

FINAL REPORT

**JOINT ADMINISTRATION-ACADEMIC SENATE TASK FORCE ON ANALYSIS OF FACULTY
SALARY EQUITY AT THE UNIVERSITY OF CALIFORNIA, DAVIS**

12 November 2013

Table of Contents

Joint Administration-Academic Senate Task Force Members _____	4
Summary of the Report _____	5
I. UC Davis' Values Regarding Salary Equity and Minimization of Salary Differences Due to Gender or Race/Ethnicity _____	6
II. Historical Context for the Work of the Task Force _____	6
III. Charge to the UC Davis Task Force _____	7
IV. Legal Definitions and Considerations Regarding Salary Differences Associated with Gender or Race/Ethnicity _____	7
V. Definitions of Salary Variables to Be Studied _____	8
A. Base Salary _____	8
B. Non-Base Salary Components _____	9
i. Off-scale salary _____	9
ii. Salary components under the Health Sciences Compensation Plan (HSCP) _____	9
iii. Externally (e.g., grant) funded salary payments _____	10
iv. Temporary summer salary and administrative stipends _____	10
VI. Overview of Analysis: Understanding Salary Variation at UC Davis _____	11
VII. Descriptive Salary Analyses: Scope and Challenges _____	12
A. Population of Faculty Members for Analysis _____	12
B. Addressing Salary Variation among Cohorts and Units _____	12
i. Impacts of cohort on base salary _____	12
ii. Impacts of academic disciplinary factors on base salary _____	13
iii. Impacts of discipline cohort on negotiated salary components _____	13
iv. Salary comparison units _____	15
v. Standardization of variables _____	16
Standardized starting base salary _____	17
Standardized off-scale salary at time of hire _____	17
Standardized starting rank and step _____	18
Standardized rate of progression through academic ranks _____	18
VIII. Proposed Models for Salary Equity Analysis _____	19
A. A Demographic Model of Salary Variaiton _____	19
B. A Structural Equation Model of Salary Variation _____	20
IX. Toward Future Analysis: Faculty Performance Metrics and Peer Evaluation _____	21
A. Overview _____	21
B. Exploring Performance Metrics as Part of the Faculty Review _____	22
C. How to Develop Informative Faculty Perfomance Metrics? _____	23
D. Potential Impacts of Voting Methods on Performance Assessment _____	25
X. Implementation Recommendations _____	29
XI. Analytical Considerations for Campus-wide Use of Faculty Performance Metrics in Salary Equity Anlalysis _____	30
A. Measurement and Analytical Challenges _____	31
References _____	32
Appendix 1: Race and Ethnicity Categories _____	33

Appendix 2: Normative Progress Year Equivalents _____	34
Appendix 3: Academic Personnel Manual (APM) Statement Regarding Components of Faculty Member Performance _____	35
A. Commentary by the Task Force on APM's Components of Faculty Performance _____	37

**Members of the Joint Administration-Academic Senate Task Force
on Analysis of Faculty Salary Equity**

Associate Dean and Professor Edward Callahan (Senate Representative; Department of Family and Community Medicine, School of Medicine)

Professor Gregory Clark (Senate Representative; Department of Economics, Division of Social Sciences)

Dean and Professor Steven Currall (Graduate School of Management), Chair of Task Force

Associate Dean and Professor Bruce Hartsough (Department of Biological and Agricultural Engineering, College of Engineering)

Lecturer with Security of Employment Stuart Hill (Department of Political Science, Division of Social Sciences)

Professor James Jones (Senate Representative; Department of Surgical and Radiological Sciences, School of Veterinary Medicine)

Professor Susan Rivera (Senate Representative; Department of Psychology, Division of Social Sciences)

Professor Leticia Saucedo (School of Law)

Vice Provost–Academic Affairs and Professor Maureen Stanton (Office of the Chancellor and Provost, Department of Evolution and Ecology), *Ex officio*

Analyst Everett Wilson (Academic Affairs, Office of the Chancellor and Provost)

Summary

The UC Davis Joint Administration-Academic Senate Task Force has developed a blueprint for assessing possible faculty salary disparities based on gender and race/ethnicity on our campus. The document begins by discussing the historical context for the analysis of faculty salary disparities, followed by a summary of legal considerations regarding discrimination based on gender and/or race/ethnicity. We explicate the distinction between faculty base salary and the components of negotiated salary. To establish equitable salary reference points, we define “salary comparison units (SCUs)” as groups of faculty within which salaries are to be compared. To promote campus-wide comparison of salaries by pooling across academic units and disciplines, we provide methodologies for standardizing current salary components and negotiated salary components at the time of hire so that each faculty member’s salary components are expressed relative to the SCU mean.

Our recommendations for analysis fall into two broad categories. First, to determine whether disparities in salary exist between faculty groups based on gender and race/ethnicity, we propose a set of descriptive, quantitative models that incorporate complexity and explanatory power. Specifically, we advocate for the development of statistical models that account for faculty demographics and changing norms of compensation during the past few decades. These descriptive analyses constitute the core of our proposed salary equity analysis for the campus, which should be repeated every other year.

Second, because it is not possible to determine whether biases play any role in determining faculty salary variation without access to objective measures of faculty performance, we recommend developing performance metrics that will be applied to faculty members in various disciplines. We suggest exploring the extent to which peer appraisals of faculty performance reflect objective metrics, gender or race/ethnicity. This task will require additional effort and resource allocation by the campus, but has the potential to reveal sources of bias in appraisal and point to practices that could minimize bias in peer evaluation, compensation and advancement. We provide detailed treatment of the complexities of such a system and the pros and cons of different approaches to measuring performance.

We conclude by providing policy prescriptions for the process whereby the campus analyzes salary disparities. We also identify methodological and statistical challenges inherent in testing for salary disparities.

The Task Force also created a supplemental report containing various analytical guidelines. This document is titled “Supplemental Report: Guidelines for Analysis of Possible Salary Disparities at the University of California, Davis.”

I. UC Davis' Values Regarding Salary Equity and Minimization of Salary Differences Due to Gender or Race/Ethnicity

Gender and racial/ethnic pay equity ensures that both women and men, regardless of racial category, are paid fairly for the work they perform. Pay equity should include equal pay for performing the same work; it also must take into account discretionary pay (such as off-scale salaries) and payments due to merit or bonuses, as well as the distribution of teaching allocations and university service. Pay equity is one aspect of satisfaction with pay, which has been shown to be associated with important organizational performance outcomes (Currall, Towler, Judge and Kohn 2005).

To adequately address the issue of gender and race/ethnicity pay equity, we must first understand whether pay disparities¹ exist. The study proposed herein is designed to accomplish that goal. In carrying out the study, it is equally important that we avoid conflicts of interest in analysis of faculty salary equity. This inherently means involving diverse constituencies in the process and seeking feedback from those constituencies along the way.

It is important to note that if consistent patterns of pay disparities based on gender or ethnicity are found, it does not necessarily imply discrimination. If differences are found, the proposed analysis is thus additionally designed to investigate which factors (e.g., starting rank and step or starting negotiated salary) may be contributing to differences. At the inaugural study's completion, best practices will involve (a) ensuring dissemination of the results of the analysis to all relevant parties; (b) introducing initiatives that would respond to any inequities revealed; and (c) ensuring that the study is repeated on a regular basis (e.g., every other year).

II. Historical Context for the Work of the Task Force

In their charge letter to the UC Davis Joint Administration-Academic Senate Task Force on Analysis of Faculty Salary Equity, Provost Ralph Hexter and Academic Senate Chair Bruno Nachtergaele provided a summary of the historical context leading to the formation of the Task Force:

"In 2007, the University Committee on Affirmative Action and Diversity (UCAAD) recommended the undertaking of a system-wide study to assess whether UC patterns of salary compensation associated with gender and race/ethnicity are in compliance with Title IX and other civil rights legislation. That report, authored by Professor Emerita Pauline Yahr (UC Irvine), a former UCAAD chair, was submitted to the UC Academic Council in June 2011 (*Analysis of UC Pay Equity by Sex and, among Men, Ethnicity*; see [Pay Equity Report](#)). In July 2011, the Academic Council forwarded the Yahr salary equity report to UC President Mark Yudof, requesting that the Administration 'provide its own analysis of the report to assist the Senate in its deliberations.' President Yudof submitted his response to the Yahr report in January 2012, which was based in part on responses from individual campuses and on two independent analyses commissioned by University of California Office of the President (UCOP). In response to feedback from both the Senate and the Administration, the UC Academic Council made its final recommendations to President Yudof, and on September 11, 2012, President Yudof outlined to Senate Chair Robert Powell the steps the University would undertake to review faculty salary equity. Importantly, at the urging of the Senate, each campus has been asked to develop its own analyses of salary."

¹ For purposes of this report, the terms "disparity" and "difference" are used interchangeably to describe the comparisons between groups based on gender, race and national origin.

Although it followed the methodologies prescribed by the American Association of University Professor's (AAUP) publication *Paychecks: A Guide to Conducting Salary Equity Studies for Higher Education Faculty* (L. Haignere 2002), the Yahr Report's findings were criticized by many in the UC system, as well as by two external compensation consultants (D.K. Ginther and D.S. Hammermesh) hired by UCOP.

The sentiments held by many about the findings of the Yahr Report were captured in the following quote from the 22 June 2012 letter from Margaret W. Conkey, chair of the UC System UCAAD to Robert Anderson, Academic Council chair:

"Most campuses opine that this study has not convincingly demonstrated that salary inequities and disparities exist due to gender and racial/ethnic discrimination, although most would agree that, minimally, the study's findings are 'troubling.' Even those opinions that most strongly doubt the efficacy of the chosen method and/or the results to date add that further attention, study and action are called for. We also note that some comments or interpretations are not completely accurate. None of these, however, are of such substance to preclude a consideration at this time of the 'next steps' that we would like to have in place by the close of this academic year."

Subsequently, UC Vice Provost Susan Carlson asked each UC campus to initiate a task force that develops a blueprint for assessing possible salary differences by gender and race/ethnicity on their campus. The initial deadline for submitting plans for analysis was 15 January 2013. All campuses submitted preliminary reports of their activities during mid-January of this year. The present document goes beyond the preliminary report to more fully represent the UC Davis blueprint for assessing possible salary disparities.

III. Charge to the UC Davis Task Force

The Provost and the Chair of the Davis Division of the Faculty Senate charged the UC Davis Salary Equity Task Force to develop a plan that: (1) tests for patterns of possible salary differences between genders and among races/ethnicities; (2) explores whether there are high - performing faculty members receiving less off - scale than their colleagues; (3) makes the findings from salary equity studies available to the university community and the public; (4) proposes campus policies, procedures and practices for ensuring that gender and race/ethnicity are playing as little role as possible in affecting compensation for our faculty; and (5) establishes timelines for implementation of the analyses and recommendations. (Vice Provost Maureen Stanton later relieved the Task Force of considering the second charge.)

The charge to the Task Force was not to conduct the salary equity analyses themselves. Rather, our charge was to develop a blueprint for analyzing and addressing possible salary disparities between genders and among races/ethnicities. Ultimately, the university will need to include reliable and valid performance measures into the analysis of salary discrepancies to determine whether any disparities due to gender and ethnicity are truly discriminatory and based on factors other than performance.

IV. Legal Definitions and Considerations Regarding Salary Differences Associated with Gender or Race/Ethnicity

The proposed study is meant to identify patterns of salary disparities, if any, that might exist among faculty at the University of California, Davis, based on gender and race/ethnicity. The existence of a pattern or patterns of salary differences based on gender and race/ethnicity do not, by themselves, indicate the presence of unlawful discriminatory practices or policies of the university that must be modified. Nor does the existence of a pattern or patterns of salary disparities indicate the presence of unlawful intentional discriminatory practices in the university.

The proposed study does not identify policies, practices or mechanisms that may affect pay disparities. Instead, the study aims to identify patterns that the campus might further investigate to determine whether there are underlying practices and policies that affect the presence of such disparities.

The Task Force recommends that the campus review future analyses of possible salary disparities with an eye toward identifying and eliminating possible practices, policies and mechanisms—whether intentional or facially neutral—that might contribute to any identified salary differences based on gender or race/ethnicity.

This report disaggregates salary information along two main axes of analysis: gender and race/ethnicity/national origin. Discrimination based on these categories involves treating someone unfavorably because of a person's gender (sex); because of a person's sexual orientation; because they are from a particular country or part of the world; because of ethnicity or accent; because they appear to be of a certain ethnic background (national origin); because the person is of a certain race; because of personal characteristics associated with race (such as hair texture, skin color or certain facial features); or treating someone unfavorably because of some combination of these factors. Discrimination can also occur as a result of the disproportionate effect of a decision or rule on a group based on gender, race or national origin.

We have chosen to limit the analysis in this study to gender and race/ethnicity. Gender categories include sexual orientation and gender identity. The race categories recommended for this study must be explicitly defined but are likely to differ from those used in the 2010 U.S. Census (Appendix 1). For purposes of this report, and in accordance with the Provost's charge to the Task Force, we recommend a series of analyses using different breakdowns of what we refer to in this report as gender and race/ethnicity. Minimally, each set of analyses should compare two to four faculty groups within a single model, as follows: white faculty members vs. those who are persons of color; white male vs. other faculty; comparisons between white, Asian and other faculty members who are persons of color; and comparisons between white, black, Asian and Hispanic faculty. We believe that analyses of comparisons among those categories will be most illuminating with respect to the patterns that might exist along gender, race and national origin lines. Depending on sample sizes, it may also be possible to examine the statistical interaction between gender and race/ethnicity. Although finer-grained analyses would theoretically be of great value, sample sizes at the campus level are likely to be too small to preserve confidentiality and permit meaningful interpretation of salary variation.

V. Definitions of Salary Variables to be Studied

Total compensation per year for faculty has both a base salary component and a negotiated salary component. We decompose salary into (1) the base salary scale component, which depends on where on the salary scale a faculty member was placed at time of hire, and their subsequent rate of advancement; and (2) all individually negotiated salary payments.

A. Base Salary

The base salary, the scale component, is typically set by unit, and not subject to individual negotiation. Links to the 17 different salary scales are listed below.

- [Table 1: Faculty - Ladder Ranks - Professor Series, Academic Year \(pdf\)](#)
- [Table 2: Faculty - Ladder Ranks - Professor Series, Fiscal Year \(pdf\)](#)
- [Table 3: Faculty - Ladder Ranks - Business/Economics/Engineering, Academic Year \(pdf\)](#)
- [Table 4: Faculty - Ladder Ranks - Business/Economics/Engineering, Fiscal Year \(pdf\)](#)
- Table 5 Scales: Faculty – Ladder Ranks – Professor Series – Health Sciences Compensation Plan:
 - [Scale 0 \(pdf\)](#)

- [Scale 1 \(pdf\)](#)
- [Scale 2 \(pdf\)](#)
- [Scale 3 \(pdf\)](#)
- [Scale 4 \(pdf\)](#)
- [Scale 5 \(pdf\)](#)
- [Scale 6 \(pdf\)](#)
- [Scale 7 \(pdf\)](#)
- [Scale 8 \(pdf\)](#)
- [Scale 9 \(pdf\)](#)
- [Table 6: Faculty - Ladder Ranks - Strict Full Time Preclinical Professor Series, Fiscal Year \(pdf\)](#)
- [Table 7: Faculty - Ladder Ranks - Strict Full Time Veterinary Medicine Professor Series, Fiscal Year \(pdf\)](#)
- [Table 8: Faculty - Ladder Ranks - Law School Professor Series, Academic Year \(pdf\)](#)

Most faculty members are employed on an academic year basis, defined as nine months of employment during the academic year. Some schools/colleges, such as the College of Agriculture and Environmental Sciences and the Graduate School of Management, have at least some faculty who are employed for the fiscal year, defined as 11-months of employment, while the Health Sciences units employ all faculty members on an 11-month basis. Salary equity analysis becomes more complex when an academic unit's faculty members are not all employed on the same term, and instead represent a mixture of either nine- or 11-month appointments. To address this issue, we recommend standardizing all faculty salaries to a fiscal year basis, as described in section VII below.

B. Non-Base Salary Components

Negotiated, additional salary payments for individual faculty members take a variety of forms. Here we outline what we believe should and should not be included in the negotiated salary component for the purposes of analyzing salary disparities associated with gender or race/ethnicity.

i. Off-scale salary

Substantial numbers of faculty receive off-scale salary payments (i.e., payments that are added on to the base salary scale). In some extreme cases, the size of the off-scale payment can meet or exceed the size of the base salary. Off-scale components are recorded in the payroll system as dollar amounts per year, regardless of whether the faculty member has a nine-month or an 11-month appointment. With few exceptions (as defined in individual MOUs), these off-scale payments are permanent components of salary. Except in the Health Sciences, off-scale is the principal, negotiated salary component.

ii. Salary components under the Health Sciences Compensation Plan (HSCP)

Salary arrangements in the UC Davis Health System are complex. All Health Sciences faculty are on 11-month appointments, but very few receive off-scale salary. Instead, other methods are used to make market-based adjustments to total annual compensation. Overall compensation for School of Medicine (SOM) faculty is composed of three elements, as follows.

1. **Total base** (base plus scale, the X + X' components) is determined by rank and step on the "regular" (REG) university ladder faculty base salary scale, whereas additional scale (X') is determined by the usual clinical or research earnings of the faculty member's Academic Programmatic Unit (APU). Each APU in the compensation plan is associated with a multiplier, numbered from 0 to 9, that increases the base by 0–100% for salary and for calculation of

retirement benefits. Multiple APUs are developed within each of the 25 academic departments (e.g., Internal Medicine, Radiology, etc.) to best capture the clusters of faculty within the department whose work overlaps the most.

2. **Negotiated additional compensation** (the “Y” component) is determined by negotiation with the chair, and reflects predictable clinical and administrative activities and/or extramural funding.
3. **Incentive/bonus compensation** (the “Z” component) is divided among individual faculty within a unit if a department has financial surplus after prudent reserves are set aside. Y is determined in negotiation with the department chair based on monies against salary that can be guaranteed with certainty prior to the academic year. Z is determined after departmental or APU earnings are determined for the year. Each department develops a Z distribution plan that prescribes rules for Z eligibility and how distribution of money is determined.

To identify among UC Davis peer faculty systemic salary differences that are associated with gender or race/ethnicity, it will be useful to combine compensation data from all campus units into a single analysis. Accordingly, our proposal for the campus-wide analysis is to treat the total base compensation for SOM faculty ($X + X'$) as comparable to the simple base salary (X) for faculty not on the Health Sciences Compensation Plan. In some versions of this analysis, the 10 APU scales could potentially be pooled across departments. All additional payments to SOM faculty (Y , Z and rarely off-scale) can be summed to calculate the negotiated salary component, analogous to negotiated off-scale salary for faculty not on the HSCP. As explained in Section VII, both base and negotiated salary components will be standardized to the means within each faculty member’s salary comparison unit to allow the pooling of data across units with very different average levels of compensation.

A further note on salary equity in the UC Davis Health system: A major goal of the Task Force is to develop salary equity analyses that can be applied to faculty across all campus units. We note, however, that in any future salary equity analysis focused on faculty in the UC Davis Health System, it will be important to disaggregate the non-base components of total compensation. For example, salary disparities may arise if faculty members of different genders or race/ethnicities consistently choose sub-disciplines that qualify them for APUs at different scales. Similarly, it will be critical to explore the impact of Y and Z salary on total salary. To explore Z adequately, medical specialty and total clinical effort as measured by relative value units (RVUs: a measure of clinical output) will need to be considered in the model, as well.

iii. Extramurally (e.g., grant-) funded summer salary payments

Faculty with extramural grant funding on academic year or 11-month (fiscal year) appointments often receive one or more summer months of salary under these grants, where the full-time monthly salary is the sum of the monthly base and off-scale components. For equity analysis, we would not include summer salary paid from external grants in the negotiated or base components of salary to be analyzed.

iv. Temporary summer salary and administrative stipends

For retention and hiring purposes, academic year faculty members are sometimes paid summer months of salary, often on a temporary basis, from university funds. In addition, faculty with part-time administrative positions such as deans, associate deans, chairs and center directors may be compensated with summer salary. Such temporary or position-contingent summer salaries would not be included in the base scale or negotiated salary components to be analyzed. Administrative stipends and temporary summer support for

administrative appointments are not included. Other summer salary that is under direct administrative control (e.g., that for a dean) should be included in analyses for possible salary disparities.

VI. Overview of Analysis: Understanding Salary Variation at UC Davis

A primary aim of the Task Force is to recommend appropriate and practical, descriptive statistical models that decompose variance in salary among faculty members in similar cohorts and disciplines across the UC Davis campus. In short, the models we recommend determine whether faculty members belonging to groups of interest are overpaid or underpaid, relative to peers of similar hiring cohort and seniority. A full descriptive analysis of possible faculty salary disparities should aim not only to measure current salary variation among designated gender or racial/ethnic groups, but also to understand how different salary components have, over time, contributed to that variation.

In the future, our proposed analytical framework can be extended, if appropriate, to encompass individuals of other groups who might be stigmatized and experience discrimination due to personal characteristics (e.g., sexual orientation or gender identity). It will not be possible to implement such an analysis, however, unless individuals of such a group choose to identify themselves as members. Upon future introduction of these personal characteristics to faculty databases, it will be important to explore the impact of these characteristics on components of total salary and progress through the academic ranks.

Purely descriptive analyses of salary variation do not seek to understand the root causes underlying associations between gender, race/ethnicity and compensation, but do attempt to trace impacts of historical and cohort factors on salary variation. We propose using two kinds of descriptive analysis for analysis of faculty salary equity at UC Davis (see details in Section VIII below).

As a first step, we can conduct a “snapshot” analysis of current, total compensation that can estimate the magnitude of salary variation among groups of interest based on simple and readily available cohort variables. We point out the limitations of this approach in our report, but include it to facilitate comparisons with other institutions that employ this approach. To achieve a better understanding of salary variation, we propose a second, structural equation model analysis that distinguishes between direct and indirect effects and that can identify specific processes and historical trends that have the greatest influence on salary disparities, and thus are potential targets for policy modification. Both of these analyses can be conducted using data that are currently available in our faculty databases.

A descriptive analysis of salary equity can identify some sources of salary variation among faculty groups, but cannot demonstrate the impacts of bias or discrimination without access to more objective measures of faculty performance. The hierarchical analysis described above will allow us to measure the impact of starting step, off-scale and rate of advancement on compensation disparities (if any) associated with gender and race/ethnicity. Because starting step, off-scale and rate of advancement are based in part on assessment of a faculty member’s potential and performance, an aspirational goal of the Task Force is to outline possible strategies for future data gathering and analysis that will help us to understand the role that appraisal of faculty performance may play in influencing salary and rates of advancement across genders and races/ethnicities. As described in Section IX below, preparing for such analyses in the future will require investment of faculty time into selecting informative, discipline-appropriate performance metrics and investment of institutional resources into additional data collection, processing and analysis.

VII. Descriptive Salary Analyses: Scope and Challenges

A. Population of Faculty Members for Analysis

There are many academic titles at UC Davis. We restrict our analysis in this report to faculty members in the ladder-rank professoriate (i.e., tenured or tenure track; hereafter referred to as “faculty”). It should be possible at a later date to extend the methods we recommend to other non-ladder academic titles on campus. Our focus on ladder-rank faculty in the present document should not be misinterpreted as implying disregard for possible salary disparities among non-ladder rank faculty. In this report, however, we are simply using ladder-rank faculty as a point of departure for analyzing possible salary disparities.

B. Addressing Salary Variation among Cohorts and Units

Variation in salary among faculty members is a result of a multitude of factors, many of which are independent of qualifications, job performance or the perception of these attributes. Although there are many ways to decompose salary, our recommended models facilitate comparisons across units by distinguishing just two components of salary: (1) base (for the School of Medicine, total base) and (2) negotiated salary components. For most ladder faculty, off-scale salary is the sole negotiated component of compensation, whereas it is the sum of off-scale, Y and Z components that constitutes negotiated salary for those on the Health Sciences Compensation Plan.

Salaries vary widely among disciplines, and the year of hire of a faculty member can have dramatic influences on both seniority and on the campus-wide pattern of using off-scale salary at hire. To enhance our ability to detect systemic variation in faculty compensation associated with gender and race/ethnicity, we should seek to control statistically for as many of these disciplinary market and demographic factors as possible. The rationale for the approach we propose is outlined in the following sections.

i. Impacts of cohort on base salary

Undoubtedly, some base salary variation associated with gender and race/ethnicity will result directly from the gradual increase in women and people of color in the professoriate over time. For example, we expect that female faculty members are, on average, somewhat earlier in their careers and were hired more recently than their male colleagues within a defined unit in which salaries can be meaningfully compared. This cohort effect is likely to be especially pronounced in disciplines that have been slower to diversify (e.g., science, technology, engineering and mathematics—STEM, as well as medicine and business).

Because progression through the ranks in the UC system is accompanied by a concomitant increase in base salary, senior faculty earn more on-scale base salary than their early-career colleagues. In addition, a faculty member hired into a senior rank after years at another institution or in a non-university position has very different qualifications and earning potential than one hired within a few years of receiving the PhD. Accordingly, analytical models designed to detect salary disparities should include, as covariates, the number of years since hire at UC Davis and the number of years between PhD and hire. Finally, we believe that some versions of salary analyses should account for years in which a faculty member has taken a substantial amount of approved family or medical leave. By doing so, it may be possible to estimate the impacts of family care responsibilities or other medical challenges on salary differences associated with gender and race/ethnicity.

ii. Impacts of academic disciplinary factors on total base salary

Faculty members in different disciplines often receive different base salaries from the university, reflecting nationwide patterns in market forces, the expectation of year-round service and historical norms. Across the UC system, there are a series of standard base salary scales, ranging from REG ladder faculty (at the low end) to the “BEE scale” for business, engineering and economics.

We define total base salary as the summed components of salary that are determined by discipline, rather than by negotiation. Generally, all faculty members within the same nonprofessional school department are either on the REG or the BEE scale, except for those with split appointments between units on different scales. In addition, professional schools at UC Davis regularly use multipliers of the REG faculty scale to derive a market-driven “supplemental” base salary for disciplines or sub-disciplines. One example of this, the APU scales used in the School of Medicine, has already been described. In the School of Veterinary Medicine, the effective base salary for ladder faculty is increased by a Strict Full-Time (SFT) multiplier that is same for all faculty members. The Graduate School of Management determines faculty members’ base salaries according to a separate scale for each academic discipline (e.g., Accounting, Business Analytics, Finance, Marketing, Organizational Behavior). Each scale is derived from national faculty salary data from the business school accrediting body. In the School of Medicine, the base salary multiplier increases with expected clinical earning power and can vary substantially among departments and even within departments.

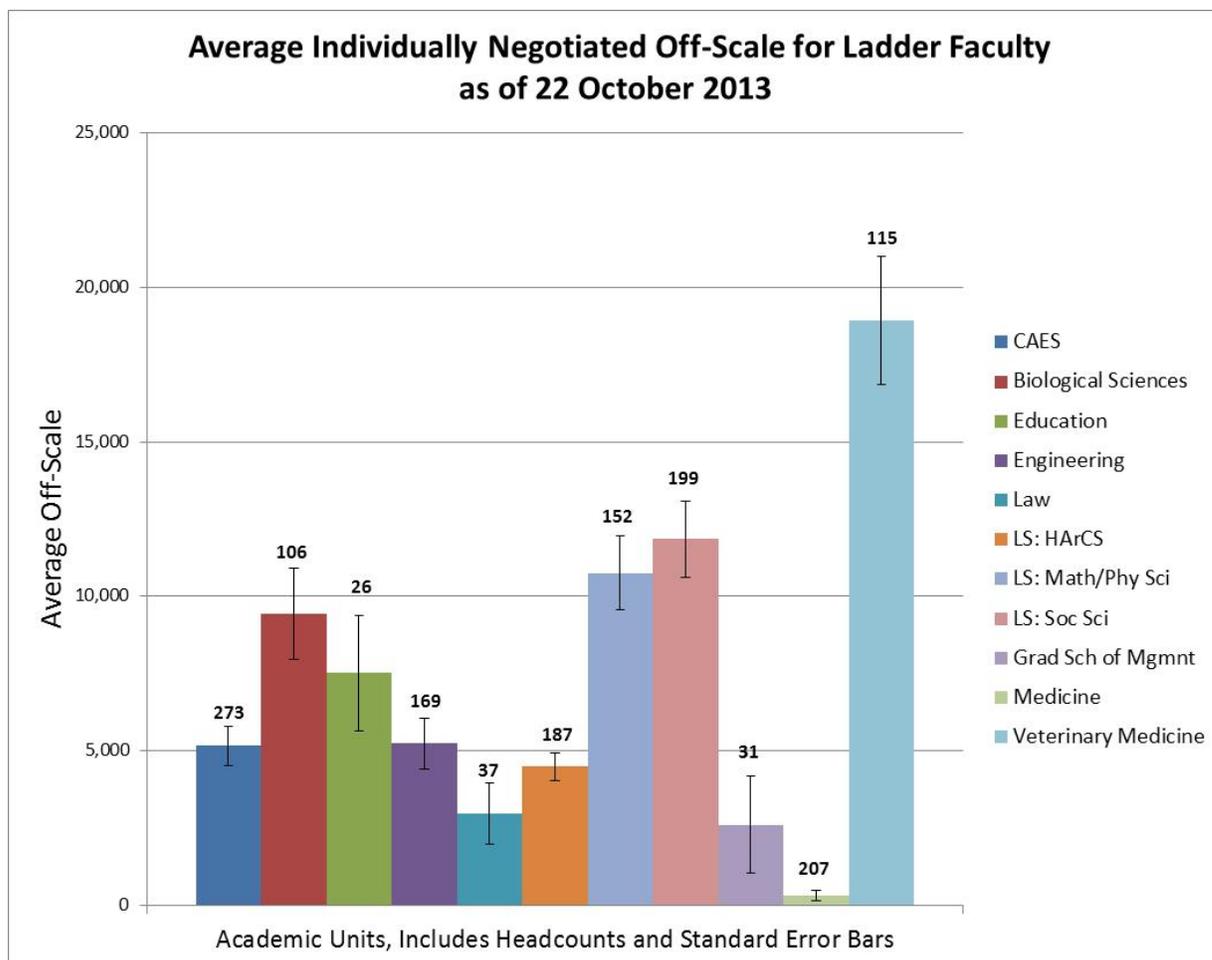
Disciplines also vary in the term of annual appointment. At UC Davis, fiscal year (11-month) faculty appointments are standard in the health sciences, but these appointments are disappearing in the other life sciences and are virtually absent in the other disciplines. In the College of Agricultural and Environmental Science and the College of Biological Sciences, recent faculty hires have academic year (nine-month) appointments, whereas almost all faculty members hired more than 10 years ago have fiscal year appointments, creating a large cohort effect within departments and colleges.

In summary, because of these discipline- and unit-specific differences in base salary scale, we recommend that the total base salary of each faculty member be converted to a fiscal-year equivalent, and expressed as a percentage of the mean within an appropriately defined unit for which salaries can be meaningfully compared.

iii. Impacts of discipline and cohort on negotiated salary components

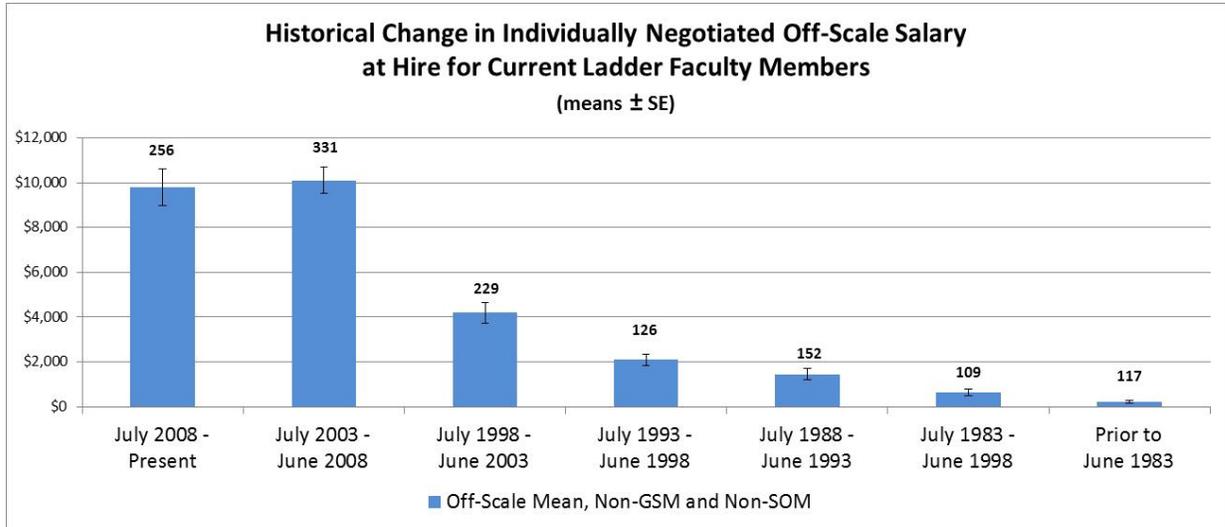
The different REG and BEE base scales reflect only a portion of the very different market forces that impact various disciplines at UC Davis. Accordingly, faculty in different units receive very different off-scale salaries, as shown in Figure 1.

Figure 1



Moreover, although the various base and supplemental base salary scales represent an effort to respond to market pressures acting on faculty in various disciplines, those scales are slow to change, and many have lagged behind salaries at comparable universities. As a result, there has been a dramatic increase over time across campus in the use of off-scale salaries to recruit new faculty and to retain faculty who are being recruited by other institutions. Thirty years ago faculty members were rarely hired with off-scale salary, whereas 73.5% of faculty hires outside of the SOM within the past four years have required off-scale adjustment for our campus to make competitive offers. (See Figure 2.)

Figure 2



Because of this increasing use of off-scale salaries during the past few decades, it is important to account for years since hire at UC Davis in analyzing possible negotiated salary differences associated with gender and race/ethnicity. Off-scale provided during recruitment and retention has generated substantial variation in compensation among current ladder faculty on the Davis campus, both within and among departments. Just as for base salary variation among disciplines and units, we therefore recommend that each faculty member’s annual off-scale salary (or other negotiated salary component) be standardized by dividing by the mean negotiated salary within an appropriately defined unit for which salaries can be meaningfully compared.

iv. Salary comparison units

To facilitate analyses across faculty members from disparate disciplines, we propose that each faculty member’s salary be standardized to the mean of an appropriate comparison unit, which we refer to as a “Salary Comparison Unit” (SCU). A SCU is defined as a collection of faculty who share the same salary reference standard. SCUs are typically represented as departments within a college, APUs in the School of Medicine, or specific disciplines or subdisciplines in a department or professional school.

The SCU averages for base and negotiated salary components, expressed on a fiscal year basis, are the references to which a faculty member’s salary is compared for analysis of potential disparities. In other words, the analyses we recommend will answer the following question, “On average do women and underrepresented minorities earn less (or more) than peer faculty in their SCU who do not belong to these groups?” The market value of a SCU can be determined by the magnitude of off-scale salary routinely offered to recruit or retain faculty members in that discipline (independent of additional off-scale added to an offer because of an individual’s exceptional performance). This market value component accounts for such factors as how disciplines vary in terms of starting salary offers at similar rank from competitor institutions. The appropriate SCU is determined not only by the magnitude of the off-scale salary offered to recruits, but also by the prevailing standards of productivity, teaching and scholarly/creative impact. Using the appropriate SCU is vital for several reasons:

1. The type of scholarship or other creative work varies dramatically across disciplines. For science, engineering and medical fields, for example, refereed journal articles are the main scholarly outputs. This holds true for most of the social sciences, law and business, as well. In the

humanities and in some social sciences, however, books are valued as the primary scholarly indicators for promotion and merits. In the visual and performing arts, exhibitions, shows and performances are the main creative outputs.

2. Disciplines using primary data and/or field data (e.g., history, anthropology) can have long horizons for studies to come to fruition, relative to laboratory-based disciplines, where data may be collected more rapidly or more directly under the scholar's control.
3. Disciplines vary dramatically in the number of journals considered outstanding, the number of articles published (and total pages published) in top journals, acceptance rates of manuscripts submitted to journals, and norms for the speed of the manuscript review process.
4. Publications from laboratory-based disciplines tend to include multiple co-authors. Non-laboratory-based fields tend to be oriented toward single-scholar publications, or publications in which there are a small number of authors.
5. Measures of impact of scholarly work vary dramatically by discipline. In many fields, citation counts may be a reasonable indicator of the impact of one's work on other scholars in the discipline. In other fields, citation counts may be largely irrelevant because there are a small number of journals, small number of members of the discipline, or differing traditions for the appropriateness of citing others' work (e.g., fiction writers rarely cite each other's work).
6. Disciplines appear to vary for average starting rank and step. In economics and business fields, for example, many junior faculty members are hired shortly after, or even before, completion of the PhD. At the other extreme, long postdoctoral periods are standard in many infrastructure-intensive experimental disciplines.
7. In its simplest form, compensation theory (e.g., Milkovich, Newman and Gerhart 2010) posits that a faculty member's salary will be determined by market pressure. Market forces also vary among academic disciplines, creating pressure to elevate off-scales to attract and retain faculty members in fields that are heavily influenced by market forces.

For some campus-wide analyses, schools, colleges or divisions may be the most appropriate comparison units for salary analysis. Such analyses could answer the question, "On average do women and underrepresented minorities earn less (or more) than peer faculty members within their college or school who do not belong to these groups?" Salary differences detected in such coarse - grained analyses may, however, reflect the effects of variation between genders and among races/ethnicities in the choice of sub - disciplines. For finer - grained analyses of relative salary differences within academic disciplines or sub - disciplines, we believe that the SCU will be best identified by the Dean of the college/school/division where the faculty member holds his/her primary academic appointment.

v. Standardization of variables

Variation in base salary due to disciplinary factors is sufficiently substantial that it could potentially mask salary inequities due to other factors. Accordingly, to maximize our ability to detect systemic salary differences based on gender and race/ethnicity among colleagues in similar disciplines and demographic cohorts, we recommend two types of standardization.

1. Each individual's salary should be standardized by dividing it by the mean for the appropriate SCU. Current total base salary, current negotiated salary and negotiated salary at hire should be

standardized separately. In the School of Medicine, this is likely to be the APU. Standardization therefore should be expressed relative to the APU means for each of those components. For details of the calculations, see below. On the main campus, the appropriate standardization unit will often be the department or the professional school.

2. To reduce the confounding effects of appointment term, we recommend expressing all base salaries for academic year faculty in terms of fiscal-year employment (11-month appointments). This conversion (see details below) ignores real differences in guaranteed, university-paid salary among faculty members by assuming that academic-year appointees have access to, and therefore receive, summer salary, but makes it easier to decompose total salary into all of its components on an annualized basis and to compare faculty members hired across the last four decades. This process would entail separating the base salary and off-scale components, then multiplying those values by the appropriate coefficient to make them equivalent to an 11-month (fiscal-year) salary (i.e., academic-year salaries would be multiplied by $\sim 11/9$)². Technically, faculty in the above-scale rank do not receive off-scale salary. Accordingly, to analyze off-scale salary at this highest professorial rank, the Office of Academic Affairs will need to reconstruct what portion of each individual's above-scale salary reflects the legacy impacts of off-scale at the time of advancement to above-scale.

The within-SCU standardizations we recommend are designed to maximize our ability to test the hypotheses that systemic salary disparities may exist between genders and among races/ethnicities for faculty across all cohorts and disciplines. This approach, however, will not detect salary variation caused by social/economic/cultural differences (e.g., family obligations) in the tendency or opportunity to pursue careers in different disciplines (e.g., becoming a neurosurgeon versus a family practitioner) or by differences, real or perceived, in academic productivity.

Standardized starting base salary: Starting base salary (SB) can be standardized to eliminate effects of factors (e.g., changes in salary scale since time of hire) by expressing starting salary in terms of the current salary scale affecting the faculty member's SCU applied to the rank and step at which they were hired. This can be expressed as:

$$SB_{rel_i} = \frac{SB_i}{\overline{SB}_{SCU}} \quad 1.$$

where SB_{rel_i} is the starting base salary of faculty member i relative to other members of the SCU, SB_i is faculty member i 's base salary at time of hire expressed in the current salary scale for their SCU and \overline{SB}_{SCU} is the mean current starting base salary for their SCU. See "Sample Calculations for Variables in Figures 1 and 2 of Final Report" in the supplemental document.

Standardized off-scale salary at time of hire: A faculty member's off-scale salary at time of hire can be standardized to the mean for their SCU. Starting off-scale salary component (SO_{rel}) offered to a faculty member at time of hire can be calculated as follows:

$$SO_{rel_i} = \frac{SO_i}{\overline{SO}_{SCU}} \quad 2.$$

² Under the existing salary scheme, however, academic-year salaries are approximately 5.4% higher than fiscal-year salaries for equivalent rank and step on a monthly basis, as well as on an annual basis if summer salary is obtained and included in the total. Therefore, it will be necessary to incorporate in analytical models an index of which scale is the basis for a faculty member's salary to correct for this difference.

where SO_{rel_i} is the starting off-scale salary of faculty member i relative to that of other members of the SCU, SO_i is the starting off-scale salary component of faculty member i and \overline{SO}_{SCU} is the mean off-scale salary offered to that faculty member's SCU currently. The latter variable can be calculated as the mean starting off-scale salary component currently paid to new appointees of the SCU and the analysis would rely on the cohort covariate to account for differences in starting off-scale salaries offered over time.

Standardized starting rank and step: All else being equal, a faculty member hired at a higher academic rank and step will have a higher current base salary than someone hired in the same year at a lower level. To allow starting step to be included as a quantitative predictor variable in statistical models of salary variation, we recommend that each faculty member's starting rank and step be expressed in normative progress year equivalents (NPYEs). In essence, beginning with Assistant Professor Step I as the baseline (at NPYE = 0), any given starting rank and step would be expressed as the number of years it would *normatively* have taken a UC Davis faculty member to progress from Assistant Professor Step I to the rank and step at hire. For example, because normative time between merits in the Assistant rank is two years, Assistant Professor Step II would have an NPYE = 2, and an Assistant Professor Step IV would have an NPYE = 6. Appendix 2 provides a table of NPYEs up to first above scale. Additional levels above scale would add four years each to the NPYE.

Standardized rate of progression through academic ranks: Faster progression through the ranks will inevitably lead to a faster increase in base salary. Moreover, faculty members who have progressed very quickly through the ranks may be more likely to be targeted for recruitment by other institutions. Accordingly, a faculty member's rate of academic progress should be included as a predictor variable in salary models.

Once ranks and steps are translated into NPYEs, it becomes straightforward to express the rate of a faculty member's progression through the ranks, relative to normative advancement, since hire. Rate of progression (RP) of a faculty member through the academic ranks can be assessed in two ways to provide an index of performance based on assessments made by the existing academic personnel evaluation system: as a fraction of normative time or as a fraction of the mean of their SCU. These can be calculated as:

$$RP_i = \frac{NPYE + Y_{accel} - Y_{decel}}{NPYE} \quad 3.$$

and

$$RP_{rel_i} = \frac{RP_i}{\overline{RP}_{SCU}} \quad 4.$$

where RP_i is the rate of progression of faculty member i , NPYE is the years at UC Davis required to reach their current rank and step in normative years, Y_{accel} is the number of years of acceleration that faculty member has experienced and Y_{decel} the number of years of deceleration. RP_{rel_i} is the rate of progression of faculty member i relative to other members of the SCU, RP_i is faculty member i 's rate of progression and \overline{RP}_{SCU} is the mean current rate of progression for their SCU. Comparison of these two indices indicates the progression of a member of an SCU relative to its other members and allows comparisons to be made between SCUs.

VIII. Proposed Models for Salary Equity Analysis

Our aim in this section is to present two demographic models for analyzing faculty salary equity that are simple, feasible and demonstrably objective. By “demographic” models, we are referring to models including gender, race/ethnicity and nonperformance factors contributing to salary, such as year of hire and years from highest degree to hire. By “feasible,” we mean analyses that could be conducted across the campus within the next two years by a statistically competent postdoc or advanced graduate student, and without the investment of significant campus resources that would be required for the collection and collation of new data. Finally, we view an “objective” model as one that includes readily measurable variables that are widely accepted as relevant to salary variation among faculty members.

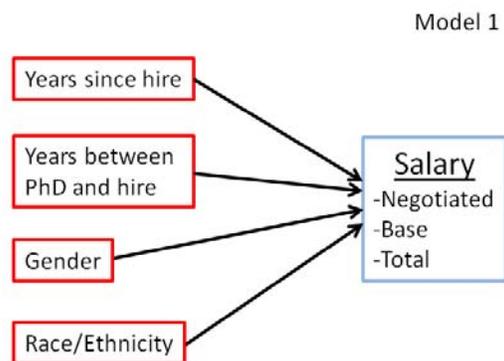
Possible metrics of faculty performance, discussed at length in Section IX below, are critically important for unequivocally demonstrating bias or discrimination in compensation, but are not included in the recommended salary equity analysis models for several reasons. First, there is considerable controversy over the utility of commonly used scholarly performance metrics such as publication rate, citation numbers and *h*-index. Second, it is unlikely that metrics given credence in one discipline will be broadly applicable to other disciplines. Third, UC Davis does not currently have systems in place to extract summary metrics from faculty records across the campus.

Despite those challenges, the Task Force members agree that it will be important to conduct future studies to determine whether peer appraisal of faculty performance during the merit and promotion process is predicted by performance metrics within a given discipline, gender and race/ethnicity and whether faculty voting practices could be changed to reduce the potential impacts of possible unconscious bias during the peer evaluation process. Those recommendations are the focus of Section IX of this report.

Because there are many questions of interest, and different statistical models are appropriate for answering different questions, there is no single approach to analyzing salary equity that is universally best in all contexts. The results from the simplest possible analysis for gender-based inequity (“Do men and women at UC Davis earn different salaries?”) would clearly be problematic to interpret, because there are many factors, ranging from career seniority to years since hire to salary variation among disciplines, that could either cause or obscure gender-based differences. Accordingly, we do not recommend that simplistic model, and do not discuss it further.

A. A Demographic Model of Salary Variation

A straightforward analysis that accounts for two often-used cohort factors is diagrammed in Model 1. With respect to salary differences based on gender, this model answers the following question: “Once we account statistically for years since the PhD and years since hire, do women and men and people of color at UC Davis *still* earn different salaries?”



This type of analysis is more informative than the simplistic model based only on identity factors (i.e., gender and race/ethnicity), especially since the influence of discipline on salary can be accounted for by the standardization methods we recommend above. However, we see three limitations with this approach for understanding the sources of salary variation: (1) it does not separate the impacts of base and negotiated salary components on total compensation, (2) it does not include historical factors that have

contributed to the current salary differences and (3) it does not test for impacts of gender or race/ethnicity on different historical processes or salary components. If the campus wishes to take actions to reduce salary inequities between groups defined by gender or race/ethnicity, it is important to understand how and when such differences arise.

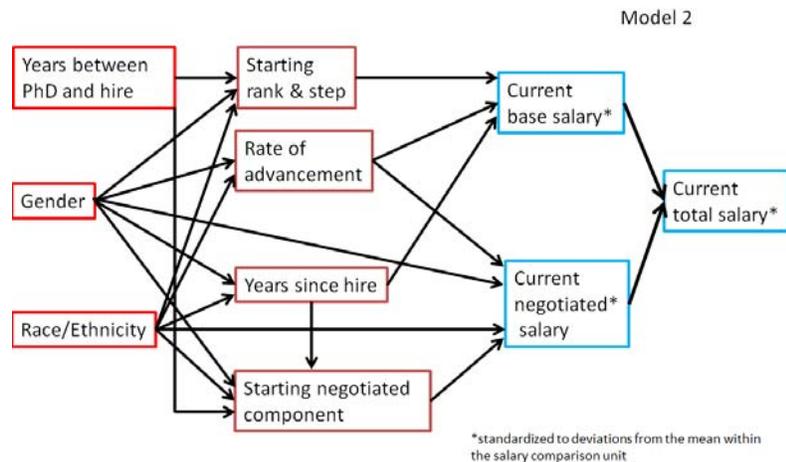
B. A Structural Equation Model of Salary Variation

To reflect the complex processes that contribute to current total salary, we propose the model shown in Model 2. This type of analysis can be conducted within the framework of structural equation modeling (SEM), allowing quantitative estimates of the hypothesized causal pathways shown in the diagram

(including both positive and negative effects). Hypothesized direct effects of one variable on another are shown by directional arrows (e.g., A → B), whereas indirect effects involve mediator variables (e.g. A → C → B).

Within the framework we recommend, salary components be standardized within SCUs, as explained above. A faculty member's current total salary (SCU-standardized) is the sum of the

standardized base salary and the standardized negotiated salary components, and so is directly affected by just these two variables. To simplify language in the following discussion, we refer just to "salary," with the understanding that this means a standardized salary component.



Separating base and negotiated salary components is critical because base salary within a given salary scale is determined entirely by three factors: starting rank and step, years since hire and rate of advancement through the merit and promotion system. Accordingly, any impacts of gender or race/ethnicity on base salary are due only to their *indirect* effects on those three causal variables. In contrast, gender and race/ethnicity could have both direct and indirect impacts on negotiated salary components. Negotiated salary reflects (1) the perceived value of a faculty member to the campus (from both the faculty member's and campus leaders' perspectives), (2) the difference between that faculty member's market value and his or her salary at UC Davis and (3) the motivation or willingness of that faculty member to test the market by seeking outside job offers. All of these factors could be directly influenced by gender and race/ethnicity.

The analytic framework illustrated in Model 2 can be used to decompose the impacts of gender and race/ethnicity on total salary through many different modeling pathways, providing quantitative assessments of questions such as the following:

- To what extent do gender and race/ethnicity impact total salary variation among faculty members within a given discipline and cohort?
- What are the relative contributions of base salary and negotiated salary on any observed salary differences between men and women or among ethnicities?
- If people of color tend to have lower base salaries than white faculty members, to what extent is this due to years since hire, starting step or rate of progression?

- Accounting for years from PhD to hire, do women and men receive different off-scales at hire?
- Does the rate of progression through our merit and promotion system affect negotiated salary, as might be expected if that progression is a predictor of a faculty member's perceived market value?
- Are gender and race/ethnicity associated with differences in the rate of progression through the merit and promotion system?

In short, understanding these interacting and historical factors should help us to address policies and practices at UC Davis in ways that will limit the impacts of structural and implicit gender and/or racial/ethnic factors possibly affecting faculty compensation.

The hierarchical model shown in Model 2 can also provide a framework for identifying individual faculty members who receive substantially more or less compensation than their peers of comparable cohort, seniority and academic progress through the UC Davis merit and promotion system. In this application, the model would probably be run without the gender and race/ethnicity variables, generating residual compensation values for each faculty member. Deans and the Vice Provost-Academic Affairs can screen these residuals to identify high-performing faculty members who appear to be undercompensated, relative to their peers. These faculty members should be considered for equity off-scale adjustments and/or career equity reviews.

The Task Force notes that this campus-wide statistical framework will be expanded to include other potentially relevant factors, including performance variables or excused family/medical leaves. For example, if appropriate data were gathered, we could add other variables to ask whether observed gender differences for starting step and rank or off-scale at hire are associated with gender differences in pre-hire publications or grants, or whether the use of work-life leaves substantially slows the rate of progression up the ranks. Although any given analysis including performance variables is likely to be controversial, one could test for the impacts of objective faculty performance metrics relevant to each of the four areas described in APM 210 within this analytical framework. Which performance metrics are accepted as relevant, however, is certain to vary dramatically across disciplines, and is expected to evolve over time.

In the following sections, we present a discussion of how one could identify discipline-appropriate performance metrics as part of a holistic faculty evaluation within faculty voting groups. We also explore how different peer voting methods may impact fair evaluation in the merit and promotion process.

IX. Toward Future Analyses: Faculty Performance Metrics and Peer Evaluation

A. Overview

Most previous analyses of academic pay equity have included only demographic rather than performance-related independent variables. For example, UCAAD's study (Yahr 2011) relied on the method suggested by the American Association of University Professors (Haignere 2002), which does not consider any measures of performance. This is understandable, due to the difficulty in readily obtaining performance data. Ferree and McQuillan (1998) noted that academic salary studies usually exclude "productivity measures," and indicate that this practice is

"probably necessary if the goal is to create usable models. ... The additional time, expense, and effort necessary to gather information about scholarly productivity, such as publications and research grants, is daunting in itself... . Even if full vitae were available and coded, the challenge inherent in deciding the worth of different journals, books versus articles, and types and amounts of grants is contentious within departments and not feasible across specialties. Scholarship is also not the only

type of productivity that is rewarded, and measures of teaching and service contributions are also poorly developed.”

Despite these challenges, performance is clearly inherent in the UC criteria for advancement, as described in the following excerpts from the APM. From APM 200: “It is the policy of the University to evaluate objectively and thoroughly each candidate for appointment, promotion, or merit increase. Promotions and merit increases are not automatic, but are based on merit.” From APM 210:

“The quality of the faculty of the University of California is maintained primarily through objective and thorough appraisal, by competent faculty members, of each candidate for appointment or promotion. Responsibility for this appraisal falls largely upon the review committees It is the duty of these committees to ascertain the present fitness of each candidate and the likelihood of the candidate’s pursuing a productive career. In judging the fitness of the candidate, it is appropriate to consider professional integrity as evidenced by performance of duties. The review committee shall judge the candidate with respect to the proposed rank and duties, considering the record of the candidate’s performance in (1) teaching, (2) research and other creative work, (3) professional activity, and (4) University and public service.” (For additional details on the APM, see Appendix 3.)

The overarching goal of salary equity analysis is to develop and maintain a system that functions without bias and that accurately advances and rewards faculty members based on their overall performance. In principle, the rank and step of a faculty member should reflect contributions in research, teaching, professional activity (e.g., enabling the success of others) and service, weighted according to how different categories of performance are valued. Upon applying the kinds of analyses discussed in Section VIII, patterns of difference in salary among gender and/or racial/ethnic groups would provide an initial indication of potential systemic bias with regard to faculty compensation. Understanding such differences, and examining the underlying assumption that the academic personnel system evaluates faculty members strictly on the basis of performance, might be achieved, in part, by explicitly measuring some aspects of faculty performance and by understanding how peer evaluation reflects those metrics.

The Task Force does not advocate the replacement of our current, holistic evaluation system with robotically applied numerical algorithms. We do believe, however, that the faculty members in each department, professional school or other voting group should begin to direct effort toward identifying measures of performance and productivity that could be used as part of a comprehensive faculty evaluation in their discipline. Without such metrics, we have no way of unequivocally identifying biases or of understanding how different types of faculty contributions are valued by peers. Moreover, if some of those metrics are included in each dossier for merit or promotion, it will become possible to assess the degree to which peer and administrative evaluation reflects those metrics, and whether current peer voting systems vary with respect to the accuracy of such evaluation.

B. Developing Performance Metrics as Part of the Faculty Personnel Process

Across UC Davis, most evaluation of faculty candidates for merit or promotion occurs by a subjective peer review of academic achievements. Every member of a voting group (a department, school or subdisciplinary group within one of those units) judges the relative contributions of his or her colleagues and votes up-or-down on the proposed action. In addition, in some units, voting peers register summary ratings of the candidate’s perceived performance in research, teaching, service and other activities. Although these evaluations are subjective, this approach draws on the expertise of those colleagues who understand and can explicitly apply the performance expectations of their discipline or subfield. Accurate evaluation by close peers is a critical part of the process.

The advantage of the current up-or-down voting model is that it solves the problem of selecting, measuring and aggregating performance measures that are complex and likely to be controversial. Yet, there are also weaknesses in this approach. First, the manner in which faculty members weigh different components of research, teaching and service is implicit and may not be transparent to all faculty candidates or administrative leaders. In addition, biases that may influence the subjective evaluation of faculty from different groups cannot be either demonstrated or disproven unequivocally without the inclusion of objective data. For example, some members of the campus community may believe that slower merit/promotion progress or less favorable peer review of female faculty, if indeed observed, is due to lower levels of productivity. Others may be inclined to believe that the academic accomplishments of faculty members from gender or racial/ethnic majority groups are systematically overvalued compared to those from underrepresented groups.

In addition to its impacts on the rate of advancement through the ranks, subjective judgment of a faculty member's value to UC Davis can influence the allocation of resources for academic work, assigned responsibilities for service or teaching, and negotiated components of compensation (e.g., off-scale salary). On our campus, off-scale salary augmentations have historically been offered primarily in two situations: at recruitment (especially in the past 10 years, see Figure 2) and in the case of a realized or potential outside job offer. Some off-scale adjustments for equity are now being considered by the deans and the Vice Provost-Academic Affairs, but these are again based principally on implicit and subjective assessments, and are not evaluated by faculty peers.

We believe that analysis of academic performance metrics could help our campus decide whether negotiated salary components and other resources are being allocated fairly by university administration (e.g., deans, the Vice Provost-Academic Affairs and the Provost).

To explore the adequacy of our traditional, implicit evaluation system, the Task Force recommends that some units on our campus undertake an effort to supplement subjective summary assessments with explicit counts and weights of the various components in each area of faculty performance. Once the quantitative accomplishments of faculty in research, teaching and service are enumerated, weights or points reflecting the performance norms of their voting unit could be applied to evaluate the worth of each of these achievements, as perceived by peers. It will then be possible to determine the degree to which the current implicit assessments of a faculty member's achievements (e.g., reflected in the up-or-down merit or promotion vote or the granting of an off-scale salary increment) are aligned with more objective performance metrics.

An advantage of a more explicit performance appraisal approach is that the voting unit's norms and value weightings are laid bare for all to judge. Once each voting unit has provided an explicit numerical template of its performance values, then future investigators can measure the association between subjective assessments and objective measures. Moreover, faculty coming up through the ranks will have access to a very explicit framework that establishes the standards for positive peer evaluation and advancement.

C. How to Develop Informative Faculty Performance Metrics?

The expectation that meritorious performance should determine peer evaluation, compensation and access to resources, without regard to gender and race/ethnicity, is an ideal on which the University of California was built. The task of identifying operational measures of faculty performance that can accurately test whether the personnel system allocates salary consistent with that ideal is a challenge that the university has yet to meet.

The use of quantitative variables to capture some elements of a faculty member's performance is highly controversial, as it is clear that no set of quantitative variables can fully reflect a faculty member's contributions in any of the areas evaluated for advancement and consequent salary increases: creative activity/research, teaching, professional competence/activities and service. Among the challenges faced in developing and applying metrics to any campus-wide analysis are the following.

- Even with respect to one evaluation area (e.g., research/creative activity), developing numerical estimates of performance is difficult. In the area of professional service, for example, simple counts of committee service, articles and grants reviewed, and panels convened would give some indication of activity, but would be insufficient to indicate leadership or degree of effort directed to these activities. Appropriate indicators of productivity, both in quality and quantity, are likely to vary dramatically among disciplines.
- Faculty are evaluated based on many different types of functions, and so metrics developed to capture some elements of faculty performance would probably need to be combined to yield an overall quantitative assessment that reflects the performance norms and the value weightings each voting unit places on different activities and achievements. To the extent that units differ in the relative values they place on different activities or performance areas, the development of informative combined variables is also likely to vary greatly among units.

For these reasons, the Task Force recommends that future development of performance metrics be done at the local, voting group level. Beyond the four performance categories identified by the APM, research/creative activity, teaching, professional competence and service, we do not recommend a uniform performance metrics across faculty voting groups. Each group should develop its own operationalization of the four performance metrics in the APM.

Development of some discipline-appropriate faculty performance metrics would allow the campus to explore a number of questions pertinent to faculty salary equity. We present a few of these issues for illustration.

- Is there a time-dependent bias in measuring the impact of research publications? For example, do faculty members who publish a high rate of first-tier journals and/or academic presses and generate impressive numbers of citations in the short run, especially in the two- to three-year timeframe for merits, receive greater recognition and pay than faculty members who publish a smaller number of pieces in comparable venues but eventually achieve the same overall citation impact?
- In fields that do not rely on journal articles and citation counts but rather examine larger, book-length manuscripts, how much weight is given to the press where publication occurs as an immediate proxy for value?
- Are there other measures of research/creative impact that may only fully emerge over time where citations are unavailable?
- How much value do colleagues place on department and university service? And, do the values given to these activities vary among units on campus?
- Do faculty of different gender or race/ethnicity have different priorities for activities related to teaching/mentorship, research/creative activity or service?

In this spirit, we present the following nonexhaustive list of examples of potential performance measures, acknowledging that all of these putative metrics have significant flaws and limitations.

Research/creative work

- Refereed publications (possibly weighted by candidate's voting unit)
- Average tier or impact factors of journals
- Presses in which publications appear

- Exhibitions and performances, including reviews of such
- Citations (excluding self-cites) published reviews
- *h*-index
- Invention disclosures and patents

Teaching/mentorship

- Evaluation scores by students (with percentages of students represented in the samples)
- Enumeration of courses taught, their level and their enrollments (e.g., student credit hours)
- Number of graduate students: current vs. completed, master's vs. PhD, major professor. vs. member
- Awards based on distinguished teaching or mentoring
- Awards received by mentees
- Evaluation by other faculty members of teaching effectiveness
- Curriculum/course innovation: numbers of new courses developed or courses significantly revised
- Awards and fellowships based on research/creative activity

Professional competence/activity

- Number of manuscripts reviewed
- Number of editorial positions (editorships, editorial boards)
- Professional association leadership roles, advisory boards
- Number of conference committees
- Professional registration/certification
- Awards related to competence/activity
- Licensure/credentials

Service

- Numbers of committees, as member vs. chair
- Part-time administrative service (e.g., dean, associate dean, chair, vice chair, director)
- Awards based on service
- Enabling success of others, such as doctoral students, postdocs and junior faculty

Given the diversity of scholarly activities and expectations among disciplines, it is hard to envision any uniform evaluation method that can be meaningfully applied across all units on campus. Indeed, simplified career metrics (e.g., the *h*-index for publications) cannot be applied to many areas of endeavor, and at best can only serve as one component of an integrated assessment of performance for faculty of similar seniority and similar sub-disciplines. For this reason, it is critical for faculty in each discipline to discuss and decide upon the best possible metrics for their field.

With the development of some discipline-specific measures of faculty performance, it will become possible to determine the extent to which peer evaluations by faculty colleagues are predicted by values of one or more of these agreed-upon performance variables. As outlined in the next section, options for performing such analyses depend on the qualitative and/or quantitative methods used to vote on merit or promotion advancements within a faculty member's department, school or other voting group.

D. Potential Impacts of Voting Methods on Performance Assessment

Psychological research suggests that alternative appraisal methods are likely to vary in the extent to which they are objective or subject to unconscious biases based on gender or race/ethnicity (Murphy, K.R. and Cleveland 1995). Based on the fact that a number of different peer voting systems are currently used on our campus, we propose a general approach that could be used to evaluate how voting method may influence faculty performance assessment.

Up-or-down voting: Members of most departments, schools, personnel review committees and other voting groups at UC Davis vote up-or-down on each proposed action, with an overall vote of “yes” indicating that the voter supports the action, and a vote of “no” (accompanied by a written rationale) indicating that the proposed action is not supported. An abstention or a failure to vote can be interpreted in a variety of ways unless specific explanatory information (e.g., ineligibility to vote) is provided in the department or committee letter. In this system, the strength of a group’s support for a proposed merit or promotion action is typically gleaned from the proportion of “yes” votes (versus “no” or “abstain”), and sometimes by the proportion of those eligible to vote who actually do so.

One could assess the strength of the relationship between up-or-down voting and discipline-specific faculty performance quality measures (however imperfect) through a fairly straightforward statistical analysis. For example, the proportion of “yes” votes registered for each recent faculty advancement action within a unit can be regressed against various performance metrics for each faculty member at the time of evaluation. To conduct such an analysis, standardized performance metrics must exist.

Including performance rating in voting: Instead of relying solely on up-or-down voting, a number of departments in the College of Agricultural and Environmental Sciences (CAES), and all units in the School of Veterinary Medicine (SVM), accompany a summary vote with a set of more nuanced assessments in which faculty peers quantitatively rate each candidate’s performance in specified academic areas. In this system, the strength of a group’s support for a proposed merit or promotion action can be estimated in two ways: by the proportion of “yes” votes (versus “no” or “abstain”) as well as by the assessments given by peers in the various areas for evaluation. In short, a “yes” vote may be stronger or weaker, depending on the performance assessments made of the faculty member. This additional information may aid those making final personnel decisions.

As for the pure up-or-down system, the correspondence between the means for these quantitative (albeit still subjective) ratings and discipline-appropriate performance metrics calculated for each faculty candidate can be assessed by regression analysis. Unlike the up-or-down system, in which the proportion of “yes” votes on a particular faculty member yields a single indicator of peer support for the action, the performance rating system generates a mean rating, a median rating and estimates of dispersion for each area in which a faculty candidate has been evaluated.

Designing and implementing analyses on the relationship between faculty performance metrics and peer assessment is beyond this Task Force’s charge, but we recommend that the campus begin planning a pilot effort. The first step will be to identify approximately 20 departments or other voting units that are willing to propose and use metrics that can aid in the appraisal of performance in research/creative activity, teaching and service during the merit and promotion process over two or more annual merit cycles. Ideally, these units will represent a range of peer voting or assessment methods.

The second step will be to statistically analyze the strength of association between peer assessment scores and the academic unit-specific metrics to (1) assess the apparent impact of different performance metrics on peer assessment and (2) to explore whether the use of ratings of candidate performance in different areas of academic endeavor enhance the apparent objectivity of the peer evaluation process. It must be stressed that a poor fit between performance metrics and evaluation outcomes could arise through a combination of two factors: the use of uninformative metrics and inaccurate (and potentially biased) peer evaluation. In contrast, if the quantitative relationship between performance metrics and assessment scores varies among groups (e.g., by gender and/or race/ethnicity), the evidence for bias in the appraisal process is strong.

There are a number of complexities that should be considered when developing and using a rating system to appraise faculty peers. For example, some faculty members may consistently be more critical when evaluating proposed merit advancements than others. One might also imagine that highly ranked departments will have high performance expectations for scholarship or that some departments place especially high value on clinical productivity or service contributions to diversity. Such assessments are extremely complex and must not be reduced to a computer algorithm.

Nevertheless, we believe that more informed decisions can be made by introducing a more rigorous decision-support system, based on the following actions:

1. Deans determine the SCUs, as an appropriate subdivision of their administrative unit, to provide "voting units" (VU) for assessment of faculty performance and impact. VUs are defined as those groups of faculty members who are authorized to vote on each other's academic personnel cases. These units might be departments or SCUs; however, they could be smaller units if faculty specialization dictates.
2. The appraisal of faculty performance can be done in one of two ways, or both. The first method is to have each faculty member *rank* the relative impact of all members of their VU other than themselves in each of the performance categories being evaluated. Those values are tabulated and the percentile score of each member within the VU is quantified and reported as the relative standing of that individual in that category within the SCU.

The alternative method is to have each member of the VU *rate* the impact in each category of each of the other members of the VU independently. This could be done on a fixed numeric scale (e.g., 1–5 or 1–9) indexed as representing poor to outstanding. For example, in the CAES the scale of such ratings varies among units (e.g., from "Poor" to "Excellent," or from "Does not meet expectations" to "Exceeds expectations"). Moreover, the areas in which peers appraise faculty performance vary among units. In CAES departments using this system, faculty members are typically rated in 3–4 areas: research, teaching, service and (often) professional competence. In contrast, faculty members in the SVM currently rate their peers' performance in many different categories, including various aspects of clinical activity and clinical teaching. Using rating methods, it would be necessary to standardize the individual assessor's values to have a mean of 5/10 among assessors, thus yielding an unbiased composite VU average for each member. Analyzing the variance within the data might be useful, as well. It is imperative to use a standardized scoring scale and wording in generating these indices, and care must be taken that the wording encompasses the array of activities in which faculty engage across diverse disciplines.

3. Because the appraisal in the previous step indicates the peer-generated index of a faculty member's impact in each category, it is necessary to further weight this value by an index of the VU's overall quality. Determination of this value might be contentious. However, this is essential, as the impact of an individual with average impact from an average VU is not equivalent to an average individual from a superior VU with higher performance expectations or standards. If the VU were a department, indices such as the latest NRC rankings (if timely) might be incorporated into determining this VU weighting factor, at least for a category (e.g., creative activity). Other metrics would be required to determine if certain VU should be weighted differently for certain categories (e.g., teaching or service). Knowledge transfer activities could be included in the characterization of a VU's overall quality. Activities to enhance diversity would be factored into determining an individual faculty member's impact in service and/or teaching.

4. An optional process that could be employed in parallel with this peer review would be to assemble various quantitative metrics (e.g., *h*-index, student evaluation scores of teaching, etc.) and incorporate them to develop a computer-generated index of performance for each of the categories. This might require assigning “point values” to different types of service activities as a way of indexing them with a digit (e.g., receiving 10 points for serving on an admissions committee with interviews but only 1 point for reviewing a manuscript). Over time, it would be possible to compare the results of the peer review and computer review and perhaps refine both.
5. As a means of evaluating the overall appraisal methodology and indices it generates, it would be possible going forward (but not *a posteriori*) to have the Committee for Academic Personnel–Oversight Committee (CAP–OC) and the Vice-Provost for Academic Affairs (VPAA) begin to assign numeric impact scores to each of the performance categories in packets they review. These values can be compared with those generated by VU and by computer algorithm. The VUs only make comparisons across a narrow range of colleagues, whereas CAP, the VPAA and the algorithm would all make comparisons across the entire faculty. Systematic variation between these entities could provide useful insight into the strengths and weaknesses of the evaluation systems. However, because CAP only evaluates personnel and conducts their equity reviews in terms of rank and step, not salary, their evaluations would only affect base salary and the extent to which that is sensitive to performance and rate of advancement. Market value, retention and other factors are independent of CAP’s purview.

Potential advantages of evaluating performance in specific areas: Based on past research, the study described is likely to provide evidence that evaluating faculty candidates with respect to defined areas of the academic productivity will increase the correspondence between informative performance metrics and peer appraisal measures, compared with the simple up-or-down voting scheme. We briefly outline the rationale for this prediction below.

Psychological research indicates that all forms of appraisal are potentially influenced by unconscious bias, regardless of the appraiser’s belief system and principles (Murphy, K.R. and Cleveland 1995). For example, numerous studies have shown that both men and women tend to undervalue the achievements of female candidates for positions requiring technical expertise or leadership qualities, even when the presented record is identical to that of a male candidate (e.g., Bowen, Swim and Jacobs 2000; Eagly and Karau 2002; Lyness and Heilman 2006). Although it is difficult to eliminate such unconscious biases altogether, research also shows that the impacts of unconscious bias are reduced if the appraiser is asked to use specific criteria to evaluate specific kinds of achievements (Murphy, K.R. and Cleveland 1995). Based on these findings, it seems likely that the reliance on simple up-or-down votes based on a gestalt of the candidate’s whole record are likely to be more subject to unintentional bias than are more explicit and detailed appraisals of specific elements of the record.

Furthermore, the simple up-or-down voting system provides less information for further evaluation by review committees and administrators, making it difficult to gauge the enthusiasm (or lack thereof) with which peers support the proposed merit or promotion action. Under the proposed UC Davis Step Plus merit and promotion system, which the campus hopes to implement within the next two years, each dossier will be evaluated for advancement in partial step increments and for potential acceleration. More nuanced peer evaluation of specific performance areas may help review committees and administrators to make more informed judgments about incremental step advancements. In addition, it is important to weigh the value colleagues place on different academic activities. Specific area ratings provided in peer assessments could be analyzed further to measure the impact of different academic activities (or the assessments thereof) on success during merits and promotions.

Finally, having the ability to appraise a faculty candidate's performance in specific areas may facilitate more honest and objective review of peers. One unintended consequence of the simple up-or-down voting scheme appears to be a reluctance to vote against a colleague's advancement, as indicated by votes that are usually more positive at the level of the department or school than at later stages of the review process. Indeed, there is evidence to suggest that binary decisions, such as up-or-down votes, may increase the likelihood that favorable decisions are made if a person meets minimally acceptable performance standards (Murphy, K.R. and Cleveland 1995, p. 157). The Task Force believes that any revisions in our peer review process should encourage more accurate and informative appraisal of faculty performance at the level of the voting group.

If the UC Davis campus moves to a peer assessment system for merit and promotion, as opposed to the existing voting system, it will be appropriate and advantageous for units to customize the set of activities they select for specific evaluation. For example, appraisal of clinical performance is relevant only in the health science disciplines. Similarly, mentorship of postdoctoral scholars is more important in STEM disciplines than in the arts and humanities. However, some areas for appraisal should consistently be included by *all* units. Minimally, these areas are teaching/mentoring, scholarship, service and contributions to diversity. A somewhat finer breakdown of areas for academic appraisal could be the following: research/creative activity, undergraduate teaching/mentoring, graduate student mentoring/training, university service, professional service, public service, contributions to diversity. Whatever shared set of performance areas are selected, these can be evaluated by review committee members, as well as by department colleagues.

X. Implementation Recommendations

The Task Force recommends the following:

1. A full, campus-wide analysis of salary equity, based on the model framework shown in Model 2, should be carried out and disseminated to the university and the public every other year. The primary focus will be on detecting differences in faculty compensation associated with gender and race/ethnicity. A secondary focus will be to understand how other demographic and historical factors influence salary variation among faculty peers, and whether these factors differentially influence compensation in faculty from different gender or racial/ethnic groups.
2. The analysis should be carried out, not by the university administration and not by the Academic Senate, but rather by a neutral third party such as a postdoctoral scholar, a member of the on-campus Statistical Laboratory consulting service or an outside statistical consultant. When other universities have used a third party to conduct equity analysis, this practice enhanced the perceived credibility of the findings in the eyes of members of the university community.
3. If systemic disparities in compensation, rate of advancement or starting step are found to be associated with gender or race/ethnicity, then the analysis has to move beyond the analysis of model 2. Performance metrics must be developed and included in future statistical analyses to determine whether these disparities in pay are due to differential levels of academic achievement or inequitable treatment. The Provost, Vice-Provost and deans will work with the Academic Senate leaders to have individual academic units identify performance metrics that represent their norms for evaluating scholarship, teaching and service. The ideas in section IX should provide guidance in developing these metrics.
4. To aid in fulfilling recommendation 3, we recommend that the Senate establish a new standing academic committee for faculty compensation, and that appointees to this committee have appropriate

expertise in quantitative analysis and social science methodologies. We also recommend that the associate Vice Provost for Faculty Equity and Inclusion be an *ex-officio* member of the committee.

5. Using the set of pilot academic units that develop performance metrics, the campus should conduct analysis to determine whether the use of performance appraisal in specific areas of faculty activities influences the relationship between that unit's agreed-upon performance metrics and appraisal scores and outcomes. Depending on the outcome of this analysis, changes might be proposed in the voting procedures used by departments and review committees.
6. A critical responsibility of deans is to ensure that faculty are fairly compensated within their units, so we anticipate that the decisions deans make in adjusting compensation will be informed by the analyses recommended in this report.

XI. Analytical Considerations for Campus-wide Use of Faculty Performance Metrics in Salary Equity Analysis

This report proposes two distinct strategies for appraising faculty performance (ranking and rating) and specifying a model that seeks to explain the differences in their salaries. Each approach carries with it strengths and weaknesses. Both approaches produce a single score for each of the three areas of faculty performance. Because they generate these scores by different means, those who implement these ranking or rating strategies can examine whether faculty members are treated fairly at different stages of the personnel process.

The overall score for each faculty member is computed as an average of the numerical assessments of that individual's VU colleagues. Generating subjective summary measures for each area of performance also simplifies specification of the statistical model of salary differences.

Research/creative productivity and scholarly impact are often expected to be the most important of the three areas of faculty performance in determining salary differentials. But the judgments of the scholarly significance of journal articles, books or research awards by one's peers may be buried in their subjective summary assessments, along with the components of teaching and service.

The justification for appraising performance with peer review is straightforward. This process replicates how the current academic personnel system operates by drawing on the judgments of faculty members who share common training. The principal difference between the current system and a more explicit empirical appraisal system is that the latter asks faculty who already fulfill this role to go one step further and summarize their judgments of each of the four areas of faculty performance with a single numerical score.

Although the model shown in Model 2 accounts for the average rate which a faculty member has ascended through the UC ranks, it does not capture any performance criteria on which faculty merit has presumably been appraised. Demographic models do not capture any of the attributes of faculty performance. These omissions in and limitations of the available data inevitably constrain how well we can specify faculty performance and how much confidence we can have in our statistical analyses of the likely causes of possible salary differentials among faculty.

The metrics used in formulating a ranking or rating model that evaluates faculty performance will be made available to the faculty, and the weighting of variables employed in that model will be published. Therefore, it is conceivable that some faculty members might be tempted to direct their activities toward those endeavors, particularly in service, that are more heavily weighted toward indicating a faculty member deserves a higher salary. For that reason, it is essential to clearly state that results of the analysis will be

only one of multiple factors that will be considered for salary equity adjustment and that such decisions will not be solely determined by a statistical model. Indeed, in presenting the advantages of introducing more rigorous quantitative methods of appraising faculty performance, we are not suggesting that the campus transition to a purely formulaic strategy for assessing faculty performance. We see a move toward more rigorous methods as development of a more systematic decision-support system that serves those who must exert judgmental discretion in final assessment of faculty performance. Such assessments are extremely complex and must not be reduced to a computer algorithm.

A. Measurement and Analytical Challenges

The promise of developing objective metrics is substantial. But we believe that the challenges of measuring the wide range of activities considered in each area of performance and determining the numerical weights is worthwhile. Generating the appropriate metrics, evaluating them and defending them across the large number of VUs in the university will certainly take much longer than subjective peer review. Before these problems are solved, investigators will have to come to grips with data that do not meet these ambitious standards.

A basic question will be how to use objective performance metrics that are only partially specified in a multivariate analysis of salary variation. Any summary index that draws upon subjective or objective data suffers from having no obvious analogue for interpreting the coefficients. A unit change in a multidimensional index of research or service has no intuitively obvious referent other than the categorical label of research, teaching or service productivity. The multivariate analysis of salary differentials across faculty using these disaggregated performance measures will face some predictable problems, such as:

1. Omitted variable bias. If key measures of performance are omitted, statistical models will be suboptimal.
2. Inherent bias in variables. A factor that could bias the outcomes of analyses is that some faculty members opt not to disclose their ethnicity as part of personnel reporting. At UC Davis, this is currently approximately 4% of the faculty. If individuals who choose not to report their ethnicity are not randomly distributed among ethnicities, particularly if those individuals are clustered in a group that might be more likely to experience ethnically based differential treatment, then their omission from the sample could bias the analytical results.
3. Inadequate measures. As investigators develop indicators to measure the omitted variables, some early efforts will suffer from a failure to capture fully the underlying accomplishments of faculty. Different VUs, for example, will have to specify their preliminary indicators of research productivity such as citation indices or some other appraisal system that accords particular weights to book-length manuscripts or other forms of creative activity. These efforts are likely to require many iterative refinements to achieve broad acceptance; the early forays may not faithfully replicate the research norms of a VU.
4. Multiple performance measures and multicollinearity. If the accomplishments of faculty members include several indicators for each area of performance appraisal, the available independent variables may be highly correlated with each other. For example, it will be difficult to identify the separate effects, for example, of three different indicators of research productivity in a scientific discipline because scores on a citation index will often be correlated with variables reflecting grant success and research awards. Multicollinearity resulting from highly correlated measures (e.g., correlations of approximately $\geq .8$) would bias the findings of the statistical analyses.

References

- Bowen, C.C., Swim, J.K., and Jacobs, R.R. 2000. Evaluating gender biases on actual job performance of real people: A meta-analysis. *Journal of Applied Social Psychology* 30: 2194–2215.
- Currall, S.C., Towler, A.J., Judge, T.A., and Kohn, L. 2005. Pay satisfaction and organizational outcomes. *Personnel Psychology* 58: 613–40.
- Eagly, A.H. and Karau, S.J. 2002. Role congruity theory of prejudice toward female leaders. *Psychological Review* 109: 573–98.
- Haignere, L. 2002. *Paychecks: A Guide to Conducting Salary-Equity Studies for Higher Education Faculty*. Second ed. Washington, DC: American Association of University Professors.
- Lyness, K.S. and Heilman, M.E. 2006. When fit is fundamental: Performance evaluations and promotions of upper-level female and male managers. *Journal of Applied Psychology* 91: 777–85.
- Milkovich, G., Newman, J., and Gerhart, B. 2010. *Compensation*, 10th edition. New York: McGraw-Hill/Irwin.
- Murphy, K.R. and Cleveland, J.N. 1995. *Understanding Performance Appraisal: Social, Organizational, and Goal-Based Perspectives*. Thousand Oaks, CA: Sage Publications.
- Marx Ferree, M. and McQuillan, J. 1998. Gender-based pay gaps: Methodological and policy issues in university salary studies. *Sociology Department, Faculty Publications*. Paper 36. University of Nebraska–Lincoln. <http://digitalcommons.unl.edu/sociologyfacpub/36>
- Yahr, P. 2011. *Analysis of UC pay equity by sex and, among men, ethnicity, 2009–10*. (UCAAD study report)

Appendix 1: Race and Ethnicity Categories: Definition of categories used in the 2010 Census

- “White” refers to a person having origins in any of the original peoples of Europe, the Middle East or North Africa. It includes people who indicated their race(s) as “White” or reported entries such as Irish, German, Italian, Lebanese, Arab, Moroccan or Caucasian.
- Persons of color include “Black or African American,” “American Indian or Alaska Native,” “Asian,” or “Native Hawaiian or Other Pacific Islander.”
- “Black or African American” refers to a person having origins in any of the Black racial groups of Africa. It includes people who indicated their race(s) as “Black, African American or Negro” or reported entries such as African American, Kenyan, Nigerian or Haitian.
- “American Indian or Alaska Native” refers to a person having origins in any of the original peoples of North and South America (including Central America) and who maintains tribal affiliation or community attachment. This category includes people who indicated their race(s) as “American Indian or Alaska Native” or reported their enrolled or principal tribe, such as Navajo, Blackfeet, Inupiat, Yup’ik or Central American Indian groups or South American Indian groups.
- “Asian” refers to a person having origins in any of the original peoples of the Far East, Southeast Asia or the Indian subcontinent, including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand and Vietnam. It includes people who indicated their race(s) as “Asian” or reported entries such as “Asian Indian,” “Chinese,” “Filipino,” “Korean,” “Japanese,” “Vietnamese” and “Other Asian” or provided other detailed Asian responses.
- “Native Hawaiian or Other Pacific Islander” refers to a person having origins in any of the original peoples of Hawaii, Guam, Samoa or other Pacific Islands. It includes people who indicated their race(s) as “Pacific Islander” or reported entries such as “Native Hawaiian,” “Guamanian or Chamorro,” “Samoaan” and “Other Pacific Islander” or provided other detailed Pacific Islander responses.
- “Some Other Race” includes all other responses not included in the White, Black or African American, American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander race categories described above.
- Respondents reporting entries such as multiracial, mixed, interracial or a Hispanic or Latino group (for example, Mexican, Puerto Rican, Cuban or Spanish) in response to the race question are included in the “Other” category. The national origin or ethnicity category includes country of origin or ancestry, regardless of race. For purposes of this study, the Hispanic or Latino category is considered a national origin category.
- The UC payroll/personnel system had not provided a category for Pacific Islander until 2011. Thus, we do not currently have accurate data on Pacific Islanders.
- “Sexual orientation” is a status that is protected from discrimination in the workplace. “Sexual orientation” in California is defined as heterosexuality, homosexuality and bisexuality. Sex discrimination can include discrimination based on a person's gender, or a person's gender-related appearance or behavior whether or not stereotypically associated with the person's assigned sex at birth.

Appendix 2: Normative Progress Year Equivalents

Rank	Step	Years At	NPYE *
Assistant	1	2	0
Assistant	2	2	2
Assistant	3	2	4
Assistant	4	2	6
Assistant	5	2	8
Assistant	6	2	10
Associate	1	2	8
Associate	2	2	10
Associate	3	2	12
Associate	4	3	14
Associate	5	3	17
Full	1	3	14
Full	2	3	17
Full	3	3	20
Full	4	3	23
Full	5	3	26
Full	6	3	29
Full	7	3	32
Full	8	3	35
Full	9	4	38
Full	A/S **	4	42

* Normative progress year equivalents

** First above scale, each above scale thereafter is four years greater than the previous above scale

Appendix 3: Academic Personnel Manual (APM) Statement Regarding Components of Faculty Member Performance and Commentary on Challenges in Measuring Performance

Statement from APM 210-1d regarding faculty performance:

“The criteria set forth below are intended to serve as guides for minimum standards in judging the candidate, not to set boundaries to exclude other elements of performance that may be considered.

*(1) **Teaching** — Clearly demonstrated evidence of high quality in teaching is an essential criterion for appointment, advancement, or promotion. Under no circumstances will a tenure commitment be made unless there is clear documentation of ability and diligence in the teaching role. In judging the effectiveness of a candidate’s teaching, the committee should consider such points as the following: the candidate’s command of the subject; continuous growth in the subject field; ability to organize material and to present it with force and logic; capacity to awaken in students an awareness of the relationship of the subject to other fields of knowledge; fostering of student independence and capability to reason; spirit and enthusiasm which vitalize the candidate’s learning and teaching; ability to arouse curiosity in beginning students, to encourage high standards, and to stimulate advanced students to creative work; personal attributes as they affect teaching and students; extent and skill of the candidate’s participation in the general guidance, mentoring, and advising of students; effectiveness in creating an academic environment that is open and encouraging to all students, including development of particularly effective strategies for the educational advancement of students in various underrepresented groups. The committee should pay due attention to the variety of demands placed on instructors by the types of teaching called for in various disciplines and at various levels, and should judge the total performance of the candidate with proper reference to assigned teaching responsibilities. The committee should clearly indicate the sources of evidence on which its appraisal of teaching competence has been based. In those exceptional cases when no such evidence is available, the candidate’s potentialities as a teacher may be indicated in closely analogous activities. In preparing its recommendation, the review committee should keep in mind that a redacted copy of its report may be an important means of informing the candidate of the evaluation of his or her teaching and of the basis for that evaluation.*

It is the responsibility of the department chair to submit meaningful statements, accompanied by evidence, of the candidate’s teaching effectiveness at lower-division, upper-division, and graduate levels of instruction. More than one kind of evidence shall accompany each review file. Among significant types of evidence of teaching effectiveness are the following: (a) opinions of other faculty members knowledgeable in the candidate’s field, particularly if based on class visitations, on attendance at public lectures or lectures before professional societies given by the candidate, or on the performance of students in courses taught by the candidate that are prerequisite to those of the informant; (b) opinions of students; (c) opinions of graduates who have achieved notable professional success since leaving the University; (d) number and caliber of students guided in research by the candidate and of those attracted to the campus by the candidate’s repute as a teacher; and (e) development of new and effective techniques of instruction, including techniques that meet the needs of students from groups that are underrepresented in the field of instruction.

All cases for advancement and promotion normally will include: (a) evaluations and comments solicited from students for most, if not all, courses taught since the candidate’s last review; (b) a quarter-by-quarter or semester-by-semester enumeration of the number and types of courses and tutorials taught since the candidate’s last review; (c) their level; (d) their enrollments; (e) the percentage of students represented by student course evaluations for each course; (f) brief explanations for abnormal course loads; (g) identification of any new courses taught or of old courses

when there was substantial reorganization of approach or content; (h) notice of any awards or formal mentions for distinguished teaching; (i) when the faculty member under review wishes, a self-evaluation of his or her teaching; and (j) evaluation by other faculty members of teaching effectiveness. When any of the information specified in this paragraph is not provided, the department chair will include an explanation for that omission in the candidate's dossier. If such information is not included with the letter of recommendation and its absence is not adequately accounted for, it is the review committee chair's responsibility to request it through the Chancellor.

(2) Research and Creative Work — Evidence of a productive and creative mind should be sought in the candidate's published research or recognized artistic production in original architectural or engineering designs, or the like.

Publications in research and other creative accomplishment should be evaluated, not merely enumerated. There should be evidence that the candidate is continuously and effectively engaged in creative activity of high quality and significance. Work in progress should be assessed whenever possible. When published work in joint authorship (or other product of joint effort) is presented as evidence, it is the responsibility of the department chair to establish as clearly as possible the role of the candidate in the joint effort. It should be recognized that special cases of collaboration occur in the performing arts and that the contribution of a particular collaborator may not be readily discernible by those viewing the finished work. When the candidate is such a collaborator, it is the responsibility of the department chair to make a separate evaluation of the candidate's contribution and to provide outside opinions based on observation of the work while in progress. Account should be taken of the type and quality of creative activity normally expected in the candidate's field. Appraisals of publications or other works in the scholarly and critical literature provide important testimony. Due consideration should be given to variations among fields and specialties and to new genres and fields of inquiry.

Textbooks, reports, circulars, and similar publications normally are considered evidence of teaching ability or public service. However, contributions by faculty members to the professional literature or to the advancement of professional practice or professional education, including contributions to the advancement of equitable access and diversity in education, should be judged creative work when they present new ideas or original scholarly research.

In certain fields such as art, architecture, dance, music, literature, and drama, distinguished creation should receive consideration equivalent to that accorded to distinction attained in research. In evaluating artistic creativity, an attempt should be made to define the candidate's merit in the light of such criteria as originality, scope, richness, and depth of creative expression. It should be recognized that in music, drama, and dance, distinguished performance, including conducting and directing, is evidence of a candidate's creativity.

(3) Professional Competence and Activity — In certain positions in the professional schools and colleges, such as architecture, business administration, dentistry, engineering, law, medicine, etc., a demonstrated distinction in the special competencies appropriate to the field and its characteristic activities should be recognized as a criterion for appointment or promotion. The candidate's professional activities should be scrutinized for evidence of achievement and leadership in the field and of demonstrated progressiveness in the development or utilization of new approaches and techniques for the solution of professional problems, including those that specifically address the professional advancement of individuals in underrepresented groups in the candidate's field. It is responsibility of the department chair to provide evidence that the position in question is of the type described above and that the candidate is qualified to fill it.

*(4) **University and Public Service** — The faculty plays an important role in the administration of the University and in the formulation of its policies. Recognition should therefore be given to scholars who prove themselves to be able administrators and who participate effectively and imaginatively in faculty government and the formulation of departmental, college, and University policies. Services by members of the faculty to the community, State, and nation, both in their special capacities as scholars and in areas beyond those special capacities when the work done is at a sufficiently high level and of sufficiently high quality, should likewise be recognized as evidence for promotion. Faculty service activities related to the improvement of elementary and secondary education represent one example of this kind of service. Similarly, contributions to student welfare through service on student-faculty committees and as advisers to student organizations should be recognized as evidence, as should contributions furthering diversity and equal opportunity within the University through participation in such activities as recruitment, retention, and mentoring of scholars and students.”*

A. Commentary by the Task Force on APM’s Components of Faculty Performance

Performance includes both quantity/productivity and quality/impact/effectiveness, as reflected in the wording in APM 210: “The review committee must judge whether the candidate is engaging in a program of work that is both sound and productive.”

For at least one of the four areas—teaching—APM 210 provides considerable guidance on across-discipline measures of productivity and effectiveness, some quantified and others not, that help committees evaluate “the total performance of the candidate with proper reference to assigned teaching responsibilities. The committee should clearly indicate the sources of evidence on which its appraisal of teaching competence has been based. ... Among significant types of evidence of teaching effectiveness are the following: (a) opinions of other faculty members knowledgeable in the candidate’s field, particularly if based on class visitations, on attendance at public lectures or lectures before professional societies given by the candidate, or on the performance of students in courses taught by the candidate that are prerequisite to those of the informant; (b) opinions of students; (c) opinions of graduates who have achieved notable professional success since leaving the University; (d) number and caliber of students guided in research by the candidate and of those attracted to the campus by the candidate’s repute as a teacher; and (e) development of new and effective techniques of instruction, including techniques that meet the needs of students from groups that are underrepresented in the field of instruction.

“All cases for advancement and promotion normally will include: (a) evaluations and comments solicited from students for most, if not all, courses taught since the candidate’s last review; (b) a quarter-by-quarter or semester-by-semester enumeration of the number and types of courses and tutorials taught since the candidate’s last review; (c) their level; (d) their enrollments; (e) the percentage of students represented by student course evaluations for each course; (f) brief explanations for abnormal course loads; (g) identification of any new courses taught or of old courses when there was substantial reorganization of approach or content; (h) notice of any awards or formal mentions for distinguished teaching; (i) when the faculty member under review wishes, a self-evaluation of his or her teaching; and (j) evaluation by other faculty members of teaching effectiveness.” Based on this guidance, it might be possible to develop a single, quantified composite of the above factors for inclusion in the aspirational model. Or, it may be that discipline-specific metrics may be necessary.

It is impossible to develop common measures of performance in research/creative work. APM 210 states the obvious: “Promotions to tenure positions should be based on consideration of comparable work in the candidate’s own field or in closely related fields.” Guidance on publications, for example, includes the following. “Publications in research and other creative accomplishment should be evaluated, not merely

enumerated. ... When published work in joint authorship (or other product of joint effort) is presented as evidence, it is the responsibility of the department chair to establish as clearly as possible the role of the candidate in the joint effort. ... Account should be taken of the type and quality of creative activity normally expected in the candidate's field. ... Due consideration should be given to variations among fields and specialties and to new genres and fields of inquiry." Each unit might develop a quantified aspirational composite of the factors deemed relevant in the specific discipline.

In contrast to the guidance on evaluating teaching performance, the APM provides little in the way of quantitative measures for professional competence/activity. "In certain positions in the professional schools and colleges, such as architecture, business administration, dentistry, engineering, law, medicine, etc., a demonstrated distinction in the special competencies appropriate to the field and its characteristic activities should be recognized as a criterion for appointment or promotion. The candidate's professional activities should be scrutinized for evidence of achievement and leadership in the field and of demonstrated progressiveness in the development or utilization of new approaches and techniques for the solution of professional problems, including those that specifically address the professional advancement of individuals in underrepresented groups in the candidate's field. ..." Some obvious measures of professional activity for some disciplines include numbers of manuscripts reviewed, positions on editorial boards, participation in organizing conferences, professional registration/certification and awards related to competence/activity.

The APM's wording on university and public service is somewhat more helpful. "The faculty plays an important role in the administration of the University and in the formulation of its policies. Recognition should therefore be given to scholars who prove themselves to be able administrators and who participate effectively and imaginatively in faculty government and the formulation of departmental, college, and University policies. Services by members of the faculty to the community, State, and nation, both in their special capacities as scholars and in areas beyond those special capacities when the work done is at a sufficiently high level and of sufficiently high quality, should likewise be recognized as evidence for promotion. Faculty service activities related to the improvement of elementary and secondary education represent one example of this kind of service. Similarly, contributions to student welfare through service on student-faculty committees and as advisers to student organizations should be recognized as evidence, as should contributions furthering diversity and equal opportunity within the University through participation in such activities as recruitment, retention, and mentoring of scholars and students." For an aspirational composite, one can imagine including the hours of effort spent each year on university and public committees and other activities, service as member vs. chair, reports authored, awards based on service, etc. As for teaching, it may be possible to use a single composite for all disciplines.

While grants are not addressed in the APM's criteria, they are considered important by other institutions when determining which of our faculty they should recruit, and therefore play a role in determining off-scale increments offered to retain these individuals. We currently have funding information in MyInfoVault (MIV), although for comparison purposes it may be of value to determine the funding per PI/Co-PI.

We believe it is important to collect and analyze data on the presence or absence as well as the magnitude of competing offers and from which type of institution they are tendered, as these influence off-scale increments at appointment as well as later. We do not at present have readily available historical data on competing offers.

Status at hire may affect both rank/step and off-scale at appointment. Status includes, in addition to years since the terminal degree, experience in various positions (years as a postdoc, in industry, etc.) and current position (faculty, postdoc, etc.).

Some measures, such as numbers of publications and citations, are inherently quantified. Others, such as qualitative evaluations by colleagues of teaching effectiveness, are not. Quantified measures have advantages over qualitative ones. They are already in numerical form. Recent studies (e.g., that were presented at an ADVANCE workshop at UC Berkeley in 2012) have shown gender bias when evaluating credentials of candidates; the same set of credentials was rated more highly if associated with a man than with a woman. And while numerical measures are more common for productivity than quality, studies of some fields have shown strong correlations between quantity of output and quality.