{ij} +$
$+ u_{0j} + r_{ij}$

These estimates indicate that a relatively well-paid faculty member would come from a department with a large percentage of males, be a full professor, have more years in rank than others in his department (but be from a relatively young department on average), have more dollars in grants than others in his department (but be from a department with lower grant dollars on average), have relatively more refereed articles, and relatively more presentations. The gender of the faculty member is not shown to be related to salary. Note that the random effects information in Table 8 indicates that the relationship between salary and refereed article production varies significantly from department to department. In over half of the departments, the relationship was negative. The coefficients ranged from -.219 to +.469.

Some of these results may need explanation. Taking years in rank as an example, we can see that the older the department is on average, the less years in rank matters.

$\beta_{4j} = \mathbf{.123} + \mathbf{-.144}\text{AVEYRS}_j$

In this data set, AVEYRS ranged from -.7 to .8, therefore the effect of YRSRANK for a given individual ranged from .2238 for faculty in very "young" departments to .0078 for fairly "old" departments. Intuitively this makes sense. In departments with mostly new faculty, merit pay decisions may not be able to be made on a great amount of evidence of performance, therefore, the chair might rely on seniority; however, in "older" departments, more long-term information exists about the productivity of the faculty and merit pay decisions may be based less on longevity.

An additional interesting finding was the significant coefficient for PCTMALE in the level-2 intercept. Not surprisingly, the model indicates that the higher percent of males that are found in a department, the higher the average salary. Percent of males in a department may, in fact, be a proxy measure for another departmental characteristic, such as quantitative level of field. Some societal or market value has been placed on this characteristic leading to higher salaries in such departments. In the future, further investigation into other possible measures to include as departmental characteristics in warranted.

Examining the random effects in Table 8, note that the residual (level 1) variance was .744 in the null model (Table 6) and has been decreased to .326 in this final model which incorporates individual characteristics of rank, seniority, and productivity.

*Comparison of Results*

With regard to the relationship between gender and salary, the MR analysis and the HLM analysis provided the same result: there was no statistically significant relationship. The models, however, did provide different interpretations. In the multiple regression model, no contextual effects were examined, and therefore, effects were fixed across departments. For example, each year in rank was associated with a .07 increase in salary in the MR model. However, the HLM model suggests that the effect of years in rank depends on the relative "age" of the department. There is no doubt that the HLM model can provide more information, however, it provides a much more complicated picture of the salary process. Is the interpretation worth the complication? If the model is more accurate, yes. The best way to determine the accuracy would be to simulate data with known characteristics and then test whether MR and HLM can recover those characteristics. The last section of the paper reports on simulation results. However, before continuing to the final section of the paper, two issues in HLM should be addressed: model fit and sample size.

*Model Fit*

An attractive feature of multiple regression is that it provides for a single indicator of model fit, $R^2$ (or adjusted $R^2$). No such single indicator exists in HLM. Some have indicated that comparing the initial residual variance in the null model with the residual variance remaining after level 1 predictors have been entered can provide a measure of proportion of variance accounted for at level 1. In our example, the residual variance dropped from .744 to .326, so about 56 percent of the "within" variance was explained, and recall that the ICC estimated that between 22 and 26 percent of the variance in salary was due to departmental effects. However, Kreft and de Leeuw (1998) in their section on analogs to an $R^2$ measure (pp.115-119) argue against the use of one overall measure of explained variance, especially with random slope coefficients. One additional measure to examine model fit is to use the change in deviance and degrees of freedom for nested models to determine whether a given model provides "significantly" better fit.

*Sample Size*

The number of observations needed for multilevel modeling should actually be approached from a multilevel perspective. To estimate level 1 effects, the number of total observations is of interest and guidelines are similar as in multiple regression. Bryk and Raudenbush (1992) suggest at least ten observations per predictor. At level 2, however, the unit of interest is the group, not the individual, and therefore guidelines refer to the number of groups, regardless of group size. Bryk and Raudenbush, again, propose ten groups per predictor at level 2, however Kreft and de Leeuw (1998) cite simulation studies which indicate that 50 to 60 groups tend to provide stable estimates of level 2 effects, but the number of groups needed depends on the size of the effect, the group size, and the intra-class correlation. An additional summary of sample size recommendations has been provided by Hox (1997).

*A Simulation Study*

As discussed earlier, the use of multiple regression on faculty data is problematic because of the nesting of faculty within departments. Researchers have attempted to model this nesting by introducing dummy variables to represent broad categories of disciplines. It is hypothesized in this paper that if the broad categories are comprised of very similar departments, in terms of salary level and percent of faculty who are male, then the multiple regression results will be similar to the HLM results with regard to the relationship between gender and salary. However, if the departments are heterogeneous in terms of salary level and the percentage of male faculty, the multiple regression results will yield misleading estimates, resulting in inappropriate conclusions about the absence or presence of salary inequities due to gender on a campus.

Assuming that rank, years in rank, and productivity variables are controlled for, suppose that Figure 3 displayed the data for four departments within one discipline cluster, the Social Sciences. The four departments might be Anthropology, Sociology, Political Science, and Economics. In two of the departments, Anthropology and Sociology, less than 50% of the faculty are male, and in the remaining two, more than 50% of the faculty are male. Also, the departments with the smaller percentage of males also have the lower average salary. Occurrences of such data are not rare. In the Appendix, the departments average minimum and maximums are listed for each cluster in the empirical data set. Note that in the SOCSCI cluster, the percent male ranges from .40 to .88 and the average department salary ranges from 57,806 to 86,205.

When a regression line is drawn as it is in Figure 3, displaying the relationship between gender and salary, it would have a positive slope, indicating gender salary inequities in favor of males. But note that within each department females are generally paid as well as males. In this case, MR results would indicate gender-based pay inequity when there appears to be none within any given department, where salary decisions tend to be made.

---

Insert Figure 3 about here

---

To understand the conditions under which MR would provide misleading estimates, a set of conditions were created under which we simulated data and then ran both MR and HLM to determine which technique would provide more accurate results.

The basic structure of the simulations was as follows: five discipline clusters were created, each consisting of ten departments, for a total of 50 departments. Within each of these 50 departments, 20 faculty observations were generated, for a total of 1,000 observations. Each observation was assigned nine variables: SALARY, YRSRANK, GRANT, REF, PRES, PROF, PERMASSC, STRVASSC, and GENDER. Within each of the 50 departments, ten of the faculty observations were assigned to be full professors, two were permanent associate professors, four were striving associate professors, and four were assistant professors, consistent with the proportions we found in the empirical dataset. The remaining variable values were generated to have specific, consistent correlations with salary. There were just three conditions which were manipulated for each department in this simulation: gender-based pay inequity, percent of faculty who were male, and the salary level.

The first parameter to be manipulated was the amount of gender-based pay inequity. Five separate conditions were simulated: no pay inequity between males and females in the same department, small positive gender inequity, large positive gender inequity, small negative gender inequity, and large negative gender inequity. Positive inequity was simulated by adding a constant to the salaries of males in a given department while subtracting a constant from the salaries of females in the same department. Negative inequity, conversely, was simulated by subtracting a constant from the salaries of males in a given department while adding a constant to the salaries of females in the same department. A small positive gender effect was defined as a .1 standard deviation *addition* to male salaries accompanied by a .1 standard deviation *subtraction* from females salaries. A large effect was defined as a .3 standard deviation addition and a .3 standard deviation subtraction.

The second simulated parameter was the homogeneity of the departments within each cluster in terms of percent male. The three conditions consisted of: homogeneity (70 percent male in all ten departments in each cluster), small departure from homogeneity (80 percent male in five departments and 60 percent male in the other five departments in each cluster), and large departure from homogeneity (90 percent male in five departments and 50 percent male in the other five departments in each cluster).

The final simulated parameter was the homogeneity of the departments within each cluster in terms of salary level. The four conditions consisted of: no difference in salary across all ten departments within each cluster, small difference (.2 standard deviations of salary were added to five departments and .2 standard deviations of salary were subtracted from the other five departments in each cluster), medium difference (addition and subtraction of .4 standard deviations), and a large difference (addition and subtraction of .6 standard deviations). In all cases, the departments with the higher percentage of males were assigned the higher average department salary.

In all, 60 sets of conditions were generated as depicted in Table 9.

Table 9

<u>Sixty Simulation Conditions</u>

| Gender Inequity | Departmen % male | No dept salary difference | Small dept salary difference (+.2/-.2) | Medium dept salary difference (+.4/-.4) | Large dept salary difference (+.6/- |
|---|---|---|---|---|---|
| No gender inequity | 70% in all | 1 | 2 | 3 | 4 |
| | 80% / 60% | 5 | 6 | 7 | 8 |
| | 90% / 50% | 9 | 10 | 11 | 12 |
| Small positive inequity | 70% in all | 13 | 14 | 15 | 16 |
| | 80% / 60% | 17 | 18 | 19 | 20 |
| | 90% / 50% | 21 | 22 | 23 | 24 |
| Large positive inequity | 70% in all | 25 | 26 | 27 | 28 |
| | 80% / 60% | 29 | 30 | 31 | 32 |
| | 90% / 50% | 33 | 34 | 35 | 36 |
| Small negative inequity | 70% in all | 37 | 38 | 39 | 40 |
| | 80% / 60% | 41 | 42 | 43 | 44 |
| | 90% / 50% | 45 | 46 | 47 | 48 |
| Large negative inequity | 70% in all | 49 | 50 | 51 | 52 |
| | 80% / 60% | 53 | 54 | 55 | 56 |
| | 90% / 50% | 57 | 58 | 59 | 60 |

One hundred data sets were generated for each of the 60 sets of conditions and then the statistical methods of interest, MR and HLM, were applied to the data. The results from the 100 replications allow us to understand the ability of each method to detect the "true" gender-based salary inequity from the generated data.

*Simulation Results*

Tables 10 through 14 provide the results from the simulation analyses. For each condition, several statistics are displayed: the average MR statistical bias, the average HLM statistical bias, the percent of times gender was non-significant in MR and HLM, the percent of times gender was significant and positive in MR and HLM, and the percent of times gender was significant and negative in MR and HLM. Statistical bias is a measure of the average deviation of a parameter estimate from its true, generated value, and is measured as

$$\Sigma(\beta' - \beta)/ r$$

where $\beta'$= estimated beta coefficient for GENDER in the respective model
$\beta$ = true coefficient (was 0 for the "no inequity" condition, .2 for small positive effect, .6 for large positive effect, -.2 for small negative effect, and -.6 for large negative effect).
r = 100 replications

In two cases, the MR estimate had a smaller bias than the HLM bias, however the difference did not exceed .01 (these occurrences are italicized in the tables). In the remaining 58 cases, HLM did as well, and often, significantly better than MR. For example, in Table 10 and Figure 4, it can be seen that with no true gender inequity, under the most extreme conditions (departments with 90 percent males had salaries .6 standard deviations above the mean for the cluster, while departments with 50 percent males had salaries .6 standard deviations below the mean for the cluster), the MR estimate was on average .671 points above the true value. The MR results indicated that males on average receive .671 standard deviations of salary more than females, when in reality, the data were generated with no gender inequities within departments. Such conclusions are unfounded, however, and an analyst blindly using MR with discipline clusters would be unaware of the potential misleading results. The statistical bias results displayed in Figure 4 are extremely similar to the results for the other four conditions of gender pay inequity, indicating that the amount of inequity has no bearing on the performance of MR.

---

Insert Figure 4 about here

---

MR seemed to provide accurate estimates if the departments within each discipline cluster were homogeneous with regard to percent male and salary. If only one condition was homogenous (percent male or salary level), then MR still provided reliable estimates. Once the homogeneity within a cluster was violated for percent male *and* salary level, even at small levels, the MR estimates grew biased toward detecting gender inequity for females.

Table 10
Condition: No Gender Inequity (β = 0)

| Percent Male | Salary Difference | Multiple Regression Bias | % ns | % sig + | % sig - | HLM Bias | % ns | % sig + | % sig - |
|---|---|---|---|---|---|---|---|---|---|
| 70% / 70% | no | -.001 | 98% | 2% | 0% | -.001 | 97% | 3% | 0% |
| | +.2/-.2 | .006 | 96% | 2% | 2% | .006 | 95% | 2% | 3% |
| | +.4/-.4 | .003 | 98% | 0% | 2% | .003 | 95% | 2% | 3% |
| | +.6/-.6 | -.005 | 98% | 1% | 1% | -.004 | 93% | 4% | 3% |
| 80% / 60% | no | -.002 | 95% | 2% | 3% | .000 | 93% | 4% | 3% |
| | +.2/-.2 | .114 | 59% | 41% | 0% | .033 | 94% | 6% | 0% |
| | +.4/-.4 | .229 | 7% | 93% | 0% | .038 | 93% | 7% | 0% |
| | +.6/-.6 | .356 | 0% | 100% | 0% | .044 | 92% | 8% | 0% |
| 90% / 50% | no | -.009 | 93% | 4% | 3% | -.014 | 92% | 5% | 3% |
| | +.2/-.2 | .226 | 6% | 94% | 0% | .077 | 81% | 19% | 0% |
| | +.4/-.4 | .459 | 0% | 100% | 0% | .104 | 74% | 25% | 1% |
| | +.6/-.6 | .671 | 0% | 100% | 0% | .076 | 83% | 17% | 0% |

It was disconcerting to note that the Type II error rate was rather large for the small gender inequity (positive and negative) conditions. We would have liked to have seen about 95% of the replications indicating that the coefficient was significant and positive in Table 11 and 95% of the replications indicating that the coefficient was significant and negative in Table 13. In fact, both MR and HLM did not perform well, and where MR provide the correct significance direction, the statistical bias was large.

Table 11
Small Positive Inequity (β = .2)

| Percent Male | Salary Difference | Multiple Regression Bias | % ns | % sig + | % sig - | HLM Bias | % ns | % sig + | % sig - |
|---|---|---|---|---|---|---|---|---|---|
| 70% / 70% | no | .001 | 14% | 86% | 0% | .001 | 14% | 86% | 0% |
| | +.2/-.2 | -.003 | 21% | 79% | 0% | -.003 | 19% | 81% | 0% |
| | +.4/-.4 | -.007 | 21% | 79% | 0% | -.006 | 16% | 84% | 0% |
| | +.6/-.6 | .002 | 21% | 79% | 0% | .002 | 13% | 87% | 0% |
| 80% / 60% | no | .007 | 6% | 94% | 0% | .007 | 8% | 92% | 0% |
| | +.2/-.2 | .103 | 0% | 100% | 0% | .019 | 3% | 97% | 0% |
| | +.4/-.4 | .232 | 0% | 100% | 0% | .038 | 6% | 94% | 0% |
| | +.6/-.6 | .358 | 0% | 100% | 0% | .048 | 4% | 96% | 0% |
| 90% / 50% | no | -.001 | 13% | 87% | 0% | .000 | 17% | 83% | 0% |
| | +.2/-.2 | .227 | 0% | 100% | 0% | .075 | 5% | 95% | 0% |
| | +.4/-.4 | .464 | 0% | 100% | 0% | .104 | 2% | 98% | 0% |
| | +.6/-.6 | .684 | 0% | 100% | 0% | .087 | 6% | 94% | 0% |

With the case of large positive inequity (Table 12), the MR and HLM results indicate the same conclusion (100% of the trials indicate positive significance), however, the MR drastically over estimates the advantage for males as seen in Figure 4.

Table 12
Large Positive Inequity (β = .6)

| Percent Male | Salary Difference | Multiple Regression Bias | % ns | % sig + | % sig - | HLM Bias | % ns | % sig + | % sig - |
|---|---|---|---|---|---|---|---|---|---|
| 70% / 70% | no | .011 | 0% | 100% | 0% | .011 | 0% | 100% | 0% |
| | +.2/-.2 | -.003 | 0% | 100% | 0% | -.003 | 0% | 100% | 0% |
| | +.4/-.4 | .003 | 0% | 100% | 0% | .003 | 0% | 100% | 0% |
| | +.6/-.6 | -.010 | 0% | 100% | 0% | -.009 | 0% | 100% | 0% |
| 80% / 60% | no | -.006 | 0% | 100% | 0% | -.005 | 0% | 100% | 0% |
| | +.2/-.2 | .120 | 0% | 100% | 0% | .035 | 0% | 100% | 0% |
| | +.4/-.4 | .227 | 0% | 100% | 0% | .035 | 0% | 100% | 0% |
| | +.6/-.6 | .351 | 0% | 100% | 0% | .041 | 0% | 100% | 0% |
| 90% / 50% | no | -.001 | 0% | 100% | 0% | .000 | 0% | 100% | 0% |
| | +.2/-.2 | .244 | 0% | 100% | 0% | .090 | 0% | 100% | 0% |
| | +.4/-.4 | .448 | 0% | 100% | 0% | .084 | 0% | 100% | 0% |
| | +.6/-.6 | .682 | 0% | 100% | 0% | .083 | 0% | 100% | 0% |

Of notable concern in this study was the finding for the condition when salary inequity exists, but the preference is toward females and not males (negative inequity conditions). In these cases, the MR results, because they are positively statistically biased, can indicate that there is indeed inequity, but it in favor of males instead of females! This result was found in five of the simulation conditions with small negative inequity (Table 13) and in one of the conditions with large negative inequity (Table 14). These values were italicized in their respective tables.

Table 13
Small Negative Inequity (β = -.2)

| Percent Male | Salary Difference | Multiple Regression Bias | % ns | % sig + | % sig - | HLM Bias | % ns | % sig + | % sig - |
|---|---|---|---|---|---|---|---|---|---|
| 70% / 70% | no | .009 | 20% | 0% | 80% | *.010* | 20% | 0% | 80% |
| | +.2/-.2 | .007 | 14% | 0% | 86% | .007 | 13% | 0% | 87% |
| | +.4/-.4 | .009 | 24% | 0% | 76% | .009 | 17% | 0% | 83% |
| | +.6/-.6 | .000 | 22% | 0% | 78% | .000 | 11% | 0% | 89% |
| 80% / 60% | no | .006 | 15% | 0% | 85% | .005 | 16% | 0% | 84% |
| | +.2/-.2 | .113 | 77% | 0% | 23% | .030 | 23% | 0% | 77% |
| | +.4/-.4 | .233 | 97% | *3%* | 0% | .040 | 33% | 0% | 67% |
| | +.6/-.6 | .335 | 49% | *51%* | 0% | .025 | 26% | 0% | 74% |
| 90% / 50% | no | .008 | 20% | 0% | 80% | .005 | 21% | 0% | 79% |
| | +.2/-.2 | .227 | 95% | *5%* | 0% | .078 | 59% | 0% | 41% |
| | +.4/-.4 | .452 | 6% | *94%* | 0% | .087 | 69% | 0% | 31% |
| | +.6/-.6 | .676 | 0% | *100%* | 0% | .072 | 55% | 0% | 45% |

Table 14
Large Negative Inequity (β = -.6)

| Percent Male | Salary Difference | Multiple Regression Bias | % ns | % sig + | % sig - | HLM Bias | % ns | % sig + | % sig - |
|---|---|---|---|---|---|---|---|---|---|
| 70% / 70% | no | -.006 | 0% | 0% | 100% | -.006 | 0% | 0% | 100% |
| | +.2/-.2 | -.004 | 0% | 0% | 100% | -.004 | 0% | 0% | 100% |
| | +.4/-.4 | .000 | 0% | 0% | 100% | .000 | 0% | 0% | 100% |
| | +.6/-.6 | -.005 | 0% | 0% | 100% | -.005 | 0% | 0% | 100% |
| 80% / 60% | no | -.007 | 0% | 0% | 100% | -.004 | 0% | 0% | 100% |
| | +.2/-.2 | .114 | 0% | 0% | 100% | .032 | 0% | 0% | 100% |
| | +.4/-.4 | .238 | 0% | 0% | 100% | .046 | 0% | 0% | 100% |
| | +.6/-.6 | .346 | 3% | 0% | 97% | .033 | 0% | 0% | 100% |
| 90% / 50% | no | .001 | 0% | 0% | 100% | .002 | 0% | 0% | 100% |
| | +.2/-.2 | .233 | 0% | 0% | 100% | .083 | 0% | 0% | 100% |
| | +.4/-.4 | .446 | 39% | 0% | 61% | .083 | 0% | 0% | 100% |
| | +.6/-.6 | .683 | 79% | 21% | 0% | .082 | 0% | 0% | 100% |

In general, the simulation results confirmed the belief that if departments within discipline clusters are not homogenous with regard to percentage of males and salary levels (controlling for all other variables), the analysis results may be biased when using MR techniques and cluster dummy variables. It is therefore suggested that researchers consider using HLM when undertaking salary equity studies with many departments or administrative units.

## *Conclusion*

It is hoped that this paper will guide others who have the onerous task of determining whether salary inequities exist between faculty groups on their campus to consider the possibility of using hierarchical linear modeling for their nested data. Without valid statistical treatment of the research data, there exists the potential that policy might be influenced by misleading conclusions.

Hierarchical linear modeling allows for the analyst to control for the contextual effects of smaller faculty units. Current methods used in multiple regression, namely the use of cluster dummy variables, can obfuscate the true relationships within a group, by ascribing group-level relationships to the individual level.

*References*


Arnold, C. L. (1992). An introduction to hierarchical linear models. Measurement and Evaluation in Counseling and Development, 25, 58-90.

Bryk, A. S., & Raudenbush, S. W. (1986). A hierarchical model for studying school effects. Sociology of Education, 59, 1-17.

Bryk, A. S., & Raudenbush, S. W. (1992). Hierarchical linear models: Applications and data analysis methods. Newbury Park, CA: Sage Publications, Inc.

Balzer, W., Boudreau, N., Hutchinson, P., Ryan, A. M., Thorsteinson, T., Sullivan, J., Yonker, R., & Snavely, D. (1996). Critical modeling principles when testing for gender equity in faculty salary. Research in Higher Education, 37, 633-658.

Birnbaum, M. H. (1979). Procedures for the detection and correction of salary inequities. In T. R. Pezzullo & B. E. Brittingham (Eds.), Salary equity (pp. 121-144). Lexington, MA: Lexington Books.

Boudreau, N., Sullivan, J., Balzer, W., Ryan, A. M., Yonker, R., Thorsteinson, T., & 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, 297-312.

Cohen, J., & Cohen J. (1983). Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Erlbaum Associates.

Diez-Roux, A. V. (1998). Bringing context back into epidemiology: Variables and fallacies in multilevel analysis. American Journal of Public Health, 88, 216-222.

Everett, J. E. (1990). Discrimination measures using contingency tables. Multivariate Behavioral Research, 25, 371-386.

Haignere, L., Lin, Y., Eisenberg, B., & McCarthy, J. (1996). Pay checks: A guide to achieving salary equity in higher education. Albany, NY: United University Professors.

Hengstler, D. D. & McLaughlin, G. W., (1985). Statistical issues and concerns in court cases. In W. Rosenthal & B. Yancey (Eds.), The use of data in discrimination issues cases (pp. 65-82). San Francisco, CA: Jossey-Bass.

Hox, J. J. (1994). *Applied multilevel analysis.* Amsterdam: TT-Publikaties.

Hox, J. J. (1997). "Multilevel modeling: When and why?" In Balderjahn, R. M. & Schader, M. (Eds.) *Classification, data analysis, and data highways.* New York: Springer Verlag.

Kalton, G. (1977). Practical methods for estimating survey sampling errors. *Bulletin of the International Statistical Institute*, 47(3), 495-514.

Kalton, G. (1983). Models in the practice of survey sampling. *International Statistical Review*, 51, 175-188.

Kenny, D. A., & Judd, C. M. (1986). Consequences of violating the independence assumtion in analysis of variance. Psychological Bulletin, 99, 422-431.

Kreft, I. & de Leeuw, J. (1998). Introducing multilevel modeling. London: Sage Publications Ltd.

Kreft, I. G. G. & de Leeuw, J. (1994). The gender gap in earnings: A two-way nested multiple regression analysis with random effects. Sociological Methods & Research, 22, 319-341.

McFatter, R. M. (1987). Use of latent variable models for detecting discrimination in salaries. Psychological Bulletin, 101, 120-125.

Millsap, R. E., & Meredith, W. (1994). Statistical evidence in salary discrimination studies: Nonparametric inferential conditions. Multivariate Behavioral Research, 29, 339-364.

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

Muthen, Bengt O. (1996). *Latent variable modeling of longitudinal and multilevel data.* Paper presented at the annual meeting of the American Sociological Association, October 1996.

Ransom, Michael (1993). Seniority and monopsony in the academic labor market. American Economic Review, 83, 221-233.

Scott, E. (1977). Higher education salary evaluation kit. Washington, DC: American Association of University Professors.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. Journal of Educational and Behavioral Statistics, 24, pp. 323-355.

Snyder, J. K, Hyer, P. B., & McLaughlin, G. W. (1994). Faculty salary equity: Issues and options. Research in Higher Education, 35, 1-19.

Toutkoushian, R. K. (1998). Using regression analysis to determine if faculty salaries are overly compressed. Research in Higher Education, 39, 87-99.

Williford, L. E. (1998). Using productivity data in faculty salary equity analysis. Paper presented at the Annual Forum of the Association for Institutional Research, Minneapolis, Minnesota.

Figure 1

Salary as a Function of Advisees, No Multilevel Effect

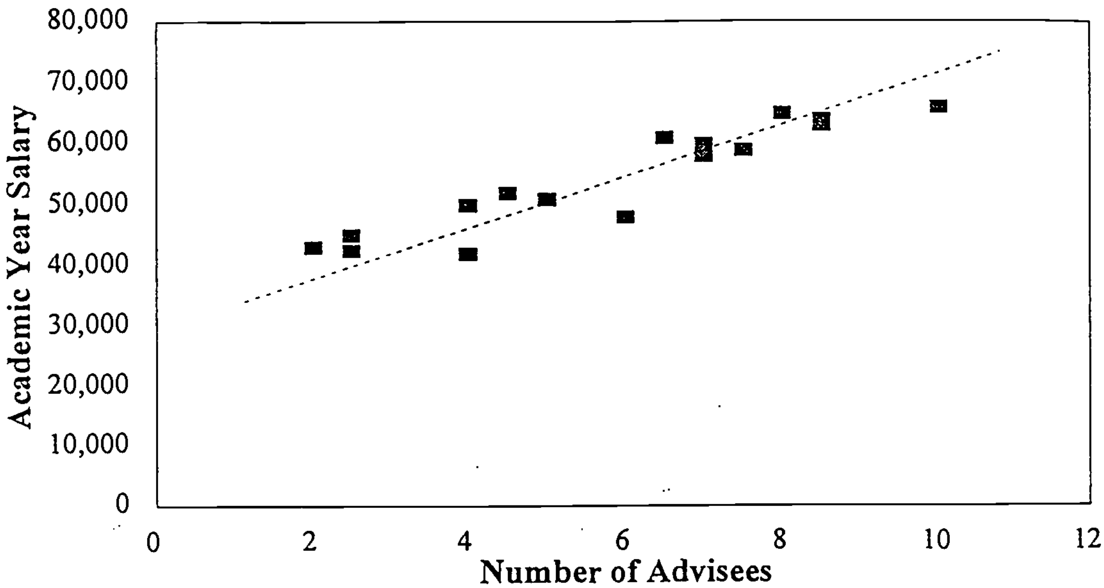



Figure 2

Salary as a Function of Advisees, Multilevel Effect Modeled

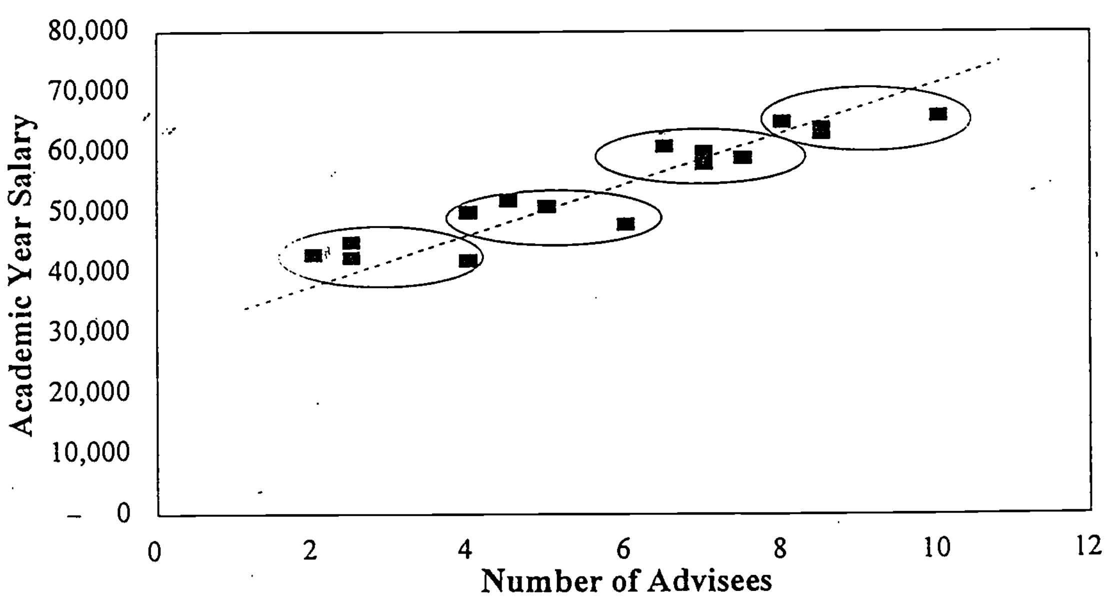


Figure 3

Salary as a Function of Gender, No Multilevel Effect

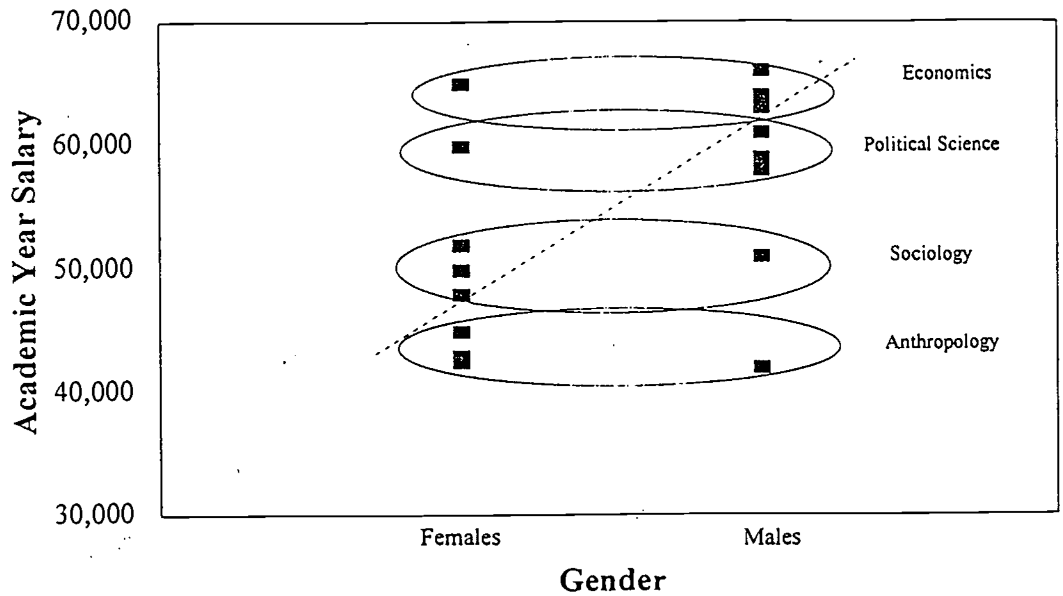
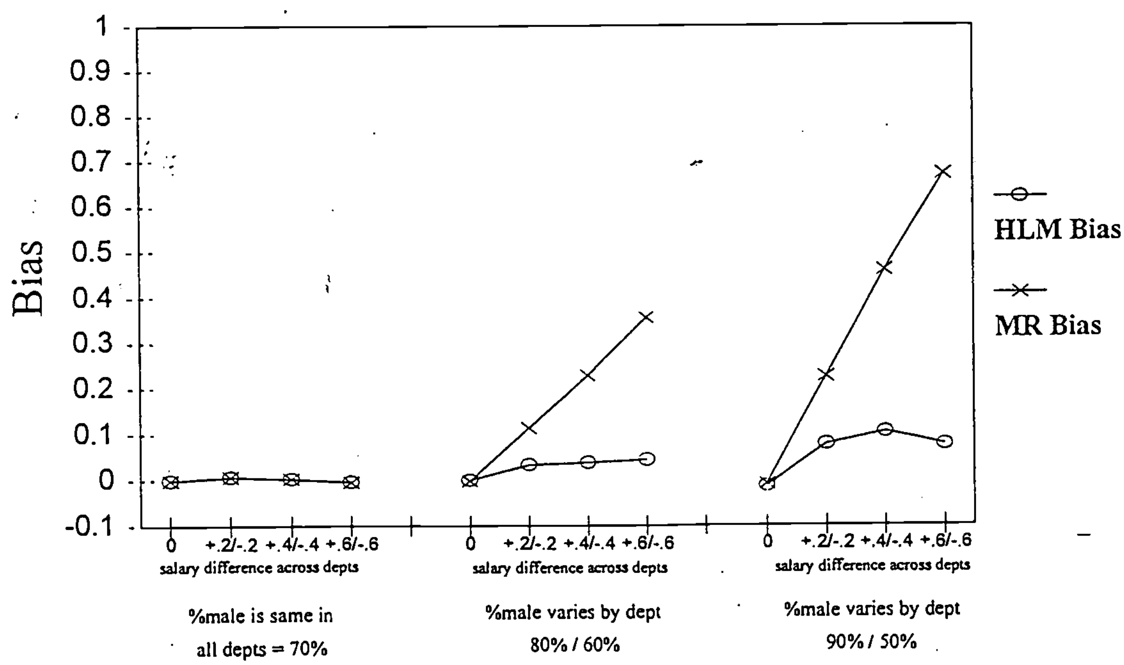



Figure 4

Statistical Bias for HLM and MR Methods -- True Gender Inequity = 0



## APPENDIX -- Cluster Averages and Range of Department Averages

| Discipline | Variable Name | Mean | St. Dev. | Range of Department Averages Minimum | Maximum |
|---|---|---|---|---|---|
| AGLIFE | N=206, Depts=10 | | | N=5 | N=39 |
| | SALARY | 60,892.11 | 16,928.35 | 53,298.00 | 77,651.19 |
| | PROF | 0.41 | 0.49 | .20 | .53 |
| | PERMASSC | 0.14 | 0.39 | .00 | .35 |
| | STRVASSC | 0.22 | 0.41 | .09 | .40 |
| | YRSRANK | 8.51 | 7.48 | 3.40 | 12.78 |
| | GRANT | 65,337.06 | 102,989.70 | 16,739.09 | 130,899.09 |
| | REF | 2.29 | 2.30 | 1.27 | 3.50 |
| | PRES | 2.67 | 2.66 | 1.18 | 4.38 |
| | GENDER | .83 | .38 | .55 | 1.00 |
| PHYSENGN | N=347, Depts=13 | | | N=5 | N=68 |
| | SALARY | 72,204.75 | 17,878.49 | 59,427.20 | 81,355.17 |
| | PROF | 0.59 | 0.49 | .29 | .74 |
| | PERMASSC | 0.05 | 0.23 | .00 | .20 |
| | STRVASSC | 0.21 | 0.41 | .12 | .40 |
| | YRSRANK | 9.82 | 8.95 | 3.60 | 14.54 |
| | GRANT | 150,802.32 | 278,753.45 | 281,890.00 | 14,341.60 |
| | REF | 4.07 | 5.91 | 1.57 | 7.31 |
| | PRES | 3.61 | 4.05 | 1.30 | 6.50 |
| | GENDER | .93 | .26 | .60 | 1.00 |
| SOCSCI | N=152, Depts=8 | | | N=5 | N=35 |
| | SALARY | 72,398.09 | 22,987.28 | 57,806.25 | 86,205.03 |
| | PROF | 0.51 | 0.50 | .25 | .62 |
| | PERMASSC | 0.12 | 0.32 | 0.00 | .33 |
| | STRVASSC | 0.21 | 0.41 | 0.00 | .34 |
| | YRSRANK | 10.27 | 8.77 | 7.00 | 13.12 |
| | GRANT | 111,762.22 | 514,861.57 | 12,799.42 | 677,989.88 |
| | REF | 1.90 | 2.52 | 0.90 | 3.10 |
| | PRES | 2.80 | 3.88 | 1.50 | 4.76 |
| | GENDER | 0.75 | 0.43 | .40 | .88 |
| HUMAN | N=257, Depts=17 | | | N=5 | N=51 |
| | SALARY | 56,587.37 | 15,514.33 | 46,751.33 | 69,164.17 |
| | PROF | 0.42 | 0.49 | .00 | .83 |
| | PERMASSC | 0.17 | 0.37 | .00 | .60 |
| | STRVASSC | 0.25 | 0.43 | .00 | .57 |
| | YRSRANK | 9.13 | 8.30 | 4.63 | 16.20 |
| | GRANT | 6,741.67 | 64,726.88 | 0.00 | 132,634.06 |
| | REF | 1.61 | 3.29 | 0.10 | 3.90 |
| | PRES | 2.28 | 2.93 | 0.86 | 6.17 |
| | GENDER | 0.67 | 0.47 | .00 | 0.84 |
| EDUCHLTH | N=131, Depts=9 | | | N=8 | N=15 |
| | SALARY | 58,178.77 | 12,499.85 | 52,273.72 | 62,979.75 |
| | PROF | 0.37 | 0.49 | .25 | .53 |
| | PERMASSC | 0.18 | 0.38 | .07 | .33 |
| | STRVASSC | 0.24 | 0.43 | .06 | .50 |
| | YRSRANK | 9.30 | 8.65 | 4.92 | 15.17 |
| | GRANT | 56,030.63 | 156,525.02 | 2,083.33 | 183,760.96 |
| | REF | 2.18 | 2.58 | 1.00 | 3.65 |
| | PRES | 3.31 | 3.37 | 1.25 | 4.38 |
| | GENDER | 0.59 | 0.49 | .33 | 1.00 |
| BMGT | N=64, Depts=1 | | | — | — |
| | SALARY | 90,015.06 | 22,020.31 | — | — |
| | PROF | 0.47 | 0.50 | — | — |
| | PERMASSC | 0.06 | 0.24 | — | — |
| | STRVASSC | 0.25 | 0.44 | — | — |
| | YRSRANK | 8.94 | 8.53 | — | — |
| | GRANT | 6,587.56 | 32,789.30 | — | — |
| | REF | 1.66 | 1.79 | — | — |
| | PRES | 1.91 | 2.43 | — | — |
| | GENDER | 0.83 | 0.38 | — | — |
| PROFCOLL | N=59, Depts=4 | | | N=11 | N=17 |
| | SALARY | 73,919.03 | 24,150.17 | 63,551.36 | 96,794.86 |
| | PROF | 0.59 | 0.50 | .45 | .71 |
| | PERMASSC | 0.07 | 0.25 | .00 | .12 |
| | STRVASSC | 0.29 | 0.46 | .21 | .36 |
| | YRSRANK | 8.12 | 7.49 | 5.50 | 9.18 |
| | GRANT | 17,481.53 | 75,694.14 | 0.00 | 53,148.59 |
| | REF | 2.75 | 10.03 | 0.82 | 5.91 |
| | PRES | 2.95 | 4.63 | 1.44 | 5.93 |
| | GENDER | 0.76 | 0.43 | .55 | 0.93 |

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

TM031963

# REPRODUCTION RELEASE

(Specific Document)

## I. DOCUMENT IDENTIFICATION:

Title: EVALUATING FACULTY SALARY EQUITY USING HIERARCHICAL LINEAR MODELING

Author(s): LAURA M. STAPLETON ROBERT W. LISSITZ

Corporate Source:

Publication Date: APRIL 19, 1999

## II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

| The sample sticker shown below will be affixed to all Level 1 documents | The sample sticker shown below will be affixed to all Level 2A documents | The sample sticker shown below will be affixed to all Level 2B documents |
|---|---|---|
| PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY<br>Sample<br>TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)<br>1 | PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY<br>Sample<br>TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)<br>2A | PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY<br>Sample<br>TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)<br>2B |
| Level 1<br>☑ | Level 2A<br>☐ | Level 2B<br>☐ |
| Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy. | Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only | Check here for Level 2B release, permitting reproduction and dissemination in microfiche only |

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

*I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.*

Sign here,→ please

| Signature: Laura M Stapleton | Printed Name/Position/Title: LAURA M. STAPLETON | |
|---|---|---|
| Organization/Address: UNIVERSITY OF MARYLAND | Telephone: 301/405-5592 | FAX: 301/314-9443 |
| | E-Mail Address: ~~301/314-9443~~ | Date: 10/16/00 |

(over)

ERIC®

# Clearinghouse on Assessment and Evaluation

University of Maryland
1129 Shriver Laboratory
College Park, MD 20742-5701

*Tel: (800) 464-3742*
*(301) 405-7449*
*FAX: (301) 405-8134*
*ericae@ericae.net*
*http://ericae.net*

May 8, 2000

Dear AERA Presenter,

Hopefully, the convention was a productive and rewarding event. As stated in the AERA program, presenters have a responsibility to make their papers readily available. If you haven't done so already, please submit copies of your papers for consideration for inclusion in the ERIC database. We are interested in papers from this year's AERA conference and last year's conference. If you have submitted your paper, you can track its progress at http://ericae.net.

Abstracts of papers accepted by ERIC appear in *Resources in Education (RIE)* and are announced to over 5,000 organizations. The inclusion of your work makes it readily available to other researchers, provides a permanent archive, and enhances the quality of *RIE*. Abstracts of your contribution will be accessible through the printed and electronic versions of *RIE*. The paper will be available through the microfiche collections that are housed at libraries around the world and through the ERIC Document Reproduction Service.

We are gathering all the papers from the **2000 and 1999 AERA Conference**. We will route your paper to the appropriate clearinghouse. You will be notified if your paper meets ERIC's criteria for inclusion in *RIE*: contribution to education, timeliness, relevance, methodology, effectiveness of presentation, and reproduction quality.

Please sign the Reproduction Release Form enclosed with this letter and send **two** copies of your paper. The Release Form gives ERIC permission to make and distribute copies of your paper. It does not preclude you from publishing your work. You can mail your paper to our attention at the address below. Please feel free to copy the form for future or additional submissions.

Mail to:

AERA 2000/ERIC Acquisitions
University of Maryland
1129 Shriver Laboratory
College Park, MD 20742

Sincerely,

Lawrence M. Rudner

Lawrence M. Rudner, Ph.D.
Director, ERIC/AE

ERIC is a project of the Department of Measurement, Statistics & Evaluation