LMU | LA
Loyola Marymount University

17 April 2012

TO: Jennifer Pate and Kevin Wetmore
Associate Professors and Co-Chairs, CCET
(Committee on the Comprehensive Evaluation of Teaching)

FR: Bryce Mason
Director of Institutional Research (IR)

RE: Analysis of New Course Evaluation Form Ratings

This memo summarizes results of analyses that addressed questions posed by CCET regarding the new course evaluation form.

A course-level data set was compiled using multiple sources. First, course-level average ratings were constructed for all questions on the new course evaluation form for each of the first four terms it was used (Fall 2009 to Spring 2011). Then, faculty characteristics were added to these ratings, including rank (assistant, associate, full, visiting/clinical, or part-time) and demographics (gender and ethnicity). In addition, using the student information system, characteristics of the courses were connected to the ratings (e.g., college, subject code, class size, average grade given, and percentage of enrollment by males). Finally, the data set was restricted to undergraduate courses. This set contained 4918 evaluated courses.

In order to simplify the analyses, the outcome variable was limited to the "overall instructor effectiveness" rating. The new form contains eight questions that measure teaching quality. IR calculated pairwise correlations for these questions and found that the correlations ranged from 0.62 to 0.93. Overall effectiveness of instruction correlated with each of the other measures from 0.78 (instructor being accessible) to 0.90 (constructive interactions). Because these correlations were so high, it seemed reasonable to use overall effectiveness of instruction as the single outcome for this study.

Question 1 – What covariates exist that might be good to control for when analyzing the data? Specifically, does "interest in this course" act as a good predictor of ratings?

IR found three strong, statistically significant covariates, including average interest in the course. First, for each one-point increase in the class's average interest rating, ratings of effectiveness of instruction rose on average 0.38 points. Second, the average grade assigned to students in the class was a significant predictor of ratings of effectiveness of instruction. For each one-point (e.g., a B to an A) increase in the average grade, effectiveness of instruction was predicted to rise by 0.14 points. Third, the age of the instructor was strongly related to overall instructor effectiveness ratings. An additional

year of age was predicted to drop the average course rating by about one one-hundredth (0.01) of a point.

To provide more clarity about the relative strength of these factors, IR conducted a beta-weighted model. A 1σ change in age, average grade, and average student interest was predicted to impact the average course rating by -0.20, +0.11, and +0.45, respectively. Therefore, interest appears about twice as strong of a predictor (for an equal change in the factor) as age, which is itself about twice as strong as the average grade assigned.

Combined, these three variables explained 28% of the variance in the outcome. All subsequent models used to test hypotheses included these three important factors. Regression output and tables used to make conclusions are presented in Appendix A.

Question 2 – Do ratings differ by rank for tenure-line faculty, or for part-time/visiting clinical faculty?

The results do not suggest that rank is related to ratings. Simple group means of average course ratings among the five groups showed some differences on the surface, with assistant professors earning the highest marks (4.47), followed by visiting/clinical faculty (4.41), part-time faculty (4.35), associate professors (4.33), and finally full professors (4.25). However, accounting for instructor age eliminated these differences.

The committee also requested a table of means by college and rank. It is provided in the appendix.

Question 3 – Is there a difference in ratings between male and female instructors?

The results do not support a difference in ratings between male and female instructors. Simple means of average instructor effectiveness show that men score roughly a tenth of a point lower than women. However, after accounting for age, this relationship vanished. Among all instructors in the study, the average age of men and women was 50.2 and 43.6, respectively.

The literature on gender differences in student teacher evaluations sometimes finds effects within the sciences, so the same analysis was conducted just among the SCSE, but no differences were found.

Question 4 – Are there differences in ratings by instructor ethnicity?

These data and methods suggest ratings differences by ethnicity. Among all faculty (tenure-line, visiting/clinical, and part-time), average course ratings for Asian and black/African-American instructors were 0.23 and 0.26 points lower than white/Caucasian instructors, respectively (Figure 1). Among tenure-line faculty, average course ratings for Asian and black/African-American instructors were 0.18 and 0.37 points lower than white/Caucasian instructors (Figure 2).

To contextualize these findings, IR and the Office of Intercultural Affairs prepared a literature review (Appendix B) on the relationships between ethnicity and student ratings.

Question 5 – Are there differences in ratings by college?

Using BCLA course ratings as a baseline, SFTV course averages were 0.22 points lower and SCSE course averages were 0.13 points lower on average. Other colleges showed no statistically significant differences.

Question 6 – Are there differences in ratings if courses had a higher proportion of students taking it to fulfill a core requirement?

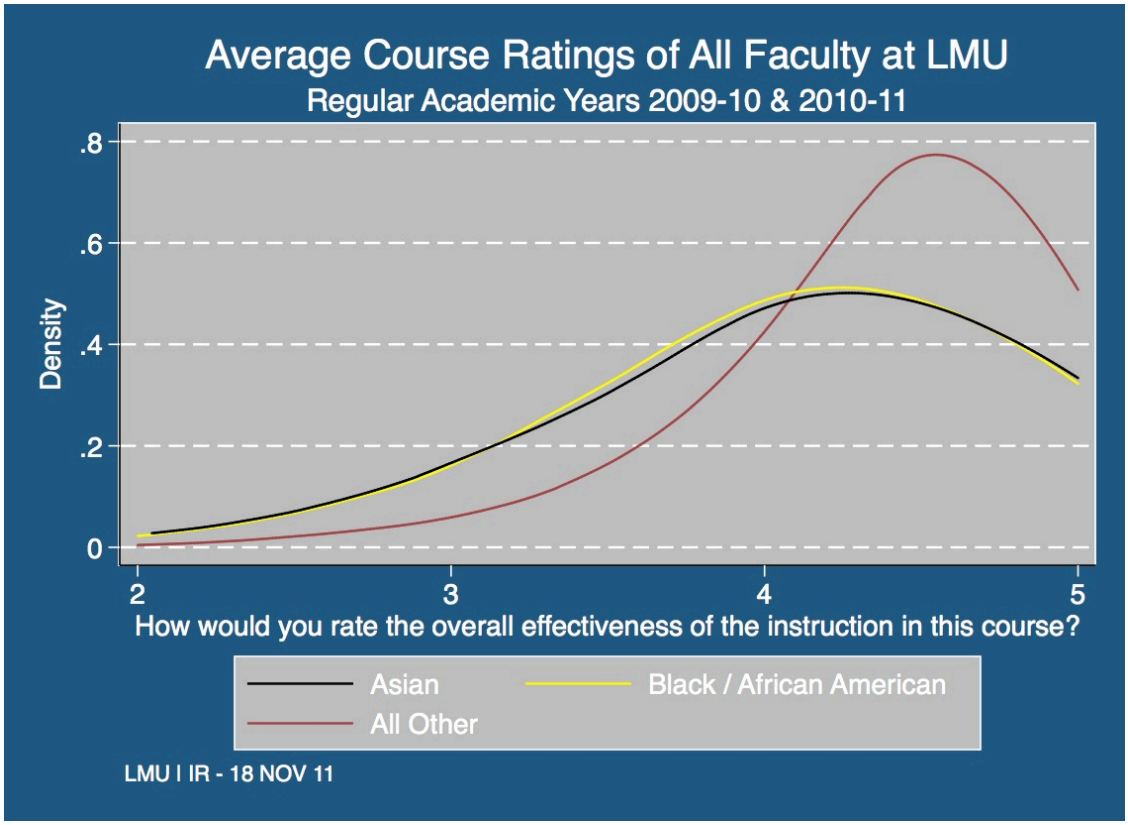
There appears to be a small positive association between the proportion of students taking a course for a core requirement and the rating. The difference between a completely non-core class (0%) and a fully-core class (100%) predicted a 0.22-point increase in the average rating of effectiveness of instruction. However, it should be noted that many students appear to misinterpret this question as "core for my college" as opposed to "university core." For example, in many business and accounting courses there were improbably high levels of students indicating they were taking it to fulfill a core requirement.

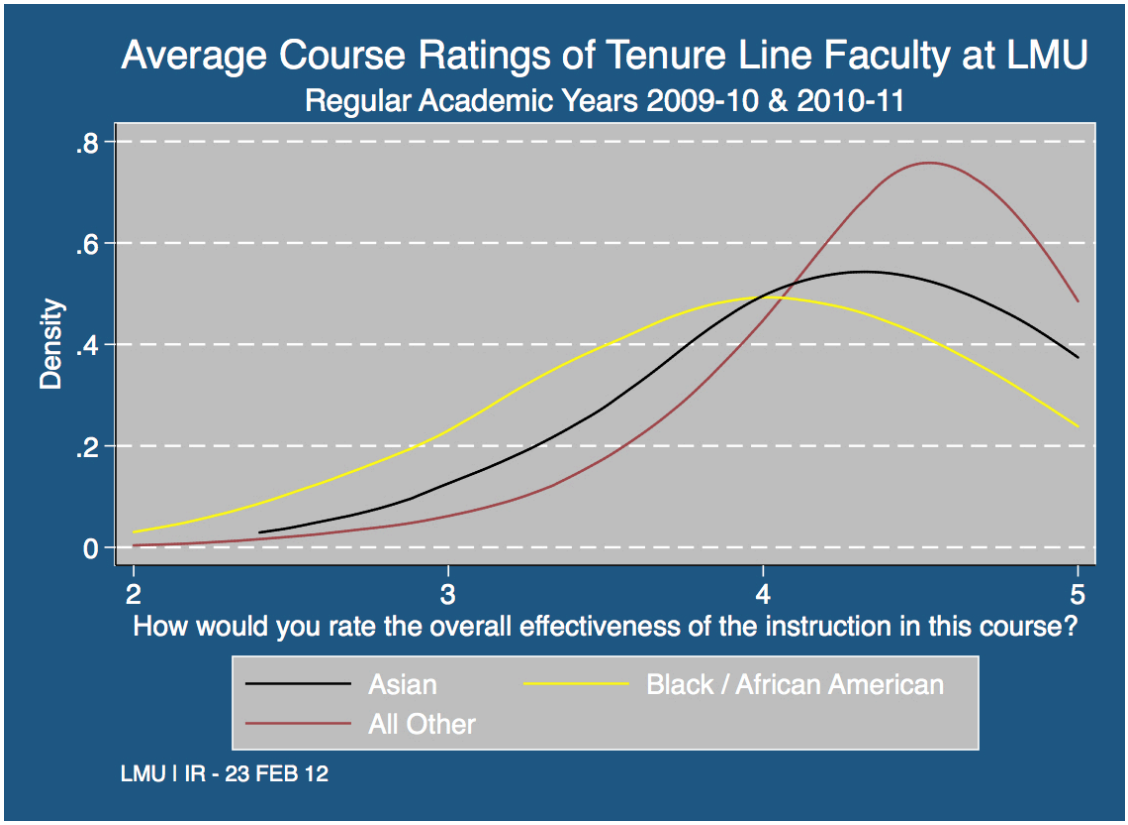
Question 7 – Are there differences in ratings if courses had a higher proportion of students taking it to fulfill a major requirement?

If there is a relationship between major-required courses and other courses, it is likely to be very modest. The difference in ratings from a completely non-major-required course (0%) to a course where everyone took it for a major requirement (100%) predicted a 0.05-point decline in overall effectiveness of instruction, and this result was only significant at the $p = 0.10$ level.

Question 8 – Are there differences in ratings if courses tended to have a higher proportion of students in that major?

Courses with more students in the major subject of the course appear, on average, to have lower overall instructional effectiveness ratings. The difference between a course with no majors in it and a course entirely composed of majors predicted a 0.13-point decline in overall effectiveness of instruction.


Average Course Ratings of All Faculty at LMU
Regular Academic Years 2009-10 & 2010-11
Density
.8
.6
.4
.2
0
2
3
4
5
How would you rate the overall effectiveness of the instruction in this course?
Asian
Black / African American
All Other
LMU I IR - 18 NOV 11


Average Course Ratings of Tenure Line Faculty at LMU
Regular Academic Years 2009-10 & 2010-11
Density
.8
.6
.4
.2
0
2
3
4
5
How would you rate the overall effectiveness of the instruction in this course?
Asian
Black / African American
All Other
LMU I IR - 23 FEB 12

Appendix A (Stata Log File)

Variable Definitions

```
outclear       Evaluation item: Learning outcomes were clear
outaddress     Evaluation item: Learning outcomes were addressed
inststu        Evaluation item: Constructive instructor-student interactions
instaccess     Evaluation item: Instructor was accessible
feedback       Evaluation item: Received feedback
challenge      Evaluation item: Challenged me to do my best work
moreinterest   Evaluation item: Course increased my interest in the subject
overallinst    Evaluation item: Overall effectiveness of instruction
attendance     Evaluation item: My attendance in course
majdept        Evaluation item: Course was in my major department
required       Evaluation item: Course was required for my major
core           Evaluation item: Course was a core requirement
interest       Evaluation item: I had an interest in taking this course

pt             Faculty characteristic: Part-time
tl             Faculty characteristic: Tenure-line
vc             Faculty characteristic: Visiting/Clinical
assistant      Faculty characteristic: Assistant professor (tl = 1)
associate      Faculty characteristic: Associate professor (tl = 1)
professor      Faculty characteristic: Full professor (tl = 1)
instmale       Faculty characteristic: Instructor is male
age            Faculty characteristic: Age during the term of instruction
upperdiv       Course characteristic: Upper division
gpa            Course characteristic: Average grade awarded
male           Course characteristic: % of enrollees who are male
minority       Course characteristic: % of enrollees who are black/African-American,
               Hispanic/Latino, or Native American
propcomplete   Course characteristic: % of enrollees who evaluated course
BA             Course characteristic: Responsible department is in CBA
CF             Course characteristic: Responsible department is in CFA
ED             Course characteristic: Responsible department is in SOE
FT             Course characteristic: Responsible department is in SFTV
LA             Course characteristic: Responsible department is in BCLA
SE             Course characteristic: Responsible department is in SCSE
```

Macros

```
`controls'     "interest age gpa"
```

Descriptive Statistics of Variables

Outcome variables

```
. summarize outclear outaddress inststu instaccess feedback challenge moreinterest
overallinst;

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
    outclear |      4918    4.531338    .3453064   2.368421          5
  outaddress |      4918    4.487513    .3932866        1.6          5
     inststu |      4918    4.483118    .4706843        1.3          5
  instaccess |      4918    4.526926    .3721561   1.789474          5
    feedback |      4918    4.308575    .5052011        1.4          5
-------------+--------------------------------------------------------
   challenge |      4918    4.322949    .4308963   1.736842          5
moreinterest |      4918    4.112362    .6073147   1.421053          5
 overallinst |      4918    4.350094    .5310424        1.1          5
         age |      4918     47.4172    12.61394   23.11507   93.26302
```

Variables of interest

```
. summarize pt tl vc assistant associate professor upperdiv gpa male minority instmale
propcomplete attendance;

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
          pt |      4918    .3538024    .4781974          0          1
          tl |      4918    .5227735    .4995319          0          1
          vc |      4918    .1234242    .3289569          0          1
   assistant |      4918    .1522977    .3593457          0          1
   associate |      4918    .1647011    .3709483          0          1
-------------+--------------------------------------------------------
   professor |      4918    .2057747    .4043077          0          1
    upperdiv |      4918    .4642131    .4987684          0          1
         gpa |      4918    3.277428    .4333211          0          4
        male |      4918    .4130829    .1917029          0          1
    minority |      4918    .2866663     .134901          0          1
-------------+--------------------------------------------------------
    instmale |      4918    .5729972     .494693          0          1
propcomplete |      4918    .8522786    .1213388   .1428571          1
  attendance |      4918    4.537113    .2444672   2.857143          5
```

Choosing an Outcome Variable

```
. *       Why we found it reasonable to study just overall effectiveness;
. pwcorr outclear outaddress inststu instaccess feedback challenge moreinterest
overallinst;

             | outclear outadd~s  inststu instac~s feedback challe~e morein~t
-------------+------------------------------------------------------------------
    outclear |   1.0000
  outaddress |   0.9286   1.0000
     inststu |   0.7916   0.8456   1.0000
  instaccess |   0.7298   0.7560   0.8306   1.0000
    feedback |   0.7911   0.8377   0.8889   0.8087   1.0000
   challenge |   0.7548   0.7981   0.7634   0.6862   0.7863   1.0000
moreinterest |   0.6982   0.7454   0.7544   0.6236   0.7565   0.7386   1.0000
 overallinst |   0.8327   0.9004   0.9001   0.7774   0.8707   0.8123   0.8061
```

Question 1

```
. gen interest_bin=round(interest, 0.2);
. table interest_bin if interest_bin>2.5, c(mean overallinst freq) format(%9.2f);

----------------------------------------
interest_ |
bin       | mean(overal~t)           Freq.
----------+-----------------------------
      2.6 |           3.52           58.00
      2.8 |           3.69           82.00
        3 |           3.93          179.00
      3.2 |           4.03          261.00
      3.4 |           4.12          338.00
      3.6 |           4.22          440.00
      3.8 |           4.33          510.00
        4 |           4.34          549.00
      4.2 |           4.39          545.00
      4.4 |           4.45          574.00
      4.6 |           4.58          647.00
      4.8 |           4.66          474.00
        5 |           4.74          207.00
----------------------------------------

. gen gpa_bin=round(gpa, 0.2);

. table gpa_bin if gpa_bin>2.0, c(mean overallinst freq) format(%9.2f);

----------------------------------------
```

```
  gpa_bin | mean(overal~t)           Freq.
----------+--------------------------------
      2.2 |           3.86          51.00
      2.4 |           4.01         128.00
      2.6 |           4.05         273.00
      2.8 |           4.20         437.00
        3 |           4.28         760.00
      3.2 |           4.34         842.00
      3.4 |           4.41         718.00
      3.6 |           4.45         736.00
      3.8 |           4.52         582.00
        4 |           4.56         359.00
-------------------------------------------
```

```
. regress overallinst `controls', cluster(inst_id1) robust;

Linear regression                                       Number of obs =    4918
                                                        F(  3,   876) =  154.08
                                                        Prob > F      =  0.0000
                                                        R-squared     =  0.2810
                                                        Root MSE      =  .45044

                            (Std. Err. adjusted for 877 clusters in inst_id1)
------------------------------------------------------------------------------
             |               Robust
 overallinst |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    interest |   .3963499   .0213268    18.58   0.000     .3544923    .4382076
         age |   -.008439   .0010027    -8.42   0.000    -.0104069   -.0064711
         gpa |   .1324764   .0276158     4.80   0.000     .0782755    .1866773
       _cons |   2.718711    .133091    20.43   0.000     2.457496    2.979926
------------------------------------------------------------------------------
```

Question 2

```
. *       By rank;
. table rank, c(mean overallinst freq) format(%9.2f);

-------------------------------------------
     rank | mean(overal~t)           Freq.
----------+--------------------------------
    Asso. |           4.33         810.00
    Asst. |           4.47         749.00
     Full |           4.25        1012.00
       PT |           4.35        1740.00
  Vis/Cli |           4.41         607.00
-------------------------------------------
```

```
. regress overallinst associate professor vc pt `controls', cluster(inst_id1) robust;

Linear regression                                       Number of obs =    4918
                                                        F(  7,   876) =   67.83
                                                        Prob > F      =  0.0000
                                                        R-squared     =  0.2828
                                                        Root MSE      =  .45003

                            (Std. Err. adjusted for 877 clusters in inst_id1)
------------------------------------------------------------------------------
             |               Robust
 overallinst |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   associate |  -.0651373   .0421848    -1.54   0.123    -.1479324    .0176579
   professor |  -.0318781   .0489377    -0.65   0.515     -.127927    .0641708
          vc |  -.0117299   .0541505    -0.22   0.829    -.1180099      .09455
          pt |  -.0569963   .0356924    -1.60   0.111    -.1270489    .0130563
    interest |   .3929917   .0212977    18.45   0.000     .3511912    .4347922
         age |  -.0080861   .0011626    -6.96   0.000    -.0103678   -.0058043
         gpa |   .1412486   .0289127     4.89   0.000     .0845023    .1979949
       _cons |   2.725663    .128135    21.27   0.000     2.474176    2.977151
------------------------------------------------------------------------------
```

```
. table crs_coll rank, c(mean overallinst freq) format(%9.2f)

------------------------------------------------------------------------
          |                              rank
 CRS_COLL |    1_Asst.     2_Asso.     3_Prof.  4_Vis/Cln.        5_PT
----------+-------------------------------------------------------------
       BA |       4.35        4.11        4.20        4.78        4.11
          |      74.00      102.00      193.00       25.00      105.00
          |
       CF |       4.45        4.44        4.39        4.64        4.53
          |     124.00      109.00      149.00       90.00      586.00
          |
       ED |                   X.XX        X.XX        X.XX        4.59
          |                   4.00        3.00        6.00       51.00
          |
       FT |       4.54        4.46        4.18        4.63        4.32
          |      61.00       68.00       46.00       13.00      152.00
          |
       LA |       4.48        4.39        4.32        4.41        4.32
          |     352.00      342.00      360.00      314.00      652.00
          |
       SE |       4.50        4.22        4.13        4.19        4.01
          |     143.00      189.00      278.00      159.00      228.00
------------------------------------------------------------------------
Note: X.XX redacted to protect privacy
```

Question 3

```
. *       By gender (and then accounting for age);
. table instmale, c(mean overallinst freq) format(%9.2f);

-----------------------------------------
 instmale | mean(overal~t)           Freq.
----------+------------------------------
        0 |           4.40         2100.00
        1 |           4.31         2818.00
-----------------------------------------

. regress overallinst instmale `controls', cluster(inst_id1) robust;

Linear regression                                      Number of obs =    4918
                                                       F(  4,   876) =  118.46
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.2810
                                                       Root MSE      =  .45048

                           (Std. Err. adjusted for 877 clusters in inst_id1)
------------------------------------------------------------------------------
             |               Robust
 overallinst |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    instmale |  -.0052445   .0273276    -0.19   0.848    -.0588798    .0483907
    interest |   .3961876   .0215032    18.42   0.000     .3539839    .4383914
         age |  -.0083854   .0010155    -8.26   0.000    -.0103786   -.0063923
         gpa |   .1321537   .0277431     4.76   0.000     .0777031    .1866044
       _cons |   2.720889   .1357631    20.04   0.000      2.45443    2.987348
------------------------------------------------------------------------------
```

Question 4

```
. *       By ethnicity as compared to white;
. table race tl, row c(mean overallinst freq) format(%9.2f);

----------------------------
          |        tl
     race |       0        1
----------+-----------------
    Asian |    4.00     4.26
          |  216.00   289.00
          |
    Black |    4.27     3.94
          |  168.00   114.00
          |
HawaiiOPI |    X.XX
          |    1.00
          |
 Hispanic |    4.41     4.47
          |  164.00   255.00
          |
     Ind. |    X.XX     X.XX
          |    5.00     8.00
          |
    Multi |    X.XX     4.32
          |    7.00    22.00
          |
    Unkn. |    4.35     4.27
          |  168.00    57.00
          |
    White |    4.42     4.36
          | 1618.00  1826.00
          |
    Total |    4.36     4.34
          | 2347.00  2571.00
----------------------------
Note: X.XX (redacted to protect anonymity)

. regress overallinst unknown multirace hisp black asian `controls', cluster(inst_id1)
robust;

Linear regression                                      Number of obs =    4918
                                                       F(  8,   876) =   70.28
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.3090
                                                       Root MSE      =  .44179

                            (Std. Err. adjusted for 877 clusters in inst_id1)
------------------------------------------------------------------------------
             |               Robust
 overallinst |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     unknown |  -.0184521   .0543986    -0.34   0.735    -.1252189    .0883146
   multirace |   .0717866   .0601749     1.19   0.233    -.0463173    .1898904
        hisp |  -.0137268   .0368914    -0.37   0.710    -.0861326    .0586791
       black |  -.2647163   .0474087    -5.58   0.000    -.3577641   -.1716685
       asian |  -.2337864   .0529116    -4.42   0.000    -.3376347   -.1299382
    interest |   .3935904   .0204141    19.28   0.000     .3535241    .4336566
         age |  -.0092133   .0009836    -9.37   0.000    -.0111437   -.0072828
         gpa |   .1239178   .0264155     4.69   0.000     .0720727     .175763
       _cons |   2.835373   .1273299    22.27   0.000     2.585466     3.08528
------------------------------------------------------------------------------
```

```
. regress overallinst unknown multirace hisp black asian `controls' if tl==1,
cluster(inst_id1) robust;

Linear regression                                       Number of obs =    2571
                                                        F(  8,   362) =   44.71
                                                        Prob > F      =  0.0000
                                                        R-squared     =  0.3188
                                                        Root MSE      =  .43085

                            (Std. Err. adjusted for 363 clusters in inst_id1)
------------------------------------------------------------------------------
             |               Robust
 overallinst |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     unknown |  -.0682355   .1040543    -0.66   0.512    -.2728623    .1363914
   multirace |   .0156278   .0699883     0.22   0.823    -.1220069    .1532624
        hisp |    .009566    .051337     0.19   0.852    -.0913903    .1105223
       black |  -.3727649   .0625566    -5.96   0.000     -.495785   -.2497449
       asian |  -.1768653     .06291    -2.81   0.005    -.3005804   -.0531503
    interest |   .3995773   .0287741    13.89   0.000     .3429919    .4561627
         age |  -.0104523   .0013615    -7.68   0.000    -.0131297    -.007775
         gpa |   .0854575   .0347638     2.46   0.014     .0170931     .153822
       _cons |   3.002776   .1687407    17.80   0.000     2.670941    3.334611
------------------------------------------------------------------------------
```

Question 5

```
. *       College differences as compared to LA;
. regress overallinst BA CF ED FT SE `controls', cluster(inst_id1) robust;

Linear regression                                       Number of obs =    4918
                                                        F(  8,   876) =   62.85
                                                        Prob > F      =  0.0000
                                                        R-squared     =  0.2955
                                                        Root MSE      =  .44608

                            (Std. Err. adjusted for 877 clusters in inst_id1)
------------------------------------------------------------------------------
             |               Robust
 overallinst |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          BA |  -.0333308   .0518382    -0.64   0.520    -.1350724    .0684108
          CF |  -.0596212   .0339948    -1.75   0.080     -.126342    .0070995
          ED |  -.0602182    .051597    -1.17   0.243    -.1614863    .0410499
          FT |  -.2166489   .0440223    -4.92   0.000    -.3030504   -.1302474
          SE |   -.133414    .038315    -3.48   0.001     -.208614   -.0582141
    interest |    .405634   .0212028    19.13   0.000     .3640199    .4472482
         age |  -.0078553   .0010173    -7.72   0.000     -.009852   -.0058586
         gpa |   .1506857   .0301859     4.99   0.000     .0914406    .2099309
       _cons |   2.652081   .1395004    19.01   0.000     2.378287    2.925876
------------------------------------------------------------------------------
```

Question 6

```
. *       Core;
. regress overallinst core `controls', cluster(inst_id1) robust;

Linear regression                                       Number of obs =    4918
                                                        F(  4,   876) =  124.77
                                                        Prob > F      =  0.0000
                                                        R-squared     =  0.2967
                                                        Root MSE      =  .44552

                            (Std. Err. adjusted for 877 clusters in inst_id1)
------------------------------------------------------------------------------
             |               Robust
 overallinst |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        core |    .216623   .0347097     6.24   0.000     .1484992    .2847469
    interest |   .4516891   .0235022    19.22   0.000     .4055618    .4978163
         age |  -.0076445   .0009937    -7.69   0.000    -.0095948   -.0056943
         gpa |   .1332633   .0273693     4.87   0.000     .0795462    .1869804
       _cons |   2.358081   .1478587    15.95   0.000     2.067883     2.64828
------------------------------------------------------------------------------
```

Question 7

```
. *       Major req;
. regress overallinst required `controls', cluster(inst_id1) robust;

Linear regression                                       Number of obs =    4918
                                                        F(  4,   876) =  116.35
                                                        Prob > F      =  0.0000
                                                        R-squared     =  0.2822
                                                        Root MSE      =   .4501

                            (Std. Err. adjusted for 877 clusters in inst_id1)
------------------------------------------------------------------------------
             |               Robust
 overallinst |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    required |  -.0534592    .032035    -1.67   0.096    -.1163336    .0094152
    interest |    .392418   .0213332    18.39   0.000     .3505479    .4342882
         age |  -.0082981   .0010026    -8.28   0.000    -.0102658   -.0063303
         gpa |   .1270739   .0276783     4.59   0.000     .0727504    .1813974
       _cons |   2.774174   .1346498    20.60   0.000       2.5099    3.038448
------------------------------------------------------------------------------
```

Question 8

```
. *       Majority in major department;
. regress overallinst majdept `controls', cluster(inst_id1) robust;

Linear regression                                       Number of obs =    4918
                                                        F(  4,   876) =  118.32
                                                        Prob > F      =  0.0000
                                                        R-squared     =  0.2881
                                                        Root MSE      =  .44825

                            (Std. Err. adjusted for 877 clusters in inst_id1)
------------------------------------------------------------------------------
             |               Robust
 overallinst |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     majdept |  -.1256546   .0312277    -4.02   0.000    -.1869446   -.0643647
    interest |   .4168711   .0222856    18.71   0.000     .3731317    .4606106
         age |  -.0080238   .0009976    -8.04   0.000    -.0099817   -.0060658
         gpa |   .1288278   .0273764     4.71   0.000     .0750968    .1825588
       _cons |   2.692646   .1318693    20.42   0.000      2.43383    2.951463
------------------------------------------------------------------------------
```

Appendix B

What Does the Literature Say about Race, Ethnicity, and Student Teaching Evaluations?

In a study of the course evaluations at LMU from Fall 2009 to Spring 2011 conducted for the Committee on the Comprehensive Evaluation of Teaching, the Office of Institutional Research found that undergraduates tended to rate Asian and black/African-American faculty lower on the global item, "overall effectiveness of the instructor." Among all faculty (tenure-line, visiting/clinical, and part-time), the average ratings for Asian and black/African-American instructors were 0.23 (0.4σ) and 0.26 (0.5σ) points lower than white/Caucasian instructors, respectively. Among just tenure-line faculty, average course ratings for Asian and black/African American instructors were 0.18 (0.3σ) and 0.37 (0.7σ) points lower than white/Caucasian instructors, respectively. This research brief reviews studies on other institutions' experiences with race/ethnicity inequities in student teaching evaluations (STE).[1]

Although there were a limited number of studies that explored the relationship between race/ethnicity and STE, they generally reported a negative main effect for faculty of color. Hamermesh and Parker (2005), who were mainly interested in estimating the effect of beauty on student ratings of professors, found a negative 0.5σ impact on STE scores for being a female minority faculty member vs. white faculty.[2] Smith (2007), who analyzed STE scores for 190 tenure-line faculty in a college of education, found that students rated black/African-American faculty 0.5σ and 0.6σ lower than white faculty in overall course value and overall teaching ability, respectively. Similarly, DiPietro and Faye (2005) noted that students tended to rate white faculty higher than Asian and Hispanic/Latino faculty. Although limited to economics classes, McPherson et al. (2009) observed a negative 0.3σ impact on STE scores for non-white faculty. Finally, in a recent large-scale study across multiple selective liberal arts colleges using STE scores from RateMyProfessor.com and student perceptions of race/ethnicity using publicly available information about the instructor, Reid (2010) found that Asian and black/African-American faculty were rated lower than White/Hispanic/Latino faculty on three dimensions (overall quality, helpfulness, and clarity) in the range of 0.1σ-0.2σ (Asian) to 0.4σ (black/African-American).

Some reasons hypothesized in the literature for inequities in STE scores for faculty of color include student prejudices, student preconceptions, language differences, and expectations of students regarding minority faculty members. Many researchers have noted that students may have communication issues with faculty whose native language is not English. Most recently, Hamermesh and Parker (2005) found a negative 0.5σ effect on STE scores for non-native English speakers, and other research has found similar results (Rao, 1995; Rubin, 1998). However, most of the literature speaks to possible prejudices and expectations of students with respect to faculty of color. One study performed controlled experiments using a fake course syllabus to tease out student preconceptions and found that students rated Anglo women professors more capable than Latina professors (Anderson & Smith, 2005). Another experiment involved an Indian

[1] A common term in the literature.
[2] Minority was not defined.

instructor giving two identical courses with either traditional or Western dress and found substantial STE score differences (Chowdhary, 1988). Further, Smith (2007) observed that among 26 measures of instructor behavior that presumably would be similar to a global measure of teaching ability, this global measure's mean score for black/African-American faculty was lower than every single one of the behavior-specific mean scores. This is supportive of the idea that students may be letting things other than instructor behavior dictate their ratings. Lastly, in a series of student interviews, Hendrix (1998) found that students in a predominantly white university did not believe that a professor's race/ethnicity influenced their perceptions of instructor credibility; however, the students simultaneously described different criteria for evaluating black/African-American instructors for courses on certain topics. Students gave more credibility to black/African-American instructors when they taught courses that had an ethnic or racial focus, and were more willing to question their credibility for courses that lacked such a focus.

Ways to address inequities in STE scores for faculty of color was rarely addressed in the literature reviewed, but many scholars pointed out that the stakes for bias in STE scores are high, because they tend to be used as a part of merit and promotion decisions (e.g., Boatright-Horowitz & Soeung, 2009; McPherson et al., 2009). McPherson et al. (2009) were the only team in the literature reviewed who outlined a framework for a policy remedy by creating adjusted ratings based on predicted vs. actual values—essentially giving credit for the average differences observed for many key factors, including race/ethnicity. Researchers that have focused on the role of faculty members' gender on STE scores, however, have recommend other actions such as including measurements of teaching techniques on STEs (Feldman, 1993; Ludwig & Meachum, 1997; Crombie et al., 2003). Chamberlin & Hickey (2001) recommend having an open dialogue about different teaching styles, how these styles are translated into STEs, and how STEs are used to make career decisions.

This literature parallels some findings at LMU, but with key differences. For example, many other institutions reported similar average score differentials (roughly 0.2σ to 0.6σ, for various measures) for black/African-American faculty. Some research also found negative impacts for Asian faculty and, similar to LMU, the size of the differential appeared smaller than for black/African-American faculty. However, none of the literature reviewed explored these differences among tenure-line vs. other types of faculty, which for LMU appeared to be an important moderating factor. A final key difference is that, while some studies found negative relationships for Hispanic/Latino faculty, LMU did not experience this type of relationship.

References

Anderson, K. J., & Smith, G. (2005). Students' preconceptions of professors: Benefits and barriers according to ethnicity and gender. *Hispanic Journal of Behavioral Sciences, 27*(2), 184-201.

Boatright-Horowitz, S. L., & Soeung, S. (2009). Teaching white privilege to white students can mean saying goodbye to positive student evaluations. *American Psychologist, 64*, 574-575.

Chamberlin, M. S., & Hickey, J. S. (2001). Student evaluations of faculty performance: The role of gender expectations in differential evaluations. *Educational Research Quarterly*, *25*(2), 3-15.

Crombie, G., Pyke, S. W., Silverthorn, N., Jones, A., & Piccinin, S. (2003). Students' perception of their classroom participation and instructor as a function of gender and context. *Journal of Higher Education, 74*(1), 51–76.

DiPietro, M. & Faye, A. (2005). Online student ratings-of-instruction (SRI) mechanisms for maximal feedback to instructors. Paper presented at the 30th Annual Meeting of the Professional and Organizational Development Network, Milwaukee, WI.

Chowdhary, U. (1988). Instructor's attire as a biasing factor in students' ratings of an instructor. *Clothing & Textiles Research Journal, 6*(2)*,* 17–22.

Feldman, K. (1993). College students' views of male and female college teachers: part II. Evidence from students' evaluations of their classroom teachers. *Research in Higher Education, 34*(2), 151–211.

Hamermesh, D. S. & Parker, A. M. (2005). Beauty in the classroom: Instructors' pulchritude and putative pedagogical productivity. *Economics of Education Review, 24*(4), 369-376.

Hendrix, K. G. (1998). Student perceptions of the influence of race on professor credibility. *Journal of Black Studies, 28*(6), 738-763.

Ludwig, J. M., & Meacham, J. A. (1997). Teaching controversial courses: Student evaluations of instructors and content. *Educational Research Quarterly*, *21*(1), 27-38.

McPherson, M. A., Jewell, R. T., & Kim, M. (2009). What determines student evaluations scores? A random effects analysis of undergraduate economics classes. *Eastern Economic Journal, 35*, 37-51.

Rao, N. (1995). The Oh No! syndrome: A language expectation model of undergraduate negative reactions toward foreign teaching assistants. Paper presented at the 79th

Annual Meeting of the International Communication Association, Albuquerque, NM., May 25-29, 1995. ERIC ED384921.

Reid, L.D. (2010). The role of perceived race and gender in the evaluation of college teaching on ratemyprofessors.com. *Journal of Diversity in Higher Education, 3*(3)*,* 137-152.

Rubin, D.L. (1998). Help! My professor (or doctor or boss) doesn't talk English. In J. N. Martin, T. K. Nakayama, L. A. Flores (Eds.), *Readings in Cultural Contexts,* 149-160. Mountain View, CA: Mayfield.

Smith, B. P. (2007). Student Ratings of Teaching Effectiveness: An Analysis of End-of-Course Faculty Evaluations. *College Student Journal*, *41*(4), 788-800.