DOCUMENT RESUME

ED 476 226 HE 035 845

AUTHOR	Haignere, Lois
TITLE	Paychecks: A Guide to Conducting Salary-Equity Studies for Higher Education Faculty. Second Edition.
INSTITUTION	American Association of Univ. Professors, Washington, DC.; United Univ. Professions, Albany, NY.
ISBN	ISBN-0-9649548-2-6
PUB DATE	2002-00-00
NOTE	110p.; Also produced by Haignere, Inc.
AVAILABLE FROM	American Association of University Professors, 1012 Fourteenth Street, NW, Suite 500, Washington, DC 20005 (members, $18; nonmembers, $28). Tel: 202-737-5900; Web site: http://www.aaup.org.
PUB TYPE	Guides - Non-Classroom (055)
EDRS PRICE	EDRS Price MF01/PC05 Plus Postage.
DESCRIPTORS	*College Faculty; Contract Salaries; Databases; *Equal Opportunities (Jobs); Evaluation Methods; Higher Education; *Research Methodology; Sex Discrimination; *Teacher Salaries

ABSTRACT

This guidebook is designed as a resource for those in the higher education community who want to conduct analyses of bias in faculty salaries or to understand and interpret the results of studies presented to them. This edition will help readers detect gender and face bias in current rank, select a salary-equity consultant, understand different perspectives on how bias occurs and ways to remedy it, and accomplish many other tasks related to ensuring equity in faculty salaries. The chapters are: (1) "Introduction to Equal Pay for Equal Work" (Lois Haignere); (2) "Considerations before Launching a Salary Study" (Lois Haignere and Donna Euben); (3) "Database Decisions and Development" (Lois Haignere and Yangjing Lin); (4) "Gender and Race Bias in Current Rank" (Lois Haignere and Bonnie Eisenberg); (5) "Gender and Race Bias in Salaries" (Lois Haignere); (6) "Small Errors with Big Consequences" (Lois Haignere); and (7) "Diagnosis Dynamics and Treatment Turmoil" (Lois Haignere). Nine appendixes contain supplemental information in essays on specific topics, such as study methodology, contract language, discrimination law, and statistical techniques. (Contains 10 tables, 9 figures, and 154 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 476 226

Paychecks

A Guide to Conducting Salary-Equity Studies for Higher Education Faculty

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

I. F. Molotsky

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Principal author: Lois Haignere

BEST COPY AVAILABLE

TE035843

Paychecks

Paychecks

A Guide to Conducting Salary-Equity Studies for Higher Education Faculty

Principal author: Lois Haignere

Second Edition

American Association of University Professors
Washington, D.C.

© 2002 American Association of University Professors, Lois Haignere, Inc., United University Professions. All rights reserved. Reproductions of excerpts for nonprofit use is hereby granted to educators, scholars, students, nonprofit educational institutions, and government. Any reproduction for commercial use without written permission is strictly prohibited.

American Association of University Professors
1012 Fourteenth St., NW, Suite 500
Washington, DC 20005

Lois Haignere, Inc.
40 Parkwood East
Albany, NY 12203-3629

United University Professions
159 Wolf Road
Albany, NY 12205

First published 1996 by United University Professions
Second edition 2002
ISBN 0-9649548-2-6

Contents

Contributors ix
Figures and Tables xi
Preface xiii
Acknowledgments xv

1. Introduction to Equal Pay for Equal Work 1
Lois Haignere

Cost of Salary Inequity 1
Unconscious Ideology 1
Objectives of This Book 2
Neglected Groups 3
White Males 3
SUNY Studies 3
Limits on Coverage 3
Demographics 4
Regression Analysis 4
Quality of Regression Results 6
Curiosity, Not Conviction 6

2. Considerations Before Launching a Salary Study 9
Lois Haignere and Donna Euben

Politics of Technical Decisions 9
Basic Requirements 10
Salary-Equity Consultants 12
Happy Endings 13

3. Database Decisions and Development 17
Lois Haignere and Yangjing Lin

Policy-Making Committee 17
Study Population 18
Data Collection 18
Variables and Possible Substitutes 19
Variables for Verification of Data 23
Cleaning of Data 24

4. Gender and Race Bias in Current Rank 27
Lois Haignere and Bonnie Eisenberg

Details and Options 27
Frequency Tables 28
Categorical Modeling 28
Limitations of Categorical Modeling 31
Event History Analysis 31
Next Steps 34

5. Gender and Race Bias in Salaries 37
Lois Haignere

Re-engineered Variables 37
Redesigned Categorical Variables: Dummies 39
Suggested Categorical Variables 39
Three Multiple-Regression Approaches 41
Recommendations 43
Two Additional Checks 45

6. Small Errors with Big Consequences 49
Lois Haignere

Standardization of the Coefficients 49
Exclusion of Certain Categories 49
Dropping of Outliers 51
Tainted Variables 51
The Significance of Significance 53
Divide and Hide 55
Stepwise Abuse 55
Dropping of Gender and Race 56
Summary 57

7. Diagnosis Dynamics and Treatment Turmoil 59
Lois Haignere

Diagnosis Dynamics 59
Pockets of Bias 60
Case Study 61
Institutional-Systemic Remedy 61
Longevity. 63
Flagging and Other Remedies 63
Toward a More Permanent Solution 66

Appendices

A. Multiple Regression and Gender- and Race-Equity in Salaries 69
Lois Haignere

Dummy Variables 71
Validity of Regression Equation 72
Interpretation of Results 72

B. The Twelve SUNY Data Sets 75
Lois Haignere

C. Salary-Equity Contract Language 77
Donna Euben

Joint Faculty-Administration Committees 77
Money 78

Minimum-Salary Scales 78
Grievance Procedures 79

D. U.S. Laws on Gender-Based Wage Discrimination 81
Donna Euben

Equal Pay Act 81
Title VII of Civil Rights Act 81
Title IX 82
Executive Orders 82

E. Activist Strategies When All Else Fails 83
Malta Levine

F. Categorical Modeling 85
Lois Haignere and Bonnie Eisenberg

Data 85
Overall Model Statistic 86

G. Redundancy Problems 87
Lois Haignere and Bonnie Eisenberg

H. Promotion and Salary Inequities Between Men and Women Faculty 89
Robert Johnson and Dorothy Kovacevich

Academic Rank As a Measure of Productivity 89
Academic Rank As a Reflection of Status 89
Conceptual and Statistical Models 90
Data 91
Findings 91
Conclusion 93

I. Locating Outliers 95
Lois Haignere and Yangjing Lin

Studentized Residual 95
Hat Matrix Diagonal Element 95
Cook's Distance 95
Comparison of the Three Outlier Measures 96

Glossary 97

Bibliography 101

Contributors

Lois Haignere, principal author of *Paychecks*, is president and sole owner of Haignere, Inc., located in Albany, New York <www.payequityresearch.com>. She is a recognized expert on salary equity at institutions of higher education, and she consults internationally on pay equity and equal pay. In her more than fifteen years of research and consulting, she has conducted hundreds of studies of salary disparities in universities, government agencies, and unions. Through her publications, she has made the statistical results of these studies understandable to a wide audience. Haignere, who holds a Ph.D. in sociology, formerly served as director of research for the Center for Women in Government at the State University of New York at Albany and director of research for United University Professions.

Bonnie Eisenberg, a statistical programmer, has ten years of experience in conducting studies of salary equity at institutions of higher education. She received an M.A. in public health from Yale University's School of Medicine in 1996 and has served as a research assistant for United University Professions and Haignere, Inc.

Donna Euben, counsel at the American Association of University Professors, staffs the Association's Committee on the Status of Women in the Academic Profession. She is a graduate of Oberlin College and Brooklyn Law School (magna cum laude), where she served as editor-in-chief of *Brooklyn Law Review*. She clerked with Justice Handler of the New Jersey Supreme Court.

Robert Johnson, professor of sociology at Kent State University, specializes in medical sociology, life course and aging studies, and the social psychology of the self. His research interests include the social and psychological correlates of physical and mental health. Johnson's recent publications have appeared in the *Journal of Health and Social Behavior*, the *Gerontologist*, and the *Journal of Aging and Health*.

Dorothy Kovacevich, professor emerita of special education at Kent State University, was a leader of her campus AAUP chapter's Committee on the Status of Women in the Academic Profession and an activist for equal pay and status for women faculty at her university. She was also the plaintiff in *Kovacevich v. Kent State University*, 224 F.3d 806 (6th Cir. 2000).

Maita Levine, professor emerita of mathematics at the University of Cincinnati, joined the faculty at Cincinnati in 1963 and received her Ph.D. from Ohio State University in 1970. She is a member of Phi Beta Kappa and Sigma Xi and a recipient of a National Science Foundation Science Faculty Fellowship. She has also served as first vice president of the American Association of University Professors, chair of the Association's Ohio conference, and president of the Association's University of Cincinnati chapter.

Yangjing Lin, senior program analyst at Automatic Data Processing, received a Ph.D. in educational administration and policy studies from the State University of New York at Albany in 1993. He formerly served as a research associate for United University Professions, where he studied salary disparities based on gender and race among higher education faculty.

Figures and Tables

Figures

1.1 Growth in the Percentage of Women Among Full-Time Faculty in the United States, Selected Years, 1974–2000

1.2 Women's Salaries As a Percentage of Men's Salaries Among Full-Time Faculty in the United States, Selected Years, 1976–2000

5.1 Accelerating Curve

5.2 Decelerating Curve

6.1 Salaries for a Two-Year College, State University of New York

7.1 Salaries Before Remedy

7.2 Salaries After Remedy

7.3 Below-the-Line Remedy

7.4 Below-the-Line Remedy Extended to All

A.1 Salary by Years of Service

A.2 Salary by Years of Service

Tables

4.1 Frequency Distribution for Hypothetical Institution

4.2 Categorical Modeling: A Pairwise Comparison for Gender Bias

4.3 Example of Categorical-Modeling Analysis of Rank

4.4 Analysis of Promotion-Probability Table: Transition from Associate to Full Professor

4.5 Event History of Promotion Results

A.1 Predicted Salaries of Male and Female Faculty in Three Disciplines

A.2 Sample Computer Output for Proxy College: Regression Analysis of Faculty Salaries

B.1 Faculty Composition by Race and Gender at Twelve State University of New York Institutions

H.1 Hazard Rate and Cumulative Proportion Not Promoted to Associate Professor, by Sex and Year Since Highest Degree

H.2 Hazard Rate and Cumulative Proportion Not Promoted to Full Professor, by Sex and Year Since Highest Degree

Preface

In 1993 United University Professions (UUP) released a study of gender- and race-based salary disparities among faculty and professional staff at the State University of New York (SUNY). Lois Haignere, one of the nation's leading experts on pay equity and equal pay, conducted the study on the twenty-nine SUNY campuses on which the UUP acts as a collective bargaining agent. The findings resulted in $2.2 million in wage adjustments for some 5,700 women and minority faculty and professional staff.

Although the immediate outcome of the study was gratifying, no one tried to pretend that the wage adjustments would make salary disparities in higher education a thing of the past. Recognizing that low pay for women and minorities inevitably undercuts all salaries, UUP president William Scheuerman approached the American Federation of Teachers for funding and support to pursue the issue. The result was the first edition of this book, *Pay Checks: A Guide to Achieving Salary Equity in Higher Education,* published in 1996.

Since that time, the UUP's research department has maintained an ongoing interest in studying equal pay. Each year, it publishes a report titled *Discounted Salaries and Dismal Promotions: A UUP Research Report on Issues of Gender Equity at the State University of New York.* The current edition of *Pay Checks,* now called *Paychecks: A Guide to Conducting Salary-Equity Studies for Higher Education Faculty,* is further evidence of the UUP's involvement in the subject. The book results from a collaboration among the UUP, the American Association of University Professors (AAUP), and Haignere, Inc.

The SUNY-UUP research, like that of hundreds of other colleges and universities, was built on techniques applied by the AAUP to identify and remedy salary inequalities. The AAUP pioneered the movement to correct salary inequities between female and male professors in the academy through its Committee on the Status of Women in the Academic Profession. The first project of the newly created committee in 1917 was a "preliminary study" of 176 institutions of higher education. It found that 47 percent of the coeducational colleges and 27 percent of the women's colleges studied "frankly admitted that women are given less salary and lower rank than men for the same work."

Over the decades, the AAUP has continued to help faculty members and administrators promote gender equity in compensation. In 1977 the AAUP developed the *Higher Education Salary Evaluation Kit,* which hundreds of administrators and faculty members have used to examine salary structures for gender inequities and ways to correct them. In addition, members of the AAUP's women's committee have offered the academic community detailed strategies for documenting and rectifying gender-based salary inequities.

Each year, the AAUP publishes a salary survey that provides comparative salary data by gender. From time to time, the Association also issues special research reports that examine women's salaries in the academy. In 1998 the AAUP published *Disparities in the Salaries and Appointments of Academic Women and Men: An Update of a 1988 Report of Committee W on the Status of Women in the Academic Profession.* The report found that substantial disparities in salary, rank, and tenure status between male and female faculty members persist despite the increasing proportion of women in the academic profession.

On the local level, AAUP chapters at colleges and universities across the country have helped establish policies to correct salary disparities. These policies are often included in faculty handbooks or collective bargaining agreements. In addition, the

AAUP supports litigation to promote salary equity on college and university campuses and offers technical advice to faculty members and administrators who want to conduct salary and promotion studies. Hundreds of colleges and universities have used the techniques pioneered by the AAUP to identify and rectify salary inequities.

Lois Haignere began her research on salaries at SUNY's Center for Women and Government, where she directed projects investigating pay equity.[1] In 1988 she became director of research at the UUP. Over the next twelve years, she analyzed data sets from SUNY campuses ranging from two-year technical schools to research universities. The diversity of the SUNY system allowed her to observe the effects of different statistical approaches on different types and sizes of institutions.

Subsequent to her work with SUNY, Haignere conducted equal pay research at other U.S. and Canadian colleges and universities. In 1998 she launched her own firm, Haignere, Inc., and has since carried out equal pay and pay equity projects for many unions and institutions of higher education, including the University of Maine Higher Education System, as well as the Canadian Human Rights Commission, the City of Toronto, and Toronto Public Libraries. For details, see <www.payequityresearch.com>.

To inform the writing of *Paychecks*, twelve SUNY institutions were selected on which to test the methods discussed. Selections were made with an eye toward including data sets ranging widely in size and content within three institutional types. Four two-year colleges, four four-year colleges, and four research universities were chosen. (For more detailed information about these twelve institutions, see appendix B.)

The contributors to this book have used their combined experience across the decades to help readers appreciate the importance of faculty salary analyses, learn how to conduct them, and understand the analyses that have been done by others.

Note

1. Pay equity, also called comparable worth, refers to the issue of paying those with traditionally female job titles, such as nurse, secretary, or teacher's aide, what men with jobs requiring comparable levels of skills and responsibilities would be paid. It is an area of gender bias not covered in this guide.

Acknowledgments

Mary W. Gray, professor of mathematics at American University, was instrumental in providing direction for this edition of *Paychecks*. The contributors are grateful to her for her thorough review of the manuscript and the numerous insights and suggestions she offered. The enthusiasm of Maita Levine, professor emerita of mathematics at the University of Cincinnati, helped initiate the collaboration that resulted in this second edition, and her commitment to the project is reflected in many ways in the final product. Members of the AAUP's Committee on the Status of Women in the Academic Profession also participated in bringing this project to fruition. The AAUP's editorial staff deserves recognition for coordinating the editing and production of the guide.

Special thanks to those who reviewed the final draft of the manuscript, including Mary W. Gray, Marica L. Bellas of the University of Cincinnati, Myra Marx Ferree of the University of Wisconsin, Thomas Kriger of United University Professions, and Herbert I. Brown of the State University of New York at Albany. Their comments were extremely helpful.

Sincere thanks also to the research assistants who worked on the project: Michelle L. Johnson, Sharon Baker, and John McCarthy.

The best interests of labor require the admission of women to full citizenship as a matter of justice to them and as a necessary step towards insuring and raising the scale of wages for all.

—American Federation of Labor Resolution
American Federationist
April 1919

ONE

Introduction to Equal Pay for Equal Work

By Lois Haignere

Institutions of higher education have played an important role in establishing the fundamental principle that discrimination based on race or sex is unfair to individuals and counterproductive for society. Yet many of the same features that give vitality to higher education leave it vulnerable to salary discrimination, however inadvertent. Academic departments, for example, enjoy great autonomy in appointing faculty members and setting their salaries. Professors believe this latitude, which is grounded in historic respect for academic freedom, market forces, and peer review, helps raise the quality of their departments and institutions.

Unlike universities, most large employers have centralized personnel procedures that lead to relatively uniform policies for hiring and salary setting, including job classification and job evaluation, across different departments and units. These uniform procedures help guard against gender and race bias in salaries.

Costs of Salary Inequity

Even relatively small differences in initial salary grow over time. If annual across-the-board salary increments average 3.5 percent, a woman or minority male who earns $1,000 less than a colleague at the outset would accumulate $84,550 less over a forty-year career. If $1,000 a year were invested or saved in an account at 5 percent interest, at the end of forty years the account would hold $210,684. For a salary difference of $2,000, this figure would more than double, owing to compounding. After forty years, the highest salary of a faculty member starting $1,000 behind her colleague would lag by $3,825, affecting her calculated pension. Pensions derived from "percent of salary" contributions to annuity plans are similarly reduced. So, too, are maximum contributions to supplemental retirement plans and any employee contributions to such plans. Add to this any bias that may occur in the allocation of discretionary or merit awards and promotions, and it is clear that systemic bias is potentially very costly to the individual.

Fairness alone justifies a review of parity in faculty salaries. But benefits also accrue to the institution and to all faculty members, men and women alike, and their families. These benefits include a sense of inclusiveness among faculty, improved academic morale, better overall salaries, lower staff turnover, and an enhanced public image.

Unconscious Ideology

Systemic discrimination seeps into salary decisions by way of unconscious ideology, which is prevalent in our society and affects everyone: men and women, members of minority groups, and those from nonminority populations.

One determinant of starting pay is what a new hire requests. It is commonly believed that some women and minority candidates ask for less than their white-male counterparts. To the extent that this is so, it may stem, in part, from their immediate experience with market bias, as evidenced by their previous salaries. We all have internalized biases that are expressed even when, intellectually, we know better.

In many contexts, the mere fact of identifying work as having been done by a woman results in its receiving a lower evaluation and a lesser reward than when the same work is attributed to a man. Research shows that both men and women judge artwork, literature, resumes, and scholarly articles attributed to women more harshly than these same items attributed to men (Goldberg 1968; Pheterson, Kiesler, and Goldberg 1971; Paludi and Bauer 1983). Moreover, for women, unlike men, a favorable performance evaluation tends not to lead to promotion and advancement. And research indicates a strong inclination among managers to discriminate against women when reviewing applications for technical-managerial jobs and making recommendations for promotion or training (Rosen and Jerdee 1974; Ruble and Ruble 1982; Gerdes and Garber 1983; Valian 1998).

Studies specific to higher education reveal that department chairs, deans, and members of faculty search committees prefer curricula vitae attached to male names over the same vitae attached to female names (Fidell 1970; Top 1991; Steinpreis, Anders, and Ritzke 1999; Davidson and Burke 2000). When search committees in one study offered positions to hypothetical candidates, female names were more likely to receive non-tenure-track, lower-rank positions; only male names were offered full professorships (Fidell 1970).

The combined impact of the internalized biases of candidates themselves, of those providing references for candidates, and of those involved in the hiring process can mean that equally qualified women and minorities are hired to do the same work as their white-male counterparts at lower salaries. Promotion and discretionary raises can also be affected.

Lest we fool ourselves that these phenomena have disappeared with time or feminist enlightenment, it has been shown that sex stereotypes persist, in spite of the women's movement, legal protections, and the professional advancement of some women (Ruble and Ruble 1982; Valian 1998; Top 1991; Toren 2000).

Over the past two decades, hundreds of institutions in the United States and Canada have conducted statistical studies looking for possible salary bias (Gray 1993). The studies have been initiated variously by bargaining units, women's groups, faculty members, and administrators.[1] The growing numbers of women and minority faculty members and the increasing availability of computerized human-resources data make it likely that more requests for statistical reviews of salaries will arise in coming years.

Objectives of This Book

From a perspective of faculty employment and labor policy, what does it mean to review salaries statistically? What are the political and legal implications? What information is gained and how useful is it? Can some statistical methods actually hide or fail to make visible gender and race bias? Will all faculty members understand the results of such reviews? The contributors to this guidebook provide practical answers to these questions based on our access to a wealth of data and our experience conducting analyses of salary bias.

The guidebook is designed to be a "full service" resource for those who want to conduct salary bias analyses or to understand and interpret results presented to them. The objective of the contributors is twofold: to present information in a way that is not too technical for nonstatistician faculty leaders and other policy makers to grasp, and to provide the technical information needed by statistical researchers to conduct multiple-regression salary reviews.

This dual purpose has led us to include a glossary and an appendix describing multiple regression as it is used to study equity in salaries (appendix A). Recognizing that our audience has a wide range of mathematical expertise, we have attempted to make this appendix understandable to anyone who can add, subtract, and multiply. It walks readers through regression modeling and provides a basic guide to the computer output of regression results. Those already familiar with multiple regression may find this appendix helpful in understanding the relative strengths and weaknesses of the different types of regression models we discuss (chapter 5), why we suggest that certain approaches be avoided (chapter 6), and the recommendations we offer for making salary adjustments to remove bias (chapter 7).

This introductory chapter and chapters 2, 6, and 7 give an overview of research methods and related remedies and discuss their political implications. Chapter 2 describes what a multiple-regression salary study can reveal; the expertise, computer resources, and time needed to conduct such a study; the process for selecting a consultant; and the value judgments and decisions involved in conducting the study. Chapter 6 analyzes the eight pitfalls that can mask gender and race bias. Too often, research decisions are touted as strictly methodological, and their subjective nature and potential political impact are ignored. Chapter 7 explores the political dynamics of diagnosing and treating salary inequities. In particular, it discusses two common views of how bias becomes embedded in salaries, the politics associated with these views, and the impact of some of the remedies that have been undertaken to eliminate salary bias. In addition, the chapter describes the methodological limits of responding to political pressures for further analyses, the remedy that was applied at the State University of New York and why it was chosen, some of the ramifications of certain remedies, and proactive steps to keep salary inequities from occurring.

Chapters 3, 4, and 5 are intended primarily for those who are directly involved in designing and implementing salary reviews, but we have tried to keep even these more technical chapters accessible by confining some of the technical information to notes and appendices. The chapters provide details on what variables are needed and how to collect and clean the data (chapter 3), methods for examining bias in rank assignment and promotion (chapter 4), and ways to conduct multiple-regression salary reviews (chapter 5).

Neglected Groups

Another goal of this guidebook is to encourage analyses where they have been neglected. Most of the reported multiple-regression salary reviews in higher education have been done at universities, with fewer such analyses having been conducted at four-year colleges and even fewer at two-year community and technical colleges. The scarcity of studies at two-year institutions may result in part from the fact that many of these colleges have salary schedules or grids that are presumed to preclude gender or race bias.[2] The relative lack of studies at two-year colleges is, however, notable in light of the large number of women and minority faculty at these institutions.

Two-year colleges are not the only neglected category. Racial minorities have also been left out of most analyses. Throughout this guidebook, we stress the importance of looking at both race and gender. In some cases, we cite studies or reports that consider only gender and we therefore do not mention race when referring to those studies. (In many salary studies, the number of racial minorities is too small to assess racial bias validly.) But the SUNY-UUP analyses consistently took race as well as gender into account whenever data on race were available. (See the preface and appendix B for details on the SUNY-UUP studies.) Chapters 3, 4, and 5 give recommendations concerning how best to combine or disaggregate race categories for specific analyses.

White Males

Testing for race and gender bias requires comparisons, and the appropriate comparison is with the race and gender category that predominates at almost all U.S. institutions of higher education: white males. Thus "white male" is mentioned often, particularly in methodological chapters explaining how to test for bias in rank assignment (chapter 4) and salary (chapter 5). We mean to assign no blame—and certainly not to kindle animosity—by these frequent references. We are simply acknowledging the fact that white males define the necessary standard—the reference category—against which other groups are measured. Rectifying salary inequities is about raising the pay of gender and race groups when the analyses indicate bias. It is never about lowering the pay of those in the standard or reference category.

SUNY Studies

The collective bargaining agreement negotiated by the UUP and SUNY in 1985 resulted in multiple-regression analyses of faculty salaries at twenty-nine SUNY campuses. The university system's Central Administration Office of Employee Relations and Personnel Operations coordinated the collection of data from the employee relations or personnel offices on each of the twenty-nine campuses. Because of time and staffing constraints, SUNY's Central Institutional Research Office chose merely to replicate and validate the findings, leaving the UUP's research department to take the lead in cleaning the data and conducting the analyses. Initial analyses often revealed the need to carry out clarifying analyses with redesigned variables or corrected data. It was not unusual for a dozen or more analyses to be run at each institution. As noted in the preface, these studies resulted in $2.2 million in wage adjustments for 5,700 women and minority academics.

In addition, analyses were conducted specifically to inform the writing of this book. Twelve SUNY campuses were selected on which to test the methods discussed here. Selections were made to obtain data sets with maximum variations of size and content within three institutional types. Four two-year colleges, four four-year colleges, and four research universities were chosen, with faculty populations ranging from 99 to 811. Specific reanalyses of salaries and rank data were conducted. The statistical methods for determining bias in rank discussed in chapter 4, and the three types of multiple-regression analyses of salaries described in chapter 5 were tested using different variables on these twelve different SUNY populations. (For more details about the institutions, see appendix B.)

Limits on Coverage

Although, as was just noted, we have conducted many analyses of faculty salaries, in general this guidebook does not report specific institutional-level findings. One exception is the Kent State University case study, which is reported in appendix H.

Also not included is information about the commonly used qualitative approach called "paired comparisons" or "counterparting." This technique involves matching white-male and female or minority faculty members on the basis of similar qualifications. At best, such case-by-case methods are cumbersome, particularly at larger institutions; when there are female or minority faculty with no peers, they are impossible. Moreover, the process of selecting appropriate counterparts generates animosity (K. Moore and Amey 1993). There are further problems when the counterparts are selected by the same administrators responsible for creating any existing inequities (Muffo, Braskamp, and Langston 1979). The paired comparisons method ignores the concept of systemic bias, which affects all women and minorities, in favor of the concept that only

a few women and minority faculty members have been affected (see chapter 7). Multiple-regression analyses, on the other hand, statistically make every conceivable paired comparison.

Demographics

Over the past two decades, the percentage of women among full-time faculty in the United States has grown steadily from only 22.5 percent in 1974 to more than 35 percent in 1999 (see figure 1.1). The continuing increase in the proportion of doctoral degrees granted to women suggests that this trend will continue (Magner 1999).

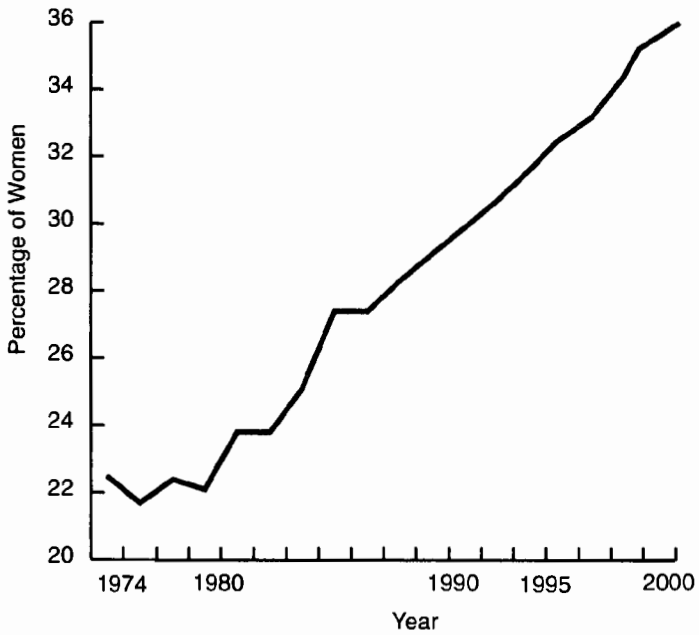
Many people in academia expected that as more women became faculty members, the gap between the average salary for faculty women and the average salary for faculty men would decrease. In fact, however, data from the AAUP's Annual Report on the Economic Status of the Profession, published each spring in the Association's magazine, *Academe*, indicate that just the opposite has occurred. Figure 1.2 shows that between 1976 and 1995 women's salaries declined relative to men's at the ranks of assistant, associate, and full professor. Moreover, between 1981 and 1984, the average salary for female faculty declined sharply relative to that for male faculty.[3]

Since 1984, the ratios have remained remarkably stable for associate and full professors. Women assistant professors, however, continued to lose ground until 1987, when the gap between their salaries and those of their male colleagues began to close somewhat. In 1999 the gap widened again and remained roughly the same in 2000. It is too early to tell if this downturn signals the beginning of a downward trend or is just a temporary fluctuation. At best, salary differences seem to be entrenched, recurring at the same magnitude year after year.

Often, upon hearing this information, people will say, "Yes, but women are paid less because they tend to (a) have less education, (b) have fewer years of experience, (c) be in lower-paid disciplines, (d) be in lower ranks, or (e) publish less often." The statistical techniques in this guidebook are designed to test the validity of these "yes, buts." As is noted later on, it is hard to test the "publish less" factor directly because of the difficulty of collecting the necessary data and of assessing quality versus quantity. If the data are available, however, regression analyses can test the validity of this explanation as well as others. Because the data on average salaries do not suggest that gender-based salary differences are going to go away, it is important to look at whether or not these differences can be explained by career attributes.

Figure 1.1

Growth in the Percentage of Women Among Full-Time Faculty in the United States, Selected Years, 1974–2000



Note: The AAUP's Annual Report on the Economic Status of the Profession, the source of data for this figure, did not provide data on the percentage of women among full-time faculty in the United States for 1978–79, 1981–82, or 1983–86.

Source: AAUP, Annual Report on the Economic Status of the Profession, 1974–2000

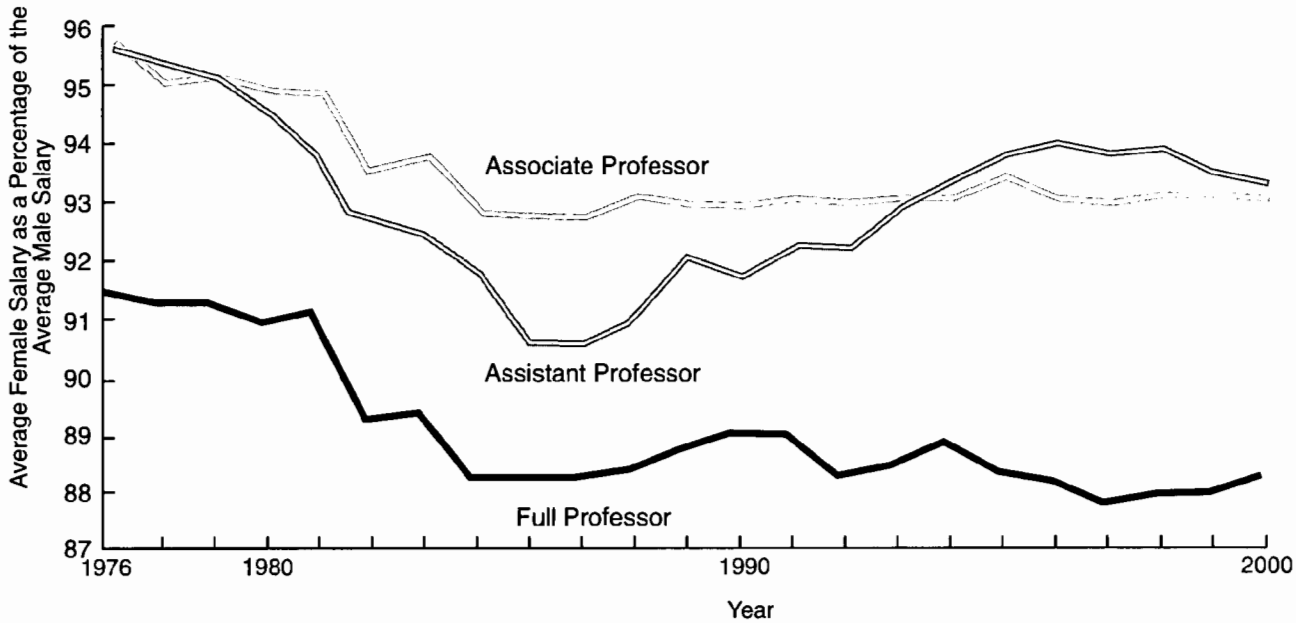
Regression Analysis

The studies this guidebook will help you conduct can be seen as tools for approximating what the salaries of women and minority faculty on campus would be in a completely gender- and race-blind society. We can ask statistically whether faculty in the category female would be paid more on average, given their career profiles, if they were in the category male. And we can ask statistically whether the salaries of those in the category male would on average earn less if they were female.

Regression analyses answer these questions by creating a line that "best fits" the data points scattered above and below it.

Figure 1.2

Women's Salaries As a Percentage of Men's Salaries Among Full-Time Faculty in the United States, Selected Years, 1976–2000



Note: The AAUP's Annual Report on the Economic Status of the Profession, the source of data for this figure, did not provide comparable figures for average salaries for all institutional types in 1979–80 or 1985–86.
Source: AAUP, Annual Report on the Economic Status of the Profession, 1976–2000

Points below the line represent individuals whose actual salaries are lower than the salaries predicted by the variables in the regression analysis. These people are being paid less than colleagues at the same institution with comparable career attributes.

In all likelihood, some points representing men will fall below the line, and some points representing women will fall above it. If, however, you add all the positive and negative distances from the line of the faculty women's scatter points and find a lower total than for faculty men, regression analysis provides a negative number (coefficient) for the variable female. That negative coefficient indicates the average amount that women's salaries would need to be increased for them to be distributed like men's salaries. In other words, this one summary number represents how much, on average, it costs a faculty member to be a woman at the institution under study.

The beauty of the answer provided by multiple regression is that it takes care of most of the "yes, buts." Conceptually, multiple regression lets us compare people with the same level of education, the same years of experience, and in the same discipline and rank, who vary only in their gender or race. Multiple-regression analyses account for variations in salaries by using a set of control or predictor variables, such as years of experience, highest degree attained, rank, and discipline. The information concerning these variables is mathematically held constant while we examine the impact of gender and race on salaries.

Publishing records are not held constant in most salary analyses. Publishing, research, service, and quality of teaching are widely recognized as measures of performance or productivity. Although it would be desirable to include performance information in multiple-regression salary reviews, the time and difficulty involved in collecting and appropriately valuing the relevant information is usually deemed prohibitive. In measuring publications, for example, do we look at quantity alone or also quality? Does the value of a journal article relative to a book depend on the discipline of the faculty member? Do the decisions of journal referees and grant-review committees incorporate gender bias?

Studies that have included performance variables have found that bias in salaries persists; rarely do these

variables have a significant effect on results (Weiller 1990; Dwyer, Flynn, and Inman 1991; Persell 1993; Toutkoushian 1994b; Kolpin and Singell 1996; Nettles, Perna, and Bradburn 2000). One explanation for this finding is that the characteristics of faculty members at institutions emphasizing research and publication are relatively uniform. People who have "made it" are remarkably alike (Gray 1990). At such institutions, rank and tenure may act as proxies for research and publication.

Most community and technical colleges do not emphasize research and publication, and the argument that women tend to teach less or less well has rarely been suggested to explain salary disparities at these institutions. As a result, noninclusion of performance variables may be less debated at these institutions than at research-oriented universities.

In considering the relative importance of performance or productivity measures, remember that regression analyses focus on variations in salaries between classes or groups, not between individuals. The factors that make a significant difference in individual salaries may not be important when the focus is on group differences. To show, for example, that some women are less qualified or productive than some men would not refute findings concerning group differences. Instead, it would be necessary to demonstrate that females, as a group, are less qualified or productive than males, as a group. Findings by Bellas and Toutkoushian (1999) suggest that differences in research productivity by gender are small after controlling for other factors, such as institutional type.[4]

Quality of Regression Results

All multiple-regression salary analyses are not created equal. It is important to know how to judge the validity of different regression equations. An analysis should be based on appropriate predictor variables. If instead of the variable "years of experience," we used "years at current residence," or if we used "highest monthly credit card debt" instead of "highest degree," the result would be a fancy equation that would not tell us much.

Multiple regression gives an estimate of how well the set of control or predictor variables—years of experience, discipline, and the like—account for the variation in the dependent variable, salary. This measure is called the adjusted R^2 (R–square). The *adjusted* R^2 takes into account the number of predictor variables relative to the number of cases (faculty members) in the data set. An adjusted R^2 of 0.75 indicates that 75 percent of the variation in salary is accounted for by the predictor variables in the equation; an adjusted R^2 of 0.55 indicates that 55 percent of the variation is accounted for by the variables.

Assuming that the predictor variables include those most commonly used (see chapter 3), most analyses of faculty salaries have adjusted R^2 values greater than 0.50, and values above 0.70 are common. Thus the variables included in most faculty salary analyses do a good job of explaining the differences between salaries. For more information on adjusted R^2, see appendix A.

Having a high adjusted R^2 does not, however, ensure the validity of your findings. In particular, care must be taken that the predictor variables used do not mask gender bias. A variable that does so is commonly referred to as a "confounding variable" in the statistical literature and a "tainted variable" in the literature on gender bias in salaries. If, for example, height were included in a salary disparity analysis, the shortness of women faculty relative to men could mask gender differences in salaries. It is highly unlikely, of course, that height would be included in an analysis of faculty salaries. But if administrative or twelve-month positions, tenure, or rank are disproportionately awarded to men at your institution, and variables for these criteria are included in the salary analysis, they may mask bias.[5] Tainted variables are discussed at greater length in chapters 4 and 6.

Curiosity, Not Conviction

Salary equity is a politically charged issue. Many individuals think they already know what the outcome of a salary-equity study will be. Some, including administrators charged with managing the salary-setting process, may believe strongly that there is no pay bias. Others, such as women and minority professors, may be just as convinced that a great deal of bias exists.

In a substantial minority of cases, study results show that gender and race have played little, if any, role in salary setting. Among the twenty-nine separate assessments of salaries carried out at SUNY campuses, for example, roughly one-third indicated no tangible gender or race bias. (See the preface for details about the SUNY study.) That some institutions are free of salary bias demonstrates that gender- and race-neutral salary-setting is possible. Yet most analyses that are properly designed to avoid methods that tend to mask bias, find evidence of gender inequity in salaries. Try to be curious about the outcome of your study without second-guessing the results in either direction.

Notes

1. Studies have been conducted at American University, Concordia University, Florida State University, Iowa State University, Kansas State University, Memorial University of Newfoundland, Monroe Community College, New Mexico State University, Oregon State University, Seton Hall

University, Simon Fraser University, Queens University, the University of Cincinnati, the University of Connecticut, the University of Hawaii at Manoa, the University of Illinois at Urbana-Champaign, the University of Maine System, the Baltimore College of Dental Surgery of the University of Maryland, the University of Maryland at College Park, the University of Nebraska at Omaha, the University of Rhode Island, the University of Western Ontario, the University of Wisconsin at Madison, and elsewhere. Some of these studies have uncovered bias, others have not. For details about them and the methods used to conduct them, see Allen 1984; Finkler, Van Dyke, and Klawsky 1989; Geetter 1988; Gray 1993, 1990; Gray and Scott 1980; Hauser and Mason 1993; Ikeda 1995; Hurley et al. 1981; Johnson, Riggs, and Downey 1987; McLaughlin, Zirkes, and Mahan 1983; Muffo, Braskamp, and Langston 1979; Brittingham et al. 1979; Ramsay 1979; Schau and Heyward 1987; Schrank 1988, 1985, 1977; and Scott 1977.

2. Although the author of this chapter has conducted no analyses of two-year colleges with salary grids, she has carried out multiple-regression salary analyses at a Canadian university that had a salary schedule with specified minimums, increments, and maximums in place for more than twenty years. Evidence of gender bias was found, and salary adjustments were made.

3. The 1980–81 edition of the AAUP's Annual Report on the Economic Status of the Profession suggests that the slippage in salary among faculty women compared with that among faculty men—despite the increasing numbers of women faculty overall—could result from two nondiscriminatory phenomena. First, as women move up in rank, they are paid less than men in the same rank because more women are at the bottom rungs of each rank's salary ladder. Second, the report notes that "women faculty may be concentrated in lower-paid disciplines even though their salaries are the same as those of men with(in) each rank in these disciplines."

The first explanation is implausible given the continuing decline of women's salaries relative to men's despite their increasing years in rank and the fact that the decline exists even at the lowest ranks, those of lecturer, instructor, and assistant professor. Indeed, the 1995–96 AAUP salary report notes, based on research controlling for age, that salary differences are not the result of women's later arrival in rank.

The second explanation regarding the concentration of women in lower-paid disciplines was undercut in 1988 by comparisons reported by the AAUP's Committee on the Status of Women in the Academic Profession showing that average salaries for women were below those of men within discipline and rank (Gray 1988). The committee assessed the equally plausible demographic explanation that the salaries of women faculty lagged behind those of men because women were being hired disproportionately at lower-status institutions. The committee found the erosion of women's salaries relative to men's within every institutional type.

Women in academe are disproportionately in lower-paid disciplines (Bellas and Reskin 1994), but there is no evidence that the sharp decline in women's salaries relative to men's from 1981 to 1984 came about because more women entered lower- rather than higher-paid disciplines during this period. Given the increasing number of women studying nontraditional fields as a result of the heightened feminist awareness of the 1970s, it is doubtful that the new faculty women of the early 1980s opted for the lower-paid disciplines in greater proportions than their predecessors.

Another suggested explanation is that before 1980, equally qualified women with new Ph.D.'s were hired at lesser ranks than their male counterparts, perhaps as instructors or lecturers rather than as assistant professors. After 1980, more women may have been hired at equal ranks, but salary differentials increased.

In any case, the multiple-regression statistical methods illustrated in this guidebook can control directly for discipline as well as rank, years in rank, and other career attributes. Most salary-regression studies show that lower salaries for women persist even after controlling for these career attributes.

4. Bellas (1992) shows that the wives of faculty men seem to provide a "(house)wife bonus," which contributes to the productivity of male faculty members. For research indicating how faculty time allocations toward research, teaching, and service differ by race, gender, and family status, see Bellas and Toutkoushian (1999).

5. If the salaries under study have been adjusted to a nine- or ten-month base and all administrative supplements have been removed, there is no need to include these potentially biased variables.

Persuading administrators and some faculty members that statistically identified disparities represent a systemic bias and not just a few cases of possible underpayment of women is the hardest part of a salary review.

—Mary W. Gray
Professor of Mathematics
American University

TWO

Considerations Before Launching a Salary Study

By Lois Haignere and Donna Euben

Deciding whether to assess faculty salaries statistically for gender or race bias (or both) involves technical and political considerations. This chapter offers an overview of some technical and policy decisions you will have to make to carry out a multiple-regression salary review. (Appendix D addresses the laws that apply to salary-equity issues.) The chapter also covers the resources you will need and the process for selecting a consultant. It concludes with a summary of seven salary studies that led to pay adjustments.

Politics of Technical Decisions

Some people mistakenly believe that applying statistics to a political issue such as salary bias eliminates subjectivity and human judgment. But it does not. Although using statistical methods may hide from public view the arena in which value judgments are made, it does not eliminate them (N. Moore 1993). Nor, we would add, does it remove their political impact.

Conducting research on bias in salaries often means involvement in administration-faculty politics. To help ensure a cooperative relationship, it is important to be proactive with the administration. Technical decisions that have political implications are best made openly by well-informed participants.

The right tool

Your first consideration is whether a multiple-regression salary review is the appropriate tool to accomplish your objectives. Multiple regression's strength is in revealing group effects. It can tell you the exact average effect of membership in a group, such as that of female, full professor, or Asian. That is why it is the method of choice for studying systemic bias. If we assume that all those in the group female are affected by the existence of gender bias, then multiple-regression is the best approach to explore the effects of that systemic bias.

But many faculty and administrators see the market and institutional processes as acting fairly except in isolated situations. They hold what is called an individual perspective on bias and discrimination rather than a systemic or structural view. (Chapter 7 provides details about these two different perspectives.) For this group, the objective of a salary study is to find the few individuals whose salaries have been affected by the rare expression of personal bias. If that is your goal, multiple regression is not appropriate. Regression is a statistical technique, not a formula for setting individual salaries.

Of course, you can obtain predicted salaries for individuals from multiple-regression analyses, but individual-level predictions are much less precise than group-level predictions. Individual salaries commonly vary because of effects not represented by variables in multiple-regression analyses. As noted in chapter 1, these analyses rarely include productivity variables; neither do they take into account such variables as being a relative of the dean or provost. Factors affecting salaries but not included in an analysis are presumed to be distributed randomly at the group level. As we pointed out in chapter 1, it is one thing to argue that an individual woman is less productive than an individual man; it is quite another to argue that women as a group are less productive than men as a group.

Decisions, decisions

The decision to conduct multiple-regression analyses of faculty salaries is only the first of many choices that will have a bearing on your results. Who, for example, will be included in the study? If you exclude non-tenure-track faculty, the results will tell you nothing about these academics, who are likely to be disproportionately women or minorities. Moreover, excluding non-tenure-track faculty may substantially reduce the size of your data set and thereby restrict your ability to use some methods. Similarly, if you decide to examine gender but

not race, will you include all males or only white males in your comparison group?

The variables you select can influence your results. As noted in chapter 1, if your institution confers administrative or twelve-month positions disproportionately on white men, and you include these variables in your analysis, they may mask bias by attributing some of the difference in salary to the confounding variable rather than to gender or race. (You may be able to avoid any need to include these potentially tainted variables by adjusting the salaries in your sample to a nine- or ten-month basis and by removing the administrative supplements.)

Current rank is even more controversial, since it is strongly related to salaries but can still incorporate gender or race bias. Elizabeth Scott (1977) argues that qualified women who are denied promotion will appear to be overpaid for their rank rather than frozen in it. Chapter 4 includes a more complete discussion of potentially tainted variables and how to assess their impact.

To further complicate the decision-making process, there are several different types of multiple-regression approaches for reviewing salaries, each giving a slightly different picture. Depending on what you want to know and the peculiarities of your data set, such as size and range of salaries, you may judge one approach better than another for your purposes. Alternatively, you may decide to conduct two or three different analyses based on different approaches to determine if the results validate each other. But if you do so and the results differ, you will need to examine why.

If this all sounds daunting, don't despair. This guidebook is designed to help you. Chapter 6 goes into further detail about policy-related decisions. Here our goal is mainly to note that your choice of population, variables, and methods will directly influence your final results.

A simpler method?

If you find that multiple regression, computerized data, and the need to control for all of those variables overly complex, consider using scattergrams. A scattergram shows the pattern of two variables, such as salary and years of experience, by putting one variable on the vertical axis of a graph and the other on the horizontal axis. Points representing the intersection of the two variables for each case (or faculty member) form a "scatter" that gives a good visual image of the relationship between the two variables. See the glossary for more information and figure 6.1 for an example of a scattergram.[1]

You can create a scattergram with the old-fashioned tools of pencil and paper. If you want to look at differences between two groups, say men and women, you can use an *x* for the dots for one group and an *o* for the dots for the other group. If you want to control for rank, you can plot separate scattergrams for each rank. If you are adept with spreadsheet software, you can forgo pencil and paper in constructing these visual representations.

If your college has a small faculty, or if you want to examine only your department or discipline, scattergrams may be the best option. That is because the validity of multiple regression in examining data sets with fewer than fifty faculty members is uncertain. For additional details about analyzing small data sets, see the discussion of faculty size under "Basic Requirements" below. Even with larger groups, scattergrams can provide preliminary information to help you decide whether a full-scale salary study is needed.

Other considerations

With computerization, the automation of personnel systems and other data sets has increased, making it easier to conduct statistical analyses. Most colleges and universities and many unions can now conduct multiple-regression analyses with minor additions or adjustments to already-existing data sets, and the growing capacity of computers to store data encourages the inclusion of more specific information. The expansion of institutional data sets increases the likelihood that institutions will conduct statistical analyses like the ones described in this guide on a variety of topics.[2] Faculty leaders and bargaining-unit policy makers therefore need to grasp basic approaches to statistical analysis.

Basic Requirements

In order to plan your study, you will need to consider the characteristics and resources of your college or university.

Faculty size

If your institution has one hundred or more full-time faculty members, you can probably validly conduct the salary analyses described in this guide. If you do have fewer than a hundred full-time colleagues, you may want to consider alternative methods, such as the scattergrams discussed above. We say "may" because of the other factors that bear on the validity of multiple-regression analyses, primarily the number of variables used and how well they account for variations in salaries.

If salary increases on your campus have tended to take place only through annual increments and promotions,

then multiple-regression analyses could predict your salaries with a relatively small number of variables: years of experience and rank, for example. Adjusted R^2 measures can help you decide, after the fact, if your data set was big enough. That is because the adjusted R^2 takes into account the number of predictor variables used relative to the number of cases (faculty members) in the data set. Unfortunately, you will not know if your data set was sufficiently large until after you have done the analysis.

If your institution has fewer than a hundred full-time faculty members, if salaries are not the result of a simple and consistent pay policy, and if you cannot afford the time and work involved in doing the analysis in order to have the benefit of adjusted R^2 hindsight, we suggest that you not conduct multiple-regression analyses.

Expertise

Although this guidebook is designed to help laypersons understand methods of salary analysis and the output that results from them, you will need someone with statistical and computer expertise to conduct the actual salary analyses. If your institution is large and you are blessed with a cooperative, trusting relationship with the administration, expertise may be easy to come by. Most universities have institutional research departments with the personnel and computer resources to conduct these analyses. Small campuses may have faculty members with the necessary expertise in departments such as sociology, psychology, economics, or statistics.

We do not mean to imply that knowledge of statistical procedures is all that is needed. An interest in the project and in helping others understand statistical procedures and output, a willingness to work cooperatively with different constituencies to design the analyses, and the trust and respect of key faculty and union policy makers and members are at least as important as expertise. It would be a mistake to ignore the political nature of faculty salary analyses when deciding who will do them.

Decision makers often determine that the expertise needed to conduct the salary study can best be acquired by contracting with an outside consultant. Hiring a consultant may bring more experience and neutrality to the project than is available from within the institution. Later in this chapter, we provide insight into consultants and the institutional processes involved in hiring them.

Computer resources

Whoever is selected to carry out the analyses will probably have access to the requisite computer resources, whether an institutional mainframe or a personal computer. Most personal computers sold today have the power and memory needed to conduct complex statistical analyses. Although multiple-regression analyses can be done with spreadsheet software, we recommend using a statistical software package. The ease of using such packages and of interpreting the results they generate outweighs their cost. Some statistical packages actually do the programming for you, prompting you concerning what variables to include.

Time

How much time will it take to conduct the analyses we describe in this guidebook? The process takes place in four main stages: data collection, data cleaning, running of the analyses, and reporting of the results.

Data collection. Your institution's automated personnel system may already contain much of the data you will need. For each variable you select, however, you must determine whether data are missing and, if so, how long it will take to collect the missing information. Some of the variables you want to include may not be available in automated form. If, for example, you have to go through the paper files in the personnel office to collect the necessary degree data, it will probably take a lot longer than if the data are available in a separate automated file, such as the one for the campus catalog. Only you can predict how long it will take to gather missing data. Our advice is to make an estimate and then increase it by half. Our experience indicates that this would be a conservative calculation of the time actually needed.

Data cleaning. Even data that appear to be complete and accurate commonly have anomalies that need to be explained or corrected. You may, for example, find faculty members whose reported initial ranks are higher than their current ranks, faculty who appear to have received their doctorates at age twenty or younger, or those whose discipline code differs from their department code. It takes time to verify or correct each anomaly. Although doing so is time consuming, it is important. As the well-worn adage says, "Garbage in, garbage out." The time needed will depend on the size of the population you are studying. We recommend that you allow, at minimum, two months for data cleaning, more if your faculty is larger than four or five hundred. Chapter 3, on developing a database, contains many suggestions for verifying data.

Running of the analyses. Once the data have been cleaned, the next step is to code the appropriate variables and program and run the analyses. This phase commonly takes more time than anticipated. Consistent

with Murphy's Law, bugs, software glitches, and hardware failures will delay progress. Running analyses may also reveal data errors that were not detected during data cleaning, making it necessary to rerun the analyses.

In addition, there is a tendency for each statistical result to generate further questions. If, for example, you use more than one method of multiple-regression analysis (we describe three types in this guide) and the results differ, you will want to know why. You may decide you want to see some analyses with and without certain variables. If the results indicate bias in salaries, additional analyses aimed at formulating appropriate wage adjustments will probably be required.

Assuming a clean data set and appropriate expertise, we *estimate* that it will take a researcher about twenty hours a week for a month, maybe two, to complete the multiple-regression analyses. Another month or two will be needed if the researcher carries out analyses of bias in rank, tenure, or other potentially tainted variables (see chapter 4). We suggest that the researcher work half time on the project, because the active involvement of the researcher will not be required continually. Given the political nature of the research, it is important to build in time for disseminating information among interested parties and getting their feedback. Of course, additional research questions may result from this exchange.

Reporting of the results. It is important to report the results of the salary study in a way that faculty members can understand. It may also be politically important to provide a draft report to the primary constituents for their review before releasing the final document to a wider audience. Allow two to three weeks for writing the draft, at least a week for the parties to provide feedback on it, and another week to redraft it in light of their comments. This process may also lead to additional analyses. Sharing information and getting feedback can be time consuming, but it helps to ensure an open process.

An open process

The salary-equity investigation should be public and open. Transparency helps gain support for the project and promotes fairness. Regular newsletters, interim reports, and student and local press coverage are good ways to educate the campus community about the effort. A salary-equity study should also involve faculty members. If the administration limits opportunities to participate, or restricts the amount of faculty involvement, then faculty members can still critique the study's findings. Such critiques will have more legitimacy if faculty members have tried to participate in the process. If, after having participated, faculty members find the study's conclusions or recommendations unwarranted, they can prepare a minority report. They can also do so if they think the study's methods were flawed or the process was too political. Again, sharing information and getting feedback can be time consuming, but it helps ensure an open process.

Salary-Equity Consultants

This guidebook presumes that someone in your union, chapter, faculty senate, or administration will conduct the salary research. For various reasons, however, an outside consultant is often sought to undertake such a study. This section offers some insight into consultants and the institutional processes for hiring them.

Competitive bids

College and university administrations rely on different mechanisms to hire outside consultants. The most commonly used is that of the request for proposal (RFP), which solicits competitive bids from interested parties. The RFP process is discussed below.

Administrations can usually find the means and money to accomplish the projects they favor. If the president's office wants to conduct a salary-equity study and knows a reliable and experienced outside consultant, the money and mechanism to make the hire will no doubt be found. Even public-sector institutions can hire noncompetitively under what is often called the "single source" provision for hiring. This provision is sometimes used when one consultant or firm can be shown to be the only appropriate source of expertise, or when the competition is determined to be inadequate.

If an outside consultant or firm will be hired through a competitive process, the first step is to write an RFP describing the proposed project and specifying the population to be studied and the data to be supplied by the institution. The RFP should ask the respondent for a detailed outline of the proposed undertaking, a timetable, a budget, a list of references, and descriptions of the methods to be used. In addition, the RFP should ask about the consultant's expertise in pay-equity studies and the qualifications and experience of the proposed project staff.

Review the proposals you receive, taking into account any regulations the university or college may have for contracting with consultants. Interview, by phone or in person, the consultants who have submitted the best proposals. Any staff of the consultants who will be working on the project should also participate in the interview. When possible, get references from faculty at

institutions where the consultants have completed studies. Members of bargaining units, faculty associations, and women's groups may be able to help with references. Finally, select the consultant who best meets your needs, and sign a contract with him or her.

Faculty role

If your chapter, faculty senate, or union decides to play an active role in the study, it must get involved early on, when the RFP is being drafted and a consultant selected. Once the study begins, it is often too late to change aspects of it you do not like.

To ensure that you get in on the ground floor, review your institution's options for hiring consultants. Who needs to approve such a hire? Can a consultant be hired without going through a competitive process? If so, how? If an RFP will be circulated, make sure that the research design includes features you consider important. As you review the draft RFP, you may find it helpful to read chapter 3 and chapter 6.

A faculty committee may want to provide names of consultants who have produced acceptable studies and ask to review the bids received to determine whether candidates have the requisite experience. Has the candidate completed similar salary studies at similar institutions? Has she completed projects to the satisfaction of faculty groups? Does he have a good track record with unions and in unionized settings? Does she understand the unique features of your institution and the proposed study?

Has the candidate submitted a clear proposal? (A confusing or overly technical proposal may be a sign that the consultant does not know the field well enough.) Does he have a staffing plan for completing the project on time and within budget? Is she committed to the notion that no faculty members should have their salaries reduced as a result of the study?

The faculty committee may also want to participate in the final selection of the consultant and review the contract before it is signed for language specifying research methods and the scope of the undertaking. In addition, the committee should ensure that an appeals mechanism exists in case the faculty wants to challenge the consultant's conclusions or recommendations.

Bottom line

A consultant works for the client. If you or your faculty group are not represented among those who hire the consultant or receive the final report, your views are unlikely to influence the research.

Consultants hired and paid unilaterally by the university or the college administration will probably be sensitive to the wishes of the administration only. If the institution involves the faculty women's committee or another faculty or advocacy committee in selecting and hiring the consultant, the consultant will understand that the methodological concerns of committee members are important to the institution. By contrast, if the consultant is ushered in and out of the president's or provost's office and meets no one else, that will speak volumes as well.

Not all consultants are hired unilaterally by the administration. Your faculty group may be able to negotiate a contract under which the institution pays for the study even though the consultant must report to both faculty representatives and the administration. If possible, the faculty group may want to contribute to the consultant's payment so that the consultant understands that the project is for two distinct clients who are fully and equally involved.

If your union, chapter, or faculty group has the resources and data, you may even want to hire the consultant unilaterally. Consultants can be expensive, but the expense may be justified if the consultant has solid experience in doing salary-equity analyses. A consultant brings the added benefit of appearing more neutral than internal faculty or institutional researchers.

Consultants—individuals and firms—abound. Look for one who puts the validity of the research above all else and who has a good history with faculty and faculty organizations.

Happy Endings

Hundreds of institutions in the United States and Canada have conducted salary-equity studies using multiple regression.[3] The analyses have been initiated variously by bargaining units, women's groups, faculty members, and administrators. We have collected most published and some unpublished reports of these studies, but we have not attempted to identify every one undertaken. We suspect that those that find bias and result in awards are more likely to be published than those that do not find bias or lead to salary adjustments.

Below we describe eight studies that resulted in awards to remove bias in faculty salaries. The amount of salary adjustments varied. In our experience, most awards range between an annual salary increment of $500 and $2,500.

Memorial University of Newfoundland

Motivated partly by a quantitative report published by the Canadian Commission on the Status of Women, Memorial University commissioned a faculty salary

study in 1973 (Rosenbluth 1967; Robson 1969). After a year of collecting and cleaning data, which included information on publications and research, multiple-regression analyses were conducted on a total of 598 tenure-track faculty members, including 104 women.

The study found average annual salary disparities between men and women amounting to $705 when rank was included in the analysis. Because of strong evidence that women were discriminated against in promotion, a further analysis was conducted omitting rank. Without the rank variable, the annual salary disparity jumped to $1,766 (Schrank 1977).

The university president rejected the faculty association's recommendation that blanket awards be made to all female faculty and instead adjusted the salaries of specific women, most of whom were in the School of Nursing or the junior division, which was made up of teachers of first-year courses. The awards amounted to roughly half of the disparity found by the regression analyses.

A follow-up study, conducted on 1982 data, again found gender disparities, but not in cases in which adjustments were made after the 1973 study. Where no adjustments had been made (the School of Physical Education and the Faculty of Arts, for example), disparities remained. Although the 1982 study resulted in no awards, the findings enabled the faculty association to negotiate a salary grid system. By gaining placement on this grid, many women, particularly those with the most longevity in the system, received substantial increases. One senior woman faculty member received a 65 percent increment.

Monroe Community College

The women's caucus of the Monroe Community College faculty association initiated regression analyses of the 1983 salaries of tenured and tenure-track teaching faculty. The study found gender bias in both salaries and promotion among a population of 99 women and 177 men.

The faculty association filed a grievance in 1985, because the college declined to correct the situation voluntarily even though sexual discrimination violated the association's contractual agreement. An arbitrator eventually persuaded the college to negotiate with the faculty association regarding possible remedies.

As a result, 99 female faculty members received salary increases of $700 and an additional amount of about $9 for each year of service. The remedy also addressed the issue of unfair promotions and provided for the monitoring of future promotions.

University of Nebraska at Omaha

In 1984 the Chancellor's Commission on the Status of Women at the University of Nebraska initiated a series of studies that looked at the status of three different groups of women: clerical and technical, managerial and professional, and teaching professional (faculty). A joint administration-AAUP committee decided such matters as the variables and the statistical methods to be used, the population to be studied, and any pay adjustments to be made if disparities were found.

Conflicts occurred over the use of starting salary as a predictor variable. The administration viewed starting salary as a reflection of historical market value. The AAUP chapter believed that the starting salaries were themselves affected by gender bias. Two studies were conducted by the AAUP to examine starting salary as the dependent variable. The results revealed gender bias in starting salary, and the university agreed to leave this variable out of subsequent analyses (see the discussion of initial salary in chapter 6).

The population to be studied was also controversial. Discussion centered on two groups: faculty hired before 1972 and temporary faculty members. The administration wanted to exclude faculty members appointed before 1972 because their salaries had been adjusted based on a 1971 study. Doing so would have left 37 percent of the faculty out of the analyses, substantially lowering the population.

The administration also wanted to exclude temporary faculty, but the AAUP chapter argued that it was legally responsible for representing the entire bargaining unit. The parties eventually agreed to include the entire faculty in the study. After more than eighteen different analyses, each woman faculty member received an award of $1,000 to remedy the disparities found (Finkler, Dyke, and Klawsky 1989).

University of Connecticut

A regression analysis of faculty salaries was conducted in 1988 in response to two factors: the requirements of a collective bargaining agreement between the university and the local AAUP chapter and the state's commitment to eliminating gender-based salary inequities.

The study found that female faculty were, on average, paid $1,806 less than their male counterparts, and that fewer female than male faculty members received salary increases indicating high merit. In addition, only 29 percent of women received salaries higher than those the analysis predicted they would have earned had they been male; 66 percent of women earned less than the salaries predicted for them. The study concluded that there was no statistical indication that only a subgroup

of women faculty had been affected by gender-based inequities, and it therefore recommended a salary adjustment for every woman on the faculty.

The total amount of money to remedy the sex-based inequity was estimated to be $470,553, or an average of $1,806 for each female faculty member. Two-fifths of the money was awarded as a flat amount to all women faculty members, two-fifths was awarded as a percentage (1.87 percent) of each woman's salary, and one-fifth was distributed to women faculty members on the basis of merit (Geetter 1988).

University of Connecticut Health Center

In 1992 a multiple-regression study of the salaries of faculty in the university's medical and dental schools was conducted on the recommendation of a committee of women faculty appointed by the administration. The study, which resembled the one carried out earlier on the university's main campus, revealed average disparities between men and women faculty of $4,731. The working group on the status of women faculty reported the results to the administration and recommended that all women's salaries be raised by $4,731 (Ferree and McQuillan 1998).

The administration did not accept the study, expressing concerns about the population included as well as the accuracy of the human resources data and the variables used. It preferred a "flagging" approach, which identifies as potentially underpaid those individuals whose statistically predicted salaries were higher than their actual salaries. The women's committee pointed out the problems with flagging, not the least of which is a lack of concern for women who are "top scholars" and are underrewarded for their productivity (see chapter 7).

Additional analyses were run to address the administration's concerns. As a result of this further study, the administration instituted biennial salary reviews and recommended to the board of trustees that salary-setting guidelines be developed. In June 1994 the administration increased the pay of women faculty across the board by $1,000 (prorated for part-time faculty). Additional awards were made to individuals based on the administration's flagging model, which showed the salaries of some of the women reviewed to be significantly below those of their male peers.

University of Hawaii at Manoa

In 1993 the University of Hawaii completed a salary-equity study as part of an affirmative action initiative proposed by the university president. The proposal responded to concerns expressed by the campus Commission on the Status of Women, other faculty members, and the Office of Federal Contract Compliance Programs.

The study assessed the effects of both gender and ethnicity on salaries. The database it relied on consisted of the 1991 personnel records of 1,004 faculty members. The results indicated that both women and persons of Japanese descent were underpaid by about 4 percent.

Although external consultants recommended an across-the-board remedy for the disparities, the university ended up implementing case-by-case reviews. Women and minority male faculty whose actual salaries fell below the median pay of comparable white-male faculty—or below the regression line—could apply for a case review. Those whose actual salaries were above the comparable white-male median had to use other routes (grievances, federal agencies, or the courts) if they wanted redress.

An all-campus faculty equity panel with 35 members, including union representatives, worked for two years deciding the cases of the 223 women and minority men who applied. (Only seven women eligible for reviews did not apply.) The total amount awarded to the 171 women who received adjustments was about $1,525,351, an average of $8,920 each. Eighty-eight minority men received about $810,331, an average of $9,208 each.

State University of New York

In 1993 United University Professions (UUP) and the State University of New York (SUNY) completed a study to assess gender- and race-based salary disparities. The study was called for by the terms of the collective bargaining agreement between SUNY and the UUP. The bargaining unit consisted of twenty-nine SUNY campuses, including four major university centers, four medical schools, thirteen four-year colleges, six technical colleges, and specialized colleges such as the Colleges of Optometry and Forestry. A joint labor-management committee selected the methods to be used and made other policy decisions, after which multiple-regression analyses were carried out on each of the twenty-nine institutions in the unit.

As a result of the analyses, SUNY made more than $2.2 million in wage adjustments went to women and minorities. These awards went to women and to those in a given racial category if average salaries for that category were found to be lower than those for comparable white males at a particular campus (Haignere, Lin, and Eisenberg 1993). The amount of money set aside within the contract for salary corrections paid only about 25 percent of the amount of the salary disparities indicated by the research.

University of Maine System

As part of the 1997–99 collective bargaining agreement between the Associated Faculties of the University of Maine (AFUM) and the University of Maine System, a joint committee was established to study the issue of gender equity in faculty salaries. Regression analyses controlling for the factors that are legitimately related to salary were conducted at the systemwide level, at all seven institutions in the system, and at some of the larger colleges within the University of Maine and the University of Southern Maine. Study findings were reported in October 2000.

The committee determined that systemic gender inequity existed when being female cost a faculty member 2 percent or more of the average male salary. Salary disparities in excess of 2 percent were found at three universities and within some of the colleges at the two largest universities. The annual salary disparities ranged from $1,438 to $3,079.

The committee recommended systematic adjustments to salaries where the 2 percent threshold was met or exceeded. The university and the AFUM reopened negotiations upon receiving the committee's report and reached agreement on salary adjustments. Roughly 80 percent of the award amount was determined by a formula containing two components, each representing roughly half of the total: a flat dollar amount for each eligible woman and an amount based on the number of years of university employment, up to twenty. If a woman's salary was already higher than those of men comparable to her in terms of department, discipline, rank, and highest degree, then the formula awarded her a reduced amount or no amount.

The remaining 20 percent of the adjustment amount was set aside for consideration of individual female salaries that, even after application of the formula, remained significantly below the earnings of men with similar qualifications and experience.

In addition to adjustments to base salaries, women faculty who had completed at least eleven years of service received a one-time payment, not added to base salary, for each year of service between eleven and twenty years. In total, roughly $400,000 was distributed to two hundred faculty members.

Notes

1. Figure 6.1 contains more information than most two-variable scattergrams because the figures on the horizontal axis (predicted salaries) are based on regression modeling information rather than on years of experience, age, or a similar single variable.

2. Institutional research departments are increasingly venturing into automated analyses that require extensive, precise data. The Office of Institutional Research (OIR) at the University of Delaware, for example, is conducting the National Study of Instructional Costs and Productivity. With a major grant from the Fund for Improvement of Post-Secondary Education, the OIR is collecting detailed information from many institutions of higher education in the United States on student credit hours and organized class sections within disciplines.

3. See note 1, chapter 1.

It's never sex bias, just lack of time in service, bad negotiating skills, bad timing, wrong field, or bad-hair day that accounts for all these "discrepancies." So say the defenders of the status quo.

—Ellen B. Kimmel
Professor of Psychology
University of South Florida

THREE
Database Decisions and Development

By Lois Haignere and Yangjing Lin

Policy-related research projects usually begin with policy decisions, and faculty salary reviews are no exception. Determining what information to include in the project database will require you to make policy decisions that will influence the outcome of your study. Who makes those decisions will vary according to the political forces that bring the project to life. At one end of the continuum, a faculty group unilaterally conducts the study; at the other, the administration does. Between these two extremes are a range of cooperative arrangements with different degrees of faculty-administration control.

This chapter is geared toward those of you involved directly in designing and implementing a salary review. It will help you determine what data you need and how to collect and clean the information to be analyzed.

Policy-Making Committee

Decisions about faculty salary reviews are often made by committee. Strive to create a policy-making committee that has representatives trusted by the different constituencies at your institution. Include men and women from among the faculty and the administration who understand the importance of addressing a perceived equal-pay problem. Of course, many of the effective faculty and administrators you want to involve will probably be busy. To encourage participation, make sure to explain that the issue at hand goes beyond potential salary improvements for minority and women professors: faculty morale and productivity, the reputation of the school, and alumni support will all be affected.

To help you determine who belongs on the policy-making committee, you may want to revisit the "Basic Requirements" section of chapter 2 and take an advance look at chapter 6. Here we list some questions with which the policy-making committees will have to grapple.

Should the study include non-tenure-track and part-time or adjunct faculty? What about nonteaching faculty and faculty from the evening division and satellite sites? How about faculty from professional schools, department chairs, and other departmental administrators?

What gender and race categories should be examined: Asian women, all minority women, Latino men? Which variables should be included: initial rank, current rank, publications, previous experience, teaching quality, discipline, and tenure? What outcome measures will be used to diagnose bias: standardized or unstandardized coefficients, residuals, adjusted R^2, or statistical significance? Each of these questions can have political implications, and each can affect the study results. How the questions are answered may depend on who is on the policy-making committee.

When other considerations are equal, it is best to appoint people who are comfortable with research and statistics. They do not need to know a lot about statistical methods, but they must not be intimidated by statistics or afraid to insist on basic explanations about the origin and the importance of the numbers. Appendix A contains information committee members will need about regression analyses.

Avoid statistical snobs—people who know a lot about statistics, but who cannot or will not explain the implications of statistical findings to others. If you end up with a "trust me, I know it all" type on your committee, you will need to read this guidebook very carefully or find someone with knowledge equal to that of your colleague to cut through the verbal fog.

Policy decisions that seem neutral to some committee members may be viewed as an affront by others. Ensuring good communication is the best way to minimize confrontation and arrive at effective outcomes. But even the best communication cannot eliminate all differences. If you cannot gain the cooperation of the administration in designing a valid study or in addressing salary inequities, you may find the strategies in appendix E helpful.

Study Population

Who should be included in the study? It might be better to ask this question in the negative: who should not be excluded from the data? Non-tenure-track or temporary full-time faculty, whose ranks tend to include many women and minorities (Chronister et al. 1997), are often left out because their jobs are deemed "different." This difference is usually vaguely defined as their having to teach lower-level courses or their being governed by separate hiring and promotion criteria. Comparable differences in teaching expectations and in appointment and evaluation standards are found across assistant, associate, and full professor ranks, making this logic unpersuasive. (This issue is dealt with in more detail in chapter 6.)

Excluding non-tenure-track faculty from the analyses will almost surely mean that they will not be considered for any salary adjustments that may result from the study. In other words, a group of low-paid and often disproportionately female and minority faculty members will be denied the opportunity to have any bias in their salaries corrected (Gray 1993; Hamermesh 1996).

Besides, multiple regression can separate out pay differences that are unrelated to gender or race, as long as rank or tenure status is taken into account. In our analyses of the twenty-nine SUNY campuses, for example, we included all full-time faculty members. To allow the equation to account for salary differences caused by tenure status, we included a variable for lecturer, which is the non-tenure-track rank at SUNY. That way, the salary differences based on gender and race within the lecturer category could be measured as part of the coefficient for the gender and race variables, while salary differences related to non-tenure-track status were attributed to the coefficient for lecturer, not to race or gender. In short, we recommend that all full-time faculty, including those not on the tenure track, be included in the database.

But what about part-time faculty? Here again, women and minorities constitute a high proportion of the population. Many must piece together an existence by teaching part time at several institutions. Unfortunately, including part-time and less-than-full-time adjunct faculty in a multiple-regression salary assessment is almost impossible. Institutions keep little automated information on such faculty, among whom the turnover rate is high. In addition, part-time faculty tend not to respond to questionnaires, making systematic collection of data on them difficult.

Another issue is that many part-time faculty are paid on a per-course basis rather than according to career attributes such as experience. Thus a retired part-time professor with many years of teaching experience may be paid no more than a graduate student. You may want to consider examining per-course pay rates for uniformity across gender and race.

If part-time faculty at your institution are paid on a salary rather than on a per-course basis, and if you have data on them that parallel the data you have on full-time faculty, include them in your analyses. To do so, you will need to adjust their salaries to make their pay comparable to that of full-time faculty. If, for example, part-time professors at your institution work half time, you would double their salaries to approximate what they would make as full-time faculty members. You will also need to include a part-time variable in your analyses.

If, as is likely, you do not have data on part-time faculty that are comparable to what you have on full-time faculty, consider extending any remedies for bias found in the full-time analyses to part-time employees, prorated to the proportion of time they work.

Data Collection

Unless your campus is way behind the times, it has an automated personnel and payroll data system. This system houses the information needed to ensure the accuracy of paychecks and usually includes date of initial hire, date of appointment to current rank, current rank, discipline or department, salary, and contract length (nine-month or twelve-month year). It may also record gender, race, and age in order to meet federal reporting requirements.

On some campuses, the database including personnel information, such as race and year of hire, may be separate from payroll data. If that is so at your institution, obtain a separate download of each database, making sure that a common identifier, such as social security number or first and last name, is attached to the information for each individual. The common identifier can be used to merge the two data sets.

Larger campuses will probably have an institutional research department that relies on the personnel and payroll data set for the faculty information it maintains. Such departments conduct statistical analyses and provide descriptive statistics to assist and advise campus policy makers. Some institutional research departments have added useful information such as initial rank, highest degree attained, and year of highest degree to the database.

As the automation of data increases, more institutional research departments are collecting data on prior faculty experience, performance measures, and the year of each promotion. We address the relative importance of years of experience at hire and proxies for that informa-

tion later in this chapter and in chapter 5. Chapter 1 discusses productivity and performance measures and the use of current rank as a proxy for them.

Of course, "low-tech" or "hard-copy" sources of data still exist. These include the personnel forms filled out by new hires, paper-based personnel records, and campus catalogs (which may contain highest degree attained, source of degree, and year of initial hire in addition to rank and department). Having been spoiled by the convenience of automated data, we shudder at the thought of the labor-intensive task of collecting and entering data from these sources. If you are lucky, you will have to rely on such sources only to check the accuracy of inconsistent data for a few individuals.

Your access to institutional data will depend, of course, on your relationship with your campus administration. The more cooperative the administration is, the easier it will be to access, collect, and verify data. Faculty groups that are not recognized bargaining units or unions may have more difficulty than recognized groups in gaining administrative cooperation, especially at institutions at which there is no public access to salary information. In the United States, faculty salaries at public institutions are technically open to public scrutiny under freedom-of-information legislation, and private colleges and universities can be forced to provide pay data by a court order. (In Canada, however, public-sector salaries can be kept secret.)

Fortunately, it is often unnecessary to go to extremes. Many institutions, even private, nonunionized colleges and universities, have been responsive to the concerns of women and minority faculty groups and have worked collaboratively with them to study faculty salaries. For the purposes of this discussion, however, we will assume that you will have little, if any, cooperation from the administration.

Under collective bargaining agreements, most unions receive information about their members, such as name, social security number, address, phone number, department, current rank, and salary. Can you "make do" with this limited data set? If you add gender and race, you can do a regression analysis of salary to see if gender or race bias seems to be present when you control for discipline and current rank. Of course, additional measures, like years in rank or years at the institution, would account for more of the variation in salaries and provide a higher adjusted R^2.

If a preliminary assessment indicates bias, administrators will probably protest that the apparent bias would disappear if you controlled for additional variables. They could be right. If women and minorities have less time at the institution and in their current rank, controlling for these variables may make the evidence of bias disappear. But then again, it may increase, or decrease just a little.

Tell the administration that you would be happy to see if the bias disappears with more control variables, and ask for information on the relevant variables. Administrators may do their own analysis. If the evidence of bias disappears, they will, in all likelihood, give you the data so you can verify its findings. An unwillingness to provide you with the data may indicate that the bias persisted or perhaps even increased in the administration's analysis.

When pressed for data, the administration may offer to do the research and show you the results instead. That is a tempting offer. Whether to accept it will depend on how much you trust the administration and the individual or individuals who will do the analyses. Meet with them. Discuss the data set they will use, the variables, and the types of analyses. Chapter 6 explains approaches that can mask bias. Are they proposing to use any of these approaches? Can you agree on the specific variables and statistical approaches to be used?

If you are at a public-sector institution and you have exhausted friendly persuasion, you may be able to acquire most of the data you need through the Freedom of Information Act (FOIA). With the passage of the FOIA in 1966, federal records became more accessible. All states have subsequently passed similar laws. The salaries of public employees as well as most job-related information such as rank (or job title) and education are covered by these laws. To find out how to request such information, consult your state's law. A limitation in using the FOIA is that personal characteristics such as race, sex, and age are specifically protected against release. You will have to collect these data through common knowledge or other means.

Variables and Possible Substitutes

In both our original analyses of the twenty-nine SUNY campuses and our reanalyses of twelve campuses conducted to inform the writing of this guidebook, we used eleven principal variables: salary, gender, race, highest degree, completion date for highest degree, years since highest degree at time of hire, date of hire at institution under study, current rank, date of promotion to current rank, contract length, and discipline.

Salary

Salary is, of course, the necessary dependent variable in a study of salary equity; there are no alternatives for it. To analyze the effect of gender and race on salary, accurate information on the base salary of each individual in

the data set is vital. Annual salaries are usually the basic unit of analysis, although monthly salaries have also been used. Temporary summer pay or temporary increments associated with administrative duties should be excluded; only base salary should be considered. An adjustment for the amount of time worked may be needed (see the discussion under "Contract length" below).

Gender

Gender is another variable for which there is no alternative. Studying the effect of gender on compensation requires the ability to distinguish men from women. Gender assumptions based on first names are accurate for about 90 percent of Anglo-Saxon names. Foreign names, initials, and gender-neutral first names require verification. Without gender information, a gender analysis is not possible.

Race

Race is also an essential variable. You may ultimately conclude that there are too few minority faculty members at your institution to warrant an analysis of the effect of race on compensation. But you will need to identify the racial minorities to make this determination.

Remember that if minority males are underpaid, their inclusion will lower the average male salary relative to the average white-male salary. Preliminary analysis can determine whether or not bias is evident in the salaries of minority men. If it is, exclude minority men from the male reference category. The comparator category for determining whether or not bias exists is properly that of white males.

Legally, certain racial groups are specified as protected classes, but some may be more important at your institution than others. If your campus is in the vicinity of a Native American reservation or near the Mexican border, for example, regional demography will influence your concerns.

Our SUNY data included a racial category called "alien," which is made up of noncitizens who are working legally in this country. Noncitizens are not a legally protected class, but many faculty members who are not citizens, most notably Asians, Latinos, and Africans, may be in a protected class. If your data include a noncitizen category, give this group special attention to determine which of its members are in racial categories that may experience discrimination.

Highest degree

Highest degree earned is an important variable. An individual's educational attainment determines many career markers, such as initial rank, initial salary, and promotions. Higher degrees garner more prestige and better pay than lower degrees, so you will want to have this variable in the model.

At two-year colleges, a higher proportion of faculty members will have a master's, bachelor's, or associate degree, or a vocational certificate, not a Ph.D. Make sure to review the distribution of degrees and code the relative distinctions at your institution.

At some institutions, the data for highest degree includes an ambiguous category with a label such as "professional degree." We have found that this label can apply to a range of degrees, including those for medicine, law, business, social work, mechanical drawing, and pottery. Categorizing professional degrees may require reviewing the data on each individual in order to ascertain the level of educational attainment (doctorate, master's, or below master's) and coding it to the appropriate category.

Most data on highest degree attained fail to note where the degree was earned. Many campuses do, however, have information on the institutions that granted degrees to faculty in a separate automated file, because it is included in the campus catalog. The status or type of institution can be coded (doctoral, comprehensive, baccalaureate, two year). If you can acquire it, such detailed information could prove valuable, particularly for initial rank or initial salary analysis. Once an individual has been working in a field for some time, the origin of a degree tends to become less important.

An alternative variable for highest degree attained is highest degree earned at time of initial hire. Some institutions collect degree data when a faculty member is appointed, but never update this information. In our SUNY study, about 20 percent of academics received their highest degrees (usually a Ph.D.) after being hired. If your information on highest degree is not current, try to update it by consulting the college catalog, deans or chairs, or the faculty members themselves.

Completion date for highest degree

Completion date for the highest degree is a time variable used to estimate experience. After receiving their highest degrees, faculty begin to acquire experience in their field. Those who finished their degrees many years ago are likely to be more experienced and more valuable to the institution than recent graduates.

Years since highest degree at time of hire

If, however, you measure experience by calculating the years since the highest degree was completed, that information will overlap with the variable for date of

initial hire (discussed below). To make the data less redundant, create a variable for previous experience by counting the years since highest degree at time of hire. Those who received their degrees before being hired will have a positive value for the variable, while those who completed their degrees after being hired will have a negative value.

Although years since degree at time of hire is indeed a possible proxy for previous experience, its use involves an assumption that once a degree is received, the recipient begins working in her field and continues in it until being appointed at your institution. But the continuity of employment after degree completion may vary by gender. Women may be more likely to take time from employment for family work. Using data sets from the twelve SUNY campuses we studied, we compared the number of years since degree with the variable for experience prior to hire. Women had fewer years of experience recorded for their years since degree than men did. As we discuss below, we have no way of knowing how carefully this information was gathered at the various campuses. Our data are, however, consistent with the hypothesis that women are more likely than men to take time out from continuous appointments. Thus, using years since degree prior to hire as a proxy for previous experience may credit women with too many years in their field.

Age is sometimes used as a substitute for years since degree at time of hire, but we do not recommend it. People tend to proceed from kindergarten through grade twelve and sometimes even through undergraduate college with little time out. But education beyond the bachelor's degree is less age-specific. Moreover, the age at which degrees are received may vary by gender and race, with women and minorities receiving degrees at older ages on average.

We suggest proxies (alternative variables) to measure previous experience because most campuses do not include information on relevant experience at the time of hire in their automated data sets. Even when the information is available, its quality varies depending on who collected and quantified it. It can easily be tarnished by gender or race bias. A man who was an administrative assistant in the business world, for example, may have his experience counted as relevant for a faculty position in the business department, yet a woman who gained similar work experience but had a traditionally female title, such as office manager, may have her experience discounted as "only clerical."

If you collect data on previous experience, we recommend that you create a list of acceptable positions and a reliable method for calculating the number of years of such experience. Although the SUNY data included a variable for relevant previous experience at time of hire, we suspect that the level of care taken in collecting and coding these data varied from campus to campus. Assessing work experience after the fact is complicated. You may have to consult paper files in the personnel office or look through vitae and resumes. This process can be time consuming; is it worth the time?

To answer that question, we performed a test on the data from our twelve SUNY campuses. We dropped the previous experience variable and reran our regression analyses for each of the campuses. The adjusted R^2 tended to decrease by less than three percentage points. Out of the twelve schools, it increased at only two campuses, and then only slightly. (See chapter 1, appendix A, or the glossary for discussion of R^2.) The effect on indicators for gender and race bias was mixed, with half of the schools showing less bias and half of the schools showing more. At only four institutions—two universities and one two-year and one four-year college—did the results indicate increases in bias in excess of $100, the largest increase being $500.

Our mixed findings led us to conclude that whether or not you take the time to collect detailed data on previous experience depends on two factors. First, if policy makers at your institution are concerned that your proxy variable for previous experience credits women with more than they actually have, you should collect the necessary data about actual previous experience.

Second, you should also do so if you are being pressured to include initial rank in your regression analysis. Our categorical modeling of this variable suggests substantial gender bias in assignments to initial rank. You will want to use the previous experience data you collect to assess whether initial rank is tainted. Chapter 6 discusses in detail why we believe it is inappropriate to include initial rank in regression analyses.

Date of hire

Date of initial hire in a faculty line at the institution under study is another measure of experience. It allows for calculation of important time variables, such as years at institution and years at institution prior to achievement of current rank. These time variables are critical to analyses of both rank and salary.

If you do not have date-of-hire information, you can approximate it if you know the number of years faculty members have been at your institution (subtract the number of years from the date of the database creation). There is a potential disadvantage to using this kind of calculation: a person entering data in January 2002, for example, may credit everyone starting in 2001 with a

year of experience regardless of whether an individual was hired in January or September 2001. Using the date of hire to calculate the variables associated with time at the institution is a more accurate approach.

What about part-time experience? Your analysis will probably include only full-time faculty because of the difficulties associated with maintaining accurate records for part-time faculty. But some of those who are now full-time faculty members may also have had part-time experience at this campus. If such experience were distributed equally among men and women, then all part-time histories could reasonably be disregarded in the salary analysis. At the twelve SUNY campuses we studied, we found that women were about four times as likely as men to have part-time experience. Not considering this experience in the analysis would tend, therefore, to make any salary disparity between men and women appear to be less than it would be when part-time history is taken into account.[1]

Of course, one year as a part-time faculty member should not be equated to one year of full-time service; part-time employment should be translated into a full-time equivalency. We took the conservative approach, treating each year of part-time service as equivalent to a third of a year of full-time experience. We did so by multiplying the part-time variable by 0.33. Thus, two years of part-time service equaled 0.66 years of full-time employment.

Do you need to go to all of this trouble? We examined our data and regression analyses with and without including part-time experience to answer that question. We found that part-time experience influences the regression results when either of two conditions exist: (1) when 20 percent or more of the faculty have part-time experience, or (2) when the average number of years of part-time experience is greater than three. If the part-time experience in your population does not meet either of those conditions, then you can probably drop it from your analyses. The difficulty is that you have to collect the data on years of part-time service to know if either of the two conditions exists on your campus. In the SUNY study we found that dropping part-time experience changed the results of the analyses at two four-year colleges and two two-year colleges. Dropping part-time service did not substantially change the results at the remaining eight schools.

Current rank

Current rank is the most hotly debated type of information included in regression analyses. It reflects the institution's assessment of a faculty member's performance and is therefore viewed by some as a proxy measure for performance. At most colleges and universities, however, the same people and processes that control salaries govern promotions to rank. If bias affects salaries at an institution, it probably also influences rank (see chapter 4). Whether or not you ultimately decide to use current rank in your analyses, you should collect rank data to assess its appropriate use and potential bias.

Because gender bias can affect current rank, it is a good idea to gather consistent, accurate information on the performance variables that determine rank. Information on these variables, which include publications, research grants, teaching, and community service, is difficult to collect and quantify. To create valid data on publications, for example, the value of books or monographs relative to journal articles would need to be coded according to how a discipline assesses them. Each journal article, in turn, would need to be weighted by the prestige of the periodical in which it appeared.[2] For grant funding, it would be best to consider both the amount of money going for overhead and the total award. Regarding teaching, should teaching load or teaching quality be included? If so, how should this information be collected?

It is unclear how these performance variables would affect study results. Because of the unwieldy size and diversity of the SUNY system, we made no attempt to collect performance data. As a result, we cannot assess the implications of substituting such data for current rank.

Date of current rank

Date of promotion to current rank is another measure of experience. Without such a variable, we have no way to distinguish between a professor who reached her current rank within the past year and one who has been in the rank for fifteen years. We would expect the fifteen-year veteran to have more experience and a higher salary.

If possible, collect dates for each promotion throughout your faculty members' careers. Having each promotion date simplifies any modifications that may be necessary if rank analyses show substantial bias. It also provides the detail necessary to conduct event history analyses of promotions (see chapter 4).

Contract length

Contract length is a variable that indicates whether a faculty member's annual salary is earned over an eleven- or twelve-month calendar year or a nine- or ten-month academic year. Perhaps the most direct way to adjust for differences in contract length is to use monthly salary rather than annual salary as the dependent variable. Your institution's automated database may

not, however, include monthly salary figures. Remember also that if salaries are earned in ten months but paid out over twelve, they are still ten-month salaries and must be adjusted accordingly. Another direct way to adjust for differences is to change the salaries of those working eleven or twelve months to reflect the amount of money they would have made if they worked only nine or ten months. Institutions that have a nine-month academic year usually adjust all calendar-year salaries to the fraction 9/11. The fraction used for adjustment varies relative to the paid vacation time awarded to those working the calendar year.[3] Some studies have included the contract-length variable in the analyses even when the dependent variable has been adjusted for academic- and calendar-year salary differences. That should not be done unless it can be demonstrated that the contract-length variable is not tainted (see chapter 4).

Discipline

Discipline is a control for market differences. Including a variable for discipline in the salary analyses ensures that pay differences related to discipline are not attributed to either race or gender bias. In essence, that means that we statistically measure the effects of gender and race by comparing faculty who are in the same discipline.

University studies have indicated that the greater the proportion of women in a discipline, the lower the average salary (Staub 1987; Bellas 1997, 1994). The relationship between the proportion of women in a discipline and the salaries its practitioners receive is relevant to the movement outside academia for equal pay for comparable worth. In this guidebook, however, we address only within-discipline salary bias (equal pay for equal work), not across-discipline salary bias (equal pay for comparable work).

Chapter 1 noted that college and university departments enjoy much greater autonomy than do equivalent divisions of many other public and private employers. The result is that the vagaries of the market and of individual administrators can affect salaries in higher education. In the public sector, entry-level professionals by and large make the same salary whether they begin their careers in the State Department or in the Department of Education. And entry-level K–12 teachers generally receive equal remuneration no matter what they teach.

In most institutions of higher education, however, faculty members are not treated in the same way. The salaries entry-level professors expect and receive vary widely according to their discipline and department. Faculty members on the same campus who teach the same number of classes and students for the same hours can be paid quite differently depending on what they teach. That is so even at two- and four-year undergraduate institutions that emphasize teaching over research.

To allow for salary differences across disciplines, we incorporate a measure for discipline into the analyses. That may sound simple enough, but this measure is not a single variable but many. Variables that are categorical (male and female, for example) as opposed to numerical (years in rank) must be entered as dummy variables. In simple terms, to specify ten academic disciplines, you would need nine discipline dummy variables and a tenth one as the default category. (See chapter 5 and appendix A for information on dummy variables.)

The number of disciplines that can be entered validly into a regression analysis depends on the number of individual faculty members in the database. If the number of variables entered comes close to the number of cases being examined, the regression model will not be valid. In research vernacular, such a situation is called "overspecifying the model." Generally, you need at least five cases (faculty members) for each independent or predictor variable. But that is a minimum; for robust findings, it is best to minimize the number of variables by maximizing the size of the discipline units used. Use only the discipline units that have a substantial impact on determining faculty salaries at your institution.

The Classification of Instructional Programs (CIP) codes developed by the U.S. National Center for Education Statistics combine departments into discipline categories that logically relate to both subject matter and market salaries. We have used CIP codes in many faculty salary studies and generally recommend their use.

Alternatively, you may choose to cluster departments into disciplines based on common administrative units. You may, for example, count departments in the same college or under the same dean as disciplines. A statistical procedure called cluster analysis can also be used to determine which departments are statistically similar based on a set of categories including salary but excluding gender and race.[4] Whichever approach you use to combine departments, it is important to eliminate or minimize clusters that contain fewer than five faculty members.

Variables for Verification of Data

In our analyses, we commonly use five categories of information (name, social security number, birth date or age, department, and initial rank) to verify individual data points that seem to be in error.

Name and social security number are useful identifiers when data problems are discovered. Without them,

you will be unable to verify or correct the data that appear to have errors. If confidentiality is important, a unique identifier for each individual can be substituted. Although the researcher need not know from whom the data are derived, someone will have to be able to trace each unique identifier to an individual in order to check on data errors and inconsistencies.

Information on birth date and age also help in screening for potential errors in the data. If, for example, variables for age and completion date for the highest degree indicate that an individual received a doctoral degree before age twenty, that individual's data should be verified. The use of birth date and other variables are discussed below ("Inconsistencies").

Department name is important to ensure that individuals are categorized in the correct discipline. It will also assist you in determining how to recode degree information for those with "professional" degrees.

Information on initial rank can help you check the accuracy of current rank, years at the institution, and years in current rank. In our database, we found many instances in which initial rank was the same as current rank and yet years at institution was not the same as years in current rank. Without initial rank data, we would not have been aware of these inconsistencies.

Cleaning of Data

Checking the accuracy of data is time-consuming drudge work and requires a tolerance for tedious details, but it is essential to the validity of your findings. Even if the data have been very carefully collected, you will probably find errors. The data-cleaning process involves finding and eliminating missing values, irregularities, and inconsistencies.

Missing values

Missing values are easy to spot but hard to eliminate. Gaps in the data usually come about because of difficulty in getting information about particular individuals. But lack of data about a faculty member effectively excludes that person from the analysis and may affect the validity of the results (N. Moore 1992). If data on a variable are missing for many faculty members, either the information should be collected and added to the data file or the variable should be excluded from the analysis.

Irregularities

Irregularities in the data are values that are impossible or illogical. If, for example, your data for educational attainment has eleven codes (1 = less than grade 7; 11 = doctoral degree), any code other than 1 through 11 amounts to an irregularity that should be corrected. If you have a collective bargaining agreement that specifies minimum salaries for each rank, use the minimums as screening criteria to flag all those who are below them. A below-minimum salary may result from a data-entry error, or it may signify a part-time faculty member mistakenly counted as a full-time faculty member—or it may call attention to a full-time faculty member who is being underpaid and should receive a salary adjustment to reach the minimum. Printing out all of the data on a case that appears to have an irregularity may allow you to make sense of it based on the person's total profile.

Inconsistencies

Although eliminating inconsistencies is more time consuming than locating missing data or removing irregularities, it is just as important. Locating inconsistencies involves comparing the data on two or more variables. The criteria for screening out inconsistencies should be decided by the members of the research team. Here are some of the criteria we have used to clean data sets.

Initial rank versus current rank. Note any cases in which a person's current rank is lower than the initial rank. A faculty member may have moved from assistant professor to a non-tenure-track rank such as lecturer to avoid the tenure clock. But it is unlikely that a faculty member who was initially an associate professor would become an assistant professor.[5] Although we have encountered cases of faculty being downgraded, apparent downgrading usually indicates an error.

Year of highest degree versus year of birth. Data indicating that a faculty member born in 1949 received a Ph.D. degree in 1960 are suspect. In screening for potential problems, we project that most people will be twenty-one before completing a bachelor's degree, twenty-three before finishing a master's degree, and twenty-five before earning a doctorate.

Total years of experience versus year of birth. If you have several different measures of experience (years in rank, years at the institution, years since highest degree at time of hire), you can create an estimate for total years of experience and compare it to the age variable. How young is too young to have acquired relevant experience? We screen any case in which a person seems to have acquired relevant faculty experience by age twenty-two.

Assistant professor rank versus years in current rank. According to the tenure system, assistant professors must achieve tenure, which is usually accompanied by promotion, before their fifth, sixth, or, most commonly, seventh year in rank. If they fail to do so, their appointment is expected to terminate. So if your data show tenure-track assistant professors who have been at the institution for more than seven years and who were not

appointed initially in a non-tenure-track rank, you probably have errors. Look for an error in current rank or years at the institution. Some assistant professors, albeit a small minority, receive tenure without promotion and remain in the rank indefinitely. You can use information on tenure to screen for this possibility.

Years in current rank versus years at the institution. Logically, years in current rank should be the same as or less than years at the institution. If, for example, your data show that a faculty member has been in the current rank for fifteen years but at the institution for only twelve, you need to investigate the three-year discrepancy. Such inconsistencies sometimes arise as a result of leaves of absence, but they should be checked.

Current rank and initial rank versus time in current rank and time at the institution. When initial rank is the same as current rank, years in current rank should be the same as years at the institution. But we have found individuals who had more years in current rank than years at the institution and vice versa. Some of these inconsistencies can be technically correct but need clarification in order to judge the probable impact on salaries. We found, for example, a faculty woman who was appointed as an assistant professor in 1976, became a nonteaching professional (academic counselor) in 1981, and was rehired as an assistant professor in 1989. Her reported information was initial rank: assistant professor; current rank: assistant professor; years at the institution: fifteen; and years in current rank: two. We recoded her years in current rank to reflect the time she spent as an assistant professor (1976 to 1981 and 1989 to 1991, for a total of seven) and her years at the institution to those when she was in a faculty rank, which was also seven years.

When initial rank is not the same as current rank, years in current rank should be less than years at the institution. Check also for instances in which faculty members have moved more than one promotional step, for example, from assistant to full professor, in only one or two years. We examine the data for anyone advancing more than one rank in five or fewer years and find that data errors account for most of these apparent moves. But there are indeed fast-track faculty who rise two or more ranks in five or fewer years.

Your goal is to rectify any gender- or race-based salary inequities that exist at your institution. To accomplish this goal, you and everyone involved must have confidence in the accuracy of the data collected. Ensuring accuracy is therefore necessary. When you have selected the right variables and gathered complete and accurate data, the results of your study will be seen as a valid basis on which to assess bias in salaries and determine any adjustments needed to correct it.

Notes

1. At the SUNY institutions we studied, we found that including part-time experience added complexity to collecting, cleaning, and coding the experience variable for years at the institution. Our data consisted only of total years of part-time experience and total years of full-time experience, leaving us to wonder whether the part-time experience occurred during the current rank or during the years prior to current rank, and whether date of hire referred to the date a person started working part or full time. These questions plagued us when there were data inconsistencies between years in current rank, years at the institution, and date of hire.

2. Toutkoushian (1994b) has made interesting use of citations as a measure of performance and found that considerable unexplained salary differences still existed between men and women. This finding may be consistent with Valian's (1998) observation that while women publish less, their publications have a higher rate of citation.

3. In our original study of the twenty-nine SUNY campuses, we did not deal with differences in contract length by adjusting the dependent variable for salary because of concerns about differences in the academic year across the campuses under examination. Instead, we entered the contract length variable in the analyses. Thus we held constant the differences in months worked with a control variable. We do not necessarily recommend this approach. Subsequent to completing the SUNY analyses, we have become aware of a potential for bias in the disproportionate awarding of calendar-year contracts, and the administrative responsibilities often tied to them, to white men (see "Problem 4" in chapter 6).

4. In chapter 5 we indicate how creating a continuous variable for discipline using average salaries circumvents the regression analysis and why such a variable should not be used.

5. Some institutions have non-tenure-track or visiting scholar senior ranks. They are common at professional schools that run clinics, but they also exist elsewhere. At these institutions, the inconsistency of having moved from a higher to a lower rank may be related to moving from the nontenure track to the tenure track.

Our study revealed that one of the major sources of differential is the rate at which women are promoted—or not promoted—and the longer period of time they remain in rank. It's not easy to pursue equity; it can be a lonely and even debilitating experience. Yet I have found that working together provides more than just better salaries. The sense of solidarity and empowerment is tremendously rewarding. I firmly believe that those of us who have been able to make gains must use our experience to help other women faculty achieve equity.

—Wendy Wassyng Roworth
Professor of Art History and Women's Studies
University of Rhode Island

FOUR
Gender and Race Bias in Current Rank

By Lois Haignere and Bonnie Eisenberg

Unlike the rest of this guidebook, this chapter does not focus on assessing bias in salaries. Instead, it explains methods for estimating whether or not bias affects the awarding of rank. Making such a determination is important because of disagreement over using current rank as a predictor variable. Warning: this chapter contains technical terms. You may need to consult the glossary for definitions.

Some argue that the current rank variable is potentially tainted and that its inclusion in the regression model will underestimate the gender or race bias in faculty salaries (Scott 1977; Allen 1984; Ransom and Megdal 1989). The people and processes involved in setting salaries at an institution probably influence rank and tenure decisions. Thus, if salary patterns are biased, promotion patterns will probably be so as well. Research shows that women and minorities disproportionately hold lower ranks and that women spend more years in lower ranks before receiving promotions (Gray 1993, 1990, 1988, appendix H). Little wonder that including variables for current rank and tenure in studies of salary equity has led to a major debate in the literature.

Most studies, however, do include current rank as an independent variable, usually without apology or justification (Schrank 1988; Gray 1991). Those who offer justification argue that salary discrimination should be defined as unequal pay for equal work. Two people in the same job level or rank have equal work; any discrimination in the awarding of rank is irrelevant to a study of salary equity. Nelle Moore adds that current rank is a strong determiner of salary and the best available proxy for productivity measures such as research, publications, and teaching.[1] In "Using Regression to Study Faculty Salaries" (1991), mathematics professor Mary Gray summarizes the predicament over whether or not to include current rank:

> The dilemma is that if rank is not used in a regression model of salaries, differences perceived to be based on gender may be just the result of differences in rank distribution between men and women. On the other hand, the differences in rank distribution may be the result of discrimination . . . resulting from placement of equally qualified men and women in different ranks.

Excluding current rank from the analyses can overestimate bias; if it is tainted, including it can underestimate bias. To illustrate the effect of a tainted variable, we ran analyses on a SUNY data set with and without a variable for initial rank, which statistical analysis had shown to be substantially gender biased. We included current rank in these analyses. When the tainted variable for initial rank was used, it masked about a third of the gender bias otherwise shown. The methods described in this chapter do not solve the "taint problem," but they can help you find out where it may exist and approximate its magnitude.

Details and Options

The level of detail available in your data will affect both your ability to test for bias in rank assignments and the statistical methods you select. On the low end of the scale, you may know only the proportion of men and women faculty at each rank. At the high end, you may have data for each faculty member indicating each career transition, including when he or she earned the highest degree, was hired, and moved from initial to each subsequent rank.

The level of detail available for current rank analyses at most colleges and universities falls between the low and the high end. Automated filing systems commonly

include data for each faculty member on the date of hire, the highest degree, the year of highest degree, current rank, initial rank, gender, and race. With this level of information, categorical modeling methods can be used to estimate the existence of gender bias across the categories of current rank.

If you are fortunate enough to have complete details concerning the career histories of your population, you can employ event history analysis to ascertain the proportion of men and women reaching a certain promotion at a specific point in their career.

Later in this chapter, we will discuss how to use categorical modeling and event history analysis to assess bias in rank. First we will illustrate how frequency tables can help you discover potential bias in rank assignments if you do not have the detailed data necessary to do categorical modeling or event history analysis.

Frequency Tables

If you know the current rank and the gender or race category of the faculty members in your population, you can create distribution tables that will give you a sense of the proportional distribution of men, women, and minority faculty throughout the ranks.[2] In table 4.1 we present a frequency distribution for a hypothetical institution to illustrate how this information can be displayed so as to show whether or not women and minorities are represented proportionally throughout the ranks.

The percentages in the table cells reveal the distribution of the different gender and racial groups across the ranks. Since, for example, there are a total of five minority faculty members, the one minority lecturer in the first cell and row represents 20 percent of the minorities at this hypothetical institution. You can see a higher proportion of white males at the senior ranks compared with any other racial or gender group. A frequency table like this one can make visible the total lack or underrepresentation of a race or gender category, such as minority females in this example. Thus frequency tables may suggest recruitment and hiring problems as well as those related to rank placement and promotion.

Unfortunately, distribution tables do not control for other variables while looking at gender and rank. These tables cannot, for example, address concerns about the "pipeline." Studies of the distribution of men and women at U.S. and Canadian institutions of higher education have found a larger proportion of women at the lower ranks and a larger proportion of men at the higher ranks. But since a faculty member cannot get into higher ranks without first passing through the lower ranks, it may be that women and minorities have just not been in the faculty pipeline long enough to be proportionally represented at the senior ranks.

Categorical Modeling

Other "yes, but" issues besides the pipeline are sometimes offered to explain the lower ranking and pay of faculty women. Some people argue that women are at lower ranks because they are less likely to have a Ph.D. and have fewer years of experience. To do statistical testing of any of these arguments, you need the data we described above as commonly available in most automated filing systems: date of hire, highest degree attained, year of highest degree, current rank, initial rank, gender, and race. If you are conducting research on bias in salaries, you probably have individual-level information, which will allow you to proceed with a categorical modeling analysis for bias in rank.

Categorical measures, such as current rank, can be used as independent variables in a multiple-regression model, but they cannot be used as the dependent variable.[3] Linear multiple-regression statistical methods require the dependent variable to be a continuous variable, such as salary. Studying the categorical variable of current rank requires a different statistical method, multinomial logit modeling; we call this method categorical

Table 4.1

Frequency Distribution for Hypothetical Institution

Race and Gender Categories	Lecturer		Instructor		Assistant Prof.		Associate Prof.		Full Prof.		Row Totals
	RANK CATEGORIES										
	#	%	#	%	#	%	#	%	#	%	#
All Minorities	1	20	0	0	3	60	0	0	1	20	5
All Whites	2	4	9	9	30	29	34	33	27	26	102
All Females	0	0	6	23	7	27	7	27	6	23	26
All Males	3	3.5	3	3.5	26	33	27	33	22	27	81
White Males	2	2	3	4	23	30	27	36	21	28	76
Minority Females	0	0	0	0	0	0	0	0	0	0	0
Rank Totals	3	5	9	8	33	31	34	32	28	24	107

Table 4.2

Categorical Modeling: A Pairwise Comparison for Gender Bias

The Odds Ratio

$$\frac{\text{Number of females in rank A}}{\text{Number of females in the next higher rank}} \quad \text{compared to} \quad \frac{\text{Number of males in rank A}}{\text{Number of males in the next higher rank}}$$

modeling for simplicity (appendix F includes technical information on categorical modeling).

Like regression modeling, categorical modeling has a dependent variable and predictor variables and can control other predictor variables while it tests the effect of gender or race. Years of experience and educational attainment are held constant by the categorical modeling procedure while it examines whether women and minorities are found in the same proportion as white men across ranks.

Categorical modeling cannot handle many predictor variables. This limitation arises from the lower range of variability in faculty ranks compared with salary. One hundred faculty members, for example, may have a hundred different salaries but only five or six different ranks (even if we count the less populated ranks of lecturer, instructor, and distinguished professor).

As a rule, reduce the number of variables in a categorical model whenever possible. Instead of including each gender and race category (white male, white female, African American male, African American female, Asian male, Asian female, Latino male, Latino female, and so on), try reducing these categories to white, minority, male, female, and an interaction term (see glossary) for minority female.

The results of categorical modeling are odds ratios, where the odds of women (or minorities) being in one rank rather than another are compared with the same odds for men (or whites). Table 4.2 illustrates the odds ratio, which can also be thought of as a pairwise comparison of adjacent ranks. Categorical modeling is actually more complicated than calculating these simple ratios, because it controls the independent variables, such as previous experience, educational attainment, and years at the institution. Also, to produce the best estimates, the model uses all available information, even for those ranks not involved directly in the particular comparison under review. (All the data for instructors, lecturers, and full professors are used to create estimates comparing assistant to associate professor, for example.)

Constraints

If your campus has a small faculty or few women and minorities, can you test for bias in current rank? The answer depends on three factors: (1) the number of faculty in relation to the number of predictor variables included in the model; (2) the number of faculty in each rank; and (3) the presence of zero cells.

Number of predictor variables. With ten predictor variables, we were able to conduct categorical modeling at all but our smallest SUNY school. At this campus, which had only ninety-nine faculty members, categorical modeling worked well with five predictor variables. Schools with fewer than a hundred faculty may have to limit the number of predictor variables or forgo categorical modeling.

Number of individuals in a rank. If a rank included in the study has fewer than three people, categorical modeling will probably not work. In such a situation, you should either combine that rank with the next closest one, or drop it from the analysis. If you choose to combine ranks, the results become more complicated to interpret. If, for example, instructors are combined with assistant professors to create a new category, the results will estimate the bias affecting promotions from instructor or assistant professor to associate professor. But how likely is it that an instructor will be promoted directly to an associate professor?

Choosing to drop a rank from the analysis may have political ramifications. In our experience, the only ranks with fewer than three incumbents have been the lowest (lecturers or instructors) and the highest (leading, distinguished, or named professors). If the low rank consists mainly of females and minorities while the high rank consists mainly of white males, you may mask bias by dropping either.

Zero cells. A zero cell is a cell with no one in it. If, for example, your school has no minority associate professors, it has a zero cell for minority associate professors. You would therefore be unable to use categorical modeling to test for race bias between associate professors and other ranks.

There is a hidden danger in zero cells. Any empty rank (zero cell) that is included in the model will be dropped from the analysis. If you review the results with the impression that the rank was included, you may misinterpret the findings. So, before starting any modeling, identify the ranks in which faculty are

found and drop all others. Among our twelve SUNY test campuses, categorical modeling worked for all but the smallest school (the one with ninety-nine faculty members) as long as there were no zero cells.

Variables for categorical modeling

Given that categorical models are sensitive to zero cells, we recommend that you include only those variables that are the best predictors of current rank. We also advise you to combine levels of categorical variables wherever possible to reduce the number of zero cells. Your population, for example, may have eight types of terminal degrees, ranging from a high school diploma to a Ph.D. But if most faculty members have Ph.D.'s, the eight levels of educational attainment can be reduced to two: Ph.D. and non-Ph.D. Below we summarize the variables we have included in our categorical models, and we discuss how the number of variables can be reduced to a minimum without losing necessary information.

Response variable. The dependent variable, known as the response variable in categorical modeling, is the variable you are studying. This chapter focuses exclusively on the response variable for current rank.

Keeping in mind the problem of zero cells, include only those ranks that have individuals in them. Avoid combining ranks unless there are fewer than three people in a rank. For the twelve SUNY schools in our sample, we combined ranks only after modeling failed owing to the presence of fewer than three individuals in a rank. At one university and one college, we combined two instructors with assistant professors. At two colleges, we combined one leading professor with full professors.

For our analyses of current rank, we included predictor variables for gender, race, highest degree attained, years since highest degree at time of hire, and years at institution. We also used quadratic terms to control for curvilinearity (see glossary and chapter 5 for discussions of quadratic terms and curvilinearity). Keep in mind the need to limit the number of variables, and include only those that most influence achieving rank at your institution.

Gender and race. As we noted above, we did not include every possible gender and race combination in our model, because we wanted to avoid zero cells. Instead, we analyzed gender and race, white faculty versus all other race categories, and male versus female. We included an interaction term for race and gender to provide information on minority women.[4] Combining all minorities may mask bias when there is substantial bias against only one minority type. Despite this potential problem, we suggest combining categories to avoid zero cells and to produce better estimates by creating larger pools of data.

Highest degree. Although our SUNY sample had nine levels of educational attainment, we found that 90 percent or more of the faculty at SUNY universities and four-year colleges had either a master's or doctoral degree. As a result, we reduced the variable to two levels: Ph.D. and non-Ph.D. At SUNY two-year colleges, most faculty had a master's degree as their highest level of attainment. Thus we used two categories: M.A. and above, and below M.A. We recommend that you examine the distribution of highest degrees at your institution and limit your variable for educational attainment to two or three categories.

Years since highest degree at time of hire. This variable can serve as a proxy for the amount of experience a person acquired before coming to the institution. It is important to give proper credit to those hired with more experience. (See the discussion of this variable and other variables for previous experience in chapter 3.)

Years at institution. We chose years at institution rather than years in current rank as a predictor variable because years in current rank cannot logically predict that current rank; the number of years a faculty member has been an associate professor is not a factor in that person getting to be an associate professor in the first place.

Quadratic terms. We suggest including a quadratic term for each variable that is a time measure (years at institution or years since highest degree at time of hire) to control for curvilinearity. A quadratic term is created by multiplying a variable by itself. Chapter 5 discusses the problem of curvilinearity and how quadratics remedy the situation. Be aware that although quadratic terms solve one problem, they cause another—redundancy. For a discussion of redundancy and solutions for it, see appendix G.

Interpretation of results

Table 4.3 shows results from a categorical modeling analysis of current rank at one of our SUNY test campuses in which all of the predictor variables listed above were held constant. For the promotional step from lecturer to assistant professor, the odds ratio for the variable female is 2.96. That means that women are 2.96 times more likely than men to be lecturers rather than assistant professors. For the same step, the odds ratio for the variable minority is 3.69: minorities are 3.69 times more likely than whites to be lecturers rather than assistant professors.

Table 4.3 shows little or no bias in the minority to white category at the two higher-rank transitions. The odds ratios for the assistant to associate professor step

Table 4.3
Example of Categorical-Modeling Analysis of Rank

	Lecturer to Assistant Professor		Assistant to Associate Professor		Associate to Full Professor	
	Odds Ratio	p-Value[1]	Odds Ratio	p-Value[1]	Odds Ratio	p-Value[1]
Female to Male	2.96	0.0741	2.62	0.0006	5.21	0.0001
Minority to White	3.69	0.1505	1.08	0.1245	0.82	0.6598

1. The p-Value is a measure of statistical significance. For details, see the entry for statistical significance in the glossary and "Problem 5" subheading in chapter 6.

(1.08) and for the associate to full professor step (.82) are close to 1. An odds ratio of 1 indicates that the two categories being compared are equally likely to be at either rank. Thus we can safely conclude that the categorical modeling results indicate little race bias in promotions from assistant to associate and associate to full professor.[5]

But what constitutes substantial bias? Unfortunately, there are no guidelines for assessing that. It would seem that the odds ratio in the female to male category for associate to full professor (5.21) indicates substantial bias. It shows that male faculty are five times more likely than comparably qualified female faculty to be full professors rather than associate professors. Odds ratios of this magnitude were rare in our SUNY analyses and seen mostly at the associate to full professor level. In our opinion, they indicate bias.

But what about the 2.62 odds ratio in the female to male category for the assistant to associate professor level? Does the finding that female faculty are two and a half times less likely than comparably qualified male faculty to be at the associate rather than the assistant rank indicate substantial bias? Our judgment would be that it does. We suggest that campus policy makers try to agree on the magnitude of odds ratios that will indicate bias before they conduct analyses.

Model Validity

It is important to know how well the predictor variables account for the differences in the dependent variable. If we substituted less valid information, such as number of siblings for years at institution, or marital status instead of highest degree, we would still get an odds ratio. But no matter what its magnitude, that odds ratio would not tell us much, because the measures we used would not account for variation in rank levels. What is needed is a measure that indicates the amount of the variation in the dependent or response variable accounted for by the independent or predictor variables. For multiple-regression analysis, this measure is the adjusted R^2. For categorical modeling, the relative quality of the model is indicated by the R-square sub-L (R^2_L). Appendix F explains this measure further.

Summary of steps

1. Condense any race or minority categories to the smallest number feasible.
2. Create a frequency table of gender and race categories by current rank to identify any zero cells and to provide a general understanding of the distribution of men and women throughout the ranks.
3. Where no zero cells appear in pairwise adjacent rank, proceed with categorical modeling using the selected predictor variables to produce odds ratios for each of the adjacent rank comparisons.
4. Calculate an R^2_L model statistic to estimate the amount of the variation in the dependent or response variable accounted for by the independent or predictor variables.
5. Consolidate the results in a table like table 4.3.

Limitations of Categorical Modeling

The categorical modeling analyses described above do not test for bias in promotion over time. They can only detect bias in current rank assignments (not promotion) at a single point in time, that is, the point at which the data are collected. They assess the relative odds of being in one rank or the other at that moment, and they answer the question: are white men with specific educational attainment and years of experience more likely to be at a higher rank than women (minorities) with the same educational attainment and years of experience?

These analyses rely only on demographic variables concerning education and experience, and the results should therefore be interpreted with caution. Still, they can indicate whether the variable current rank is tainted based on the predictor variables entered, and they can suggest where rank assignments create glass ceilings for women and minority faculty.

Event History Analysis

If you are fortunate enough to have data showing how men and women have progressed through the ranks over time, you may want to conduct event history analyses. Event history techniques can help you answer questions

such as: what is the probability of an assistant professor being promoted to an associate professor after four years of being in rank? How do these probabilities differ for white men and women and for minorities?

Like categorical modeling, event history analyses compare the odds of women and minorities to the odds of white males while controlling for years of experience and education. Data for every faculty member in the database are used, no matter when the person was eligible for the rank transition being studied. Event history's advantage is that it provides results that directly assess promotions rather than just rank assignments.

The difficulty is that the data needed to conduct the analyses are not commonly available. The automated databases of most colleges and universities maintain year of appointment information for only the most recent appointment, which is current rank. If that is the situation on your campus, you can do categorical modeling analyses but not event history analyses.

To conduct event history analyses, you need to know the year each faculty member received each promotion at the institution. The twelve SUNY campuses on which we tested the methods described in this guidebook did not have the promotion details needed to do event history analyses. Fortunately, we were able to test the impact of different event history approaches on a non-SUNY database. See appendix H for a case study conducted at Kent State University using event history analyses. Appendix H also provides technical information and discusses the impact of common expectations concerning the timing of promotion.

Population and time frame

Because faculty members that are hired at the senior ranks have been promoted elsewhere or have negotiated rank in conjunction with hiring, we recommend that you restrict your study population to those hired at the junior ranks of lecturer, instructor, or assistant professor. These faculty are more fully subject to the promotional mechanism of your institution.

Likewise, it makes sense to restrict the time frame you examine to the period when a faculty member is eligible for the promotion being studied. In the event history vernacular, the analysis is best confined to the time during which the individual is at risk of experiencing the event.

Still, we know that events that happen before a person is eligible for promotion can affect how soon she or he gets promoted. Having completed postdoctoral work or having been an instructor, for example, may influence how quickly an assistant professor is promoted to associate professor. It is therefore tempting to include these years in the dependent variable when we examine promotion to associate professor.[6] But including years as an instructor in the dependent variable years in rank for an analysis of promotion from assistant to associate professor would be to consider years when the possibility of promotion to associate professor is remote or impossible. Confining your analysis to the years when faculty are truly eligible to experience a promotion does not mean that you have to ignore prior career investments. Predictor variables such as years of experience at time of hire and years at institution prior to current rank allow the event history analysis to control for the impact of such periods. In this way, postdoctoral work or experience gained as an instructor can be considered in predicting the odds of being promoted from assistant to associate professor. But only the years when faculty are eligible to be promoted to associate professor are included in the dependent variable.

Person-year unit of analysis

The unit of analysis for most of the methods discussed in this guidebook is the individual person. Each record or row of data corresponds to an individual faculty member and includes his or her salary, current rank, years in rank, gender, race, discipline, and so on. When conducting event history analyses, however, the basic unit is a person-year, rather than a person. That means that a faculty member who has been at an institution for ten years would contribute ten records to an event history database, one record for each year at the institution. The data you obtain for gender and race analyses will probably be for persons, not person-years. One way to reorient your thinking to person-years is to physically construct a database in which you create a record for each person-year. The data for a faculty member who has been at your institution for ten years would be reentered so that the information would constitute ten records or rows of data. Each record would contain a code for each variable needed to test for bias in promotion rates. (See the list of predictor variables for categorical modeling given earlier in this chapter.)

Unfortunately, you cannot simply copy an individual's record and reenter it ten times. Although there are variables that do not change with time (race, gender, and years since degree at time of hire), other variables (years in rank and years at institution) change each year. Furthermore, some faculty members may not have received their Ph.D. degree until after they became assistant professors. If the data for an individual indicate that she received a Ph.D. in 1978 but became an assistant professor in 1975, her first three person-year records will have non-Ph.D. as her

educational attainment. For the fourth person-year, the highest degree code will change to Ph.D.

In addition to creating the predictor variables for every year an individual contributes to the database, you will need to create a transition variable. All of the faculty members who receive a promotion (or transition to a higher rank) in a given year score 1 on the transition variable. All of the faculty members who do not receive a promotion in that same year score 0 on the transition variable. An individual hired in 1992 as an assistant professor and promoted in 1998 would have a transition variable equal to 0 for his person-years from 1992 to 1997. In 1998 the code would change to 1.

Probability tables

If you have complete career history data, you can construct promotion probability tables to assess the odds of being promoted at each year in rank and the proportion of unpromoted faculty who have reached a certain year in a specific rank. The figures for the cells in the probability table are calculated using an equation based on logistic or Cox regression analyses that include only years in rank, the gender or race category you are examining, and the transition variable for the promotion. Probability tables, like frequency tables, do not control for other variables like Ph.D. or years of experience at time of hire.

The beauty of probability tables is that for each year in rank examined, the data for all faculty who have ever reached that year in rank are considered, whether they reached it fifteen years ago or last year. All faculty members who have reached their sixteenth year as an associate professor, for example, can be studied to compare the difference in the proportion of unpromoted men and women after sixteen years as an associate professor.

Table 4.4 shows the relative probability at one institution of male and female faculty members being promoted to full professor during each year they serve as an associate professor. The annual promotion rate indicates the probability of a faculty member experiencing a promotion from associate to full professor at each year. At year 9 as an associate professor, for example, women faculty have a 4 percent chance (.041) of being promoted to full professor, while male faculty members have about double the chance (8 percent or .078) in the same year.

Table 4.4
Analysis of Promotion-Probability Table: Transition from Associate to Full Professor

	Females		Males	
Year in Rank	Annual Promotion Rate	Cumulative Proportion Unpromoted	Annual Promotion Rate	Cumulative Proportion Unpromoted
1	0.007	0.989	0.013	0.979
2	0.011	0.979	0.021	0.959
3	0.016	0.963	0.031	0.928
4	0.022	0.941	0.043	0.885
5	0.029	0.912	0.055	0.830
6	0.129	0.783	0.066	0.764
7	0.039	0.743	0.075	0.689
8	0.042	0.701	0.079	0.611
9	0.041	0.660	0.078	0.533
10	0.038	0.622	0.072	0.460
11	0.033	0.589	0.063	0.397
12	0.027	0.562	0.051	0.346
13	0.020	0.542	0.039	0.307
14	0.014	0.527	0.028	0.279
15	0.009	0.518	0.018	0.261
16	0.006	0.512	0.011	0.250
17	0.003	0.509	0.006	0.243
18	0.002	0.507	0.003	0.240
19	0.001	0.506	0.002	0.238
20	0.000	0.506	0.001	0.237

Note: Event history statistical methods are commonly used to examine certain hazards such as illness and death. Most event history studies therefore report "hazard rates" and "proportions surviving." For applicability to promotions, we have renamed the hazard rate the "promotion rate" and the proportion surviving the "proportion unpromoted."

The cumulative proportion unpromoted reports on faculty remaining in the lower rank of associate professor.[7] By year twenty, the cumulative proportion of unpromoted females is .51 (rounded from .506), while the cumulative proportion of unpromoted males is .24 (rounded from .237). After twenty years at the associate professor rank, then, more than half of female faculty members have not been promoted to full professor. By comparison, less than one-quarter of male faculty members have not been promoted to full professor after twenty years at the associate rank. See appendix H for additional information on promotion probability tables.

Logistic and Cox regression

If you have created a database in which person-year is the unit of analysis and each record is one year of a

faculty member's career, you can analyze this data using a basic logistic regression procedure such as that of Proc Logistic in the SAS statistical analysis package. As long as you have correctly coded variables such as years in rank, years at institution, and perhaps highest degree, and changed them accordingly with each person-year, you need not worry about time-dependent variables.

Table 4.5
Event History of Promotion Results

	Assistant to Associate Professor			Associate to Full Professor		
	Coefficent	Odds Ratio	p-Value	Coefficent	Odds Ratio	p-Value
Female to Male	-0.2197	0.803	0.309	-0.6824	0.505	0.182

Note: The control variables are years at institution, years of experience at time of hire, highest degree, and the quadratic terms for the "years" variables. p-Value refers to statistical significance (see the glossary entry for statistical significance and "Problem 5" subheading in chapter 6).

If your database unit of analysis is the individual (not person-years), the analysis can be accomplished using a statistical technique called Cox regression.[8] Cox regression can reconstruct a database with individual records so as to analyze person-years. But you will need to have someone on hand who understands and can program the Cox regression procedure. There are complexities. One is that the procedure needs to correctly interpret cases that are never promoted, called censored cases. Another is that time-dependent variables require special programming for proper entry into the Cox regression calculations.

The computer output for logistic and Cox regression provides a coefficient for the gender and race categories you use. The exponential value (see glossary) of this coefficient is the relative odds of those in one such category (e.g., women) being promoted compared with the odds for those in the default category (e.g., men). Table 4.5 illustrates gender results from a logistic regression analysis including the coefficient, odds ratio, and statistical significance (p-value) values for the variable female.

The coefficients resulting from logistic regression are actually log odds. For ease of interpretation, the log odds are converted to antilogs. In table 4.5, for example, the assistant to associate professor transition coefficient is -0.2197. The exponential of this number yields an odds ratio of 0.803. This odds ratio tells us that female faculty are, on average, 80 percent as likely as comparably experienced male faculty to be promoted to the rank of associate professor in any one year. For the promotion from associate to full professor, the odds ratio indicates that female faculty are, on average, about 51 percent as likely as male faculty with comparable experience and education to be promoted to full professor in any one year.

Summary of steps

1. Determine if you have the level of detail needed to conduct event history analyses. Do you have complete career history information for the individuals in the database? Does your database contain the year each faculty member received each promotion at the institution?
2. Restrict your study population to those fully subject to the promotional mechanism of your institution. At most colleges and universities, that means including only those hired at the junior ranks of lecturer, instructor, or assistant professor.
3. Restrict the time frame being studied, that is, the dependent variable years in rank, to the period when a faculty member is eligible for the promotion being studied. Use the independent variables to control for influences before that period, such as time spent as a lecturer or in postdoctoral study.
4. Transform the individual-level database to a person-year database. The change can be accomplished either by constructing a database in which you create a record for each person-year or by using Cox regression.
5. Construct promotion probability tables that allow you to view differences that may exist in the odds of being promoted at each year of academic pursuit. Such tables indicate the proportion of unpromoted faculty who have reached a certain year in a specific rank.
6. Conduct logistic or Cox regression computer analyses for each promotion you are studying. These analyses estimate the relative odds of promotion for the gender and race categories you enter.

Next Steps

So far, we have described how you can use frequency tables, categorical modeling, and event history analyses to diagnose gender or race bias in current rank. Now we will consider your alternatives for analyzing bias in faculty salaries relative to your findings concerning discrimination in current rank.

Findings of little bias

Four of the twelve SUNY schools we studied using categorical modeling showed little bias in current rank.

Indeed, gender and racial bias in rank assignment is not universal, and some institutions have developed relatively equitable promotion processes. If your results suggest no bias in current rank, then you can include it in your regression analyses of salaries without fear of recrimination.

You may, however, want to analyze your institution's initial rank assignments. Three of the four SUNY campuses with relatively little bias in current rank displayed notable evidence of bias in the assignment of initial rank.

Findings of notable bias

If you have found bias in current rank and want to proceed with multiple-regression analyses of salaries, what are your options? You can omit current rank from the analyses and assume your results overestimate salary disparities based on gender or race. We do not, however, recommend this option. Current rank is widely seen as related to job level and as a legitimate determinant of salary. If you leave it out, many academics will dismiss the results of your salary analyses despite the care you have taken to demonstrate that current rank incorporates bias. But if you include the tainted variable in your analyses and still find gender or race bias in salaries, these results cannot be easily discounted. Although there may be evidence of bias in current rank assignment, current rank directly reflects the institution's formal recognition of individual status and performance.

Another option is to substitute performance data for current rank. If your current rank data are tainted, one solution would be to replace them with comprehensive data on publications, research, and teaching. The literature notes that it is difficult and time consuming to gather and score this information (Scott 1977; Gray 1990), so we do not suggest that you embark on collecting it unless someone has already automated a substantial proportion of the data. Even if you do have such data available, building appropriate measures may be complex. In measuring publications, for example, are you interested only in quantity or also quality?[9] How should you rate a newspaper article versus a book review versus a paper in a refereed journal versus a book? How can you assess teaching effectiveness?[10] Is it necessary to score the productivity variables differently for different disciplines?

Yet another option if you have found bias in current rank is to clean the variable up by excluding the biased categories. It may be that your analyses of current rank will indicate bias across only two or three ranks. If you remove the distinctions between these tainted ranks, you will remove the bias and the potential to mask bias in salaries. If, for example, you find women are substantially more likely to be instructors than assistant professors but are appropriately represented in all other ranks, you can create a new current rank category that combines instructors and assistant professors.[11]

Unfortunately, this approach has a major drawback. If most faculty members are in the ranks you combine, you will eliminate a substantial amount of current rank information. This situation may arise if you find bias between the associate and full professor ranks, for these two ranks often contain 70 percent or more of the faculty at a given institution. Combining the two would blur important job level distinctions for most of the faculty in your analyses.

Still another option is to include current rank and assume that the results underestimate bias. Although it is troubling to include a variable that can mask bias in salaries, we prefer this option over that of leaving current rank out of the regression analyses. The only compensation for doing so is that it justifies the assumption that any gender or race bias indicated by the regression analyses underestimates the amount of discrimination that actually exists.

Given the options we have described, is it worth conducting analyses for taint in rank? Yes, for three reasons. First, you may find that current rank is not biased. Second, if it is biased, you can choose from among the four options we have noted (or other options that you may design), and assess your multiple-regression results accordingly. Third, you may choose to use the results on bias in current rank to diagnose where glass ceilings may be blocking the promotion of women and minorities.

Among the twelve SUNY institutions we studied, we observed different patterns depending on institutional type. At two-year colleges, for example, a glass ceiling seems to exist for women at the associate professor rank, but once a woman breaks through and achieves the rank, she is just as likely as a man to be promoted to higher levels. At four-year colleges, women consistently have lower odds of getting into the rank of assistant professor than they have of getting into senior ranks once they have achieved the assistant professor rank. Women faculty at universities seem to continue to have substantially lower odds of reaching senior ranks even after obtaining the assistant professor rank. Having the odds ratios for the ranks at your institution may enable you to flag and change institutional processes that create barriers.

Notes

1. Studies that control directly for productivity still find gender bias in assignments of current rank (Downey et al. 1985; Long, Allison, and McGinnis 1993).

2. Note that for most institutions of higher education, these data are available through the AAUP's Annual Report on the Economic Status of the Profession.

3. There are other potentially tainted categorical variables. The most likely culprits are administrative positions and initial rank. As chapter 6 explains more fully, there are good reasons to exclude both of these variables from salary analysis. Tenure is also likely to be a tainted variable. If you have tenure data, you can conduct categorical modeling similar to the analysis described for current rank in this chapter.

4. The interaction term is created in the following way: all minorities are assigned a value of 1 for the dummy variable minority, and all females are assigned a value of 1 for the dummy variable female. The interaction term is obtained by multiplying the minority variable by the female variable and creating a dummy variable with a code of 1 for all minority females and a code of 0 for everyone else.

5. Such a conclusion, of course, assumes an adequate number of cases in all of the cells being compared, and no zero cells.

6. Unlike categorical modeling, event history analysis examines each rank transition separately. The chance of experiencing a transition from assistant to associate professor is a separate analysis from the chance of transition from associate to full professor.

7. To understand the relationship between the annual promotion rate and the cumulative proportion unpromoted, subtract the promotion rate for each year from the cumulative proportion unpromoted column of the previous year to get the cumulative proportion unpromoted for that year.

8. If you are using the SAS statistical package, the procedure that will run Cox regression is called Proc PHREG.

9. Schrank (1988) raises the question whether the decisions of journal referees and grant review committees incorporate gender bias.

10. Allen (1984) suggests that variables measuring teaching quality are probably unnecessary because it has never been suggested that women are less effective teachers than men.

11. A complication of this approach is that the years in current rank variable must be modified to reflect the time spent in the newly created combined rank, but the dates of promotion to previous ranks may not be available.

Currently, widespread gender discrimination in wages in universities throughout this country is supported and sustained by our tax dollars. To eliminate pay inequity, universities should conduct independent wage studies to diagnose gender bias. By correcting salary disparities shown in these studies, universities will affirm a commitment to equal pay in the workplace.

—Norma Sadler
Professor of Elementary Education and Specialized Studies
Boise State University

FIVE

Gender and Race Bias in Salaries

By Lois Haignere

This chapter takes up where chapters 1 and 2 left off on the use of multiple regression to assess salary equity. If you have skipped those chapters, you may want to read them now, particularly the conceptual overview. Better yet, look at the discussion of multiple regression in appendix A. Even readers familiar with multiple regression may find this appendix helpful in understanding the relative strengths and weaknesses of the different types of regression models. Again, many of the terms used in this chapter are defined in the glossary.

The strength of using multiple regression to study salary equity is that it sorts out how race and gender affect salaries as if all else—highest degree, years of experience, and discipline—were equal. As a result, multiple regression is commonly acknowledged to be the most useful statistical tool for studying salary equity, and there have been many multiple-regression reviews of faculty salaries.

But not all multiple-regression approaches are alike. The first section of this chapter discusses some ways of modifying predictor variables to improve their use in multiple-regression salary analyses. The second section describes three regression model approaches that are commonly reported in the literature and summarizes their strengths and weaknesses. The final section presents our findings and recommendations concerning the three regression approaches, based on their impact on the results of salary studies for four universities, four four-year colleges, and four two-year colleges.

Re-engineered Variables

Chapter 3 enumerates the variables that are desirable in a data set. Here we suggest ways of adapting some of these variables before entering them in multiple-regression analyses.

Redundancy in time variables

Experience and longevity variables all involve time, and overlap between time variables should be avoided. Years in current rank and years at the institution are often used as indicators of seniority and experience. Using both introduces substantial redundancy.[1] Therefore, we recommend redesigning years at the institution to be years prior to current rank, which would be the time between initial date of hire and appointment to the current rank. This eliminates the overlap with years in current rank.

Years since highest degree at time of hire and years of experience prior to hire are both measures of previous experience. We recommend using both, although they are somewhat redundant, because each has a disadvantage that is lessened by use of the other. Years since highest degree at time of hire is the more reliable of the two measures because this variable represents a set event that can be confirmed: graduation. However, some people may not acquire professional experience for every year since highest degree. Our data suggest that women are credited with less experience than men relative to their years since highest degree at time of hire. But estimates of relevant work experience prior to hire are subjective, and employees and personnel offices frequently disagree over this issue. (See the discussion of this variable in chapter 3.) Nevertheless, we believe a years of experience prior to hire variable should be used if you have it. The problem of overlap between these two variables is reduced because some people acquire previous experience prior to their year of highest degree, and some do not get their degrees until after they are hired.

If your campus has no automated data on years of experience prior to hire, should you collect them? To assess the importance of this variable, we ran regression

analyses on salary with and without the years of experience prior to hire variable. Without this variable, the results for eight of the twelve SUNY analyses did not change substantially either in terms of the R^2 or the bias indicated. For the remaining four schools—one two-year and one four-year college and two universities—the results did indicate an increase in both gender and race bias when experience prior to hire was omitted. The increase in the amount of bias shown at these four schools was not, however, substantial, the largest increase being $500. Whether this limited improvement in precision merits the time and work involved in collecting the relatively subjective data on years of experience prior to hire is a judgment call for policy makers. As noted in chapter 3, however, it is important to collect previous experience information if there is pressure to include initial rank in the regression analyses. Data on experience at the time of hire are important in statistically assessing whether or not initial rank is a "tainted" variable and, therefore, will mask bias in salaries.

Age is rarely used in salary-equity analyses, as long as better measures of experience and longevity are available. To the extent that women interrupt their careers more often—or for more years—than men, the use of age could overestimate gender differences in salaries. Moreover, our findings indicate that inclusion of the age variable does not add more information (raise the R^2) or change the estimates of gender and race bias. In light of these factors, and to minimize redundancy, we recommend leaving age out of the model.

Curvilinearity: The problem and some solutions

Multiple regression assumes that there is a linear relationship between the independent and dependent variables. According to this premise, for each year of additional experience, there is a corresponding unit of increase in salary. Unfortunately, salary does not always change in a truly linear way across time. In fact, there are opposing nonlinear pressures, particularly when a union represents the faculty. Percentage raises that are negotiated by the union cause salaries to grow at a faster rate as raises build on raises over time. This effect, when graphed, produces an accelerating curve (see figure 5.1). The other effect is that, as faculty acquire more years of experience, their salary growth may slow because they are reaching the end of the promotion ladder and the related raises (Gray 1991). (See figure 5.2.) If these two pressures offset each other, curvilinearity may be minimized so that *linear* multiple regression correctly measures the relationships between the time variables and salaries. If they do not offset each other, either type of curvilinearity can adversely affect the predictive power of multiple regression.

Figure 5.1

Accelerating Curve

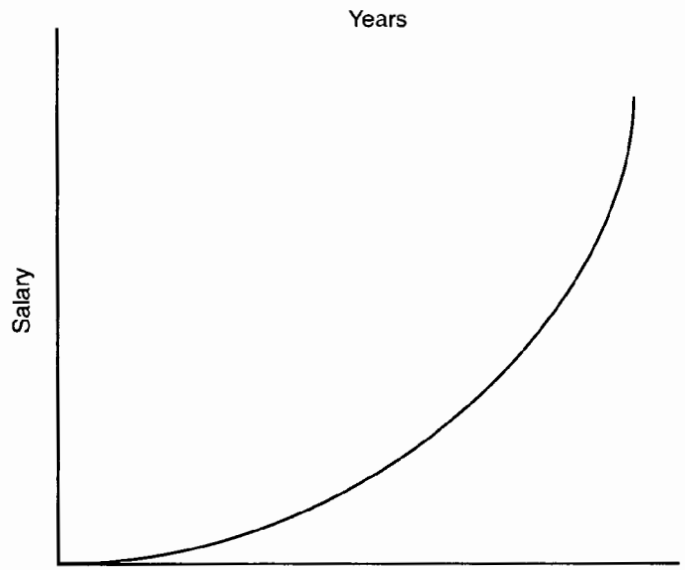
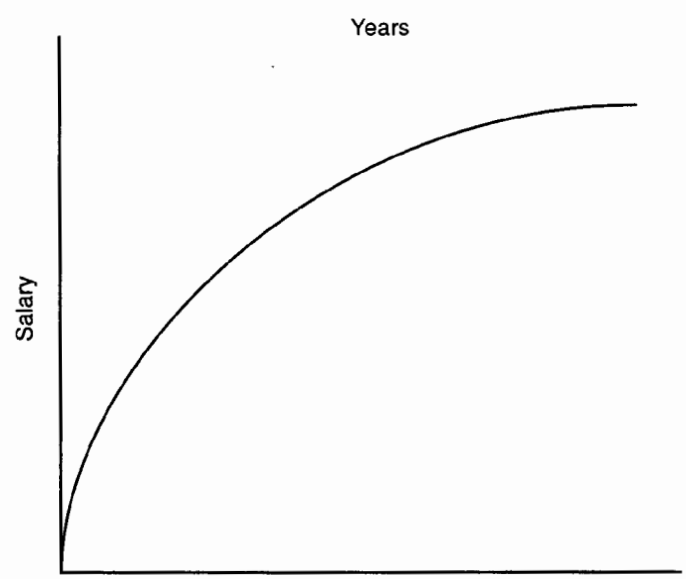


Figure 5.2

Decelerating Curve



The solution to the problem of curvilinearity, as we indicated in the previous chapter, lies in adding to the regression analysis a quadratic term for each time-related variable. A quadratic term is the square of a variable, the variable multiplied by itself. The methods used in the regression do not change (Pedhazur 1982; Hays 1991); the regression is just run with the quadratic term included. The new equation represents both the linear aspect (the time-related variable itself) and the curvilinear aspects (the quadratic term for the time-related variable) of the data. If the regression result for the quadratic term is negative, the data have some element of a decelerating curve. If the result is positive, the data have some element of an accelerating curve (Gray 1991). The amount of curvilinearity is shown by the amount of variance accounted for by the quadratic term. You can examine this by noting the amount of improvement in the R^2 when the quadratic term is entered.

As is so often the case, this solution creates another problem: variable overlap or redundancy. A variable and that variable multiplied by itself are very similar.[2] This problem can easily be solved by centering the original linear variable. Centering involves subtracting the mean of the variable from each measure of that variable, so that the new mean is equal to zero (see appendix G). Quadratic terms calculated on centered variables have substantially less overlap. Another solution is to drop the quadratic term, if it does not add anything to the analysis.

Redesigned Categorical Variables: Dummies

Variables like rank, gender, race, and discipline are categorical, rather than continuous. That is, they have a finite number of categories that are probably not equal in size. In contrast, continuous variables like years and dollars deal in quantities and have units that are equal in size (see the glossary).

Special steps need to be taken before categorical variables are included in multiple-regression analyses. Specifically, a categorical variable must be transformed into a set of dummy variables (see appendix A). Each dummy variable has only two possible values, 0 or 1. For example, for the variable female, all women are coded 1, and all others are coded 0; for the variable associate professor, we assign the value of 1 to those who are associate professors and the value of 0 to all others. Where there was originally one variable for current rank, there are now several current rank dummy variables, that is, full professor, associate professor, assistant professor, instructor, and lecturer. Therefore, the transformation to dummy variables involves a substantial increase in the number of variables.

For each categorical variable, the total number of dummy variables entered into the regression equation is one less than the total number of categories. The reason is that all necessary information is present when you enter all but one of the dummy variables. For example, in gender there are two categories, male and female. If we have a dummy variable for female, then we do not need one for male. If you are not female, then you are male. When there are more than two categories, the same situation applies. If there are only four race categories and you are not African American, Latino, or Asian, then you must be white, so we do not need to enter a white variable in the regression equation.

The dummy variable category that is excluded from the regression analysis becomes the default or reference group and is crucial in interpreting the parameter estimates of the other dummy variables. This is the group to which all other groups are compared. For example, if white is the default category, the parameter estimate for Asian tells us how much more or less money Asians make than whites with the effects of all the other variables in the model held constant. The Latino parameter estimate tells us how much money Latino faculty make compared to whites. To see how much Latinos make compared to African Americans, just subtract one parameter estimate from the other. (This discussion presumes that the parameter estimates are *un*standardized. See chapter 6.)

Since creating dummy variables substantially increases the number of variables in the regression equation, it is important to know how many variables are too many. The answer depends on the size of the faculty data set. Gray (1993) indicates that as few as five cases for each variable are acceptable, if statistical significance is not a concern. (See also Blalock 1979 and the discussion on statistical significance in chapter 6.) This would mean that with a faculty population of 250, you could use 50 variables—more than we have ever used. The difference between the R^2 and the adjusted R^2 is one way to assess the effect of the number of variables. The adjusted R^2 takes into account the number of variables, whereas the R^2 does not. Both measures are provided by the computer output of the regression model. As a general principle, trim the number of dummy variables to the smallest number needed.

Suggested Categorical Variables

You can follow these guidelines for including categorical variables in your study.

Current rank

Combining rank categories is not recommended, unless you do so specifically to erase the bias embedded in

them (see chapter 4). Different ranks are rewarded at different salaries. Therefore, we suggest entering even rank categories with small populations separately in the regression model.

Highest degree

You will probably find, as we have, that most faculty members have either a master's or a Ph.D. for their highest degree. In this case, you can reduce the number of dummy variables for highest degree to two or three. At the SUNY university centers and colleges, we reduced nine categories to two: Ph.D. or equivalent and non-Ph.D. Some disciplines have degrees that are designated "terminal" even though they are not Ph.D. degrees. If the information concerning these degrees is available, you may want to use a separate variable for terminal nondoctorate as well.

At SUNY two-year colleges, where the majority of faculty members had master's degrees, we used three variables: above master's, master's, and below master's. Some data sets include professional degrees. As we indicated in chapter 3, we recoded these to appropriate levels. In most cases, outside of medical and law school analyses, medical and law degrees were recoded as equivalent to Ph.D. degrees. M.B.A. and M.S.W. degrees were recoded as master's degrees.

Discipline

While there are a limited number of ranks or levels of educational attainment, there are usually many departments, which may add too many dummy variables to the regression model. Therefore, as indicated in chapter 3, it is best to combine departments into disciplines. As previously mentioned, the federal government has developed Classification of Instructional Programs, or CIP, codes for this purpose. Alternatively, you may combine departments into discipline groups based on common administrative units such as those in the same college or under the same dean, or through the use of a statistical technique called cluster analysis (see glossary).

Instead of creating dummy variables, some researchers transform discipline into a continuous variable by using salaries. This approach uses the average salary for each discipline within the institution, or the average market salary for each discipline, to create a continuous variable. We caution against this approach. When we conducted analyses to observe its impact, we noted that the R^2 declined measurably when such continuous variables were used. Using averages ignores how departments vary—for example, whether the department is a new one with many junior faculty or a more established department with many senior faculty. Using average market salaries ignores whether the discipline is one that brings great prestige to the institution and, therefore, is allowed to pay top dollar to attract the best in the field, or whether the discipline is one less well regarded nationally. Entering each discipline as a separate dummy variable allows the regression equation to assign the appropriate value for each discipline, given the faculty salaries paid in that discipline by the institution being studied.

Race and gender categories

Whether to combine race and gender categories is a complicated question. Several factors need to be taken into account: political consequences, potential masking of bias, and the number in each category. Politically, each racial constituency usually wants its own analysis, and the women in each group also want their own analysis. The racial categories we analyzed at SUNY were African American, Latino, and Asian. Other categories might be appropriate for your study. For example, schools in New Mexico and Alaska may want to include Native Americans.

Different dynamics affect each group, so breaking each category out is reasonable. But doing so can mean fewer than ten people in many race-gender categories, depending on the size of your institution (see table B.1 in appendix B). Even though you are using a population rather than a sample, results based on larger groups are more convincing than those based on one or two faculty members. If the regression results show ten or more African American males at a salary disadvantage, you can be more confident that this group is being subjected to systemic bias than if there are only one or two African American males on the faculty.

Statistically, there is a danger of masking bias when different race-gender groups are combined. For example, at two of the SUNY universities, Asian male faculty members as a group experienced a salary advantage relative to comparable white males, whereas African American male faculty members were disadvantaged compared with white males. Combining these two minority male groups would have masked the salary disparity experienced by the African American males. The opposite pattern occurred at the four SUNY colleges. Asian males as a group had substantially lower salaries than white males, while African American males on average had a salary advantage over white males.

If you have several small race-gender groups and you are wondering if it is all right to combine them, first try running regression analyses with the groups separate. This allows you to assess the degree to which

masking could be a problem if you do combine these groups.

What about minority women? Should they be included as minorities, as women, or separately as their own race-gender category? There may be more incentive for combining minority women than minority men. Judging from all twelve of the schools in our analyses, there are fewer minority women than minority men on faculties. The results from the twelve SUNY schools indicate that minority women's salaries are more similar to those of white women than to those of the men in their same minority category, and that minority women experience greater salary bias than their male counterparts. Our findings are consistent with other studies that indicate that minority women's ranks and salaries are below those of minority men (Exum 1983; Fairweather 1991; Tuckman 1979). Therefore, if you want to combine minority women to create a larger pool, we suggest combining them with each other or with all women, but not with their minority male counterparts.

Finally, there is the question of whether to conduct race analyses at all. Unfortunately, most of the studies reported in the literature assess salaries for gender bias but not for race bias. Of the studies that have examined race, some have found that minorities are paid as well as, if not better than, white men, while others have found the opposite (Exum 1983; Fairweather 1991; Tuckman 1979). We encourage you to conduct assessments for both race and gender bias, bearing in mind that individual characteristics can unduly influence analyses where there are only a few individuals in a minority category.

Ignoring race may mask gender bias, since lower-paid minority men could lower the male average salary. In other words, if the average salary for male minority faculty is less than the average salary for white males, the total male average will be lower than the white-male average, thereby decreasing the gap between women's and men's salaries. Our assessment indicated that omitting race information did not substantially change the amount of gender bias shown, or the adjusted R^2, at most SUNY institutions we studied. At one college, however, the gender bias shown dropped by one-third when minority men were included with white men. We urge you retain the race variable unless the number of minority males is very small.

Three Multiple-Regression Approaches

The three types of multiple-regression models reviewed here are the ones most commonly reported in the literature.[3] You need not choose among them. You can use all three to examine the consistency of the results. If the results agree, the information provided here on the differences between the three approaches may be of limited interest. If the results differ, or if you decide to limit your study to a single approach, this section will be helpful. First we describe each approach along with its primary advantages and disadvantages; then we describe our general findings relative to each approach, and our resulting preference.

Total population—actual salary analysis

The first approach uses the total population of faculty in deriving the regression equation. Men and women, minorities and nonminorities are included in the analyses. The dependent variable is salary in actual dollars. Gender and race are accounted for by entering the dummy variables for each race-gender category, such as Latino male, Asian female, and so on.

An advantage of this model is that the regression coefficients (parameter estimates) can be directly and easily interpreted in real dollar amounts. For example, the coefficient for the variable white female is the average salary difference between the default category, white male, and white female, with all other variables held constant. If the coefficient is -3,000, white females are paid $3,000 less than white males with comparable predictor variable scores. If the dummy variable African American male is entered, then the average salary difference between African American males and the default category, white males, is indicated by the coefficient for this variable. Appendix A explains this regression approach thoroughly.

The total-population regression analysis can mask inequity because it assumes that every factor that influences white-male salaries affects female and minority salaries at the same rate. As a result, some discriminatory differences get averaged away. For example, if, on average, males are rewarded $1,000 for having a Ph.D. and females are rewarded $500, this model masks this inequity because it looks only at the average reward for Ph.D. Later in this chapter, we indicate how to compare the results of the total-population analysis with the white-male analyses to reveal whether women and men are being paid differently. If this problem exists, one way to eliminate it is to use the "white-male population" approach.

Natural logarithm of salary analysis

The second model we tested was a total-population model using the natural logarithm (ln) of salary instead of actual salary as the dependent variable. Taking the natural log of salary means that you are no longer studying dollar units and that the parameter estimates can no longer be directly interpreted in dollars. Rather,

they become proportions by which a salary is changed when an independent variable increases by one unit; multiplied by 100, they become percentages (Hodson 1985; Gray 1991; Becker and Toutkoushian 1995; Halvorsen and Palmquist 1980). For example, if white males are the default group and the coefficient for the variable female is -0.0634, then the average female in the population is making 6.34 percent less than the average white male.[4]

The natural log of salary transformation can create a more "normal" (see glossary), less skewed and less curvilinear, distribution. It is generally felt that the more normal the shape of the data distribution, the better, particularly if you are using a random sample to make inferences about a population, rather than studying the total population as is commonly done in faculty salary studies (see "Problem 5" in chapter 6). The improvement in the distribution of the data so that it is a "better fit" for linear regression analyses is usually indicated by an increase in the adjusted R^2. Gray (1991) notes that the "log model generally allows us to get a better fit to the data, but at the sacrifice of simplicity."

Using the natural log of salaries lowers the highest salaries, bringing them closer to the mean and the rest of the distribution (Hodson 1985), and reduces the effect of very high salaries on the regression results. If there are extremely high-paid individuals who make more than ten times as much as most people in the salary analysis, using log salaries could be very important in reducing the impact of these individuals' salaries on the regression results. The log of salaries also pushes the lowest salaries further from the mean and the rest of the distribution, thus tending to increase their impact on the regression results.

The log of salary is conventionally used in the field of economics, and faculty salary-equity studies frequently attract the attention of faculty economists. Economists' affection for the natural log of salary is related to their focus on returns for investments and frequent research on populations with an extremely wide range of salaries. When focusing on return for investment, it makes sense to report the results in percentage rates. To report results in actual dollars means that you always have to know the baseline. If an investment nets $100, you need to know the size of the original investment. On the other hand, if an investment nets 30 percent, you know it has done well. Hodson (1985) notes that economists find this useful because they can report findings such as "each year of investment in education brings a 13 percent increase in earnings."[5] But if you are studying the effect of gender and race on salaries, the logic of "investments" disappears. It makes no sense to say that for each year of investment in being a white male, earnings increase by a certain percentage. Gender and race are not investment choices.

Because using logs lessens the gap between the mean and the high end of the salary distribution, they are commonly used in studies of the general population where the highest salaries may be a hundred times or even a thousand times higher than the lowest salaries. But the occupational and institutional context of faculty salaries restricts their variations to a much narrower and less skewed range than in the general population. Observing this, Ferree and McQuillan (1998) concluded that "converting salaries to logarithms offered more cost in loss of ready interpretability than it was worth" (23). We suggest that you weigh the importance of conducting the log salary analyses based on the range of salaries. You can do this by dividing the highest faculty salary by the lowest. If the result is greater than ten, say from a lowest salary of $30,000 to a highest salary of more than $300,000, you should conduct log analyses. If you do have such a wide range of salaries, we suggest you also visit the section on outliers in chapter 6 and appendix I on how to locate outliers.

White-male-population salary analysis

The third approach, the "white-male-only" model, has been widely promoted because it solves the total-population problem of masking different rates of pay for women and men (Braskamp, Muffo, and Langston 1978; Gray and Scott 1980; Scott 1977).[6] *The Higher Education Salary Evaluation Kit* by E. L. Scott (1977) recommends the white-male analysis because it provides "an estimate of what the salary of the woman or minority faculty member would be if she or he were a white male with the same attributes and experience."

To apply this method, you calculate a regression equation using only the white-male faculty. There is no need to use race-gender variables because there is only one gender and race in the analysis. This "white-male" equation is used to predict what the salaries for women and minorities would be if their career attributes were rewarded in the same way as those of white males. (See appendix A for how to calculate a predicted salary using regression parameter estimates.) The difference between each woman or minority's predicted salary, based on the white-male equation, and his or her actual salary is called that person's salary *residual*. The average residual for a race-gender category measures the difference between the actual salaries of those in the group and what they would have been paid if their race had been "white" and their gender "male." A negative average residual indicates that the actual salaries of the

category in question (for example, Latino men) are lower than their predicted salaries based on how white males are paid. A positive average residual means that the race-gender category's actual salaries are higher than their predicted salaries using the white-male equation.

Gray (1993, 1991) indicates that while this model is theoretically better because it shows what females would be paid if they were male, it may be more difficult to use. A problem occurs any time there are no white males in a category where there are women and minorities, making it impossible to derive an accurate predicted salary for these women and minorities. For instance, if there are no white males in the school of nursing, then there is no parameter estimate for this category and the salary worth of this discipline will depend on the default discipline. Any category with no parameter estimate is treated the same as the default category.[7]

To get around this problem, Scott (1977) suggests agreeing on another discipline that is similar to the discipline with no males in it. For instance, the pharmacy white-male coefficient could be used in calculating the predicted salaries for nursing faculty.

The smaller the white-male population, the more likely you are to have disciplines or other categories (for example, a rank like instructor, or an educational attainment like master's degree) that have women and minorities but no white-male counterparts. Even if there are men in all categories, there may simply not be enough white males to create a valid regression model. We advise that there be at least five white males in each discipline, rank, and degree category to apply this method; otherwise, having one or two uncharacteristic white males in a category can throw the validity of the white-male analysis into doubt (see "Headaches" below). If there are fewer than fifty white males among the faculty, the accuracy of the white-male regression results are questionable. As is the general rule for regression modeling, there should be at least five cases (faculty members) for each independent variable in the analysis. Since twenty or more variables are common in faculty equity studies, a count of a hundred or more white males is desirable for these analyses.[8]

Recommendations

If you had only a single type of multiple-regression analysis available to you, you would probably want it to be the first method we described in this chapter, the total-population salary analysis. It is easy to understand. The unstandardized results represent the actual salary variations. It does not exhibit the frequent problems experienced with the white-male analysis caused by too few or uncharacteristic white males. The results are in actual dollars and, consequently, more readily interpreted and understood than log of salary.

But you may nonetheless find it worthwhile to try all three models. If all three confirm a similar level of bias, then they validate each other. (In three of the twelve faculty data sets we analyzed, one two-year and two four-year colleges, our results were consistent across all three analyses.) If the findings differ somewhat, you will need to assess why, but in the process you will gain understanding of the salary anomalies that exist at your institution and how they relate to gender and race. Reporting three estimates of bias also avoids overconfidently reducing a complex issue to one summary number.

Headaches

If white males are paid more for a Ph.D. or for a year of experience than women and minorities, this source of bias is hidden in the averaging done by the total-population approach. How prevalent is this problem, and how does the white-male population approach correct for it?

Our analyses of the twelve SUNY institutions showed that the white-male results supported the total-population results in six of the twelve analyses. In other words, at half of these institutions, there was no substantial differences by gender and race in the rewards faculty received for career attributes such as Ph.D.'s or years of experience. In fact, when differential rewards existed, our findings indicated that the white-male results were most likely to be based on anomalies in the salaries of white men that are accentuated by the white-male population approach.

The results of the white-male analyses showed substantially higher bias than the total-population salary analyses for only two of the twelve SUNY data sets. In both cases, the higher estimate of bias was based on the aberrant salary of one white male. At one college, a white male speech pathologist was paid $13,000 more than the average for all the faculty members in this discipline.[9] As a result, the predicted salaries of the women and minorities in this department were substantially higher using the white-male method, and their actual salaries seemed to be biased by comparison. When we dropped this discipline from the analysis, the amount of bias indicated by the white-male analysis was the same as that for the total-population and log of salary results. At a two-year college, an inconsistency resulted from a white male in construction technology whose educational attainment was below a master's degree but who was paid more highly than most of the Ph.D. faculty members. The result was that the women and minorities

whose highest degree was below a master's appeared to be low paid by comparison, and their predicted salaries, based on the white-male analysis, were substantially higher than they would have been as a result of the total-population analysis.

At one two-year college and one university, the white-male analyses indicated substantially lower bias than did the total-population salary analyses. Once again, the presence of one white male with an atypical salary caused the misleading white-male results. At this two-year college, there was only one low-paid white male in public service and early childhood education. By comparison, the nine women and minorities in the department looked overpaid. At the university, there was only one white male in the school of nursing, and his pay was relatively low. By comparison, the twenty-three women in the school of nursing appeared to be overpaid.

We concluded that when the number of white males in a category is fewer than five, and when the white-male results differ from the total-population results, the total-population analyses are more likely to accurately reflect the amount of bias present in salaries.

At two other SUNY institutions the white-male analyses were problematic because there were no comparators in disciplines where there were women and minorities. Thus, six of the twelve white-male results are problematic. These problems were not confined to the smaller schools. Three of the six institutions (one university and two colleges) had more than two hundred white-male faculty members. If our twelve SUNY analyses are characteristic, a white-male analysis at any one institution has a fifty-fifty chance of displaying incongruent findings that take time and resources to investigate. Before deciding to conduct a white-male analysis, we recommend that you check for disciplines or degree categories at your institution with no white males—or fewer than five white males—and consider whether you have the time and resources to investigate any anomalies that may result.

The argument that white-male analyses avoid averaging away greater rewards given to white males for degrees and experience implies that the white-male analyses reveal higher bias than total-population analyses. We have not found that to be true.

Improved analysis

When the white-male results show inconsistencies, we believe that using total-population analysis is more logical and defensible. Regression equations using the entire faculty population include more information and, thus, may be more accurate than the regression equations based only on males. For example, a male-only salary regression equation might include only two or three faculty members in a particular discipline, but the all-faculty salary regression at that campus may include two or three additional faculty members in that discipline. Thus, the all-faculty regression would have twice as much information concerning how salaries in that discipline vary relative to salaries in other disciplines.

If you are concerned about whether a total-population analysis is masking differences in the way men and women are paid, this problem can easily be assessed. Place the regression-equation computer output for the white-male analysis and the one for the total-population analysis side by side.[10] Compare the coefficients (parameter estimates) for each of the variables. If you find a substantial difference for a category, examine how many white men are in that category. If there are four or less, perhaps their salary-related rewards for this attribute are very dissimilar from the others in this category. If you find that the differences noted between the white-male and total-population results are not due to one or two aberrant salaries, you can correct for the difference by adding an interaction term to the total-population analysis (Gray 1993).

An interaction occurs when the effect of one variable changes across levels of another variable. Suppose that the effect of not having a Ph.D. changes across gender. An interaction term variable is created by multiplying the two variables together.[11] Create an interaction term by multiplying the Ph.D. variable by the gender variable, and include this interaction variable in the analysis. The parameter estimate for the interaction term tells what the difference is between men with a Ph.D. and women with a Ph.D. For statistical purposes, it is important not to add any more interaction terms than are needed.

Complexities

The primary disadvantage of using log of salary as the dependent variable is the complexity it adds to interpreting the results. When dollars are converted to log dollars, the intervals between the values of different salaries change. One person earning more than another will also have more log dollars, but the distance between the log salaries of the two people grows smaller as the size of their salaries increases. Complexity of interpretation is more important in reporting findings than it is in deciding which analyses to do. As indicated earlier, if the range from the lowest to the highest faculty salary is greater than a factor of ten, we suggest you conduct log of salary analyses. But when it comes to writing the report of your findings, assuming that the results do not differ substantially, consider reporting on

the salary results instead. Our preference for reporting the salary rather than log of salary results is based on simplicity and on the logic of reporting on what people actually take home in their paychecks.

Using the natural log of salary usually results in slightly higher adjusted R^2 values, commonly in the range of two or three percentage points. At the twelve SUNY schools we studied, institutional type did not seem to play any role in the amount or direction of the change in R^2 values. Closer examination indicated that the slight improvements in the adjusted R^2 values were probably primarily due to the reduction of the positive skew of the dependent variable at the universities and colleges.

Two Additional Checks

Once the analyses have been completed, two important additional tests remain to be conducted: a test for outliers and an examination of the residuals.

Outliers

Although we do not generally recommend dropping outliers, it is best to check for them (see "Problem 3" in chapter 6). Assessing the impact of outliers is a good validity check on your data and results. For a discussion of methods of checking for outliers, see Appendix I. As noted there, we recommend the Cook's Distance (CD) approach. The advantage of Cook's Distance is that it pinpoints those cases that greatly influence the parameter estimates rather than just aberrant cases on one of the variables. Cases that are not distorting the results need not be examined or eliminated before you certify the accuracy of your analyses.

The developers of the Cook's Distance procedure indicate a CD cutoff value of 1.0. Out of twelve SUNY-campus analyses, we found only one case that exceeded the CD cutoff value. A check of this professor's data showed him to have been a full professor for only two years, but he had actually been a full professor for seventeen years. The erroneous "two years in current rank" referred to the time at which he stepped down from a dean's position. If your analyses include cases with CD values above 1, review their data for accuracy. You may find errors that explain their outlier status.

If you note CD values that are substantially higher than those of most of the population, examine these for potential data errors. For example, if the highest CD value is 0.63 and the next highest CD value is 0.15, the data for the case with the CD of 0.63 should be checked for accuracy. If you are assessing a small population of faculty salaries (under a hundred), consider how representative that case is. At one small college, such a case turned out to be a non-Ph.D. male instructor who had received tenure thirty-five years ago but never progressed in rank. After running the analysis with and without this outlier, the joint faculty-administration committee at this institution decided he was not representative of the population and left him out of the analysis.

Although the detection of outliers is a recommended statistical procedure, dropping them is as much a political decision as a methodological one. Drop outliers only if they distort the real picture.

Residuals: The ungreat unknown

Regression analysis looks at the variation in the dependent variable (salary) and all of the information provided by the predictor variables, and creates an equation for the straight line that best predicts your institution's salaries. But real-world salaries do not dutifully line up on the straight line provided by the equation. They may cluster tightly around it; how tightly is indicated by R^2. But even when the R^2 indicates that 70 or 80 percent of the variation in salaries has been accounted for with the predictor variables, there is scatter around the line (see figure A.2 in appendix A). This scatter of 20 or 30 percent variation in salaries is called the unexplained variation or residual. It results from factors you have not measured or cannot measure precisely. Some examples might be unmeasured productivity, administrative favoritism, and being hired in a financially dry year.

The residual or unexplained variation is the difference between a person's actual and predicted salary from his or her actual salary. As explained in appendix A, the regression equation provides a formula that can create a predicted salary for each faculty member based on the predictor variables. All of the widely used statistical computer software programs provide the predicted salary for each individual based on the regression equation (when properly asked, of course). They also provide the residuals or unexplained variations. You can create a scattergram of the residuals by plotting each faculty member's actual salary on the vertical axis and the unbiased predicted salary on the horizontal axis. For examples, see figures 6.1 and 7.1.[12]

When you use the white-male regression equation and arrive at the women and minorities' predicted salaries by plugging in their career attributes (the predictor variables), the result is a direct measure of what women and minorities' salaries would have been if they were white men. But the predicted salaries and, by extension, the residuals provided by the total-population regression analysis output include the salary bias variation related to the female and minority variables. Regression analyses let us know how much of the variation in

the dependent variable, salary, is accounted for by the independent and predictor variables. We provide predictor variables that are seen as both legitimately related to salary (experience, degree, discipline) and variables that are seen as illegitimate predictors of salary (gender and race). The regression equation that results does not distinguish between these. It tells us mathematically which variables predict the women and minorities' salaries. If it finds that, after accounting for all the variables that we consider legitimately related to salaries, women and minorities' salaries are best predicted by subtracting something from their salaries, it includes that subtraction in the equation that calculates the women and minorities' predicted salaries. It is our job to remove that part of the equation reflecting gender or race bias.

If, for example, the coefficient for female is -1,000, the salaries of the women faculty members are predicted to be $1,000 less, on average, than if they were male faculty. The predicted salaries for the female faculty must be corrected by the amount of the coefficient for female: it would be necessary to add the $1,000 back to each female faculty member's predicted salary. Once you have corrected the predicted salaries to remove any gender and race effect, you can subtract them from actual salaries to get a corrected residual. Alternatively, as described above, you can create a scattergram of the residuals by plotting each faculty member's actual salary on the vertical axis and the "corrected-for-bias" predicted salary on the horizontal axis.

There are both political and methodological reasons for examining the residual or unexplained variation in salaries (see chapter 7 for more on the political implications). You may observe a pattern of the residuals that suggests an additional variable not included in the analysis. Including that new variable could account for more of the variation in salaries, thus raising the R^2. For example, you may note a cluster of faculty paid more than their predicted salaries, all of whom are named or distinguished professors. By including this as a category you may improve your analyses. Or you may discover a group of faculty whose actual salaries are much lower than their predicted salaries and who are all in a particular subunit or discipline. As Ferree and McQuillan (1998) note, such a result could point to a pocket of bias caused by a particularly prejudiced administrator.

Notes

1. In technical terms, redundancy means high correlations that make it difficult, if not impossible, to determine the separate effects of the two predictor variables on the dependent variable.

2. To determine whether variable overlap is a problem, run a correlation matrix to show how much each variable and its quadratic term overlap. If any correlation reaches 0.9 or above, you have a problem.

3. There are, of course, many variations in the literature. See Toutkoushian (1994) for a statistical discussion of the alternative approaches and their potential impact on salary adjustments. In the next chapter, we discuss why we believe two variations, one that is called stepwise regression and one that omits the gender and race variables from the equation, should not be used. There is also the complex Oaxaca (1973) approach, which uses both male and male regression analyses, but, in practice, the results are very similar to using just the white-male analysis.

4. When interpreting percentage rates, be aware that the rate for one group is not the inverse of the rate for the other. Since the percentage rate is based on the average salary of the group, there is no common denominator. In actual dollars, if women need to make $1,000 more to be equal to men, then men would need to make $1,000 less to be equal to women. However, with percentage rates, the numbers are not the same. If men make $1,000 and women make $900, then a 10 percent decrease in male salaries would constitute equality, but an 11.11 percent increase in female salaries would be needed to achieve equality.

5. Hodson points out a very important implication resulting from the use of natural logarithm of salary and the associated rates of return, like the rate of return for a year of education. A rate of return is always calculated based on the mean of the group being analyzed. For example, if separate analyses are done for men and women, the calculated rate of return is based on increments above their separate average logged salaries. Because women make substantially less, on average, than men, women can appear to gain more, even though their actual salaries are increased by less. Hodson (1985) notes that whereas men make more in dollars ($66), which translates to a 0.54 percent rate of return, women make less ($46), but this translates into a 1.25 percent rate of return, because of the much lower average earnings of women. Thus, when differences in mean earnings between groups are substantial, the analysis of logged salaries may create apparent differences which are artifactual and very difficult to unravel.

6. For simplicity of explanation, this section assumes that the dependent variable used in white-male population analysis is salary as opposed to natural log of salary. Although there are added complexities, it is, of course, possible to conduct a white-male population log of salary analyses.

7. The impact on the predicted salaries of the nursing faculty of having no white-male counterparts can be high or low depending on which discipline is in the default category. That is because any group or category with no parameter estimate (coefficient = 0) is treated the same as the default category. If

the discipline left out of the equation is psychology, then the nursing faculty women will be paid like white-male psychologists. If the default discipline is the highest paid, the predicted salaries of the nursing faculty will be based on their being in the high-paid discipline; if the default discipline is the lowest paid, the predicted salaries of the nurses will be based on what they would make in this lowest-paid discipline.

8. If your primary analyses use a white-male regression approach, it is advisable also to conduct a female-only regression analysis to check whether it underpredicts male salaries. It is possible, although in our experience not probable, for the male model to predict higher female salaries and for the female model to predict higher male salaries (Gray 1991).

9. Speech pathology was the only department in the health professions discipline.

10. Alternatively, a female- or minority-only analysis can be run and its coefficients compared with the white-male line coefficients. Some researchers recommend running the female or minority regression and assessing the male faculty salaries with this model. That is sometimes suggested to assess whether the female-only model indicates that the male faculty are underpaid (Gray 1991), just as the white-male-only model indicates that female faculty are underpaid. Although this potential is important to check, in reality it is rarely the case that men are found to be underpaid using the female model (Gray 1990). At most institutions of higher education, men still constitute the bulk of faculty and are the basis for institutional norms; it therefore makes more sense to fit the women to the men's model than vice versa.

11. Using interaction terms involves complexities. If the non-gender variable involved has more than two categories, the default category used changes the interpretation. This is particularly noteworthy if you are doing interaction terms between gender and a category like discipline or college with the objective of deciding the relative salary bias in each discipline or college. When interaction terms are used, the interpretation of the regression coefficients changes (see Jaccard, Turrisi, and Wan 1990).

12. With regard to white-male equation analyses, Ferree and McQuillan note a second approach to examining the residuals. The residual standard deviation units can be arrayed in a bar chart that will approximate a normal curve. Creating one array for the white males and another for the women and minority categories may display "humps" in the residuals that indicate pockets of bias that are not systemic. (See chapter 7 and Ferree and McQuillan 1998 for more information.)

I have always believed that contemporary gender discrimination within universities is part reality and part perception. True, but I now understand that reality is by far the greater part of the balance.

—Charles M. Vest
President
Massachusetts Institute of Technology

SIX
Small Errors with Big Consequences

By Lois Haignere

For many readers, this chapter will be the most important in this guide. In designing a statistical study, you will have to make many decisions. Too often, these decisions are touted as strictly methodological, and their subjective nature and potential political impact are ignored.

Even if you are one of the policy makers involved, you may be concerned that those conducting the salary-equity analyses for your institution will, wittingly or unwittingly, design them in a way that masks gender or race bias. This chapter helps alert you to what to guard against and what questions to ask. Some common circumstances with which this chapter will assist you include the following.

- The administration has handed you its study, expecting you to accept the results as "objective truth."
- Institutional researchers, or others with more statistical knowledge than you, have suggested a research design for your approval.
- You are a member of a joint faculty-administration planning committee to design a salary-equity study, and you are suspicious of some suggested approaches.
- You are conducting your own analyses, and you want to know what procedures to avoid.

Problem 1: Standardization of the Coefficients

The regression coefficients (also called the parameter estimates) are the values that indicate how much influence each variable has on the dependent variable, salary. Most statistical computer packages provide *un*standardized regression coefficients by default, because they are easier to interpret. Getting standardized coefficients requires extra computer programming to override the default. *Un*standardized results indicate the average dollars associated with each variable. (See appendix A for a detailed explanation and an example of computer output.) An unstandardized coefficient of -1,500 for female faculty indicates that, on average, women faculty are paid $1,500 less than comparable male faculty.

Standardization of the coefficients does not mask gender and race bias per se, but it does change the results of the regression analyses from a form that almost anyone can understand to one that only researchers and statisticians can interpret.

To obtain standardized coefficients, both the independent and dependent variables are changed to standard deviation units. The standardized coefficient for a particular independent variable represents the amount of change in that variable's standard units needed to produce a change in the dependent variable of one standard unit—a far cry from the direct dollars interpretation of the unstandardized coefficients. Standardized coefficients can be converted back to dollar equivalents.[1] However, this should not be necessary, since the analysis can simply be rerun specifying that the output be in *un*standardized units.

Problem 2: Exclusion of Certain Categories

Some higher education institutions direct their research where they may be least likely to find gender and race bias in salaries, and then declare that they just can't find it. They do so by systematically excluding the academic groups where women and minorities predominate. For part-time faculty, exclusions may be a necessary evil (see "Study Population" in chapter 3). But all too often, the full-time non-tenure-track faculty are unnecessarily excluded without any methodological or theoretical justification.

Exclusions based on track or rank

Non-tenure-track teaching ranks are often excluded for reasons that, if explained at all, are described vaguely as either different teaching and research expectations (Gray 1993) or different hiring and personnel policies

(N. Moore 1992).[2] Some institutions of higher education have an array of non-tenure-track faculty, including "visiting," "research," and "clinical" assistant, associate, and full professors. But most campuses confine non-tenure-track faculty to one or two lower ranks, such as instructor or lecturer. At the SUNY institutions, most non-tenure-track faculty are lecturers, long-term teaching faculty who rarely have research obligations. But in regard to teaching, their work is much the same as that of tenure-track faculty. Most variations in teaching expectations and hiring and personnel policies that exist between non-tenure-track and tenure-track teaching faculty are no greater than those that exist across most ranks. Just as rank differences are methodologically accounted for by the inclusion of dummy variables for the different ranks, non-tenure-track salary differences can be measured by the regression equation through the addition of a non-tenure-track dummy variable.

There may be different research expectations for the faculty who are on tenure track than for those who are not, but research expectations are not prevalent on all campuses. In fact, teaching rather than research is the primary focus at many two- and four-year colleges. Even at research universities, the expectations regarding research may depend as much on discipline as on whether faculty are on track for tenure. If a department is recognized as one of the best in its field, the research expectation for those hired into it is often greater than that for faculty in less prominent departments.

Moreover, differences in policy regarding research, teaching, hiring, and other personnel issues are taken into account by multiple-regression analyses. Multiple regression can separate out pay differences that are unrelated to gender and race as long as the separate rank and discipline categories are named in the analyses. By entering a variable for each discipline and rank (see the discussion of dummy variables in appendix A), we enable the regression equation to properly attribute salary differences related to discipline or rank to the discipline or rank variables and not to gender or race.

We can do the same for non-tenure-track ranks. All full-time faculty members, including those not on tenure track, were included in our SUNY analyses. As indicated in chapter 3, to account for non-tenure-track salary differences, we entered the variable lecturer as a synonym for "nontenure track." This means that any gender or race salary differences *within* the lecturer category can be captured by the equation and attributed to the coefficients for the gender and race variables. Any salary variations due to different teaching or research expectations or personnel policies for lecturers will be attributed to the coefficient for the variable lecturer.

To assess whether dropping non-tenure-track faculty improved the quality of the analysis, we reran the analyses, excluding those in the rank of lecturer at the nine SUNY schools that have this non-tenure-track rank. Neither the adjusted R^2 measures nor the amount of gender or race bias shown changed substantially when we dropped lecturers. At two campuses (one university and one four-year college), the amount of bias increased modestly, suggesting that, at these schools, those in the non-tenure-track lecturer rank experienced less bias than those in the tenure-track ranks. At two campuses, the amount of bias measured did not change. At the remaining five campuses—one two-year college, two four-year colleges, and two universities—the amount of bias decreased slightly, suggesting that the non-tenure-track faculty at those schools probably experienced more salary bias than those in the tenure-track ranks.

Nevertheless, it is important to examine the regression results carefully for indications that the non-tenure-track ranks could be masking bias in the tenure-track ranks or vice versa. This lesson was made clear by a study that we conducted recently at a private-sector college on the East Coast. The results indicated a modest amount of gender bias in salaries. But a comparison of the parameter estimate of the total-population and white-male equations showed major differences for the non-tenure-track variable. Interaction-term analyses indicated that the bias was greater among the non-tenure-track faculty. When we dropped the non-tenure-track faculty from the analyses, the indication of salary bias disappeared. As is true at many institutions, there were too few full-time non-tenure-track faculty to support a separate analysis within the non-tenure-track ranks, but the interaction-term results provided clear evidence of salary bias in the salaries of those in the non-tenure-track ranks.

The tendency to exclude non-tenure-track faculty in salary analyses may be due in part to an assumption held by many higher education policy makers that there is less bias at the lower-status levels, where women and minorities predominate. There is also an inclination to believe that recent hires are treated more equitably than less recent hires, and that the non-tenure-track ranks include a higher proportion of recent hires. Our analyses suggest that these commonly held assumptions are often not supported by study results.[3]

The most important consequence of excluding non-tenure-track faculty from the salary analyses, as we indicated in chapter 3, is that they will be ineligible for any salary adjustments that may result from the study. That means that the lowest-paid faculty, who are likely to be disproportionately minority and female, are

denied the opportunity to have any bias in their salaries assessed and corrected. We know of no theoretical or methodological justification for this unfortunate exclusion. All full-time faculty, including the non-tenure-track ranks, should be included in the multiple-regression salary review.

Potentially legitimate exclusions

Different campuses and institutions rely on an array of titles, positions, and funding mechanisms to provide for the diversity of staff needed. We have conducted analyses at institutions that have research and clinical faculty lines (sometimes called rank modifiers) that span the assistant, associate, and full professor ranks and also include both tenure-track and non-tenure-track faculty. These are all considered faculty lines and should be included in the faculty salary study.

By contrast, campuses also commonly have nonteaching grant-funded positions, such as research assistants and associated positions. Those holding these non-tenure-track research assistant and associate positions have much in common with faculty: they may hold Ph.D.'s and work right along with faculty in a specific discipline. But they are usually not in faculty lines and do not teach, and their salaries may be determined in whole or part by outside funding sources rather than by the higher education institution. Here you may need to sort carefully so that your study includes only teaching faculty, both on and off the tenure track, whose salaries are determined by the campus administrations.

Exclusions based on discipline

At some institutions the faculty members who specialize in teaching students basic reading, writing, and mathematics skills are clustered in a special department. At SUNY institutions these skills are taught in Equal Opportunity Programs and Equal Opportunity Centers. Segregating the remedial education function in a separate organizational unit can result in excluding these faculty members from a salary-equity analysis.

Here too, regression analyses can sort out the effects of gender and race from those due to membership in a separate discipline unit. All that is needed is inclusion of a dummy variable for remedial education. Again, no theoretical or methodological reason exists for excluding remedial educators. All full-time faculty should be included in the multiple-regression salary review.[4]

Problem 3: Dropping of Outliers

Trimming the outliers, those whose pay is way out of line with that of others, sounds like a statistically legitimate, innocent thing to do. If your school has a Pulitzer Prize-winning professor who was recently hired at a very high rate of pay, why should she be allowed to mask bias against other minority women? If your school has kindly kept on staff, at a token salary, a severely injured white-male professor so as to continue providing his health insurance, why should he be allowed to mask bias? If the gender and race of these examples were reversed, the logic would not change. The dropping of such outliers should be discussed and agreed upon by all parties. Appendix I explains three statistical tests for locating outliers.

The problem comes when more statistically lazy or politically expedient methods of "trimming" outliers are used. If the hands-on researchers inform you that they have dropped all "outliers" with salaries more than two standard deviations from the mean, this is not as innocent as it may sound. By definition, approximately 2.5 percent of the faculty will be more than two standard deviations above the mean and another 2.5 percent will be two standard deviations below it. If the analysis involves four hundred faculty members, that means dropping about twenty people—ten with actual salaries much higher than their predicted salaries and ten with actual salaries much lower than their predicted salaries.

If gender and race are randomly distributed among these two groups, dropping them from the analyses will probably make no difference. Such randomness was not the case at any of the twelve SUNY institutions in our analyses; rather, those on the high-paid side were often white males, and those on the low-paid side tended to be females and minorities. The practice of dropping the outliers is predicated on the assumption that they distort the real picture. If the so-called outlier faculty do not distort the picture but are simply the highest- and lowest-paid cases in the picture, there is no real reason to drop them; doing so can peel off a chunk of the gender and race bias.

Problem 4: Tainted Variables

The choice of predictor variables included in multiple-regression analyses is critical in determining the accuracy of the results. If you include a predictor variable that is biased, then some of the salary estimate that might properly be attributed to being female or minority will instead show up as part of the biased or tainted variable's estimate.

For example, suppose height were included in a salary-disparity analysis where gender bias exists. Because women are shorter on average than men, any gender differences in salaries could largely be explained by the inclusion of height as a predictor. The salary

disparity related to gender would be proportioned out by the regression analysis to both height and gender, masking the true magnitude of the gender bias.

While no one would include a height variable in a salary study, there are several common variables that can be tainted. Chapter 4 discusses in detail how to test the categorical variable current rank for any taint that might mask bias. We also discuss what to do if you find that current rank is tainted. Our conservative recommendation is that you still include it as a variable in the regression analyses, but assume that the results underestimate the amount of bias. In this section, we warn against inclusion of three variables that are likely to mask bias even though it may not be possible to demonstrate statistically that they are tainted: current or previous administrative appointment, initial rank, and initial salary.

Current or previous administrative appointment

Because being appointed to dean, assistant dean, department head, or another administrative position is much more common for male than female faculty (Johnsrud and Heck 1994), we thought it best to check this variable for taint before including administrative appointment data in our SUNY analyses. Unfortunately, we found that for most of the twelve SUNY schools, the statistical checks for bias in this variable were not reliable. The categorical modeling approaches detailed in chapter 4 gave low R^2_L measures, indicating that the variables in the model (years at the institution, years of experience prior to hire, highest degree, gender, and race) were poor predictors of receiving an administrative appointment. Perhaps these variables will be better predictors of administrative appointment at your institution. Or your data set may include other variables that are more directly related to who is appointed to administrative positions.

At some institutions, the extra pay associated with administrative appointments is awarded only while an individual is in that office, and is dropped afterward. If your institution is one of these, you can simply drop the extra salary stipend of all current administrators, running the analyses on base salary.

For most institutions, however, the situation is not that simple. Many campuses in theory remove the stipend when the faculty member no longer holds an administrative position, but in practice many faculty keep some or all of the administrative salary increment after they step down. Other institutions openly allow those who were once in administrative positions to continue to be rewarded at a higher rate, based on their history in an administrative position. It may also be that those who have formerly served in administrative positions are disproportionately awarded discretionary raises.

Although in our SUNY studies we were unable to document gender bias in the awarding of administrative positions using categorical modeling, we did not include this information in our regression analyses. We did this based on the substantially higher proportion of white-male faculty that had received administrative appointments. In our opinion, if former administrators are allowed to keep the extra salary stipends associated with their previous administrative roles, including this variable is likely to mask bias in salaries.

Initial rank

If women are hired at lower initial ranks than men with similar credentials on a systemic basis, and if the variable initial rank is included in the analysis, then initial rank, like height, could be interpreted as explaining salary disparities more appropriately attributed to gender.

Studies reveal that department chairs, deans, and members of faculty search committees prefer curricula vitae attached to male names over the same vitae attached to female names (Fidell 1970; Steinpreis, Anders, and Ritzke 1999; Davidson and Burke 2000). Fidell found that the descriptions attached to male names were, on average, 10 percent more likely to be judged as deserving appointment at the tenured level (associate and full professor) than those attached to the female names.

Consistent with Fidell's findings, our categorical modeling results at SUNY institutions indicated substantial gender bias in initial rank assignment. Women were consistently at least two to four times less likely than comparable men to receive senior rank (associate or full professor). At two universities and one two-year college, women were more than six times less likely to be awarded initial ranks of associate and full professor than men with similar highest degrees, years since highest degree at hire, and years of experience prior to hire. Racial minorities were more likely to be hired as lecturers and instructors rather than assistant professors.

These findings of bias in initial rank are consistent with the results of a study at the University of Hawaii. Male faculty there "have more than twice the chance of women of being hired initially into upper ranks" (Hauser and Mason 1993, 10). Both the Hawaii study and the analyses of twelve SUNY institutions lacked data on the relative supply of male and female candidates, information that would ideally be included in any study of bias in initial rank. Nevertheless, in our opinion, the levels of bias shown for initial rank were striking and suggested bias in hiring practices.

To test the effect of omitting initial rank, we ran regression analyses with and without this variable. Although for many of the twelve SUNY institutions, the amount of salary bias indicated did not change substantially, at one university the bias indicated by the white-female coefficient decreased by about 30 percent when initial ranks were included in the model. This school was also the university with the greatest amount of gender bias in the assignment of initial rank. This result offers a clear case of a tainted-rank variable masking salary bias.

Based on this information, we omitted initial rank from our SUNY analyses. We would also note that it is illogical to include initial rank in accounting for current salaries. The rank assignment that should account for current salary is current rank.

Initial salary

Although we have not experienced pressures to include initial salary in current salary analyses, other researchers have. In negotiating the variables to be included in a study examining gender bias in salaries at the University of Nebraska, for example, the bargaining unit representatives discovered that the university administration intended to include starting salary as a variable to explain existing salaries (Finkler et al. 1989). The problem with including such a variable involves both the likelihood that there is gender bias in starting salaries and the substantial redundancy or overlap between initial salary and current salary variables. In all likelihood, this overlap will mean that most of the variance in current salaries will be explained by the variance in starting salaries, leaving little to be explained by any of the other predictor variables, including gender and race.[5]

If the administration on your campus suggests such an approach, you may find it helpful to know how this issue was handled by the AAUP bargaining unit representatives at the University of Nebraska. The administration claimed that starting salaries were not gender-biased and, therefore, could be used in the analysis. The basis for this claim was a regression analysis with starting salary as the dependent variable. However, this analysis included only faculty members hired since a previous adjustment for discrimination in 1972. This limitation meant that only 266 of the total 422 faculty members were included. The results of this regression analysis showed a disadvantage for women of $1,231 in starting salaries, with a significance level of 0.0522 (see glossary). The administration claimed that since this finding was not statistically significant (the significance level was not at or below 0.05), the variable starting salary was not gender-biased and could be used in the current-rank regression analysis.

The bargaining unit researchers responded that although a significance level of 0.0522 meant that the salary disadvantage could technically be considered statistically insignificant, it was not evident that there was gender equality in starting salaries. They conducted two additional analyses of starting salaries. The first analyzed the initial salaries of *all* 422 current faculty, rather than just the 266 hired after 1972. That analysis found gender bias at $1,038 with a significance level of less than 0.02, which is highly significant. The second study looked at all of the 600 faculty hired since 1974 regardless of whether or not they were still on staff. This analysis found gender bias at the level of $810 with a statistical significance level of less than 0.01, indicating a very high degree of significance. The administration dropped its insistence on the inclusion of starting salaries.

There are two morals to this story. First, multiple regression can assess gender bias in starting salaries just as it can in current salaries. Unless starting salaries can be shown to be bias free, they should not be used as a predictor variable for current salary. Second, amounts that would not reach statistical significance in smaller data sets may become statistically significant with larger populations or data sets. Whereas bias at the level of $1,231 fell short of statistical significance when the data set contained only 266 faculty, bias at the level of only $810 was statistically significant when the data set included 600 faculty. This is important background information for our discussion of the "significance of significance."

Problem 5: The Significance of Significance

If our results concerning salary bias are "significant," they are assumed valid. The question is, if regression results indicating salary bias are not statistically significant, are they invalid? The answer is no; to understand why, we must look at the technical meaning of statistical significance.

Significance levels are also called probability levels. Probability is a less misleading term, because what is being measured is the probability of replicating the findings in another sample. Probability levels were developed for use with inferential statistics. Inferential statistics make inferences about a whole population based on a sample. If we wanted to draw inferences or conclusions that extended beyond the faculty members in our data set, significance levels would be important. If we took a random *sample* of faculty members on a campus and found a salary difference related to gender or race, we would use a significance test to estimate

whether this finding was due to chance or whether we could expect the same difference if we selected another sample or examined the entire population.

Most faculty salary studies directly examine the entire faculty population at a higher education institution, not a sample of that population. Nelle Moore (1993) notes that faculty salary-equity studies should always involve total populations: "There is no reason to draw a sample when the complete data are available" (119). Any salary difference found in that population is an *actual* salary difference between the white male and female or minority faculty categories.

When you study a population of faculty members, there is no sampling error, because there is no sample. But when you ask the computer to do a multiple-regression analysis, it does not know that a population, rather than a sample, is being studied. It has been programmed to provide the sampling error and related levels of statistical significance on the presumption that it is looking at a sample. Since the figures are right there, it is difficult to ignore them. But the focus should be on substantive rather than statistical significance.

If women or minorities are, on average, affected by even a few hundred dollars, this could be a "significant" amount to them. At one of the four-year SUNY colleges with a faculty of only 117, the regression result for being female was -1,488, with a significance level of 0.06. In other words, on average, female faculty members were paid $1,488 less than male faculty, all other independent variables held constant. Is this a substantial amount of money for those involved? Given that $1,488 is an actual difference within a population and not an inferential finding from a sample, should a salary disparity of this magnitude be ignored?

It is particularly important to understand that statistical significance is directly tied to the size of the database. (Note the example of the University of Nebraska study in the previous section.) The smaller the data set, the less likely that the results will be statistically significant. A regression analysis on a data set with fewer than two hundred faculty members is unlikely to yield a statistically significant finding. With large data sets, a small dollar difference can be statistically significant but of little practical significance.

The size of the race-gender category being examined is also important. Even when you have a large data set of, for example, over a thousand faculty, if the number in the group being assessed is very small, the amount of bias indicated has to be substantial to achieve statistical significance. If there are twenty Asian women professors, and the results indicate bias in their salaries, this result is more likely to reach statistical significance than if there are only two Asian women. We do not have to be statisticians to understand that. The salary disparity calculated from a population of two Asian women is more likely to be affected by unusual and random factors than it would be based on a population of twenty Asian women.

The implications of this connection between the number of faculty in a category and statistical significance bear noting. An institution with hiring practices that keep the number of women or minorities low could conceivably use the resulting lack of statistical significance to exonerate them of salary bias. This possibility underscores the suggestion that significance levels should be considered just one piece of information in the interpretation of the results. If, for example, there are very few Latino men on the faculty and the results indicate a substantial but not statistically significant bias in their salaries, this result should be interpreted in light of patterns for this campus. If there are few women and minorities and the analyses show that they are paid less than comparable white men, there may well be bias against the Latino men. If, on the other hand, this college has hired many Latino women and other minorities and if these groups are paid relatively fairly, the results regarding Latino men might be considered to be chance findings.

The issue of database size in relation to statistical significance can become the focus in struggles between those who do and do not want the study to indicate salary bias. For example, in court cases the defendant university will argue that deans or even department heads determine salaries and that, therefore, smaller units should be studied separately. Those suing the university for salary discrimination will argue that the provost or president has the final say on salaries and the oversight responsibility to make sure that bias in salaries does not exist, so the analyses should be conducted at the university level.[6]

In our opinion, the appropriate approach is to recognize the limited role of statistical significance when a population is being studied and, therefore, to look at the substantive importance of the results. A small salary difference at a large university may not be important even though it is statistically significant. A large salary difference at a small college may be important even if it is not statistically significant. If eliminating a salary discrepancy could improve morale, recruitment, retention, and fairness, asking if it is statistically significant is the wrong question.

Nelle Moore notes that statistical significance has been confused with substantive importance in part because of the normal usage of the word *significant*.

Noting that significance testing provides an illusion of objectivity built on fantasy (Carver 1978), she concludes that it has no meaning in salary-equity studies of institutional populations. She notes, "There is nothing random about the data, about the hiring process, or about the awarding of salaries. There is no sampling procedure. There simply is no context within which the use of significance tests could be considered appropriate" (N. Moore 1993).

Gray (1991, 1993) is less emphatic about ignoring significance levels. She agrees that no question of inference exists when you study a population. However, she suggests that statistical significance still be noted as an indicator of whether the observed difference could be the result of random variations in faculty salaries. For example, salary increments for promotion may be smaller in a low-funding year. Being both a lawyer and a statistician, Gray notes that many courts have moved from "complete disregard of the concept of statistical significance" to an almost "mystical reliance" on the 0.05 significance level. She also indicates that there are judicial decisions to suggest that the intent to discriminate can be inferred from sufficiently large statistically significant differences (Gray 1993).

We recommend that tests of statistical significance be used as one piece of information in weighing the importance of the results. The importance of any salary differences found should be evaluated in light of the general pattern of the findings and should not be strictly a statistical decision. We agree with Snyder et al. (1994) that the absence of statistical significance should not be viewed as proof of the absence of bias. The authors note that if there is a pervasive pattern of negative residuals for the female variable across ranks and colleges, there is cause for concern despite an absence of statistical significance.

Problem 6: Divide and Hide

The objective of most multiple-regression studies of faculty salary equity is to check for systemic gender and race bias at the institutional level. But in studies at some institutions, the total institutional analysis (the forest) is skipped over in favor of separating the tenure track from the nontenure track or senior ranks from junior ranks. At this tree level, the decision is to look at separate colleges or schools within the institution. Within a set of disciplines or a large discipline, the suggestion is made that the problem is really due only to a few departments. At the department level (the twigs), the problem is attributed to a few individuals. So we get down to the leaves, the individual female and minority faculty members.

Why bother to use regression analysis if you have discarded the idea that discrimination is systemic and affects all those in the group?

Failure to understand that regression analyses mathematically control for discipline and rank may lead policy makers to accept the misguided practice of studying only subsets of the faculty population. Studies of bias in faculty salaries include the variables discipline, rank, degree, and years of experience in order to control any salary differences related to these variables while gender and race are examined. Thus, by design, these analyses do not address different jobs, but rather very similar ones, for example, associate professors in discipline x. The result is an equal-pay-for-equal-work study, where gender and race differences *across* disciplines or ranks are statistically controlled. There is no reason to conduct separate studies of the different disciplines and ranks to accomplish this.

The problem with subdividing the population into discipline or rank groups is that the data sets become smaller and smaller. As a result, the regression analyses may be less valid. As previously noted, for regression analysis, it is recommended that there be at least five observations (faculty members) for each independent variable in the model (Blalock 1979). With fewer people, the analysis has less information to go on. This will be reflected in the deterioration of the significance levels reported in the computer output. Even if significance levels are ignored, the significance of the results will be less clear when the data set and the numbers in the individual race-gender groups are small. For valid results, stick with data sets that have at least five faculty for each independent variable.

There is some diagnostic appeal to looking for the institutional subsets in which bias is more prevalent. See chapter 7 for a discussion of when and how to validly assess subunits like certain ranks and disciplines for salary inequities.

Problem 7: Stepwise Abuse

The statistical procedure called stepwise regression produces results that at best are difficult to interpret and at worst can mask a substantial gender or race salary disparity simply by ignoring gender and race.

There are two types of stepwise analyses: forward and backward. The forward stepwise procedure is the more problematic for salary studies. In this procedure, the researcher enters a long list of variables whose effects are unknown, to discover which meet the test of a specified level of significance. The procedure selects the predictor variable that has the strongest relationship with the dependent variable and enters it on the first

step. Next, it searches for the second variable based on the second strongest correlation with the dependent variable and enters that variable on the second step. This process continues until the procedure finds no more variables that are related to the dependent variable at the level of statistical significance specified. Such fishing expeditions or exploratory studies are sometimes necessary, but Pedhazur (1982) and N. Moore (1993) note that the researcher should be responsible for selecting the variables, not the computer.

Stepwise procedure has its proper uses, but in salary studies it is commonly used to conceal gender and race relationships. If the forward stepwise procedure is used with a restrictive level of significance, the stepwise process may stop before the gender and race variables enter the analysis.[7] The race and gender variables are then declared "insignificant," and salaries are declared bias free. This procedure is not designed for use when variables that are predictive of the dependent variable are as well recognized as they are for salary analyses. Moreover, since stepwise regression relies blindly on statistical significance, it should be avoided for all of the reasons indicated in our discussion of significance levels.

If those conducting the research at your institution insist on using a stepwise analysis, make sure that they use a backward stepwise procedure. Pedhazur (1982) notes that the use of the forward stepwise procedure is problematic and that the backward is preferred. Backward stepwise regression starts with all of the variables in the analysis and drops those variables that do not reach the significance level specified, one at a time. Thus, the computer output for the initial step of this analysis allows you to view the results for the gender and race variables even if these variables are dropped in later steps for failure to meet the specified significance level. Unless their intent is to hide the magnitude of the gender or race relationship by using stepwise procedures, the administrations should not object to using the backward (rather than forward) stepwise procedure and providing the computer output for the initial and subsequent steps.

While backward stepwise procedures are more acceptable than the forward approach because they do not hide the effect of gender and race, the results are difficult to interpret. Because they do not meet the significance levels required, some of the dummy variables do not enter the analysis. The smaller your faculty, the more likely some dummy variables will be dropped by the analysis. This means that some ranks enter the analysis and others do not. What does it mean, for example, when lecturers and leading professors are left out of the analysis because there are too few individuals in these categories to reach statistical significance? For one thing, it means that both ranks are treated like the default or reference current-rank category, that is, the dummy variable rank that has been left out of the analysis or set to zero. If full professor is the default, both lecturers and leading professors are treated as full professors. If instructor is the default, both lecturers and leading professors are treated as instructors. Similar problems occur when a highest degree or discipline dummy variable is dropped by the backward stepwise procedure.

In our opinion, there is no reason to use stepwise procedures—forward or backward—to assess gender and race bias in faculty salaries. Faculty salary-equity studies are explanatory, not exploratory, analyses.

Problem 8: Dropping of Gender and Race

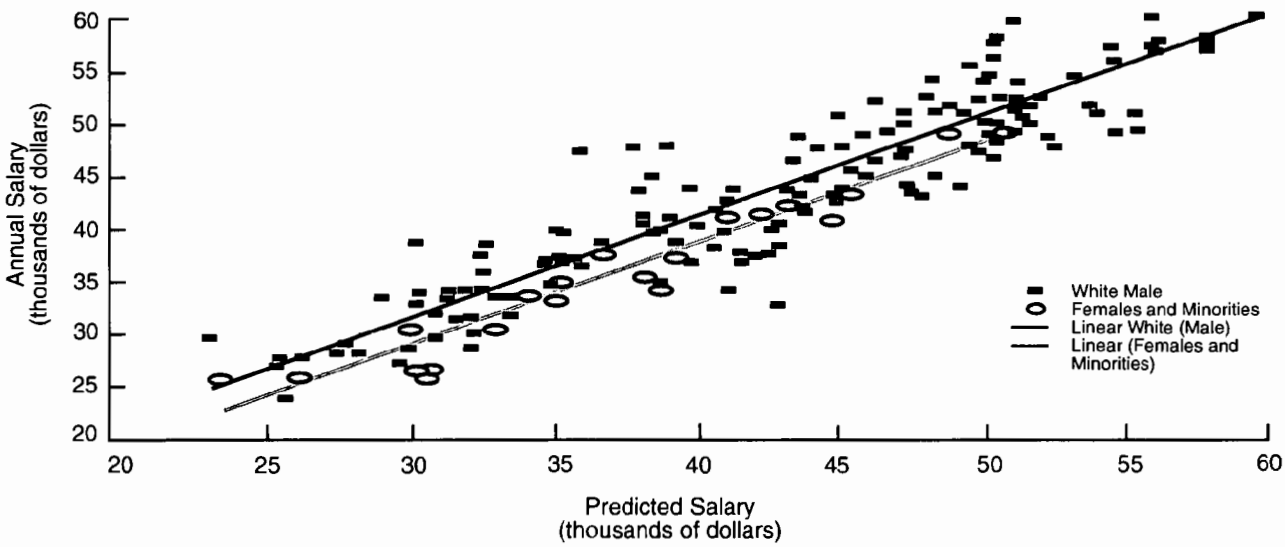
Stepwise regression is not the only way that race and gender variables get excluded from multiple-regression salary studies. Some salary reviews, even when they purport to check gender and race bias, leave these crucial variables out of the analyses. The common excuse for doing so is that since race and gender should not be allowed to play a role in determining pay, they should not be used to create the *predicted salary* of the individual faculty member (see appendix A).

This sounds good in theory, but in practice the predicted salaries are not the actual salaries. If we could get rid of the effects of gender and race on actual salaries just by leaving these variables out of regression analyses, how wonderful that would be. What we need is to know how much of the variation in actual salaries is attributable to race and gender. If the variables are omitted from the equation, the degree to which they explain actual salaries is never determined. It is easy enough to remove race and gender effects from predicted salaries. For example, if being female is found to cost $1,500 on average, you simply add this amount to each woman's predicted salary. Alternatively, you can calculate all faculty members' predicted salaries as if they were white males.

In addition, omitting these variables from the regression model embeds any gender or race bias that may be present against white men, making them appear statistically to be overpaid. This happens because, in the absence of any information on gender and race, the regression procedure takes the entire population into account, averaging the salaries of all faculty. Compared to this overall regression line, most white men will appear to be overpaid. By contrast, when race and gender variables are entered, white men are compared to their own average line. Half of the white men are above the line and half below it, so their average residual is zero. This is true for each of the race and gender groups.

Figure 6.1

Salaries for a Two-Year College, State University of New York



Note: This figure is a scattergram. See glossary for definition.

Figure 6.1 helps to make this clear. This graph shows a scatter of actual salaries against predicted salaries for a SUNY two-year college. Note that most of the female faculty are in the lower part of the scatter and most of the male faculty are in the upper part.[8] The dark line represents the regression line of the male scatter, and the light line, that of the female scatter. If the gender variable had not been entered in this analysis, the regression line would be between these two lines, pulled down by the lower salaries of the women faculty. More of the male faculty scatter would then be above the average line, making more men appear "overpaid." If we gave all females enough money to bring their average salary up to the average for the total population, general scatter for the female faculty would still be below that of the male faculty, indicating that there was still bias in pay.

Without the race and gender variables, the regression procedure creates an overall pay line that, because of the lower salaries of female faculty, makes the male faculty look overpaid and fails to provide the proper estimate to remove the bias in the female faculty pay.

Summary

In this chapter I have discussed eight ways methodological decisions can affect the outcome of the analyses so as to mask gender or race bias. This list is far from exhaustive. Every new research design brings with it the potential for subjective methodological decisions that influence the results. I hope, however, that this chapter has alerted you to issues to guard against and questions to ask concerning the research design and results.

Notes

1. If you are stuck with interpreting standardized coefficients, divide the standard deviation of the dependent variable by the standard deviation of the independent variable. (The multiple-regression computer output from most statistical software packages provides the standard deviations for the independent variables.) Then multiply the result by the coefficient of the independent variable. This lengthy, unnecessary process gives you the unstandardized regression coefficient dollar estimates.

2. Chronister et al. (1997) report that the proportion of non-tenure-track faculty has increased to 27.3 percent, with non-tenure-track women faculty now making up more than half of the faculty at two-year and master's-level institutions.

3. Our regression analyses examine whether faculty hired within the last five years at a specific institution are less likely to have gender or race bias in their salaries. We usually find little evidence of less bias in the salaries of recent hires. By contrast, Toutkoushian (1998), using 1988 and 1993 data from the National Center for Education Statistics, finds evidence that the wage gap for younger women is smaller than that for older women in academe. Given the difference in methodologies and the populations being studied, there could be a number of explanations for these different findings.

4. If your institution separates out remedial educators, thereby providing an institutional mechanism through which they can be paid less than others in their discipline despite their need for special teaching skills, you may be able to use the occasion of this study to correct their salary inequities. At the request of Equal Opportunity Program members at a SUNY university, we assigned these educators to the discipline they teach, rather than the discipline of "remedial education." The average difference between their predicted and actual salaries gives an estimate of the degree to which they have been systemically underpaid.

5. Gray (1990) noted that a justification offered for using starting salary as a variable to predict current salary is that if raises have been equal for men and women, there is no current discrimination. She cites a 1986 Supreme Court decision that employers have a continuing obligation to equalize salaries and that giving equal raises is not sufficient.

6. Ferree and McQuillan (1998) note that competing implicit conceptions of discrimination also play a role in these struggles. See chapter 7.

7. Most computer procedures set a nonrestrictive significance level recognizing that fishing expeditions are best when they allow many variables, even those with low predictive power, to enter the analysis. SAS, for instance, has a default significance level of 0.50 for forward stepwise procedures and 0.10 for backward stepwise procedures. If, however, the researchers so choose, they can set much more restrictive significance levels. That was done for a faculty study at a university in Ontario. Forward stepwise regression was used with a significance level of 0.05 rather than the common, less restrictive significance default of 0.50. Not surprisingly, the gender variable did not enter this analysis. Lois Haignere was asked by the faculty association to comment on this study. She used institutional data from SUNY to demonstrate that a less restrictive significance level would have, in all likelihood, revealed a gender effect.

8. The only minorities in this data set are four Asian males.

At first, we were surprised that our argument for an across-the-board settlement—that it would be fairer to the best women on the faculty—was rejected. Ultimately, we imputed this resistance to both the administration's tacit view of discrimination as an individual process and its desire to keep salary information and decisions in its own hands.

—Myra Mary Ferree
Professor of Sociology
University of Wisconsin

—Julia McQuillan
Professor of Sociology
University of Nebraska

SEVEN

Diagnosis Dynamics and Treatment Turmoil

By Lois Haignere

The researchers analyzing faculty salaries at your institution are likely to reach a point at which they believe the job is done and a diagnosis of whether or not salary inequities exist has been made. Other parties in the process may have other ideas. In this chapter, we note the practical side of the political dynamics of diagnosis as the parties progress from research results to remedy.

Diagnosis Dynamics

Before looking at remedies, faculty and administration should attempt to come to an understanding of the meaning of bias findings.

Two views of bias

Underlying many debates over how to study and correct faculty salary disparities are assumptions about how discrimination comes to be embedded in salaries in the first place. Ferree and McQuillan (1998) have described the two primary conceptualizations of discrimination as the *institutional* and *individual* perspectives. The institutional perspective views discrimination as systemic, generally affecting all those in the women or minority category in question. The individual perspective sees discrimination as resulting from isolated personal prejudices that cause pockets of salary disparity.

According to the institutional-systemic view, the basic reason for gender bias in salaries is that women and women's work have traditionally been undervalued. There is a pervasive cultural attitude that women are second-class citizens and, by extension, their work is worth less than that of men. This cultural devaluing of women and their work permeates all realms of our society—our psychological, political, and economic existence. "The market, as it functions in the daily lives of people, is not independent of the values and customs of those who participate in it" (Kessler-Harris 1990, 117). Paying women less than men for equal work was not made illegal until 1963; the acceptability of paying women less remains an implicit social norm.

"Gender is present in the processes, practices, images, and ideologies, and distributions of power in the various sectors of social life" (Acker 1992, 567). Gender distinctions are created and maintained by everyday social interaction and accumulated past practices and policies. A basic tenet of the institutional-systemic perspective is that historical and ongoing prejudice becomes embedded in institutional processes, and the resulting policies and practices undervalue most, if not all, women workers. The purpose of a faculty salary study is to identify and to propose institutional solutions for systemic biases in the salaries of women and minorities.[1]

By contrast, the individual view of the potential for gender and race bias in salaries is that the market tends to reward human capital fairly. Thus, a year of education or experience or the attainment of a higher rank will be equally rewarded in the salaries of women, minorities, and white men. Intervention is rarely needed because the market is generally fair. Isolated personal prejudices can exist, however, causing pockets of salary disparities. The purpose of a salary study under the individual perspective is to find the few individuals whose salaries have been affected by personal prejudice and adjust their salaries accordingly. Depending on the

findings, a secondary objective may be to remove the prejudiced person(s) from hiring and salary assignment responsibilities.

Articulating these two different conceptualizations of the source of salary bias early in your institution's process of examining salaries can facilitate discussions about the methods you use and the remedies you recommend.

Fit between perspectives and parties

You may find, as Ferree and McQuillan (1998) did at the University of Connecticut Health Center, that these perspectives tend to be associated with the parties involved in the salary discrimination studies: the faculty and the administration. Ferree and McQuillan reported that the faculty members on the health center committee believed that salary inequity existed on the institutional level, as the result of gender and race stratification that was historical, pervasive, and ongoing. Ferree and McQuillan interpreted the administration's behavior to be a reflection of their belief that gender- and race-based salary inequities did not exist, or if they did, that they stemmed from the isolated prejudices of a few individuals.

Note that these are not necessarily competing or mutually exclusive perspectives. Holding the view that historical and systemic gender or race bias is transferred to salaries through societal and institutional processes does not rule out also believing that biased individuals can exacerbate bias in salaries in their particular departments or colleges. Faculty members and administrators can, and frequently do, hold both perspectives.

There is also national variation in whether the institutional-systemic perspective is espoused. In our experience, both administrators and faculty members in Canada, New Zealand, and other Commonwealth countries are more likely than their U.S. counterparts to hold the institutional-systemic view of salary bias. In the United States, both faculty members and administrators tend to espouse the individual perspective. For example, at an East Coast public research university, the women faculty had previously won the right to use the white-male regression equation to calculate their individual predicted salaries for comparison with their actual salaries. They believed this information empowered them individually, despite the fact that the university analyses showed salary inequities for women as a class that were not being addressed.

Fit between perspectives and study results

In our experience, if bias is found at the university level, administrators may push for analyses of the nooks and crannies (colleges, departments) in hopes of making it disappear or confining it to a few specific subunits. On the other hand, if bias is not found at the university level, women and minority faculty are likely to lobby for further analysis of subunits to see if pockets of bias exist.

Overall, however, the tendency in the United States concerning choice of perspective is consistent with the experience of Ferree and McQuillan: administrators are far less likely to espouse the institutional-systemic perspective on bias than are faculty. Like Ferree and McQuillan, we have experienced many settings in which administrators press for further analysis conducted at smaller and smaller institutional subunits (see "Problem 6" in chapter 6).

Despite the administration's affection for the individual perspective, never have we had a brave administrator say, "Well, it is gratifying to know that we have no evidence of bias in salaries at the university-wide level, but I suspect we have bias in salaries in the college of XYZ. Would you please check there?" Such observations seem to come only from the faculty. Frequently, they come from the faculty even if there is substantial evidence of systemic bias at the institutional level.

Pockets of Bias

Pockets of bias may exist. Determining their effect as distinct from that of institutional bias is a difficult and sensitive problem.

Institutional-level checks

In the best-case scenario, there is a common appreciation among faculty and administration of the methodological fit between multiple regression's focus on class differences and the logic of the institutional-systemic perspective. A common agreement on this perspective would mean that all parties accept the institutional-level results. This does not mean that you fail to check whether the individual perspective and pockets of bias can explain your findings. Ruling out the presence of pockets of bias may be accomplished in two ways: through use of interaction terms and through examination of the pattern of the residuals.

As explained in "Improved analysis" in chapter 5, you can check for interactions that indicate that your results may be due to pockets of bias rather than systemic bias. You may get a clue that findings differ based on discipline or tenure status by examining the coefficients for these categories and comparing the white-male with the total-population results. If you observe substantial differences, first check the number of faculty in the suspect categories. Might the differences be due to one or two aberrant salaries? Are there more than five faculty members in the categories involved? Next,

rerun the total-population analysis with a race-gender interaction term (see chapter 5) and observe the specific impact. For example, you might surmise that average lower salaries for women arise from salary bias among non-tenure-track faculty and does not exist among tenure-track faculty. To test that possibility, enter into the analysis an interaction term for non-tenure-track women. If the coefficient for this interaction term is negative and substantial while the coefficient for the female variable (which now represents all female faculty except the non-tenure-track women) is much less negative or even positive, the salary bias you observe relates primarily to non-tenure-track faculty.

The second method of checking for pockets of bias is to examine the pattern of the residuals using the white-male faculty regression equation. Ferree and McQuillan (1998), who recommend this method, note that:

> If the average difference between men's and women's residuals is explained by some individual women being grossly and unfairly underpaid—as the individual model suggests would be the case—there will be a bimodal curve, with a second "hump" near the tail end of the negative residuals. Such a "hump" would represent a cluster of women with abnormally low salaries who pull down the average for the whole group (14).

If you find such an abnormal picture when you plot the residuals, further analysis may reveal that the faculty members in this "hump" are disproportionately in a certain school, discipline, or hiring cohort. In turn, you may find that this effect relates to a prejudiced administrator or administrative process that has a constricted impact.[2] (See "Residuals: The Ungreat Unknown" in chapter 5 for more information on residuals.)

Subunit analyses

If no agreement exists on the institutional-level findings and pressures exist for subunit analysis, the challenge is to proceed in a methodologically correct way. Multiple-regression statistical analyses are group-level analyses of systemic bias. They should not be applied to the individual level. The general rule of five cases for each independent variable should be respected. Many colleges within large universities are large enough to sustain separate regression-model analyses and still have five cases (faculty) for each independent variable. However, just having the numbers necessary to analyze subunits does not, in and of itself, legitimate such analyses. Consider the reality of how salaries and salary increments are assigned at the institution. Do the subunits actually have substantial input in determining salaries? If so, then analyses at the subunit level may be appropriate.

Case Study

The SUNY-UUP joint faculty-administration committee met on a regular basis for over a year to decide how best to adjust for salary bias revealed by multiple-regression studies at twenty-nine SUNY institutions. At first, both the union and the administration assumed that salary adjustments would be individually determined. The question that began to haunt us was: what level of data accuracy was needed before we were justified in giving money to one woman and not another, or to one African American man and not another, or to several white women but only one Latino man?

Despite close attention to data accuracy and diligence in data cleaning (see chapter 3)—plus a general pride in the quality of our data—we did not believe our data were accurate enough to estimate individual-level salary disparities. If we had to validate the data for every individual woman and minority man at twenty-nine SUNY institutions, we would still be doing the study today.

We began to see the wisdom of an institutional-systemic approach to awards, similar to the one used in adjusting faculty salaries by the University of Connecticut at Storrs in 1988. There is the advantage that errors are assumed to be random, eliminating the need to validate every piece of data studied. Although consideration of the institutional approach was initially motivated by methodological concerns, the other advantages to this approach became apparent, including the extension of the adjustment to all women, including the superstars; the adjustment of the female scatter to more closely approximate the male faculty scatter; the fit between multiple-regression class or group results and the adjustments of salaries for all those in the group; and lower total costs.[3] These advantages are discussed more fully in the next section.

We concluded that using group rather than individual disparities provided the best approximation of what the salaries of women and minorities would have been in a completely gender- and race-blind society. Thus, we gave all of those in the gender and minority categories who had a negative coefficient the same award based on the amount of negotiated money available.

Institutional-Systemic Remedy

The institutional approach assumes that the effect of gender and race on salaries is systemic, affecting all those in a

Figure 7.1

given gender and race category. In other words, the undervaluing of workers based on gender and race affects the "superstars," the "duds," and the average performers. Why should the highly productive females have actual salaries that are lower on average than the highly productive males? Similarly, why should the substandard women be paid less, on average, than the substandard men? Gray (1990) states that "discrimination affects the salaries of the best, the poorest, and the average woman faculty member"(7). Any remedy should address the entire class.

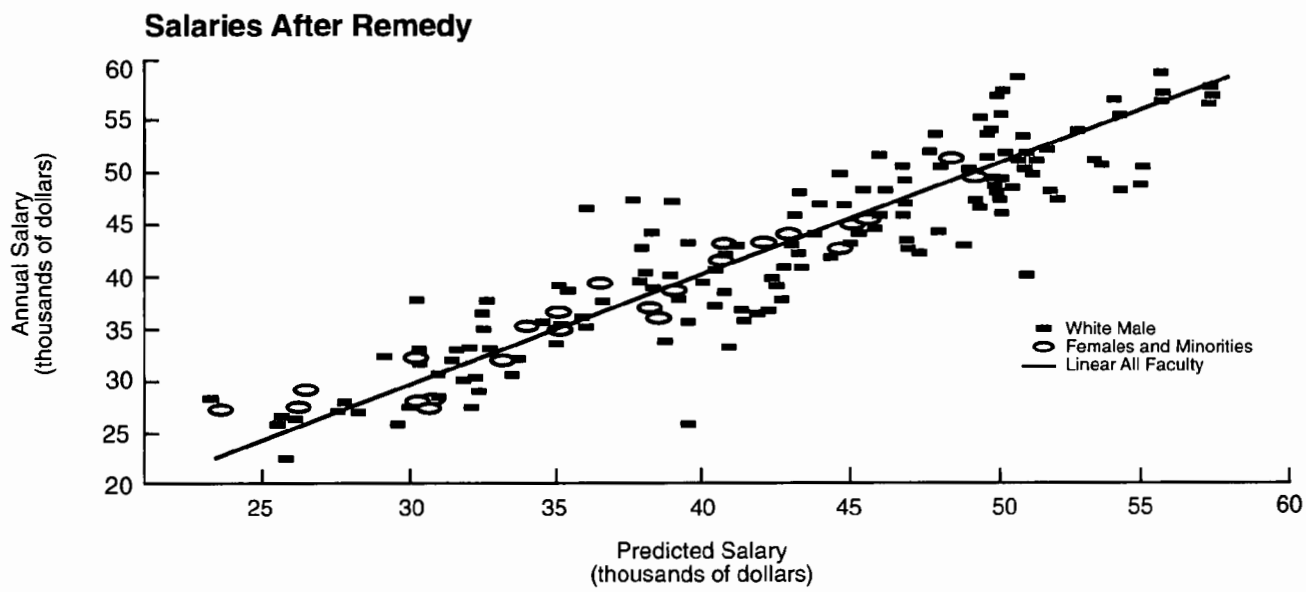
In fact, an emphasis on group or class differences, rather than individual differences, is a more appropriate use of multiple-regression statistics, because multiple-regression results, like averages, indicate class, rather than individual, differences (Gray and Scott 1980). For example, suppose the regression equation indicates that women faculty members receive $1,200 less a year on average than comparable white-male faculty members after controlling for rank, discipline, years of service, and the other predictor variables. This does not mean that there are no faculty women who are paid above the average received by comparable men. Neither does it mean that there are no white men paid less than women or minorities. What it means is that it is less likely that white men make less than comparable women and minorities and that it is less likely that women and minorities make more than comparable white men.

Applying the group approach to salary awards means that the distribution of women and minorities' residuals (or the scattergram of their actual and predicted salaries) will be more similar to that for white men. The highest-paid women and minorities will have salaries more like the highest-paid white males, and the lowest-paid women and minorities will have salaries more like the lowest-paid white males. Figures 7.1 and 7.2 show the effects of this approach on a SUNY two-year college.

Figure 7.1 plots the actual salaries (vertical axis) against the regression-predicted salaries (horizontal axis) for each faculty member. Each rectangle represents a male faculty member's predicted and actual salaries, and each oval represents a female or minority faculty member's predicted and actual salaries. The scatter for the women and minorities is lower than that for the men, and separate lines representing the general trend of the scatter (the line of "best fit") have been plotted for each group. Raising the salaries of all those in the women and minorities category by the total amount of their negative coefficient has the effect of moving the female-minority best-fit line up to coincide with the male line (figure 7.2). The scatter around that line will persist so that relatively equal proportions of the white-male scatter and women and minorities' scatter are above and below that line.

While the group approach creates equalization across gender and race groups, it does not change the distribution of salaries within these groups. Most white males in the UUP bargaining unit expected the gender and race adjustments and knew that they came out of

Figure 7.2



specially negotiated money that could not be used for any other purpose and would not draw away from annual increments or discretionary awards. However, women and minorities were expecting some corrections in their salaries. Awarding increments to some women and minorities but not others would have created a difficult situation. For the twenty-nine SUNY-UUP institutions, awards were based on institutional-level results. If the multiple-regression analyses indicated no bias for a particular race-gender group at a particular college or university, the individuals in that group did not receive a salary adjustment. When bias was found for women or a minority male group at an institution, each individual in that group at that institution received the same salary increase.[4] Under the group approach, no women or minority faculty saw others in their same race-gender category leap ahead of them in salary. The group approach to remedy recognizes the underpaid women and minority stars and confronts the cultural pressures to limit them to salaries that are average or below those of white men.

Longevity

The most senior women and minority faculty members may have suffered more bias simply because of the compounding effect of time.[5] Gray (1990) recommends adjusting for seniority either by introducing an across-the-board adjustment with a seniority bonus or by basing each individual's adjustment on the number of years at the institution. The seniority bonus approach could, for example, give a bias increment to all faculty in an underpaid race-gender category and, in addition, a longevity bonus to those with more then ten years of service to the institution. Alternatively, the total remedy could be based on years of service. For example, if the regression results indicate that, on average, each person in a race-gender category is underpaid by $1,000 and the average time at the institution is ten years, then each female or minority can receive $100 for each year at the institution. Thus, a faculty member who has been at the institution for five years would receive $500, and someone who has been there for fifteen years would get $1,500.

A percentage increase is sometimes suggested as a way of correcting for the compounding effect of bias over time. The presumption is that the highest-paid individuals have been at the institution longest and therefore should be awarded proportionately higher bias corrections. We do not recommend this approach. As multiple-regression studies demonstrate, many factors other than longevity contribute to high pay. A person hired last year as a full professor in a prestigious discipline could receive a much higher award than the many women and minorities in disciplines that are low paid (Bellas 1994).

Flagging and Other Remedies

The group-award approach just described is popular with researchers because of its consistency with the multiple-regression method. It is popular with faculty organizations because of its broad equitable approach to

the problem. But it is not very popular with deans, department heads, and others who have been responsible for the salary-setting process.[6] These administrators prefer to view gender and race bias as a rare occurrence and, therefore, to "flag" those women or minorities whom the analyses show to be underpaid relative to their predicted salaries. Gray (1990) notes:

> Any suggestion that one use the regression model merely to identify individuals whose salaries are below what is predicted for them and raise their salaries to the predicted amount must be resisted; even more invidious is the notion that these, and only these, cases will be examined to see whether the low salaries are "justified" (7).

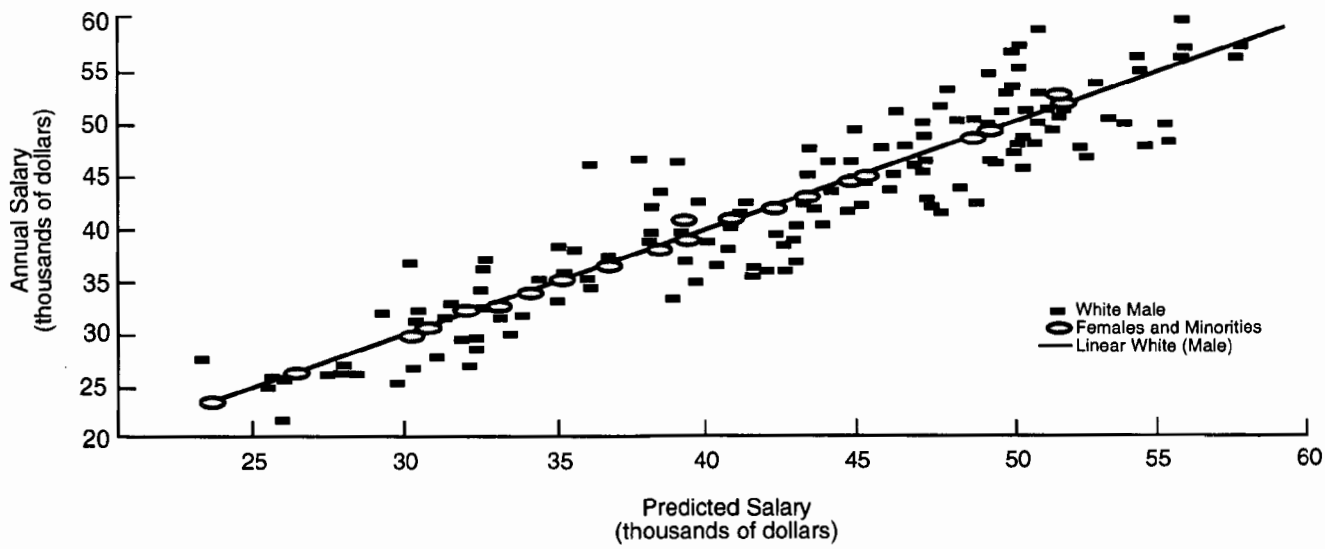
Salary bias identified by multiple regression is by definition not individual, but pertains to class or systemic differences (Gray and Scott 1980). Accordingly, it is controversial to base remedies on the individual-level predicted salaries provided by the multiple regression. Multiple-regression results, like averages, indicate class, rather than individual differences. A class can be any group membership such as a rank, discipline, highest degree, gender, race, or hiring cohort—but not an individual. Flagging, which uses multiple regression to focus on individuals and individual corrections, is, therefore, inappropriate.

The most common flagging approach locates all the women and minorities whose actual salaries are lower than their predicted salaries and raises their salaries to the predicted levels. In other words, the salaries of women and minorities whose residuals are negative (who are below the white-male line) are brought to the prediction line. Only white males remain below the line. Those white males whose actual salaries are below their predicted salaries do not receive any adjustments and, therefore, end up being paid less than the adjusted women and minorities (see figure 7.3).

Any remedy that involves only those whose predicted salaries are below their actual salaries is misguided. When the regression coefficient for any group or class studied is negative, everyone in that group is, on average, paid less than everyone in the default group. For example, if the default rank is associate professor and the variable for assistant professor has a negative coefficient, this indicates that, on average, all assistant professors are paid less than associate professors. To assume that being an assistant professor affects only those assistant professors that are paid below the associate professor line misuses this finding. Similarly, if being in a liberal arts rather than a natural science discipline is shown to, on average, cost a faculty member $400, it is wrong to assume that the only liberal arts faculty members affected are those paid below the natural science discipline average.

Figure 7.3

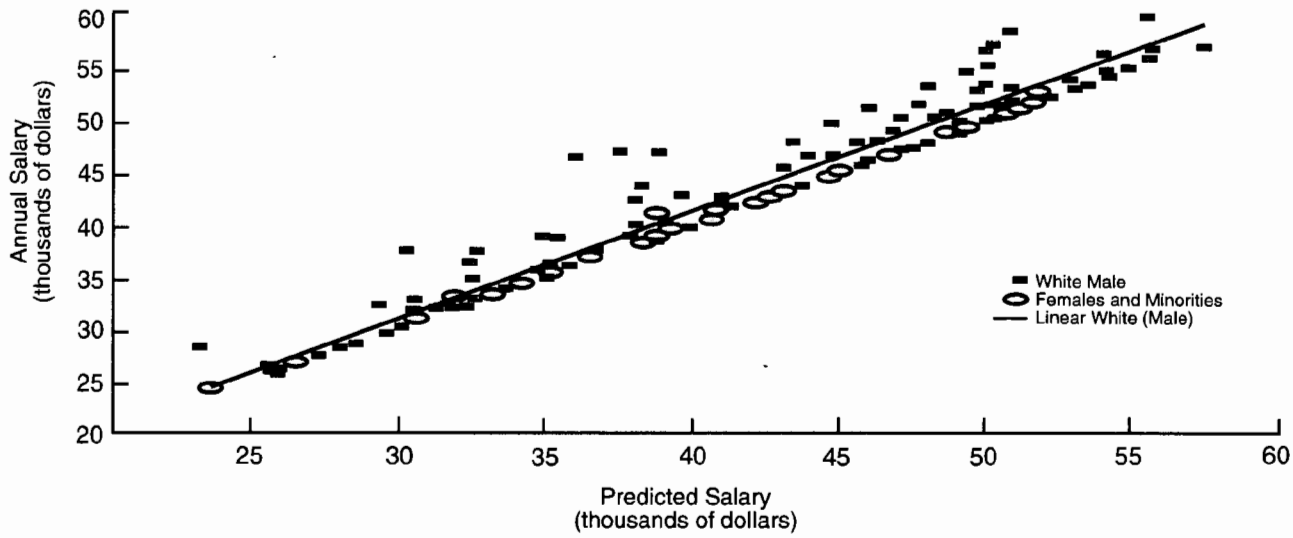
Below-the-Line Remedy



74

Figure 7.4

Below-the-Line Remedy Extended to All



Moreover, a number of practical problems arise with the application of this flagging approach. The most obvious one is that leaving all of the white males below the line while raising the women and minority faculty members' salaries to the line increases the potential for reverse discrimination allegations. This can lead to a second problem. Faculty organizations sometimes attempt to raise the salaries of all those below the line to the line. Such an adjustment aggravates the gender bias in salaries rather than eliminating it (figure 7.4). Raising salaries of the many white males below the line lifts the regression line itself, so that a substantial majority of the female and minority faculty members are paid below that line. This could lead to chasing the line—an activity like chasing your tail.

Another variation on this flagging approach is to "allow" all the women and minorities whose actual salaries are below their predicted salaries to apply for individual case reviews. Case reviews can involve pairing an individual woman or minority with a comparable white male or small group of comparable white males to illustrate the need for adjustment (Holmes-Rovner et al. 1994). Case reviews are lengthy processes, necessitating the development of criteria for comparing faculty members, and focusing attention on the issue and related controversy for an extended period of time. Such comparisons tend to become accusatory, competitive, and contentious, perhaps leading to recrimination, defensive re-action, and exacerbation of any race or gender animosity.

Whether or not case reviews involve paired comparisons, they have the obvious drawback of using the same decision makers and institutional structures that created the discrepancy in the first place, perhaps even requiring self-incrimination. And what happens to monies that are not awarded? Does the administrative unit that does not award them retain them? (See Snyder, Hyer, and McLaughlin 1994.)

Case reviews assume that bias is individual, not systemic. Under this assumption, no reason exists to conduct a multiple-regression analysis. Statistical methods do not adequately address the individual level. Even if they did, the data available for most salary analyses are not adequate or appropriate for suggesting remedies for individual cases of salary disparity. Individual faculty salary disparities are better explored through qualitative approaches.

If you use multiple-regression analyses and find indications of gender or race bias in faculty salaries, consider a class-based remedy consistent with that statistical method. Remedies that are distributed equally to all those in the affected group can be applied easily, efficiently, promptly, and without prolonged attention to the issue.

Remedy approaches that do not include the women and minorities at the top risk confirming the stereotype that women and minorities are low performers. Many highly successful minorities and women may acquiesce to such an approach because they feel apologetic about having more power, status, and rewards than others in their gender and race groups. Given that they are already better off, they may be reluctant to insist on the

real value of their work and to compare themselves with white men. But fairness is more than just bringing up the bottom. When elite women and minorities get paid more fairly relative to white men, they make it easier for all others in their race-gender group to be more fairly treated.

Toward a More Permanent Solution

When members of the AAUP's Committee on the Status of Women in the Academic Profession first proposed, over twenty years ago, that the average salaries for men and women faculty be published separately, they optimistically presumed that "unmasking salary discrimination would be the prelude to eliminating it" (Gray 1988). But the problem persists. Both faculty and administrators assume there is less bias in the salaries of more recent hires. Too often this assumption is false.

In chapter 1 we noted that prejudice is resistant to change because it is based on "unconscious ideology"—a set of beliefs that we accept implicitly, but which we are not aware of because we cannot conceive of alternative interpretations of the world (Top 1991; Bem and Bem 1970).

The author of this chapter experienced the resilience of unconscious ideology directly in explaining a faculty salary-equity study and related salary awards to a faculty audience. A particularly unhappy white-male associate professor of physics explained that he and a woman faculty member had been hired at the same time six years earlier with what he admitted were equal credentials. They had both been promoted in the same year, four years after hire. He had been paid $1,300 more annually than she until the gender-equity adjustment. Under the remedy, she would be making $300 more than he was making. In the six years since their hire, he had made $7,800 more than she. At $300 a year, it would take her twenty-six years merely to catch up. He had not been bothered when the inequity was in his favor, but when the tables turned in a nonstereotypical direction, he felt the injustice keenly.

Statistical analyses have little impact on such unconscious ideology. Gray notes that "[t]he courts' refusal to accept well-constructed statistical models as evidence of discrimination is not based on statistics, but on an unwillingness to accept that statistics can prove what the decision maker does not want to believe" (1993, 154). Confronted with analyses of faculty salaries showing evidence of discrimination, often the very faculty whose stock in trade is to persuade others of the efficacy of statistics refuse to believe what is presented to them (Gray 1993).

We hope the information we have provided helps you become more proactive in your efforts to identify and, if appropriate, remedy salary and promotion disparities. Attention to the fairness of the many interrelated aspects of academic life that influence the campus reward system, including recruitment, hiring, promotion, tenure, counteroffers, and the enticement of faculty "stars," can help combat unconscious ideology.

Formal procedures that protect employees in general protect against the subtle influence of gender and race bias. Salary schedules that stipulate entry-level salaries or attach specified criteria for salary setting deter unconscious prejudice. Any standardization of the processes used in awarding initial salaries, discretionary salary awards, research resources, promotions, and other honors can help prevent gender and race bias.

The methods we describe in this guide tend to be used as one-time fixes for many higher education institutions. They focus attention on the problem for a short time. Although salaries are adjusted at a moment in time to correct some of the historical impact of unconscious ideology, the subtle drift toward gender and race bias continues. Periodic multiple-regression salary reviews and diagnostic analyses of rank assignments should be conducted, and they should be independent of the salary setting and rank assignment processes.

We know how difficult it is to reach agreement on the assessment of bias in salaries or rank once, let alone repeatedly at regular intervals. But a powerful deterrent to the subtle drift of gender and race bias is "the spotlight." Higher education institutions such as American University, the University of Wisconsin-Stout, and North Carolina State University have implemented a process of annual or other periodic review and adjustment. Other institutions set the time period for the next review when they finish the current study; the University of Maine System, having just completed studies of its seven institutions, has scheduled the next study to be in three years.

Of the twelve SUNY institutions used to test the methods for this guide, two (one two-year and one four-year college) showed no substantial gender or race bias, which indicates that salary bias is not inevitable. We believe that it is possible for higher education institutions to take the steps necessary to eliminate what systemic gender and race bias exists in faculty salaries and to maintain equitable salaries. We congratulate those that have done so and hope that this manual encourages others to achieve gender- and race-neutral faculty salaries.

Notes

1. As indicated in chapter 1, the research described in this guidebook does not address difference in pay that is related to

the underpayment of disciplines where women and minorities predominate. Discipline is one of the control variables. Therefore, comparisons are constrained by the regression modeling to those within discipline.

2. Snyder, Hyer, and McLaughlin (1994) suggest a variant of this approach as an alternative to conducting regression analyses on very small subsets. Once you have examined bias at the institutional level, you can use the predicted salaries and compare the average white-male and female or minority residuals within rank and discipline subsets. This approach could reveal the "humps" Ferree and McQuillan (1998) describe, if the residuals are calculated using the white-male equation and not the total-population equation.

3. Toutkoushian (1994a) notes that the across-the-board approach to salary adjustments is likely to be less expensive than individual adjustments because of restrictions on lowering salaries in the remedy process. For example, if there are four faculty women of whom three are underpaid by $1,000 and one is overpaid by $1,000, the average underpayment is $3,000 - $1,000/4 = $500. To give each of four faculty women $500 costs $2,000. If the actual salary discrepancy were paid to each individual with a negative residual, the cost would be $3,000.

4. To address concerns about very low outliers, we agreed that those whose residuals were more than one-and-a-half standard deviations below the mean would be submitted to each institution for special discretionary award consideration.

5. Just prior to the SUNY-UUP salary-equity awards, some longevity corrections had been made. As a result, the gender and race awards did not address longevity.

6. The researchers involved in two studies, Virginia Commonwealth University and the University of Hawaii at Manoa, recommended the group-award approach. In both cases, administrators chose instead to proceed with a "flagging" approach.

APPENDIX A

Multiple Regression and Gender- and Race-Equity in Salaries

By Lois Haignere

This appendix provides an introduction to the interpretation of regression statistics for salary-equity studies. Recognizing that it will be read by an audience with a wide range of mathematical knowledge, we have attempted to make it understandable to those who are not familiar with statistical techniques.

To begin with a simple example, assume that we are interested in finding out how some variables relate to body weight. These variables are shoe size, hours of exercise each week, eye color, fast-food meals consumed, height, and make of automobile. If we used multiple regression to relate these characteristics to body-weight data, we would expect some to be more strongly associated with body weight than others. We would probably find that make of automobile and eye color had no relationship to body weight. The amount of exercise each week might be negatively related to body weight—as exercise goes up, body weight goes down. Height, shoe size, and fast-food meals might be positively related to body weight—as they go up, body weight goes up. Among these positively related variables, we would probably find that height is more strongly related than shoe size and fast-food meals.

The particular strength of multiple regression is that it can isolate the effect of one of these variables while controlling for all of the others. For example, it can control statistically for height, shoe size, and fast-food meals while examining the impact of hours of exercise each week. Conceptually, we can compare a group of people of exactly the same height, wearing the same size shoes, and eating the same number of fast-food meals each week, differing only in their amount of exercise.

Instead of body weight, we are interested in explaining variations in faculty salaries. In particular, we want to estimate the effect of variables like gender and race while controlling for other important salary-related variables, like years of service and discipline. To explain how multiple regression works, we begin by considering how just one variable, say, years of service, explains differences in salaries. If we plotted the years of service against salaries, we would probably see a scatter plot similar to figure A.1. Even a casual glance at

Figure A.1

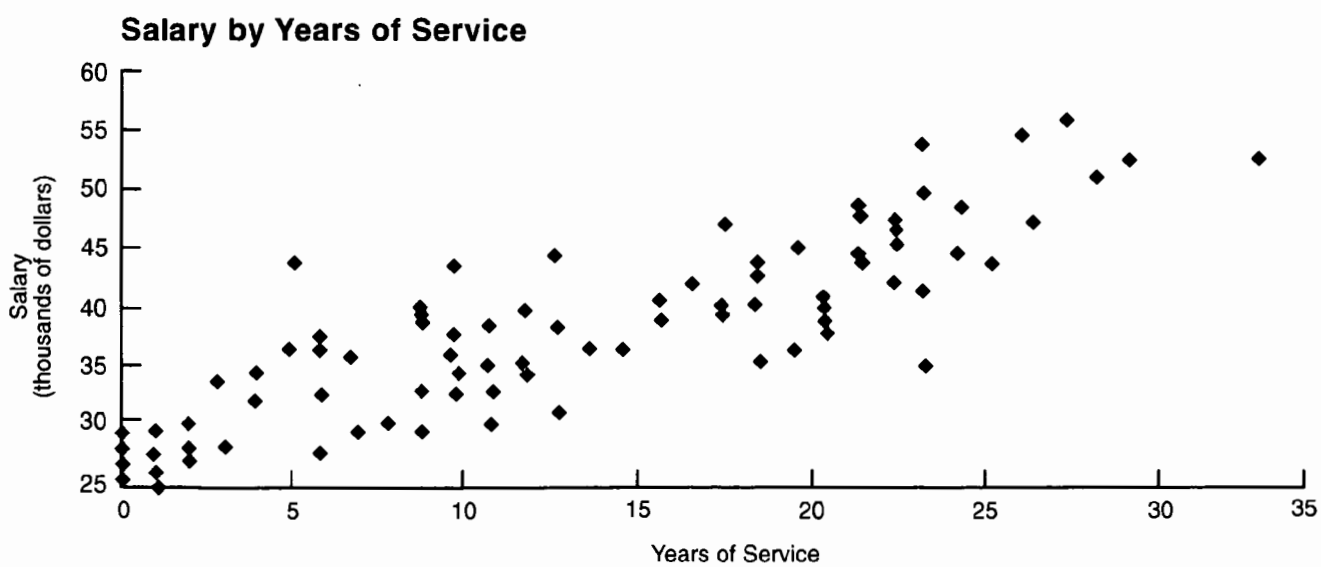


Figure A.2

this figure indicates that salary increases as years of service rise. Note, however, that the relationship is not perfect; every increase in years of service does not result in an equal jump in salary. If the relationship were perfect, all points would fall on a straight line.

To describe this relationship statistically, we could provide an equation that would estimate how large a difference in salary we would expect, on average, for individuals who differ by one year in their years of service. This is done by fitting these points with the line of "best fit" (figure A.2).

"Best fit" is a statistical criterion, indicating that the line is drawn to minimize its distance from the points scattered around it.[1] In other words, the line is as close to all points as a straight line can be. The slope of this line indicates the predicted change in salary for each year of service. For example, if we draw a straight line up from five years of service on the horizontal axis of figure A.2 until we reach the line of best fit and then draw a line over to the vertical axis, we will find the average *predicted* salary for faculty members with five years of service.

We do not have to have a graph or line of best fit in front of us to be able to predict the salary of those with five years of service. Regression analysis provides us with a formula representing the straight line on figure A.2. This line can be described by just two pieces of information: (1) the intercept, that is, the place the line starts on the vertical axis; and (2) the slope of the line (called the regression coefficient), which is the average increase in salary for a one-unit (year) increase in length of service.

This formula is: predicted salary = intercept point + slope of the line x years of service. It is the same as the formula we learned for a straight line in basic algebra: $Y = a + bX$, where Y is the predicted salary, a is the intercept value, b is the slope of the line, and X is years of service.[2] Thus, for any number of years of service, we can easily arrive at the predicted salary.

Assume, for example, that the regression formula tells us that the starting point of the regression line (the intercept or a) is $29,000, and the slope of the regression line is $800. We can figure out that a faculty member with five years of service is predicted to have a salary of: Y = $29,000 + ($800 x 5 years of service) = $33,000. This example is a simple two-variable linear regression. Salary is the dependent variable and years of service is a predictor or independent variable.

Since we want to know about the effects of many variables on salary, we use *multiple* regression. Fortunately, the equation for multiple regression is a straightforward extension of the two-variable equation. Suppose we are looking at just two predictor variables, years of service and years in rank. The multiple-regression procedure might tell us, for example, that with the introduction of this new variable, our intercept has changed to $31,000, the unstandardized regression coefficient (the slope of the line) for years of service has changed to $700, and the unstandardized regression coefficient for years in rank is $800. For a

faculty member with five years of service, two of which have been in his or her current rank, the predicted salary (Y) would be: Y = $31,000 + ($700 x 5 years of service) + ($800 x 2 years in rank) = $36,100. But what happens when we try to include variables such as discipline, which is not a quantity? How can this become a part of our calculation?

Dummy Variables

The two independent or predictor variables we have thus far used in the example, years in rank and years of service, are continuous variables. That is, they take on a series of values, equal distances apart; each additional year of service or year in rank is equivalent to any other year of service or year in rank. Such variables can be entered into regression analyses in their current form. But many of the independent variables commonly used in studies of salary equity are not measured in equal intervals; that is, they do not have numeric value. Special steps must be taken to include them in the multiple-regression analysis.

Discipline, gender, race, and rank are called categorical variables; they represent categories rather than quantities. Some categorical variables (discipline, gender, and race, for example) cannot be ordered; others (like rank) have an order, but the differences between levels are not necessarily equal. For example, we do not know if the value difference between the ranks of instructor and assistant professor is the same as the value difference between the ranks of associate professor and full professor, or whether the rank of full professor is worth twice as much as assistant professor and four times as much as instructor. Similarly, we have no basis for deciding that being in the business and management discipline is worth twice as much as being in the education discipline, but only half as much as being in the computer and information sciences discipline. Regression analysis can actually tell us these relationships if we transform these variables by making them into what are called dummy variables.

Dummy coding is a way of quantifying variables that are basically qualitative or categorical in nature. For group membership variables (race, sex, rank, and the like) you need to convert each category within the variable into a separate variable. Each of these new dummy variables can have only two values: 0 or 1. For example, for the variable female, all women are coded 1, and all others are coded 0; for the variable assistant professor, we assign the value of 1 to those who are assistant professors and the value of 0 to all others. The transformation to dummy variables, therefore, involves an increase in the number of variables. Where there was originally one categorical variable called current rank, there are now five dummy variables, one for each rank category. Where there was originally one gender variable, there are now two—one for male, coded 1 and 0; and one for female, coded 1 and 0.

Of course, saying someone is female (female = 1) is the same as saying she is not male (male = 0); we do not need both categories. When entering a group membership variable into the regression analysis, one of the dummy categories is omitted. This is because you convey all of the information contained in the codes of the original variable with one less than the number of categories. If there are five categories of rank, anyone who is coded as zero in four categories must be in the fifth. The selection of the particular category to be omitted from the regression analysis does not affect the analysis, but it is simplest to consider the omitted category to be the logical reference. Thus for pay-equity studies it makes sense, for example, to choose white males as the reference group, omitting male as a gender category and white as a race. Similarly, it may pay to choose a well-understood rank category like associate professor as the omitted reference category, rather than lecturer, which is a rank whose use varies across institutions.

The estimate for the omitted category is represented by the intercept. For example, if the category male is omitted for gender, the category associate professor is omitted for rank, and the category social sciences is omitted for discipline, the salary at the intercept will be the estimate for the average salary of male associate professors in social sciences with zero years of service and zero years in rank. To calculate the average salary for any other group, the regression coefficient for that group is added to the intercept value. (In the case of a negative regression coefficient, the sum will be less than the intercept, because adding a negative amount to a number results in subtraction, thereby reducing it.)

How can the dummy variables be part of regression equations?[3] Our dummy variables each have only one of two values: 0 or 1. These values must be multiplied by a regression coefficient, a number we derive signifying the effect of that dummy variable on predicted salary. If we find that being female has an average effect on salary of -$900, then -$900 is the regression coefficient for female. In effect, being female (1) adds to the equation (1 x -$900) or -$900. This coefficient has no effect on the reference category male, since the intercept already represents the salary for those in the default categories.

Suppose we include the dummy variables for gender and discipline and leave out the reference categories of male and social sciences. Say that our multiple-regression equation indicates an intercept of $33,000

Table A.1

Predicted Salaries of Male and Female Faculty in Three Disciplines

Male in business

Intercept		Years Service		Years Rank		Business		Male		Predicted
$33,000	+	(3 x $700)	+	(3 x $900)	+	$2,500	+	0	=	$40,300

Female in business

Intercept		Years Service		Years Rank		Business		Female		Predicted
$33,000	+	(3 x $700)	+	(3 x $900)	+	$2,500	+	-$900	=	$39,400

Female in social science

Intercept		Years Service		Years Rank		Soc. Science		Female		Predicted
$33,000	+	(3 x $700)	+	(3 x $900)	+	$2,500	+	-$900	=	$36,900

Male in fine arts

Intercept		Years Service		Years Rank		Fine Arts		Male		Predicted
$33,000	+	(3 x $700)	+	(3 x $900)	+	-$400	+	0	=	$37,400

and these regression coefficients: years of service = $700; years in rank = $900; fine arts = -$400; business = $2,500; female = -$900. Using this intercept and those coefficients, Table A.1 estimates the salaries of male and female faculty members with three years of service and three years in rank in the disciplines of business, social science, and fine arts.

Remember that the categories of social science and male are the defaults and, thus, the intercept represents the salary for faculty members in these categories. This is why nothing is added or subtracted for these categories in the formula. You can see by these examples that the parameter estimate (or unstandardized coefficient) for the dummy variable female is a measure of how much on average it costs a faculty member to be a woman, assuming that all the other variables in the equation are held constant. Similarly, dummy variables for race, such as African American and Latino, can indicate the average effect of each race category.

Validity of Regression Equation

It is important to know how to judge the validity of different regression equations. Returning to the body-weight example, we could run a regression equation with variables like eye color and make of automobile that do not strongly relate to the dependent variable. The result would be a fancy equation that would not tell us much. Multiple regression provides an estimate of how well the set of independent or predictor variables (eye color or shoe size) accounts for the variation in the dependent variable (individual body weight). This measure is called the adjusted R^2 (R-square). An adjusted R^2 of 0.75 indicates that 75 percent of the variation in salary is accounted for by the predictor variables in the equation; an adjusted R^2 of 0.55 indicates that 55 percent of the variation is accounted for by the variables.

Another way of conceptualizing this is in terms of the scatter of points around the "best fit" line in figure A.2. The smaller the scatter of observed points around the line represented by the regression equation, the better the prediction and the closer the adjusted R^2 is to 1. If there is no association between the predictor variables and the dependent variable (that is, the scatter is random and does not tend to form a line), the adjusted R^2 = 0. In the social sciences, an adjusted R^2 below 0.3 is generally thought to indicate little or no association. Those in the range of 0.4 to 0.6 are considered to indicate moderate association. Those above 0.7 are considered to show strong association, indicating that most of the variations in the dependent variable have been accounted for by the independent or predictor variables.

Interpretation of Results

Table A.2 is an example of typical computer output from a multiple-regression analysis of faculty salaries for an institution we call Proxy College. At the top of that illustration, the adjusted R^2 results are reported as 0.8211. This means that 82.11 percent of the variation in salary is accounted for by the variables in the equation. The remaining 17.89 percent could be due to random factors, measurement error, or variables left out of the equation. An adjusted R^2 of this magnitude is an indication that the variables in the equation explain most of the variation in salaries.

Table A.2

Sample Computer Output for Proxy College: Regression Analysis of Faculty Salaries

Variable	DF	Sum[a]	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEPT	1	335	29495	994.62906029	29.654	0.0001
YR_RANK	1	2927	544.348571	42.44607444	12.824	0.0001
YR_SERV	1	4988	336.390498	48.57062401	6.926	0.0001
ASST	1	81	-5447.273440	634.67422069	-8.583	0.0001
FULL	1	134	5951.380714	455.91364220	13.054	0.0001
MASTERS	1	101	-539.921096	808.87423663	4.005	0.0001
BACHELORS	1	4	-1076.643425	1641.2392210	-0.656	0.5123
AGIRESRC	1	11	5032.979425	1053.0875729	4.779	0.0001
ARCHENVR	1	9	3988.706165	1135.4294293	3.513	0.0005
BUSINESS	1	9	6457.295117	1170.5822575	5.516	0.0001
BIOLOGY	1	14	4456.221675	967.48669527	4.606	0.0001
AREASTDI	1	13	4976.437719	1004.8530098	4.952	0.0001
COMUNCTN	1	8	441.642358	1159.3391086	0.381	0.7035
COMPUINF	1	10	2922.576103	1067.8918658	2.737	0.0066
EDUCATIN	1	8	1422.662763	1155.0865960	1.232	0.2191
ENGNERIN	1	15	2393.906709	936.39183011	2.557	0.0111
FINEARTS	1	6	2380.802276	1340.5263086	1.776	0.0768
FORGNLAN	1	3	3548.019256	1724.3069450	2.058	0.0405
HEALTPRF	1	5	1738.377402	1395.4345988	1.246	0.2138
HOMECNMY	1	5	1588.998793	1376.3579156	1.154	0.2492
LAW	1	5	1356.105647	1378.8903794	0.983	0.3262
LETTERS	1	5	4060.422238	1467.0322010	2.768	0.0060
LIBRARY	1	8	791.285924	1178.9245073	0.671	0.5026
MATH	1	14	473.654141	947.47812117	0.500	0.6175
PHYSICS	1	6	568.258532	1281.6642459	0.443	0.6578
PSYCLOGY	1	8	1243.279501	1146.2883723	1.085	0.2790
PUBSERVC	1	9	1476.943558	1106.7881700	1.334	0.1831
THEOLGY	1	17	466.892901	908.79433574	0.514	0.6078
FEMALE	1	117	-1016.832795	389.18698941	-2.613	0.0694

Note: Dependent variable: salary. Dummy variable defaults: male, social sciences, Ph.D., associate professor.
a. SAS output usually includes this column elsewhere with other simple statistics. We have transplanted it into this chart so that it is beside the parameter estimate (unstandardized regression coefficient column).

To illustrate the common appearance of multiple-regression computer output, table A.2 includes columns titled Standard Error, T for HO, and Prob > T even though (you will be happy to note) these three columns can be ignored by most faculty salary analyses. They are important for inferential statistics, which make inferences about a population based on a sample. Faculty salary studies are typically not based on samples. Most include the entire population of faculty at a given institution, so interpretation of inferential statistics is not needed or meaningful. (See the discussion of "Problem 5" in chapter 6.)

The left-hand column in table A.2 identifies the independent (predictor) variables. The next column, DF, indicates the degrees of freedom. Each variable has one degree of freedom associated with it. The next column, Sum, is the sum of the variable for all cases in the equation.[4] For dummy variables, the sum tells the number of cases in that category. We see that there are 81 assistant professors and 134 full professors included in the equation.

The next column is headed Parameter Estimate. The specific type of parameter estimate shown in this column is the unstandardized regression coefficient that we have been describing. A single unit change in the variable results in a change in predicted salary that is shown by the parameter estimate. As previously indicated, when dummy variables are used in a regression equation, one category for each group-membership variable must be omitted from the equation. In table A.2, the omitted variables are listed in the table note. In this case, they consist of male for gender, social science for discipline, Ph.D. for educational attainment, and associate professor for current rank. With these omitted categories, the intercept, which is listed in the first row, would represent the salary for a male associate professor in a social science discipline whose highest degree is a Ph.D. This also explains why these variables are not found in the variable list of the first column.

We can look down this column to the regression coefficient (labeled Parameter Estimate in table A.2) for Yr_rank, and see that it is 544.348571. This means that if the individual's years in rank are greater than zero, we would multiply those years in rank by 544.348571 and add that amount to the intercept to get a more accurate estimate of his or her salary. If the individual is not an associate professor, but an assistant, we would add -5,447 (the unstandardized regression coefficient for assistant professor) to the individual's salary to improve our estimate. (The addition of a negative number actually amounts to subtraction.) The unstandardized regression coefficient for the variable female shows us that, even when controlling for all other factors in the equation, women at Proxy College are paid an average of $1,017 less than men. Again, this is indicated by the negative unstandardized regression coefficient.

To see if you understand this output, calculate the predicted salary for a full professor with a Ph.D., three years in rank, and ten years in service in the discipline of business. You should get a predicted salary of $46,895 if this faculty member is a male and $45,878 if the person is a female (rounding to the nearest whole number).

Notes

1. The line of "best fit" also creates an average residual of zero. In other words, the average difference between the actual and predicted scores is zero.

2. Another way to describe the intercept is the value of *Y* (salary) when the value of *X* (years of service) is zero. Thus, with our current example, it is what the average faculty member is paid if he or she has no years of service.

3. We have included only two dummy variables to keep the example simple. In an actual regression analysis of salary, other dummy variables such as current rank and highest degree would probably also be included.

4. SAS output usually includes this column elsewhere with other simple statistics. We have transplanted it into this chart so that it is beside the parameter estimate (unstandardized regression coefficient) column.

APPENDIX B

The Twelve SUNY Data Sets

By Lois Haignere

In selecting the twelve SUNY institutions on which we tested the methods explained in this guide, our first criterion was attaining the broadest possible cross section of different-sized data sets within each institutional type. A second consideration was the accuracy of the data and the ease of correcting any belatedly discovered data errors. The four selected two-year colleges included both the largest (244 full-time faculty members) and the smallest (99). The four selected four-year colleges ranged in size from 421 to 117 faculty members, and the four university centers, from 811 to 477. This variety in faculty size and institutional type has enabled us to observe the effects of the different statistical approaches on a range of institutional types and sizes.

As indicated in chapters 1 and 2, the AAUP publishes its Annual Report on the Economic Status of the Profession in the March–April issue of its bimonthly magazine, *Academe*. According to the descriptive categories used in these reports, the four SUNY university centers are "doctoral-level institutions," characterized by significant doctoral-level education, granting a minimum of thirty doctoral-level degrees. Although two of the SUNY universities include health science centers, the faculty at these centers were not included in these analyses.

Three of the four colleges are "comprehensive institutions," characterized by diverse post-baccalaureate programs, granting a minimum of thirty post-baccalaureate degrees. The fourth college is a "general baccalaureate institution," characterized by a primary emphasis on undergraduate bachelor's degree education. The four two-year colleges are classified as "two-year institutions with academic rank," characterized by conferring at least 75 percent of their degrees below the bachelor's-degree level. The SUNY two-year institutions are technical colleges, which resemble community colleges in that the faculty's main responsibility is teaching. Therefore, salary determinants at the two-year colleges more closely resemble those at community colleges than those at institutions with strong research missions.

Table B.1 displays the gender and race composition of the twelve faculty data sets. Clearly, white males are the predominant race-gender category at every institution. As a result, and because they are presumed not to experience bias, white males form the baseline comparison group for assessing salary equity. Testing for race and gender bias requires comparisons, and the appropriate comparison is the white-male reference category.

Table B.1

Faculty Composition by Race and Gender at Twelve State University of New York Institutions

	Total	White		African American		Latino		Asian	
Universities	**Faculty**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
Albany	634	430	122	7	10	15	10	31	9
Binghamton	477	323	95	10	5	11	2	29	2
Buffalo	811	560	124	25	6	7	3	79	7
Stony Brook	639	447	97	15	5	6	5	54	10
Colleges									
Buffalo	421	286	95	10	6	5	4	11	4
Geneseo	233	154	63	1	3	1	1	8	2
Oswego	311	211	71	6	4	4	1	12	2
Purchase	117	64	41	2	1	2	0	5	2
Two-Year Inst.									
Alfred	175	148	23	0	0	0	0	4	0
Canton	99	72	26	0	0	0	0	1	0
Cobleskill	132	88	34	2	2	1	1	3	1
Farmingdale	244	147	75	6	3	3	3	6	1

Appendix C

Salary-Equity Contract Language

By Donna Euben

Contract language addressing salary equity varies a great deal from contract to contract. While one contract may contain a detailed article establishing a joint faculty-administration salary-equity committee, another contract may simply include a new salary schedule that does not explicitly mention salary equity but that incorporates equitable salary rates. Some collective bargaining agreements include no salary-equity provisions, because such efforts are included in a memorandum of understanding, which is a separate side agreement. Contract clauses covering the following areas may be designed to help achieve salary equity: the makeup of, and procedures for, faculty-administration salary-equity committees; the amount of money available; the establishment of minimum salary floors; and the grievance procedure.

Regardless of the final language negotiated, your salary-equity contract provision should clearly establish that no professor's pay will be downgraded as a result of a salary-equity adjustment program. For example, the 1989–92 collective bargaining contract of the AAUP chapter at St. John's University stated that the information gathered for salary adjustments "shall be used solely for the purpose of determining the needs for upward salary adjustments. In no event shall the salaries of any faculty members be lowered as a result of these procedures, nor shall the foregoing procedures be used for any purposes other than salary adjustments."

The specific contract language that you bring to the table will depend on your union's salary-equity policy, as well as the particular circumstances in your university or college. Some or all of the following suggestions may apply to your situation.

Joint Faculty-Administration Committees

If you decide to bargain for a salary-equity study, the most important contract language is a clause that establishes the faculty union's role in this process, generally through a joint union-administration committee. The language should specify the committee's membership, selection mechanism, goals, the scope of the committee's work, the process the committee will use to reach decisions, the timetable for the committee's work, and a dispute-resolution procedure.

Often helpful in such joint committee discussions is salary data from the institution. The AAUP's annual salary survey, published in the March–April issue of the AAUP's bimonthly publication, *Academe*, provides a wealth of information on faculty salaries by institution, including wage differentials based on gender, which may be helpful to you in negotiations. For example, the 1995–98 collective bargaining agreement of the University of Cincinnati's AAUP chapter provided:

> *Equity Study*. In order to assure that the University has an equitable salary and rank structure, the parties to the contract hereby establish a joint committee to study these issues. The committee will be composed of four (4) individuals: two (2) selected by the Administration and two (2) selected by the AAUP. The Committee shall be named no later than the beginning of Fall Quarter 1993. The Office of Institutional Research should be charged with the responsibility to assist in this project; the report from the Salary Equity Committee details a comprehensive model for such a study and should be considered the guide for this study. The final report and recommendations from the Committee will be submitted no later than June 1, 1994, to the parties to the contract.

The United University Professions (UUP)–State University of New York (SUNY) collective bargaining agreement (1985–88) provided as follows:

> A joint SUNY-UUP committee consisting of three members appointed by the Chancellor and three members appointed by the President of UUP shall be established to review perceived salary disparities, including, but not limited to, retention and recruitment problems and salary disparities by title based on comparisons within or outside the University. The Committee's first priority shall be to investigate, consider and review instances where the salary of any

employee with ten or more years of University service is below the average salary for such employee's rank. Joint recommendations to correct demonstrated salary disparities shall be made by the committee to the Director of the Governor's Office of Employee Relations.

Money

Another step—either before or after the completion of a salary-equity study—is to negotiate contract language that sets aside a specific amount of money and delineates the controlling formula. The UUP-SUNY agreement established a "disparity fund":

> The State agrees to make every effort to address demonstrated salary disparities which may exist within the unit. To this end, the State shall apportion an amount of One Million ($1,000,000) dollars in the second year of this Agreement and Two Million ($2,000,000) dollars in the third year of this Agreement . . . to be used to correct demonstrated salary disparities. . . . The unexpended portion of any year's apportionment shall be carried over into the succeeding year and added to the apportionment for the succeeding year. These funds may not be used for across-the-board increases on a University-wide basis.

The salary-equity provision in the 1993–96 collective bargaining agreement of the Eastern Michigan University's AAUP chapter established a percentage amount:

> *Salary Differential Pool.* EMU-AAUP and the University recognize and are concerned that on occasion salary differentials may be created among faculty in departments owing to external market conditions. In order to address these concerns, an amount equal to seven-tenths (7/10) of one percent (0.7 percent) of Faculty base salaries, calculated as of February 1, 1994, shall be set aside and shall be distributed among faculty immediately following the 1994–95 salary adjustment . . . in accordance with the formula developed by the University and EMU-AAUP. Additionally, an amount equal to five tenths of one percent (0.5 percent) of Faculty base salaries, calculated as of February 1, 1995, shall be set aside and shall be distributed among faculty immediately following the 1995–96 salary adjustment . . . in accordance with the formula developed by the Association and the University.

Minimum-Salary Scales

Collective bargaining language that establishes minimum-salary scales often helps to mitigate salary disparities by limiting the salary gap (at least among the lowest paid faculty in each rank) that often emerges between male and female faculty. For example, the 1995–97 collective bargaining contract of the Cleveland State University's AAUP chapter included the following provision:

> *Minimum salary.* Effective retroactively to September 19, 1994, the following minimum salaries will be implemented: (1) Instructor-$28,000, (2) Assistant Professor-$34,000, (3) Associate Professor-$42,000, (4) Professor-$52,000.

Other helpful collective bargaining language ensures that existing faculty are brought up to these minimums. For example, the 1996–98 contract of Kent State University's AAUP chapter provided:

> *Salaries and Promotion: Academic Years 1996–97 through 1997–98:* . . . Salary Floors–Beginning Academic Year 1996–97. As a means of assuring appropriate entry-level pay at both academic ranks, the Employer will establish salary minima on a University-wide basis. The minimum annual contract salaries for Faculty members at the professional academic ranks of Instructor and Assistant Professor . . . shall be as follows: Assistant Professor (with terminal degree) $37,500 (nine months), $45,833 (twelve months). Instructor (nonterminal degree) $31,000 (nine months), $37,888 (twelve months). The salaries of all current Faculty will be brought up to these minima. No one's salary or rank will be reduced.

Similarly, the recent UUP-SUNY collective bargaining agreement (1995–99) provided for such salary minimums and that not only a new hire but "[a]n incumbent promoted on or after the effective dates . . . shall receive not less than the minimum basic annual salary for the rank or grade to which that incumbent has been promoted."

In addition, such minimums can be set for new faculty positions, which provide the union with the opportunity to ensure that equity is established at the outset.

The contract may spell out provisions for obtaining union agreement with the creation of new positions and the assignment of these positions to salary ranges.

Grievance Procedures

To the extent the law allows, salary disparities based on sex, race, or national origin, and so on should be subject to the grievance and arbitration procedures in a union contract. While it is difficult to recommend language that would be adaptable to all contracts because grievance procedures vary, the following points apply in many situations.

Any claims that arise under a fair-employment clause that prohibits discrimination on the basis of sex, for example, should be grievable and arbitrable, and should be broad enough to include dealing with salary inequities. In addition, the grievant should not be precluded from seeking other available external remedies. The current law is unsettled in this area, and it is important to confer with a lawyer who practices in your jurisdiction to determine any legal restrictions regarding the interplay between your contract and antidiscrimination laws.

A contract provision should require the employer to furnish payroll and other information, including after-salary adjustments and the source or type of adjustment, to the union in order to bargain or to administer existing contracts without barring disclosure of the information to the membership. (Decisions of the National Labor Relations Board entitle private-sector unions to this information. In the public sector, the information is available through bargaining laws and "freedom of information" laws.) For example, the 1989–92 collective bargaining agreement of the St. John's University AAUP chapter provided: "The Administration shall provide the AAUP-FA with a complete list of all salaries of continuing faculty members in the affected discipline after the adjustments have been made, identifying the adjustments made."

In the end, the ultimate language to which you agree is specific to your local or chapter and the higher education institution. The above suggestions are provided to foster discussion about the most effective ways for your faculty organization to address salary-equity issues in collective bargaining.

Appendix D

U.S. Laws on Gender-Based Wage Discrimination

By Donna Euben

The purpose of this guidebook is to encourage faculty and administration to work together in developing and implementing salary-equity studies and recommendations. Nevertheless, salary-equity studies are sometimes undertaken for, or used in, litigation in which professors challenge salaries set by administrations as discriminatory. The application of federal laws to salary equity in the academy can raise a myriad of complex legal issues. This appendix briefly reviews the federal laws that address gender-based salary inequity on higher education campuses. There are four legal bases for challenging such disparities under federal law in higher education: the Equal Pay Act, Title VII of the Civil Rights Act, Title IX of the Education Amendments, and Executive Order 11246.[1]

Equal Pay Act

The Equal Pay Act of 1963 (EPA), which is part of the Fair Labor Standards Act (FLSA), 29 U.S.C. Sec. 206(d), provides:

> [N]o employer [who is covered by the FLSA] shall discriminate . . . between employees on the basis of sex by paying wages to employees in such establishment a rate less than the rate at which he pays wages to employees of the opposite sex in such establishment for equal work on jobs the performance of which requires equal skill, effort, and responsibility, and which are performed under similar working conditions.

However, the EPA allows salary differences when such differences are based on: (a) a seniority system; (b) a merit system; (c) a system that measures earnings by quantity or quality of production; or (d) a differential based on any other factor other than sex.

Congress passed the law to fight the "ancient but outmoded belief that a man, because of his role in society, should be paid more than a woman" and to ensure that "'equal work will be rewarded by equal wages'" (*Corning Glass Workers v. Brennan*, 417 U.S. 188 [1974]). In EPA cases, the issues are whether the jobs in which an alleged pay disparity exist are "substantially similar," and whether any of the four exceptions mentioned above apply. The EPA is enforced by the Equal Employment Opportunity Commission (EEOC). (See D. Green, "Application of the Equal Pay Act to Higher Education," 8 *Journal of College and University Law* 203 [1981–82].)

Title VII of Civil Rights Act

Title VII of the Civil Rights Act, 42 U.S.C. Sec. 2000e et seq., prohibits discrimination based on race, color, religion, sex, or national origin in employment, including hiring, promotion, and dismissal. It provides, in relevant part:

> It shall be an unlawful employment practice for an employer—
> (1) to fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment, because of such individual's race, color, religion, sex, or national origin; or
> (2) to limit, segregate, or classify his employees or applicants for employment in any way which would deprive or tend to deprive any individual of employment opportunities or otherwise adversely affect his status as an employee, because of such individual's race, color, religion, sex, or national origin.

Wage discrimination based on gender constitutes sex discrimination under Title VII. In 1972 Congress amended the law to apply to all public and private educational institutions, including colleges and universities. Generally, Title VII cases, unlike those arising under the EPA, require proof of intent: "Under Title VII, in all but a few cases, the burden of proof remains with the plaintiff at all times to show discriminatory intent. In contrast, 'the Equal Pay Act creates a type of strict liability in that no intent to discriminate needs to be shown'" (*Fallon v. Illinois*, 882 F.2d 1206, 1213 [7th Cir. 1989]). Title VII complaints are filed with the EEOC or the state human rights commission counterpart.

Title IX

Title IX of the Education Amendments of 1972, 20 U.S.C. Sec. 1681 et seq., prohibits sex discrimination, including in employment, in all education programs or activities receiving or benefiting from federal financial assistance. (See *North Haven Board of Education v. Bell*, 456 U.S. 512 [1982].) In so doing, it also offers protections to women who work at educational institutions that receive federal funds. Title IX is administered by the Office for Civil Rights of the Department of Education. (See AAUW Legal Advocacy Fund, *A License for Bias: Sex Discrimination, Schools, and Title IX* [2000].)

Executive Orders

Executive Order 11246, as amended by Executive Order 11375, prohibits discrimination "because of race, color, religion, sex, or national origin," and mandates affirmative action for minorities and women. The order applies to federal government contractors and subcontractors, including colleges and universities. Such contractor agreements must include an equal opportunity clause, and contractors must file compliance reports with the federal contracting agency. Some state and local government contractors are exempt from coverage. However, "educational institutions and medical facilities" are excluded from that exemption. The executive order is administered by the Office of Federal Contract Compliance Programs of the Department of Labor. (See Annotation, "Right to Maintain Private Employment Discrimination Action Under Executive Order 11246, As Amended, Prohibiting Employment Discrimination by Government Contractors and Subcontractors," 31 *American Law Reports Federal* 108 [2000].)

Note

1. For a more thorough examination of the legal issues involved in salary-equity litigation in higher education, see Euben 2001.

APPENDIX E

Activist Strategies When All Else Fails

By Maita Levine

If you find that you can not gain the cooperation of the administration in designing the study in a valid way or addressing any salary inequities that may be found, you may find the list below helpful.

- Invite all faculty members, especially women, to a meeting to discuss the study.
- Use campus grievance procedures.
- Request meetings with individual members of the administration and the board of trustees to discuss solutions to the problem.
- Bring the issues to the bargaining table if you are on a campus engaged in collective bargaining.
- Obtain the names and addresses of alumni from the university development office and mail copies of your report or grievance to them.
- File a complaint and ask for an investigation by your regional office of the U.S. Department of Labor, Office of Contract Compliance.
- Publicize the results of the study as widely as possible, both on campus and in the community.

These activities assume an increasingly adversarial relationship. Since studying and correcting salary inequities involves reordering existing priorities, some of the more extreme strategies listed may be required.

APPENDIX F
Categorical Modeling

By Lois Haignere and Bonnie Eisenberg

Selecting an appropriate statistical approach is important. We chose to use as our categorical modeling method the most generalized logistic regression model, based on maximum likelihood tables, as opposed to (a) a logistic regression for an ordinal response variable, which satisfies the assumption of proportional odds; (b) a logistic regression for adjacent levels of an ordinal response variable; or (c) a discriminant analysis.

The logistic regression for ordinal response variables, which satisfies the assumption of proportional odds, was not used because in ten of our twelve data sets, the levels of the response variable (current rank) were not parallel with regard to predictor variable effects; that is, proportional odds could *not* be assumed. That would mean, for instance, that having previous experience or a Ph.D. has the same predictive value for full professors as it does for assistant professors. We could have used the logistic regression if we had included interaction terms, but that would have produced too many predictor variables in the model. To discover if all current ranks at your school are affected in the same way by the predictor variables, you can perform a score test for the proportional odds assumption.

A logistic regression for adjacent levels of an ordinal response variable did not work with our data, since the statistical software that we used allowed only for the weighted least squares method of estimation with this type of regression. This form of estimation is more sensitive to very small data pools than the maximum likelihood tables method. Even a faculty of more than eight hundred, when broken down by current rank and then further subdivided according to each of the predictor variables, produced very small data pools. So we chose the method least sensitive to small numbers.

When used for studying potential gender or race bias, discriminant analysis produces results that are more difficult to interpret than does the method chosen. Rather than generating an odds ratio for women and men for each pair of adjacent ranks, discriminant analysis requires two separate analyses to produce interpretable results, one for men and one for women. These analyses report the percentage found above, at, and below the predicted ranks. After these analyses are done, you must still make the comparison of these percentages for men and women. Thus, the results are not easily interpreted, requiring much manipulation before any conclusions can be made. In addition, discriminant analysis does not work with categorical *predictor* variables such as gender or race, although it can be used if these variables are translated into dummy variables.

The software that we used, SAS CATMOD, led to two complexities worth noting. First, each rank step required a separate model to produce results that were easily interpreted. Since the last level of the dependent or response variable was the one to which all others are compared, we needed to run a separate analysis for each pair of adjacent ranks. The first model would compare all ranks to assistant professors in order to analyze the promotional step from instructor to assistant professor; the second model would compare all ranks to associate professors to analyze the step from assistant to associate professor, and so on. Although there was an SAS option to analyze adjacent ranks in one model, this option used the weighted least squares method of estimation, which will not work on the very small pools of data found in current rank analyses.

The second complexity worth noting was that the results were in the form of the *natural log* of the odds ratios, not the odds ratios directly. Therefore, we had to add programming to create the inverse natural log of the estimate produced for each predictor variable before the odds ratios could be easily interpreted.

Data

To analyze gender bias in current rank, you must have enough men and women in all of the rank levels. If you have ranks that lack either men or women (zero cells), then you cannot use categorical modeling to investigate gender bias for these ranks.

To identify the zero cells in your data, produce a frequency table like table 4.1 that shows men, women, whites, and minorities in each current rank. If you have faculty members in each category (no zero cells), then you can perform categorical modeling that will provide you with information about bias for every promotional step. If some current ranks lack women or minorities, or white males for that matter, then you need to determine

whether your zero cells are critical ones that preclude using categorical modeling.

If you are analyzing both gender and race bias, the two must be examined separately for zero cells. As table 4.1 shows, it is possible to have critical zero cells for minorities that block race analyses but still have enough women to allow gender analyses.

The distribution of women faculty in table 4.1 also illustrates that zero cells will eliminate only some of the promotional steps from a study. The absence of female lecturers stops the model from computing gender bias from lecturer to instructor. But there are women in the instructor, assistant, associate, and full professor current ranks, so three pairwise comparisons are possible: (1) instructor to assistant professor; (2) assistant to associate professor; and (3) associate to full professor.

If there are zero cells, but not in every pairwise comparison, you will need to decide whether categorical modeling will provide enough insight to be worthwhile. For example, if the only step categorical modeling can assess is from instructor to assistant professor, you may decide that this information is not valuable enough to warrant running a categorical model.

Overall Model Statistic

It is valuable to know whether the predictor variables chosen for your categorical model are any good at predicting the response variable (the dependent variable). In a current rank analysis, this translates into: do the control or predictor variables truly influence current rank at your institution? If the control variables do not have substantial influence on rank assignments at your school, then any bias suggested by the findings can be minimized by the argument that the logistic regression model cannot accurately predict rank with the variable information available to it. The overall model statistic R^2_L (R-square sub-L) provides an indication of how well the predictor variables are predicting rank, so that you can have some measure of confidence in your results.

A categorical model analysis of current rank without any predictor variables shows the total amount of variability found in rank. This measure, called D_o (D-Naught), is the last iteration (or final estimate) of the -2 log likelihood found in the maximum likelihood analysis table. The amount of variability remaining after the predictor variables are included in the model is called the D_m (D-M, which stands for model). D_m is the last iteration of the -2 log likelihood produced for the model that includes the predictors. G_m is simply the difference between D_o and D_m, and represents the amount of variability accounted for by the predictor variables. Finally, the R^2_L (the proportion of variability accounted for by the predictor variables) is calculated by dividing D_o (the total amount of variability) by G_m (the variability accounted for by the predictor variables).

For categorical models, R^2_L values in the middle range indicate the most reliable results. A value for R^2_L of 0.8 is questionable because it is too close to perfect. At the other extreme, R^2_L values below 0.2 indicate that the variance accounted for is very little.

APPENDIX G

Redundancy Problems

By Lois Haignere and Bonnie Eisenberg

Statistical modeling tries to determine mathematically how much of the dependent variable (current rank or salary) differences are attributable to each predictor or independent variable. When two predictor variables overlap or measure the same underlying dimension, the results may be unreliable because the procedure cannot reliably determine which of the two predictor variables should get the credit for explaining certain current rank or salary differences. This problem is called redundancy or multicollinearity.

Redundancy is relative. There may be total redundancy, as when the same information is labeled differently; age and time since birth date would be completely redundant. Accordingly, one of the variables has to be dropped. There may be partial redundancy, as with the two variables time since degree at hire and experience prior to hire. If we use both of these variables in analyses, we test carefully to see if the redundancy between them could make the categorical modeling results unreliable. The redundancy between these two variables was not a problem at any of the twelve schools studied, due in part to the fact that some people acquire previous experience before they receive their highest degree, and some people do not get their degrees until after they are hired.

There is, however, a source of redundancy that we frequently find has to be addressed. Recall that to control for curvilinearity a quadratic term variable is used for each time variable in the analysis. To do this you square the variable. Not surprisingly, there is redundancy between each variable and its squared term, between time at the institution and (time at the institution)^2, for example. Initially, we tried dealing with this problem by dropping one of the two redundant variables. It is accepted practice to drop the quadratic term if it is not statistically significant. However, to determine whether or not the quadratic is significant or, alternatively, which variable is less important and, therefore, should be the one dropped, requires running the analysis twice, once to determine relative importance so as to decide which variable to drop and a second time to get the results.

We find it easier to just "center" these variables. Centering a variable is simple, and the redundancy is sufficiently reduced so that dropping a time variable or its quadratic term is no longer necessary. Just subtract the mean for that variable from each of the measures of that variable. For example, if the mean or average number of years a person has been at the institution is 5.4, then 5.4 is subtracted from each faculty member's years at the institution. The effect is that those with fewer than the average number of years at the institution (for example, 4) will have a negative score ($4 - 5.4 = -1.4$) and those with more than the average (say, 7 years) will have a positive score ($7 - 5.4 = 1.6$). The mean for this centered variable is zero (if there are no rounding errors). The quadratic or squared term is recalculated based on the newly centered variable.

At the twelve SUNY campuses we studied, centering eliminated all but one redundancy problem between time-related variables and their quadratic terms.

APPENDIX H

Promotion and Salary Inequities Between Men and Women Faculty

By Robert Johnson and Dorothy Kovacevich

For nearly a century now, studies have documented disparities in salary and rank between the men and women who teach in our nation's colleges and universities. Laws forbidding discrimination in salary or conditions of employment date back to Title VII (1972) of the Civil Rights Act of 1964 and the Equal Pay Act of 1963, yet the effects of discriminatory practices continue to trouble most institutions of higher learning in the United States. In 1990 the American Association of University Professors' Committee on Women and Minorities reported that most campuses were far from achieving pay equity, especially for women faculty members with many years of service. At the end of the 1990s, institutional discrimination on the basis of sex remained a primary source of salary inequity for women in the workforce. Evidence shows that women are hired and promoted less often and at slower rates than men (Johnson and Herring 1989). The combined effects of the market's devaluation of sex-integrated occupations and institutional discrimination by sex are often compounded when differences between academic disciplines are used to measure legitimate (nondiscriminatory) market factors. This problem is widely recognized by researchers who attempt to model salary inequities by sex (Bellas 1994).

Academic Rank As a Measure of Productivity

One attempt to resolve the problem of measuring productivity uses academic rank as an indicator of total productivity, on the assumption that people who attain higher ranks are more productive in research, teaching, and service than those who do not. Rank, it is argued, represents factors—the quality of a professor's research, teaching, and service—that go unmeasured in other models of salary inequity. The idea is that once you introduce rank and the sex difference in salary disappears, then there is no inequity. This notion has obvious flaws. Teaching, research, and service are not rewarded equally across ranks, nor across colleges or campuses. Some campuses emphasize their teaching mission over research. Others place equal stress on teaching and research, while yet others focus primarily on research. Few campuses or colleges emphasize service to the same degree as research or teaching. Rewards for teaching, research, and service can also vary across ranks. At the assistant professor level, for example, research productivity often carries more weight than teaching or service. On the other hand, teaching may become more important at higher ranks at the same institution.

A related problem is that rank is a minimum threshold measure of productivity, not a continuous measure. It cannot account for differences in levels of productivity that might distinguish the faculty member who meets the minimum criteria for promotion from the professor who far exceeds that performance threshold. Moreover, rank cannot account for increasing expectations of scholarly performance as the thresholds change for each new cohort of faculty. And rank most certainly cannot account for the variations in productivity that occur after promotion to full professorship. For a further discussion of the use (or misuse) of rank as a measure of productivity in regression models that predict salaries, see chapter 4.

Academic Rank As a Reflection of Status

Using rank to judge individual productivity has obvious flaws—yet promotion, rank, and tenure remain important predictors of salary. If they do not measure productivity, what accounts for their effects on salary? The simple answers are often overlooked. Promotion is a status transition, and rank and tenure represent achieved status. Status itself, regardless of differences in productivity, is something valued and rewarded in our society. We confer status on people we value highly (or who have attributes we esteem), and we withhold status from individuals we value less highly (or who have attributes we do not esteem). Thinking about promotion and rank in terms of status is more suitable than conceiving of rank as a measure of productivity. In a society that values men over women, the status of rank will be conferred more often on men than women, and withheld more often from women than men. Such sex bias is a strong argument against using a status attainment variable as a measure of productivity. Doing so

becomes particularly troublesome when the evaluators of women seeking to attain status are men, especially if these men rely on traditional sex-biased values in their evaluations.

Conceptual and Statistical Models

Some studies of salary inequity, assuming that the promotion process is probably not free of bias, have attempted to quantify the level of discrimination in the process. A study by Behling et al. (1994) applied a multinomial-ordered probit model to rank, which predicts the probability of a given faculty member being in a given rank. Using the model as described by Maddala (1992), the study concluded that the variable of rank therefore contained significant sex bias.

Other studies (Raymond, Sesnowitz, and Williams 1990) have attempted to estimate the conditional probabilities of attaining a given rank at a given time, making adjustments for those who do not meet minimum requirements (faculty who have fewer than five years in rank, for example). The adjustments simply restrict the sample to individuals who have met the minimum criteria.

Unfortunately, neither approach considers the timing of the promotion or allows for "early promotion," something believed to be more common among male faculty members than among females. An alternative to both approaches, which does adjust for the timing of the promotion, is the proportional hazard linear regression model. The "Data" and "Findings" sections of this appendix describe how the model was used to analyze promotion patterns at Kent State University. Before touching on the Kent State study, however, it is necessary to explain the hazard rate, which is the dependent variable in the model. The hazard rate can be thought of simply as the likelihood of an event happening at any given time. For example, how likely is it that, after ten years, a female faculty member will be promoted to the next rank? That likelihood is the hazard rate.

The hazard rate is usually examined in the context of descriptive tables known as life or survival tables. The hazard rate is depicted by the expression $h(t)$, which is the risk of an event (the hazard) occurring at time t, or the probability of an event occurring at time t in an at-risk population. The hazard rate is formally defined by the following equation, where G is a number of individuals from the population j at any given time, t is the time, Δt is the change in time, $(t + \Delta t)$ is a time later than t, and t_0 is the time at which all members of j were at risk (when $G=j$) (Tuma and Hannan 1984).

$$h_j(t \mid t_0) \equiv \lim_{\Delta t \downarrow 0} \frac{G_j(t \mid t_0) - G_j(t+\Delta t \mid t_0)}{G_j(t \mid t_0)\Delta t}$$

The equation reads "the hazard rate in a population j at time t, given t_0, is defined as the limit as the change in t goes to zero of the ratio of (those in G at t, given t_0, minus those in G at t plus the change in t, given t_0) to (those in G at t multiplied by the change in t)." If we let Δt (the change in time) get smaller and smaller and take the limit of the function above, we have a formula for the hazard rate. This hazard function gives the instantaneous probability of an event occurring at a specific time t.

Consider, for example, the proportion of faculty members who have not been promoted by time t, but who will be promoted between t and $(t + \Delta t)$. This proportion is the number in the group G_j at t who are promoted between t and $(t + \Delta t)$, divided by the number of faculty members eligible for promotion at time t. In the models examined in this chapter, the event used for the hazard rate is promotion in rank. It will be calculated separately for men and women to compare the probability of promotion over the course of a twenty-year career. The hazard rate is a useful indicator not only of the probability of promotion, but also of the timing of promotion. For even though men and women may have the same probability of occupying each rank, they may be promoted at different times in their careers.

When 50 percent of a group under study has experienced an event, the median survival year has been reached. It is based on another useful indicator of survival analysis, the cumulative proportion surviving (known as the survival function) at any given time. For the purposes of this chapter, the surviving proportion is made up of faculty members who were not promoted. The survival function depends on the hazard rate, and can be roughly thought of as its inverse function, although neither the survival function nor the hazard rate is ever negative. That is, as a hazard rate increases with time, the survival function exhibits a faster decline, although it will not increase as the hazard rate decreases. The survival function remains constant if the hazard rate is zero.

Among faculty members, there are several normative expectations regarding the survival function and hazard rates for promotion to associate and full professor. Promotion to associate professor is expected to occur at the end of the tenure process, which usually takes place five or six years after service to the university begins. Some faculty members receive credit for earlier career activity (postdoctoral studies or tenure-track service to another university, for example), while others may be promoted before five years for extraordinary scholarly performance. Similarly, when scholarly performance is limited, tenure may occur, but promotion may be delayed for several years. Nonetheless, based on the

principle of "up or out," most faculty members are promoted to associate professor at the time they receive tenure. Two patterns in the hazard rate for promotion to associate professor are therefore expected: (1) peak hazard rates should occur five or six years into college or university service, and (2) most promotions should occur earlier rather than later over the course of a twenty-year career.

Promotion to full professor is also a normative experience, which often follows a five-year period spent as an associate professor, a period commonly referred to as "time in rank." Promotions to full professorships do not, however, occur at a substantial rate until after about ten years of service, including an initial five-year period, promotion to an associate professorship, and five additional years at the rank of associate professor.

Moreover, it is not expected that all faculty members will attain the rank of full professor. Even over the course of twenty years of service to the university, many faculty members will not perform enough scholarly activity to merit such a promotion. Two additional patterns are therefore expected: (1) peak hazard rates of promotion to full professor should begin after a minimum of ten years of service, and (2) there should be a much smaller cumulative proportion of promotions to full professorships compared with promotions to associate professorships.

Hazard rates can also be specified as the dependent variable to be estimated in general linear models of proportional hazards. The regression estimates obtained in the "Data" section of this chapter are defined by the Cox (1972) proportional hazards linear regression model. Independent variables are used to predict the hazard rates. The model is represented in the equation:

$$h(t)=e^{a+\Sigma bixi}$$

In the data described in the following section, regressions of the hazard rate for promotion to associate professor are performed among all faculty who were hired at the assistant professor rank. Regressions of the promotion rate for promotion to the full professor rank are performed among all faculty members who were hired at the assistant professor rank and later promoted to the rank of associate professor. In each model, one variable is sex. If the estimate for sex is negative and significant, it will mean that females are less likely than men to be promoted at any given time.

Data

The Office of the Associate Vice President for Academic and Student Affairs at Kent State University provided the data in this section to the Kent State University chapter of the American Association of University Professors under the provisions of the faculty collective bargaining agreement. The database contains information on each faculty member in the collective bargaining unit for the Kent campus and for seven regional campuses of Kent State University. Full-time administrators, including those who hold faculty rank (for example, department heads), are excluded from the analysis, which is based on 833 full-time, tenure-track faculty whose main responsibilities are teaching and research. There are 636 faculty members on the Kent campus (449 men and 187 women) and 197 faculty members on the seven regional campuses (130 men and 67 women). With regard to rank, there are 18 instructors (6 men and 12 women), 326 assistant professors (181 men and 145 women), 287 associate professors (212 men and 75 women), and 202 full professors (180 men and 22 women).

The information on each faculty member provided by the administration included date of birth, sex, race, salary, rank, years in rank, year hired, department, the calendar years of promotion to each rank, year tenured, year of highest degree, and highest degree. These are the variables used in the analyses described above to determine the existence and magnitude of sex inequities.

Findings

Table H.1 shows the hazard rates and cumulative proportion of female and male faculty members *not* promoted to associate professor in each of the twenty years following receipt of the highest degree.

The analysis was conducted only for faculty members hired at the assistant professor or instructor level. Faculty members hired at the associate or full professor level were excluded, because they were promoted elsewhere or negotiated a rank at time of hire and were therefore not fully subject to Kent State's promotion mechanism. The median survival time among the 464 men eligible for promotion to associate professor was 9.55 years, while the median survival time among the 229 women eligible for the same promotion was 16.93 years. The 7.38-year gap between men and women in time spent in a lower rank means that the women in the sample had a lower probability of promotion in each of the ten years after attainment of the highest degree. On top of that, the men continued to get promoted to associate professor at high rates even after the normative five- to six-year probationary period, while the rate of promotion among the women dropped off drastically after the sixth year. In other words, the men continued to receive promotions even when they failed to follow

Table H.1

Hazard Rate and Cumulative Proportion Not Promoted to Associate Professor, by Sex and Year Since Highest Degree

Year Since Highest Degree	Hazard Rate for Promotion		Cumulative Proportion Surviving[a]	
	Female	Male	Female	Male
1	.01	.01	.99	.99
2	.00	.03	.98	.96
3	.02	.04	.97	.92
4	.02	.06	.95	.87
5	.09	.10	.87	.79
6	.08	.14	.80	.68
7	.06	.10	.75	.62
8	.03	.15	.73	.53
9	.05	.12	.69	.47
10	.10	.09	.62	.43
11	.02	.12	62	.38
12	.05	.10	.58	.34
13	.02	.04	.57	.33
14	.08	.03	.52	.32
15	.03	.05	.51	.30
16	.03	.03	.50	.29
17	.04	.06	.48	.28
18	.00	.04	.48	.27
19	.00	.04	.48	.26
20	.06	.02	.46	.25

a. The cumulative proportion surviving is the proportion of faculty members who have not been promoted in each year since receiving their highest degree.

Table H.2

Hazard Rate and Cumulative Proportion Not Promoted to Full Professor, by Sex and Year Since Highest Degree

Year Since Highest Degree	Hazard Rate for Promotion		Cumulative Proportion Surviving[a]	
	Female	Male	Female	Male
6	.01	.01	.99	.98
7	.00	.01	.99	.97
8	.00	.02	.99	.95
9	.00	.03	.99	.93
10	.01	.03	.98	.90
11	.02	.06	.96	.85
12	.03	.04	.93	.82
13	.01	.02	.92	.80
14	.01	.05	.91	.76
15	.02	.05	.89	.72
16	.02	.03	.88	.69
17	.04	.03	.85	.67
18	.04	.02	.81	.66
19	.03	.04	.79	.64
20	.03	.02	.77	.62

a. The cumulative proportion surviving is the proportion of faculty members who have not been promoted in each year since receiving their highest degree.

the normative pattern for time spent on probation, whereas the women seemed marked if they were not promoted "on time." The repercussion of not being promoted is important for salary, because an associate professor earns on average $10,615 more than an assistant professor.

The same discrepant patterns of promotion between men and women showed up for faculty members eligible for promotion to full professor (see table H.2). But for both the men and the women in this group, the median survival time for promotion was over twenty years, suggesting that fewer than half of all academics can expect to receive full professorships in their first twenty years of service. Nonetheless, among the actual proportion surviving in the Kent State sample after more than twenty years, men still fared better than women: only 54 percent of the men remained in the lower ranks, while 74 percent of the women did. This is important for salary, because a full professor earns on average $26,224 more than an assistant professor and $15,609 more than an associate professor.

The Cox (1972) regression model predicting probability of promotion to associate and full professor at each time interval since receiving the highest degree includes the variables sex, whether or not the faculty member has earned a Ph.D. degree (as opposed to another degree such as a master's), whether or not the faculty member is on the Kent campus, and the number of years of service to Kent State University. The results as reported showed that the coefficient for sex in both equations was negative and significant at $p<.05$ (one-tailed test). This meant that women were less likely overall than men to be promoted to either the rank of associate or full professor. The exponential value for the coefficients are the relative odds ratio for promotion of women compared to men. For the associate professor level, this value is .70, indicating that women are only 70 percent as likely as men to be promoted in any given year. For the full professor level, the value of the exponent or the coefficient is .66, indicating that women are

only 66 percent as likely to be promoted in any given year since their highest degree.

Conclusion

The Kent State analysis revealed that female faculty members at the university were promoted later and at a lower rate than their similarly situated male colleagues. Discrimination against female faculty members probably accounted for most of these promotion inequities. The prevalence of discrimination can be detected in failure to promote or to promote on time, and its most serious consequence is individual salary inequities in the $10,000 range.

APPENDIX I

Locating Outliers

By Lois Haignere and Yangjing Lin

The practice of dropping outliers is predicated on the assumption that they distort the real picture. As indicated in chapter 6, sometimes outliers get dropped in salary-equity studies when they do not distort the picture but simply are the highest and lowest cases in the picture.

There are a number of diagnostic measures that can be used to detect outliers, but three are of particular interest for salary equity studies. They are:

- the studentized residuals, which locate dependent variable deviant cases;
- the hat matrix diagonal element, which locates predictor (or independent) variable deviant cases; and
- Cook's Distance (CD), which locates cases that are deviant relative to both dependent and independent variables and substantially affect the parameter estimates or distort the results.

The following is a brief discussion of each of these three measures.[1]

Studentized Residual

The studentized residual is a measure that identifies an outlier in the dependent variable. It is created as follows:

- A regression model is run without one individual's information.
- The results of this regression are used to create a predicted salary for the individual who was left out of the regression.
- The difference between the actual salary and the predicted salary (the salary residual) for that individual is calculated.
- This residual is then standardized by dividing it by its standard error. The resulting value is the studentized residual for that individual.

This measure determines observations that have unusually high or low dependent variable data (salary). The suggested cutoff value to identify outliers ranges from the absolute value (negative or positive) of two to the absolute value of three (Stevens 1984). Larson (1994) recommends using a table based on the number of model parameters and the sample size (incorporated in an SAS macro program) to determine this cutoff value. Based on this macro program, the cutoff values for our twelve SUNY schools for the studentized residual were generally greater than three, and relatively few of the faculty in these analyses exceeded the cutoff value, that is, were found to be outliers. If you observe individual cases that are outliers based on the studentized residual values, remember that this does not mean that the observation in question substantially affects the parameter estimates or distorts the results (Stevens 1984).

Hat Matrix Diagonal Element

The hat matrix diagonal element is a measure that identifies an outlier in predictor (or independent) variables. The cutoff value for the hat diagonal element depends on the number of predictor variables (p) in a regression model and the number of cases (n) in the study. For a data set that has more than ten predictor variables and more than fifty cases, it is suggested that $2p/n$ be used. For example, if there are eleven predictor variables and sixty cases, then the cutoff value would be $2 \times 11/60 = 0.4$. Any value that is greater than the number derived from this formula is considered an outlier. Like the studentized residual value, a large hat diagonal element does not mean that the observation identified substantially affects the parameter estimates or distorts the results (Stevens 1984).

Cook's Distance

Cook's Distance (CD) is a measure indicating the combined effects of the dependent variable, as well as the independent variables, on the change of the regression parameter estimates that would occur if a given observation were dropped from the regression model. Outliers identified by this method substantially affect the parameter estimates and may distort the results. High CD values are more likely to be found in small rather than large data sets.

The cutoff value for CD is about one (Cook and Weisberg 1982, 118). In our experience, those cases that

have a substantially higher CD value than the rest of the population can also potentially have significant effects on the parameter estimates even though the CD value is less than one. A case in point is a white female faculty member at a two-year college. Her CD value is 0.52, which is about ten times higher than the other people at her school (the average CD value for the analysis of this school is 0.05). Dropping this highly paid woman from the regression analysis substantially increases the bias indicated for white females.

Comparison of the Three Outlier Measures

In our experience, using either the studentized residual or the hat diagonal element alone can be misleading because neither necessarily identifies cases that affect the parameter estimates. If the estimates are not affected, then the case is not distorting the results and need not be eliminated. At the twelve SUNY schools, we found that the percentage of cases reaching the cutoff value for the hat diagonal element (about 10 percent) was higher than for the studentized residual (about 5 percent). By contrast, at all twelve SUNY schools there was only one case that was found to exceed the CD cutoff value. This means there was only one outlier that had excessive influence on the parameter estimates.

Using the studentized residual or the hat diagonal value alone to flag and drop outliers is a questionable procedure. By the very nature of the studentized residual statistic, positive and negative outliers are identified. In salary-equity studies, most of the high-paid outliers are white males, while most of the low-paid outliers are females or minorities.

Observations flagged by Cook's Distance should be scanned for data errors. A case in point is an individual from one of the SUNY two-year colleges who was identified as an outlier by a CD of 1.66. We carefully examined his information regarding salary, current rank, highest degree, and years in current rank. After checking all the possible information, we found that this individual had seventeen years in his current rank as a full professor instead of two years, as was recorded in our data. He had been a dean, as well as a full professor, for fifteen years. He stepped down from the deanship but continued to be a full professor. The two years in current rank referred to the time after he left the dean's position. Adjusting his years in current rank dropped his CD value to 0.08, indicating that he is not an outlier.

Dropping the deviant cases flagged by Cook's Distance (greater than 1.0) may be appropriate once their data have been carefully confirmed as accurate. These are cases that are very different from the general population being studied and excessively influence the parameter estimates. Nonetheless, remember that while statistical methods can be used to detect outliers, dropping them is as much a political decision as a methodological one. The practice of dropping outliers is predicated on the assumption that they distort the real picture. Don't drop them unless they do. (See chapters 5 and 6 for additional discussion.)

Note

1. For a more technical discussion of using influence statistics, see Stevens 1984; Chatterjee and Hadi 1988; Belsley, Kuh, and Welsch 1980.

Glossary[1]

Adjusted R^2

A measure of how much variability in the dependent variable is accounted for by the predictor (independent) variables in a regression analysis. As the adjusted R^2 gets closer to 1.0, the predictor variables account for more of that variability. The term is referred to as "adjusted" because it accounts for the number of predictors. (For a more in-depth discussion of this topic, see appendix A.)

Antilog

See **Log Odds.**

Categorical Modeling

A statistical procedure used to analyze the relationship between predictor variables and a categorical variable. It is a generalized logistic regression model that uses maximum likelihood tables to estimate these relationships.

Categorical Variable

A variable that classifies subjects into a limited number of categories. Even if we assign numbers to each category or level, the categorical variable does not become continuous because the intervals between levels are not equal. Rank provides a good example. Even if we assign lecturer = 1, instructor = 2, assistant professor = 3, associate professor = 4, and full professor = 5, we do not know if the difference between being an instructor (rank = 2) and an assistant professor (rank = 3) is the same as the difference between being an associate professor (rank = 4) and a full professor (rank = 5), or whether an associate professor (rank = 4) is worth twice as much as an instructor (rank = 2) and four times as much as a lecturer (rank = 1).

Cell

Number of Individuals by Race in Each Rank

Race	Lecturer	Instructor	Assistant	Associate	Full
Black	**3**	5	7	3	4
Asian	2	4	2	8	5
Hispanic	1	0	0	2	1
White	25	63	572	641	250

Each piece of information found within a table or matrix is called a cell. In the matrix (table) in this example, each number represents a cell. The bolded 3 is the cell that identifies the number of black lecturers in this institution. All zeros in the table identify categories that have no individuals in them and are called "zero cells." Hispanic instructors and Hispanic assistant professors both constitute zero cells.

Censored Cases

Censored cases are used in event history analyses. Cases that are censored are ones that contain incomplete or missing information. Typically, the information is missing or incomplete because the individual fails to experience the event during the time period in which the analysis occurs. For example, if the event we are studying is promotion from the rank of associate professor to full professor during the time period 1990–97 and the individual has not been promoted during this time, this individual case becomes censored. By censoring the data, this individual contributes seven person-years to the database, even though she or he has not experienced a promotion.

Cluster Analyses

Any of several multivariate procedures designed to determine whether cases are similar enough to fall into groups or clusters.

Coefficient (Parameter Estimate)

The regression coefficient, which is also called parameter estimate, is a number that indicates how an independent variable affects the dependent variable—for example, how years of experience or gender or race affect salary. The coefficient also represents the slope of the line for the independent variable relative to the dependent variable. An unstandardized regression coefficient in an analysis of salary can be interpreted in dollar equivalents. If the coefficient for years of experience is 200, on average, each year of experience increases salaries by $200. See appendix A for a detailed explanation of how coefficients are multiplied by variables in regression equations to calculate predicted salaries. In studies of the effect of gender and race on salaries, the coefficients for the gender and race variables are the results, or findings. For an explanation of the *standardized* coefficient and its common misuse, see "Problem 1" in chapter 6.

Continuous Variable
A variable that is numeric, with equal intervals between levels of the variable. Salary is an example of a continuous variable. For salary, we know that each additional dollar of salary is the same as all other dollars. The interval between one level of salary and the next level is equal no matter what level is being reviewed. The difference between $1 and $2 is the same as the difference between $40,000 and $40,001.

Cox Regression
A logistic regression procedure that reconstructs individual records to person-years.

Critical Zero Cells
Empty cells that preclude categorical modeling as an analytical tool. If gender bias in the awarding of academic ranks is being examined, critical zero cells occur when every other rank lacks females. If race bias is being examined, then critical zero cells are found when every other rank lacks minorities.

Curvilinearity
A relationship between two variables that, when plotted on a graph, forms a curve rather than a straight line or linear relationship.

Default Category
The group that is not represented by a value of 1 in any of the dummy variables. For example, if you have a dummy variable for female (female = 1, male = 0), then the default category is male. As a result, the regression coefficient for the female variable in a salary-regression analysis represents what females are paid compared to males. In an analysis where all of the race-gender groups are represented by dummy variables, then white males are the default group, and the coefficients represent what each minority-gender group is paid compared to white males.

Dependent Variables
The focus of an analysis whose values are predicted by or "depend" on the independent or predictor variables (Vogt 1993). Changes in the independent variables are thought to affect the dependent variable. In an analysis of whether salary allocation is biased by gender, the dependent variable is salary.

Dummy Variable
A variable that is used to represent the different levels of a categorical variable. A dummy variable has a value of either 0 or 1. The number of dummy variables needed to fully represent a categorical variable is the number of levels of that variable minus one. For example, for a variable with two levels, such as gender, you will need one dummy variable, female, which you can code 0 for males and 1 for females. Because anyone who is not female is male, you do not need a variable for male. If the variable has three levels, you will need two dummy variables, and so on. See appendix A for a further discussion of the use of dummy variables.

Event History Analyses
Methods for studying the movement over time of subjects through successive states or conditions. The goal of event history analyses is to study changes from one state to the next, such as from associate to full professor. Event history analyses are a useful adjunct to pay studies for diagnosing gender or race bias in the promotion process.

Exponential Value
The exponential value of a coefficient (or of a number) is the power (or exponent) to which the base of natural logarithms, e, is raised. For example, if the coefficient is equal to c, then the exponential value is e^c, where e=2.718. When the coefficient (or number) is negative the exponential of this number yields an odds ratio, for example, $e^{-0.22} = 1/e^{0.22} = 0.803$.

Independent (Predictor) Variable
A variable that is used to predict changes in a dependent variable, based on a change in its own value. Some possible independent (predictor) variables that may affect the dependent variable salary are gender, race, years of experience, educational level, and discipline.

Interaction Term
A term that describes the effect of one independent (predictor) variable on the dependent variable across levels of a second independent variable. For example, if males are paid $200 for each year of previous experience, but females are paid $100, then there is an interaction between gender and years of previous experience. The interaction term is the product of the two independent variables.

Log Odds
An odds ratio that is based on the logarithm or the exponent of a base number indicating the power to which that number must be raised to produce another number. For example, the log of 100 is 2 because 10^2

(10 x 10) equals 100, and the log of 1,000 is 3 because 10^3 (10 x 10 x 10) equals 1,000. The antilog turns the relationship around. For example, antilog 2 = 100 and 3 = 1,000.

Logistic Regression
A type of regression analysis used when the dependent variable is categorical and scored 0, 1. It is commonly used to predict whether or not an event (such as promotion or business failure) will happen.

Multicollinearity (Redundancy)
The condition that exists when one predictor variable is highly correlated with another predictor variable or group of variables in the analysis. In general, correlations above 0.80 are suspect, and correlations above 0.90 are red flags. These high correlations mean that the variables overlap so much that it is difficult, if not impossible, to determine their separate effects on the dependent variable. Also called redundancy.

Normal Distribution
A distribution of scores that when graphed produce a symmetrical, bell-shaped curve, centered around the mean (average). Statistics are often tested using normal distributions.

Outlier
A case in the data that has extreme values on one or more variables in the analysis.

Parameter Estimate (Coefficient)
A number that indicates how an independent variable affects the dependent variable—for example, how years of experience or gender or race affects salary. In multiple-regression analyses, this measure is also called the coefficient. See **Coefficient** for more information.

Predictor Variable
Used interchangeably with Independent Variable. *See* **Independent Variable.**

Quadratic Term
A variable that has been squared (multiplied by itself). It is used to solve the problem of curvilinearity in statistical models that rely on linear relationships between predictor and dependent variables. Time-related variables are often thought to have curvilinear attributes, and are frequently included in models in both the linear form and the quadratic form.

Redundancy: *See* **Multicollinearity.**

Response Variable
The dependent variable in categorical modeling. In an attempt to predict academic rank based on other career characteristics, rank is the response variable.

SAS
A statistical analysis software package that is commonly used by social scientists.

Scatter
The distribution of scores above and below a regression line. A large amount of scatter means that the regression model is not very good at predicting scores on the dependent variable. In this case, the addition of other independent variables may reduce the scatter and improve the predictive power of the model.

Scattergram
A pattern of points that indicates the relationship between two variables. Each point represents where one unit of the analysis—for example, a faculty member—is on the two variables. For the purposes of studying salary, we frequently put salaries on the vertical axis and another variable, like years of experience, on the horizontal axis so that the scatter represents the relationship between salary and the other variable. (See appendix A, figures A.1 and A.2.) The more the points tend to form a straight line, the stronger the relationship. Also called "scatter diagram" and "scatter plot."

Statistical Power
The probability of being able to say that one variable affects another. If the effect of one variable on another is very small, then more power is needed to detect the effect, just as you would need a higher power microscope to see a virus than you would need to see an ant. Aspects of the data set, such as the number of variables, variance, and sample size, affect power.

Statistical Significance
A judgment criterion based on the probability that a sample comes from a specific population. The number can range between 0 and 1; a low number for statistical significance indicates that the data are unlikely to be ascribable to chance. The alpha level, which is usually set at either 0.05 or 0.01, is the point at which we reject the null hypothesis that our sample came from the same population that we are comparing it to; that is, it is the point at which we consider our results

to be statistically significant. (See also "Problem 5" in chapter 6.)

Tainted Variable
In bias studies, a variable that is likely to have discrimination embedded in it and thus mask or suppress the gender or race effects. In research vernacular, such variables are called confounding or suppressing variables because they obscure or conceal causal relationships between other variables.

Time-Dependent Covariate
A variable whose value can change over time. For example, highest degree is commonly a time-dependent covariate, since individuals in the database who are hired with a master's degrees commonly subsequently receive Ph.D. degrees. Time-dependent covariates are used in event history analyses. Special programming is needed to assure that time-dependent covariates are properly recognized by the statistical software.

Zero Cell
Any cell of a matrix that contains no observations. In some analyses, for example categorical modeling, a zero cell can curtail the analyses or cause misinterpretations of the results.

Note
1. Some of the definitions in this glossary are borrowed from the *Dictionary of Statistics and Methodology* by W. P. Vogt (1993).

Bibliography

Acker, J. 1989. *Doing comparable worth: Gender, class, and pay equity.* Philadelphia: Temple University Press.

Aitchison, J., and J. Brown. 1963. *The lognormal distribution.* Cambridge, Mass.: Cambridge University Press.

Allen, J. 1984. *Manual for determination of academic salary discrimination against women.* Ottawa: Canadian Association of University Teachers.

Baldus, D., and J. Cole. 1980. *Statistical proof of discrimination.* New York: McGraw-Hill.

Barbezat, D. 1987. Salary differentials by sex in the academic labor market. *Journal of Human Resources* 22:422–28.

———. 1991. Updating estimates of male-female salary differentials in the academic labor market. *Economics Letters* 36:191–95.

Barnes, D. 1983. *Statistics as proof: Fundamentals of quantitative evidence.* Boston: Little, Brown and Co.

Barrett, R. 1995. *An analysis of gender differences in salary at the University of Montana.* Unpublished manuscript.

Bartholet, E. 1982. Application of Title VII to jobs in high places. *Harvard Law Review* 95:947.

Becker, G. 1975. *Human capital.* 2nd ed. New York: Columbia University.

Becker, G., and B. Chiswick. 1966. Education and the distribution of earnings. *American Economic Review* 56:358–69.

Becker, W., and R. Goodman. 1991. The semilogarithmic earnings equation and its use in assessing salary discrimination in academe. *Economics of Education Review* 10:323–32.

Becker, W., and R. Toutkoushian. 1995. The measurement and cost of removing unexplained gender differences in faculty salaries. *Economics of Education Review* 14:209–20.

Behling, D., et al. 1994. *Statistical analysis of 1993–94 salary data.* Bowling Green State University Faculty Senate. Unpublished.

Bell, R. 1992. Career mobility: Does gender matter? *Thought and Action* 8:39–60.

Bellas, M. L. 1992. The effects of marital status and wives' employment on the salaries of faculty men: The (house)wife bonus. *Gender and Society* 6 (4):609–22.

———. 1994. Comparable worth in academia: The effects on faculty salaries of the sex composition and labor market conditions of academic discipline. *American Sociological Review* 59:807–31.

———. 1997. Disciplinary differences in faculty salaries: Does gender bias play a role? *Journal of Higher Education* 68:299–321.

Bellas, M. L., and B. Reskin. 1994. On comparable worth. *Academe* (September–October) 80:83–85.

Bellas, M. L., and R. K. Toutkoushian. 1999. Faculty time allocations and research productivity: Gender, race and family effects. *Review of Higher Education* 22 (4):367–90.

Belsley, D., E. Kuh, and R. Welsch. 1980. *Regression diagnostics: Identifying influential data and sources of collinearity.* New York: Wiley.

Benjamin, E. 1987. Improving the status of women: A faculty perspective. In *International Association of University Presidents: VII Triennial Conference Proceedings* 65, 67.

———. 1999. Disparities in the salaries and appointments of academic women and men: An update of a 1988 report of Committee W on the Status of Women in the Academic Profession. *Academe* (January–February) 85:60–62.

Bergmann, B. 1989. What the common economic arguments against comparable worth are worth. *Journal of Social Issues* 45:67–80.

Blalock, H. M. 1979. *Social Statistics.* New York: McGraw Hill.

Blinder, A. 1973. Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources* 8:436–55.

Braskamp, L., J. Muffo, and I. Langston. 1978. Determining salary equity: Policies, procedures, and problems. *Journal of Higher Education* 49:231–46.

Brittingham, B. E., T. R. Pezzullo, G. A. Ramsay, J. V. Long, and R. M. Ageloff. 1979. A multiple regression model for predicting men's and women's salaries in higher education. In *Salary equity: Detecting sex bias in salaries among college and university professors,* edited by T. R. Pezzullo and B. E. Brittingham. Lexington, Mass.: Lexington Books.

Broder, I. E. 1993. Professional achievements and gender differences among economists. *Economic Inquiry* 31:116–27.

Burns, M. 1992. Women, part-time faculty, and illusion. *Thought and Action* 8:13–28.

Cabrera, A. F. 1991. Logistic regression analysis in higher education: An applied perspective. In *Higher education: Handbook of theory and research,* vol. 10, edited by J. C. Smart. New York: Agathon Press.

Calathes, W. 1991. Perpetuating social inequalities. *Thought and Action* 7:137–54.

Carver, R. P. 1978. The case against statistical significance testing. *Harvard Educational Review* 48:378–99.

Chatterjee, S., and A. Hadi. 1988. *Sensitivity analysis in linear regression.* New York: Wiley.

Chronister, J. L., B. M. Ganeneder, E. Harper, and R. G. Baldwin. 1997. Full-time non-tenure-track faculty: Gender differences. *NEA Higher Education Research Center Update* 3 (5):6.

Cox, D. R. 1972. Regression models and life tables. *Journal of the Royal Statistical Society* 34:187–202.

Davidson, H. K., and M. J. Burke. 2000. Sex discrimination in simulated employment contexts: A meta-analytic investigation. *Journal of Vocational Behavior* 56 (2):225–48.

Dean, J., and R. Clifton. 1994. An evaluation of pay equity reports at Canadian universities. *Canadian Journal of Higher Education* 24:87–114.

Deaux, K. 1976. Sex: a perspective on the attribution process. In *New directions in attributions research,* edited by J. Harvey, W. Ickes, and R. Kidd. Hillsdale, N.J.: Erlbaum.

DeSole, G., and M. Butler. 1990. Building an effective model for institutional change: Women's committees as catalysts. *Initiatives* 53:1–10.

Downey, R., M. Riggs, P. McIntyre, and D. Hoyt. 1985. *Salary and rank at KSU: A study of institutional policies.* Unpublished manuscript.

Dwyer, M. M., A. A. Flynn, and P. S. Inman. 1991. Differential progress of women faculty: Status 1980–1990. In *Higher education: Handbook of theory and research,* vol. 10, edited by J. C. Smart. New York: Agathon Press.

Euben, D. R. 2001. "Show me the money": Salary equity in the academy. *Academe* (July–August) 87:30–36.

Farber, S. 1977. The earnings and promotion of women faculty: Comment. *American Economic Review* 67:199–206.

Ferber, M. A., J. W. Loeb, and H. M. Lowry. 1978. The economic status of women faculty: A reappraisal. *The Journal of Human Resources* 23 (3):385–401.

Ferree, M., and J. McQuillan. 1998. Gender-based pay gaps: Methodological and policy issues in university salary studies. *Gender and Society* 12 (1):7–39.

Fidell, L. 1970. Empirical verification of sex discrimination in hiring practices in psychology. *American Psychologist* 25:1096–97.

Finkelstein, M. 1980. The judicial reception of multiple regression studies in race and sex discrimination cases. *Columbia Law Review* 80:737–54.

———. 1984. The status of academic women: An assessment of five competing explanations. *The Review of Higher Education* 7:223–46.

Finkler, D., D. Van Dyke, and J. Klawsky. 1989. *How statistics, law, and politics influence the evaluation of gender salary disparity in higher education.* University of Nebraska at Omaha. Unpublished manuscript.

Fisher, F. 1980. Multiple regression in legal proceedings. *Columbia Law Review* 80:702–36.

Fogel, W. 1986. Class pay discrimination and multiple regression proofs. *Nebraska Law Review* 65:289–329.

Formby, J., W. Gunther, and R. Sakano. 1993. Entry level salaries of academic economists: Does gender or age matter? *Economic Inquiry* 31:128–38.

Geetter, J. 1988. *Report from Salary Inequity Committee.* University of Connecticut. Unpublished report.

Gerdes, E., and D. Garber. 1983. Sex bias in hiring: Effects of job demands and applicant competence. *Sex Roles* 9:307–18.

Goldberg, P. 1968. Are women prejudiced against women? *Transaction* 5:28–30.

Gray, M. 1990. *Achieving pay equity on campus.* Washington, D.C.: American Association of University Professors.

———. 1991. Using regression to study faculty salaries. *Thought and Action* 7:55–100.

———. 1993. Can statistics tell us what we do not want to hear? The case of complex salary structures. *Statistical Science* 8:144–79.

Gray, M., and Committee W. 1988. Academic women and salary differentials. *Academe* (July–August) 74:33–34.

Gray, M., and E. Scott. 1980. A "statistical" remedy for statistically identified discrimination. *Academe* (May) 66:174–81.

Guppy, N. 1989. Pay equity in Canadian universities, 1972–73 and 1985–86. *Canadian Review of Sociology and Anthropology* 26:743–58.

Haignere, L. 1990. Salary equity four cell. In *Collective bargaining in higher education: The 1990s,* edited by J. M. Douglas. National Center for the Study of Collective Bargaining in Higher Education and the Professions, Baruch College, City University of New York, April 1990.

———. 1993. A study of equal pay for faculty at twenty-nine SUNY institutions: Labor/management negotiation, aggravation, experimentation, and implementation. *Proceedings of the Twenty-first Annual Conference of the National Center for the Study of Collective Bargaining in Higher Education and the Professions* 184–215.

Haignere, L., Y. Lin, and B. Eisenberg. 1993. *Report of the women and minority salary disparity analysis.* United University Professions. Unpublished manuscript.

Halvorsen, R., and R. Palmquist. 1980. The interpretation of dummy variables in semilogarithmic equations. *American Economics Review* 70:474–75.

Hamermesh, D. 1996. Not so bad: The annual report on the economic status of the profession. *Academe* (March–April) 82:14–108.

Hansen, R., and V. O'Leary. 1985. Sex-determined attributions. In *Women, gender, and social psychology*, edited by V. O'Leary, R. Unger, and B. Wallston. Hillsdale, N.J.: Erlbaum.

Hauser, R. M. 1980. On "stratification in a dual economy": A comment on Beck, Horan, and Tolbert, 1978. *American Sociological Review* 45:702–12.

Hauser, R., and K. Mason. 1993. *Faculty pay equity study: gender and ethnicity pay equity analysis.* University of Hawaii at Manoa. Unpublished manuscript.

Hays, W. 1988. *Statistics.* 4th ed. New York: Harcourt Brace.

Hengstler, D., and G. McLaughlin. 1985. Statistical issues and concerns in court cases. In *The use of data in discrimination issues cases*, edited by W. Rosenthal and B. Yancey. San Francisco: Jossey-Bass.

Hodson, R. 1985. Some considerations concerning the functional form of earnings. *Social Science Research* 14:374–94.

Holmes-Rovner, M., et al. 1994. Compensation equity between men and women in academic medicine: Methods and implications. *Academic Medicine* 69 (2):131–37.

Hurley, R., et al. 1981. Female faculty equity study: University of Maryland. *Resources in Education.*

Ikeda, K. 1995. *Evaluation research issues in women faculty pay equity adjustments: Evaluation of across-the-board over case-by-case equity adjustments.* University of Hawaii at Manoa. Unpublished manuscript.

Ishii-Kuntz, M. 1994. *Ordinal log-linear models.* Thousand Oaks, Calif.: Sage.

Jaccard, J., R. Turrisi, and C. Wan. 1990. *Interaction effects in multiple regression.* Newbury Park, Calif.: Sage.

Jaccard, J., and C. Wan. 1996. *Interaction effects in multiple regression, II: Multiple indicator and structural equation approaches.* Newbury Park, Calif.: Sage.

Johnson, C., M. Riggs, and R. Downey. 1987. Fun with numbers: Alternative models for predicting salary levels. *Research in Higher Education* 27:349–62.

Johnson, R. J., and C. O. Herring. 1989. Labor market participation among young adults: A life history analysis. *Youth and Society* 21:3–31.

Johnsrud, L., and R. Heck. 1994. Administrative promotion within a university. *Journal of Higher Education* 65:23–44.

Kessler-Harris, A. 1990. *A woman's wage: Historical meanings and social consequences.* Lexington: University Press of Kentucky.

Kolpin, V. W., and L. D. Singell. 1996. The gender composition and scholarly performance of economics departments: A test for employment discrimination. *Industrial and Labor Relations Review.*

Konrad, A. 1991. Faculty productivity and demographics. *Thought and Action* 7:19–54.

LaNoue, G. R., and B. A. Lee. 1987. *Academics in court: The consequences of faculty discrimination litigation.* Ann Arbor: University of Michigan Press.

Lee, J. 1995. Faculty salaries, 1993–1994. *The NEA 1995 Almanac of Higher Education* 7–20.

Lewis, D., and W. Becker. 1979. *Academic rewards in higher education.* Cambridge, Mass.: Ballinger.

Long, J. S., P. D. Allison, and R. McGinnis. 1993. Rank advancement in academic careers: Sex differences and the effects of productivity. *American Sociological Review* 58:703–22.

Luna, G. 1990. Understanding gender-based wage discrimination: Legal interpretation and trends of pay equity in higher education. *Journal of Law and Education* 19:371–84.

Maddala, G. S. 1983. *Limited-dependent and qualitative variables in econometrics.* Cambridge, Mass.: Cambridge University Press.

Magner, D. K. 1999. Record number of doctorates awarded in 1997. *Chronicle of Higher Education,* 8 January 1999, A14.

Massy, W., and R. Zemsky. 1994. Faculty discretionary time: Departments and the academic ratchet. *Journal of Higher Education* 65:1–22.

McLaughlin, G., M. Zirkes, and B. Mahan. 1983. Multicollinearity and testing questions of sex equity. *Research in Higher Education* 19:277–84.

Melani v. Board of Higher Education of the City of New York, 561 F. Supp. (S.D.N.Y. 1983).

Mellers, B., and M. Birnbaum. 1989. *Determining group and individual bias.* Talk delivered to the American Sociological Association.

Michael, R. T., H. I. Hartmann, and B. O'Farrell, eds. 1989. *Pay equity: Empirical inquiries.* Washington, D.C.: National Academy Press.

Monroe Community College Faculty Women's Caucus. 1985. *Statistical Study Monroe Community College Faculty Salary Schedule Report.* Unpublished report.

Moore, K., and M. J. Amey. 1993. Making sense of the dollars: The costs and uses of faculty compensations. *ASHE-ERIC Higher Education Report No. 5.* Washington, D.C.: George Washington University, School of Education and Human Development.

Moore, N. 1992. *Evaluation of faculty salary equity models.* Farmington, N.Mex.: San Juan College.

———. 1993. Faculty salary equity: Issues in model selection. *Research in Higher Education* 34:107–25.

Morrison, D., and R. Henkel. 1969. Significance tests reconsidered. *American Sociologist* 4:131–40.

Muffo, J. A., L. Braskamp, and I. W. Langston, IV. 1979. Equal pay for equal qualifications? A model for determining race or sex discrimination in salaries. In *Salary equity: Detecting sex bias in salaries among college and university professors*, edited by T. R. Pezzullo and B. E. Brittingham. Lexington, Mass.: Lexington Books.

Nelson, R. L., and W. P. Bridges. 1999. *Legalizing gender inequality: Courts, markets, and unequal pay for women in America*. Cambridge, Mass.: Cambridge University Press.

Nettles, M. T., L. W. Perna, and E. M. Bradburn. 2000. *Salary, promotion, and tenure status of minority and women faculty in U.S. colleges and universities.* Washington, D.C.: National Center for Education Statistics.

Oaxaca, R. 1973. Male-female wage differentials in urban labor markets. *International Economic Review* 14:693–709.

Paludi, M., and W. Bauer. 1983. Goldberg revisited: What's in an author's name? *Sex Roles* 9:387–90.

Pedhazur, E. 1982. *Multiple regression in behavioral research: Explanation and prediction*. 2nd ed. Chicago: Holt, Rinehart and Winston.

Persell, C. H. 1983. Gender, rewards and research in education. *Psychology of Women* 8 (1):33–47.

Pezzullo, T. R., and B. E. Brittingham. 1979. *Salary equity*. Toronto: Lexington.

Pheterson, G., S. Kiesler, and P. Goldberg. 1971. Evaluation of the performance of women as a function of their sex, achievement, and personal history. *Journal of Personality and Social Psychology* 19:114–18.

Polowy, C. 1975. *Collective bargaining and discrimination issues in higher education*. Academic Collective Bargaining Information Service. Monograph 1.

Ramsay, G. A. 1979. A generalized regression model for predicting discrimination: A demonstration using constructed data. In *Salary equity*, edited by T. R. Pezzullo and B. E. Brittingham. Toronto: Lexington.

Ransom, M. R., and S. B. Megdal. 1989. *On the status of women in the academic labor market in the affirmative action area*. Unpublished manuscript.

Raymond, R. D., M. L. Sesnowitz, and D. R. Williams. 1990. *Further evidence on gender and academic rank*. Kent State University. Unpublished.

Raymond, R. D., M. L. Sesnowitz, and D. R. Williams. 1988. Does sex still matter? New evidence from the 1980s. *Economic Inquiry* 26:43–58.

Reed, L., R. Douthitt, B. Ortiz, and D. Rausch. 1988. *Gender differences in faculty retention at the University of Wisconsin-Madison*. Unpublished report.

Regan, C., and J. F. Volkwein. 1993. *The influence of organizational variables on disparities between male and female faculty salaries*. Paper presented at the Eighteenth Annual Conference Association for the Study of Higher Education.

Reskin, B. F., and P. A. Roos. 1990. *Job queues, gender queues: Explaining women's inroads into male occupations*. Philadelphia: Temple University Press.

Reskin, B. F., D. Liddy, L. Haignere, L. L. Francis, and R. Steinberg. 1992. Salary-setting practices that unfairly disadvantage women faculty. *Academe* (July–August) 78:32–35.

Robson, R. 1969. A comparison of men's and women's salaries in the academic profession. *Canadian Association of University Teachers Bulletin* 17:50–75.

Rosen, B., and T. Jerdee. 1974. Influence of sex role stereotypes on personnel decisions. *Journal of Applied Psychology* 59:9–14.

Rosenbluth, G. 1967. The structure of academic salaries in Canada. *Canadian Association of University Teachers Bulletin* 15:19–27.

Ruble, D., and T. Ruble. 1982. Sex stereotypes. In *In the eye of the beholder: Contemporary issues in stereotyping*, edited by A. Miller. New York: Praeger.

Schau, C., and V. Heyward. 1987. Salary equity: Similarities and differences in outcomes from two common prediction models. *American Educational Research Journal* 24 (2):271–86.

Schrank, W. 1977. Sex discrimination in faculty salaries: A case study. *Canadian Journal of Economics* 10:411–33.

———. 1985. *Sex discrimination in faculty salaries at Memorial University: A decade later*. Report submitted to the president of Memorial University and the Executive Committee of the Memorial University of Newfoundland Faculty Association.

———. 1988. *Multiple regression analysis as a method of ascertaining salary anomalies*. OCUFA workshop on the compensation of female academic staff.

———. 1989. Multiple regression analysis and the determination of gender discrimination in university faculty salaries. *Canadian Association of University Teachers Bulletin* 36:5.

Scott, E. 1977. *Higher education salary evaluation kit*. Washington, D.C.: American Association of University Professors.

Smart, J. 1991. Gender equity in academic rank and salary. *The Review of Higher Education* 14:511–26.

Snyder, J., P. Hyer, and G. McLaughlin. 1994. Faculty salary equity: Issues and options. *Research in Higher Education* 35:1–19.

Staub, K. 1987. *Level of female participation: An overlooked factor in salary differences among faculty disciplines*. Presentation at the Annual Conference of the Southern Association for Institutional Research.

Steinberg, R. J., and L. Haignere. 1987. Equitable compensation: Methodological criteria for comparable worth. In *Ingredients for women's employment policy*, edited by C. Bose and G. Spitze. Albany: State University of New York Press.

Steinpreis, R. E., K. A. Anders, and D. Ritzke. 1999. The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles* 41 (7–8):509–28.

Stevens, J. 1992. *Applied multivariate statistics for the social sciences*. 2nd ed. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Stolzenberg, R. 1980. The measurement and decomposition of causal effects in nonlinear and nonadditive models. In *Sociological methodology*, edited by K. Schuessler. San Francisco: Jossey-Bass.

Strober, M., and A. Quester. 1977. The earnings and promotion of women faculty: Comment. *American Economic Review* 67:207–13.

Thompson, K. 1992. Piecework to parity: Part-timers in action. *Thought and Action* 8:29–38.

Top, T. 1991. Sex bias in the evaluation of performance in the scientific, artistic, and literary professions: A review. *Sex Roles* 24:73-106.

Toren, N. 2000. *Hurdles in the halls of science: The Israeli case*. Lanham, Md.: Lexington Books.

Toutkoushian, R. K. 1994a. Issues in choosing a strategy for achieving salary equity. *Research in Higher Education* 35 (4):415–28.

———. 1994b. Using citations to measure sex discrimination in faculty salaries. *Review of Higher Education* 18 (1):61–82.

———. 1998. Sex matters less for younger faculty: Evidence of disaggregate pay disparities from the 1988 and 1993 NCES surveys. *Economics of Education Review* 17 (1):55–71.

———. 1999. The status of academic women in the 1990s: No longer outsiders, but not yet equals. *Quarterly Review of Economics and Finance* 39:679–98.

Treimann, D. J., and H. I. Hartmann. 1981. *Women, work, and wages: Equal pay for jobs of equal value*. Washington, D.C.: National Academy Press.

Tuckman, B., and H. Tuckman. 1976. The structure of salaries at American universities. *Journal of Higher Education* 62:51–64.

Tuma, N., and M. Hannan. 1984. *Social dynamics: Models and methods*. New York: Academic Press.

Valian, V. 1998. *Why so slow? The advancement of women*. Cambridge, Mass.: MIT Press.

Vogt, W. P. 1993. *Dictionary of statistics and methodology*. Newbury Park, Calif.: Sage.

Weiller, W. 1990. Integrating rank differences into a model of male-female faculty salary discrimination. *Quarterly Review of Economics and Business* 30:3–15.

West, M. 1994. Gender bias in academic robes: The law's failure to protect women faculty. *Temple Law Review* 67:67–178.

———. 1995. Women faculty: Frozen in time. *Academe* (July–August) 81:26–29.

———. 1989. *Salary Equity Study*. Virginia Commonwealth University. Unpublished report.

AAUP

Haigmere, Inc.
Equal Pay and Pay Equity Research

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

NOTICE

Reproduction Basis

[X] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

[ ] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").