UC San Diego General Campus and SIO Ladder-Rank Faculty
2012 Pay Equity Study
Summary of Methodology, Goals and Outcomes

Methodology

This study applies nonlinear regression models to ladder-rank faculty salary data based on: gender, ethnicity, department, years since highest degree, and years since ladder-rank appointment at UC San Diego. Differing sets of regression variables may be included in models to examine specific factors. The model is related to methods from labor economics and was applied for similar pay equity studies in 2006–07 and in 2001–02 by the UC San Diego Senate/Administration Gender Equity Task Force. The initial model was developed by Professor Mat McCubbins (Department of Political Science) for this Task Force, and the model structure was endorsed by the UC San Diego Academic Senate. For the current study, the nonlinear regression model was fitted to salary data from July 2007, July 2008, July 2009, July 2010, July 2011, and October 2011.

Once fitted, the numerical model parameters can be examined to determine whether salaries of specific groups of faculty have a statistically significant variation from the norm. The fitted model provides computed predicted salary levels for each individual, which permits identification of faculty with apparently anomalous salaries compared with the norm across the discipline, moderating for years since highest degree and years since ladder-rank appointment at UC San Diego.

Model structure

The model structure of the regression analysis is based on but extends that of the *Gender Equity Report 2004-05, Part 1: Gender Equity Analysis* by Mat McCubbins and uses two base points--years since highest degree and years since ladder-rank appointment at UCSD--in addition to controls for gender, ethnicity, and department, the first two of which are optional. It is a nonlinear regression model, explained below, and has access to faculty data from July 1997 to October 2011.

The central model has the following structure for faculty member j in department i with gender k and ethnicity l at year t.

$$\log(w_{j,t}) = c_{0,i} + c_{1,k} + c_{2,l} + c_3 p_{j,t} + c_4 p_{j,t}^2 + c_5 y_{j,t} + c_6 y_{j,t}^2 + c_{7,t} + e_{j,t},$$

where:
$w_{j,t}$ is the nine-month-equivalent salary of faculty member j in year t,
$c_{0,i}$ is a constant coefficient per department, where member j is a member of department i,
$c_{1,k}$ is a coefficient for gender, where member j is of gender k (there are two genders),

$c_{2,l}$ is a coefficient for ethnicity, where member *j* is of ethnicity *l* (there are twelve ethnicities),
$p_{j,t}$ is the number of years since highest degree,
$y_{j,t}$ is the number of years since ladder-rank appointment at UCSD,
$c_{7,t}$ is a coefficient to account for scale changes etc. by year in the study relative to the first year,
$e_{j,t}$ is the model fitting error.

The logarithm used is the natural logarithm, i.e., base *e*.

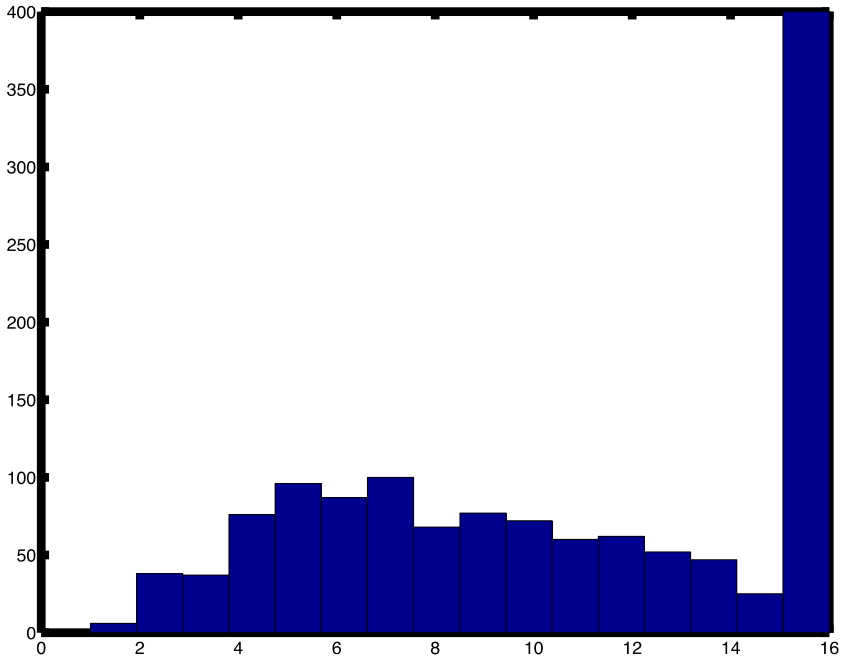
For Part 1, the model has: one coefficient per department (SIO is counted as a single department), twelve ethnicity coefficients (if included), one gender coefficient (if included), two degree-year coefficients, two appointment-year coefficients, and as many yearly coefficients as there are in the model fit. The model structure is determined by entries in the file DefineConstants.m, which requires the start year, stop year, ethnicity flag, and gender flag.

For Part 2, the model has: one coefficient per department (SIO is counted as a single department), two degree-year coefficients, two appointment-year coefficients, and as many yearly coefficients as there are in the model fit. The model structure is determined by entries in the file DefineConstants.m, which requires the start year and stop year.

The construction of the matrices used in the regression and the regression itself are performed by code in ConstructRegrMats.m and MakeRegressionData.m. Using the same matrices, the matlab command mvregression yields the same answers.

The base data set contains 13,699 records for 1,303 individuals, with a distribution of years present in the record given in the following graph. The mean duration is 10.49 years. Note that July 2011 and October 2011 count as separate years.

UC San Diego General Campus and SIO Ladder-Rank Faculty
2012 Pay Equity Study
Summary of Methodology, Goals and Outcomes

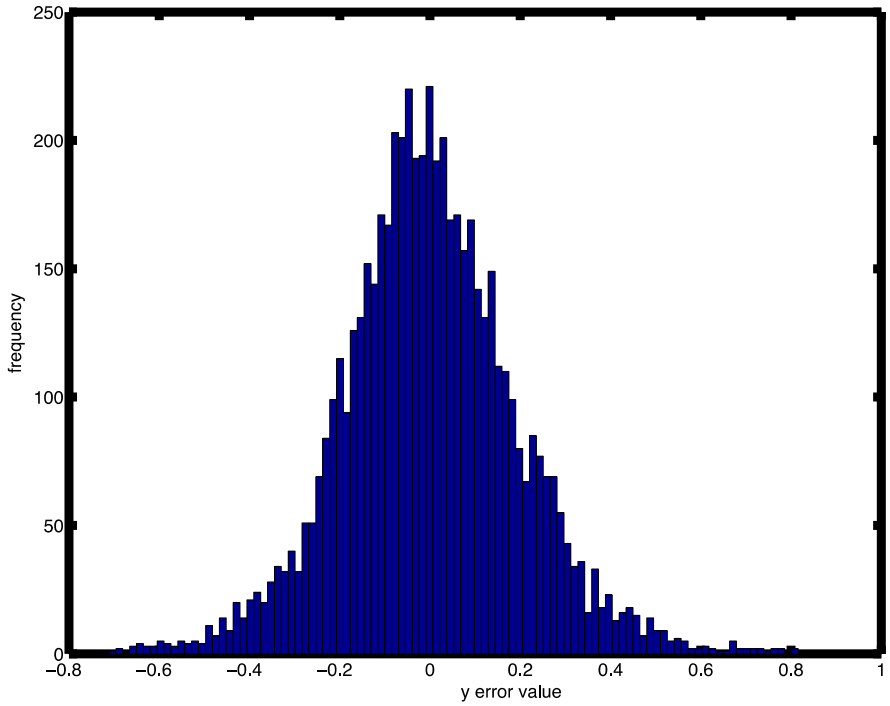


This type of study is a longitudinal analysis. The presence of multiple records for the same individuals in the study means that successive records should not be considered as independent. The records associated with a single individual should be clustered. The effect of this is not to alter the coefficient estimates, but it needs to be considered in interpreting the statistical significance of the quality of fit. Rather than having 13,699 independent records, the model fitting has used 1,303 independent sets of records.

Predictions based on the model

Once the model is fitted, it may then be used to produce predictions for each person in each year, using the equation above with $e_{j,t}$=0. Matlab, as part of its analysis tools, permits the calculation of residuals from the model – the effective $e_{j,t}$ values. This is presented below.

UC San Diego General Campus and SIO Ladder-Rank Faculty
2012 Pay Equity Study
Summary of Methodology, Goals and Outcomes



Note that the residual errors are calculated in the log-domain. In this case, they have zero mean and standard deviation of 0.1952. The density depicted has a longer tail on the high-salary end than on the low-salary end, as is to be expected. Since exp(0.1952)=1.2155, the one-sigma error corresponds to roughly a 22% salary change, additive log-domain values being multiplicative in the dollar domain.

Study Goals

Part 1 is intended to help us determine whether the salaries of women and/or underrepresented minority faculty, as a group, vary in a statistically significant way. This analysis examined the sample size and sign, magnitude, and variance of coefficients $c_{1,k}$ for gender and $c_{2,l}$ for ethnicity.

Part 2 is intended to identify individual faculty members with a negative residual 10% or more below the predicted salary for further study. Part 2 will rerun the regression analysis without the factors of gender and ethnicity. Faculty identified will be reviewed by the deans in more detail to determine the cause(s) for a salary 10% or more below what the model predicts. Faculty will be excluded for further review if they are full-time faculty administrators, are at the assistant level, are members of the Health Sciences Compensation Plan, or have had one or more no-changes in the last three review cycles, or if an advancement effective 7/1/12 or 7/1/13 brings their salaries within 10% of what the model predicts.

UC San Diego General Campus and SIO Ladder-Rank Faculty
2012 Pay Equity Study
Summary of Methodology, Goals and Outcomes

Study Outcomes

Part 1

Below is sample output from a run with gender included, ethnicity excluded, for the years July 2007 to October 2011.

Regression coefficient values

AnH	=	10.85855899
Bio	=	11.02479367
Eng	=	11.13322482
IRPS	=	11.22630931
PhyS	=	11.04916173
SIO	=	11.01959396
SocS	=	10.93906048
Econ	=	11.40826906
Rady	=	11.67019219
Gndr	=	-0.05221527
YrDeg	=	0.03234484
YrDeg^2	=	-0.00022560
YrUc	=	-0.00381390
YrUc^2	=	0.00002782
2008	=	0.05508555
2009	=	0.06032061
2010	=	0.06372578
2011	=	0.06992348
2011.25	=	0.09486175

This coefficient vector reveals a number of features.

- The divisional offsets appear to be ranked in the correct order;
- The gender coefficient is negative, indicating that women, on average, are earning less than men;
- There is a positive effect of the passage of time since degree reducing in effect because of the negative YrDeg^2 parameter;
- There is a diminishing salary with time since UCSD appointment which increases in negative impact over time;
- There is a net growth of salaries over the years since July 2007, with significant jumps in (October) 2008 and in October 2011.

A separate program TweakYr was written to add a specified number of years to years-since-highest-degree for female candidates only to understand how many years were represented by the negative gender coefficient. It was found, with this data set, that adding 2.19569 years [2 years and 71 days] changed the gender coefficient from its current value to effectively zero.

It is not clear why this is, however we are lending support to a National Science Foundation Advance PAID-Research grant that is conducting a longitudinal study about how disadvantages for women, minorities, and lesbian, gay, and bisexual faculty in STEM fields accrue over time.

The sample sizes were inadequate to make any statistically reasonable conclusions with regard to ethnicity.

The predicted salary model, as determined in Part 1, did not identify any statistically significant trends indicating that salaries of women and/or URM faculty fall below the predicted level more often than their similarly situated peers.

Part 2

While the study identified faculty whose salary is lower than that predicted for the average faculty member with the same regression variables, the regression model intentionally relies exclusively on quantitative measures and therefore could not address the potential legitimate reasons why an individual's residual might be negative or positive.

Therefore the deans were asked to review the faculty identified in the regression analysis as more than 10% below the predicted salary level to determine if the salary level could be explained by legitimate academic factors.

The salaries of 994 general campus and SIO LRF faculty members were analyzed.

153 faculty members were identified in the study but subsequently "excluded" from further review based on one of the following reasons:

- At the Assistant professor level
- Holds a full-time administrative appointment
- Has had a "no-change" in one or more of last 3 reviews
- A 7/1/12 salary change brought the individual to within 90% of predicted salary
- Participates in HSCP

78 faculty were identified for further review by the deans:

- 17 women and 61 men

- 3 African American, 12 Asian, 3 Hispanic, 59 white, and 1 unknown

18 faculty were subsequently removed from the "flagged" group based on 7/1/13 review outcomes that either resulted in a no change (3) or brought their salary to within 10% of predicted salary (15). One faculty member flagged has resigned effective 7/1/13.

Generally the deans were able to identify legitimate and academically based factors for why the salary level was below what our model predicted. However after considering the dean's feedback and weighing various options for how best to address individual cases, the EVC proposes to provide a modest but meaningful adjustment to all 59 faculty.

Faculty at the associate rank will receive a $4,000 market off scale adjustment. Those at the full professor rank will receive a $6,000 market off scale adjustment.

Chancellor Khosla has graciously agreed to fund these parity adjustments effective July 1, 2013.

The deans have been provided the final lists of faculty eligible for these parity adjustments and CAP has been asked to review the individual adjustments proposed in Part 2.

Final adjustments will not be announced or approved until after CAP has reviewed and provided feedback on the proposals.