Correlates of Reputation Analysis

The reputational quality of a program is a purely subjective measure; however, it is related to quantitative measures in the sense that quality judgment could be made on the basis of information about programs, such as the scholarly work of the faculty and the honors awarded to the faculty for that scholarship. Therefore, it may be possible to relate or to predict quality rankings for programs using quantitative measures. It is clear that predicted quality rankings would also be subjective and that the accuracy of such predictions may change over time.

One way to construct such a relationship is to do a least squares multilinear regression. The dependent variable in the regression analysis is represented by a set of average ratings, $r_1, r_2, \ldots, r_N$ for N programs in a particular field. The predictors or independent variables would be a set of quantitative or coded program characteristics that are represented by a vector, $\mathbf{x}_n$, for program n. The analysis would construct a function $f(\mathbf{x})$ which provides a *predicted* average rating $f(\mathbf{x}_n)$ for program n. In this case the relation between r_n and $f(\mathbf{x}_n)$ would be

$$r_n = f(\mathbf{x}_n) + e_n = a_1x_{1,n} + a_2x_{2,n} + \ldots + a_mx_{m,n} + a_{m+1} + e_n \qquad (1)$$

where $x_{1,n}, x_{2,n}, \ldots, x_{m,n}$ represent the m quantity or coded characteristics for the program n in the field, and e_n, is the residual or the amount by which the predicted average rating varies from the actual average rating for that program. If the prediction is "good" then the residuals are relatively small. The coefficients a_j are determined by minimizing the sum of the squares of the differences $r_n - f(\mathbf{x}_n)$.

While a single regression equation is generated using quantitative data and the reputational score, the selected raters of the program provide a certain amount of variability. This variability can be shown in the following manner: Associated with each coefficient a_i is a 95%-confidence interval $[L_i, U_i]$, and by randomly selecting values for the coefficients within their confidence intervals, a predicted average rating $\acute{r}_n$ can be generated for program n. A measure of how close the set of $\acute{r}_n$ ratings is to the r_n ratings can be calculated by

$$\| \acute{\mathbf{r}} - \mathbf{r} \|^2 < p\, s^2\, F, \qquad (2)$$

where $\acute{\mathbf{r}} = (\acute{r}_1, \acute{r}_2, \ldots, \acute{r}_N)$, $\mathbf{r} = (r_1, r_2, \ldots, r_N)$ and $\| \ \|^2$ denotes the sum of squares of the components of the difference vector. The bound on the inequality, $p\, s^2\, F$, is a constant that is derived from the regression analysis.

p = m, the number of nonconstant terms in the regression equation,
s^2 is the "mean square for error" given in the output of a regression program, and
F = the 95% cutoff point for the F-distribution with p and n-p degrees of freedom.

By repeating the random selection of coefficients many times, a collection of coefficients can be determined that satisfies inequality (2), and the upper- and lower-bounds of this

collection defines an interval $[L'_i, U'_i]$. For coefficients in these intervals a range of predicted ratings can be generated.

From the practical point of a program trying to estimate the quality of its program, a few years after a reputational survey is conducted, it could use a linear regression equation with coefficients in $[L'_i, U'_i]$ to generate a new range of ratings based on current program data, or if data for all programs in the field were available, a new interquartile ranking of programs could be obtained.

The following is an example where this method is applied to the 1995 ratings of programs in Mathematics.

Mathematics

Using the STATA statistical package and applying a forward stepwise, least-squares linear regression on a large number of quantitative variables which characterized publications, citations, faculty size and rank, research grant support, number of doctorates by gender and race/ethnicity, graduate students by gender, graduate student support, and time to degree, the following seven variables were identified as being the most significant:

(ginipub) Gini Coefficient for Program Publications, 1988-92: The Gini coefficient is an indicator of the concentration of publications on a small number of the program faculty during the period 1988-92.

(phds) Total Number of Doctorates FY 86-92

(perfull) Percentage of Full Professors Participating in the Program

(persupp) Percentage of Program Faculty with Research Support (1986-92)

(perfpub) Percentage of Program Faculty Publishing in the Period 1988-1992

(ratiocit) Ratio of the Total Number of Program Citations in the Period 1988-1992 to the Number of Program Faculty

(myd) Median Time Lapse from Entering Graduate School to Receipt of Ph.D. in Years

Results of a regression analysis are shown below. About 95% of the variation is explained by these variables, where $R^2 = 0.8304$.

```
      Source |       SS       df       MS              Number of obs = 139
-------------+------------------------------           F(  7,   131) =  91.60
       Model |   112.36003     7   16.0514329           Prob > F      =   0.0000
    Residual |    22.954789  131     .175227397        R-squared     =   0.8304
-------------+------------------------------           Adj R-squared =   0.8213
       Total |   135.314819   138   .98054217          Root MSE      =    .4186
```

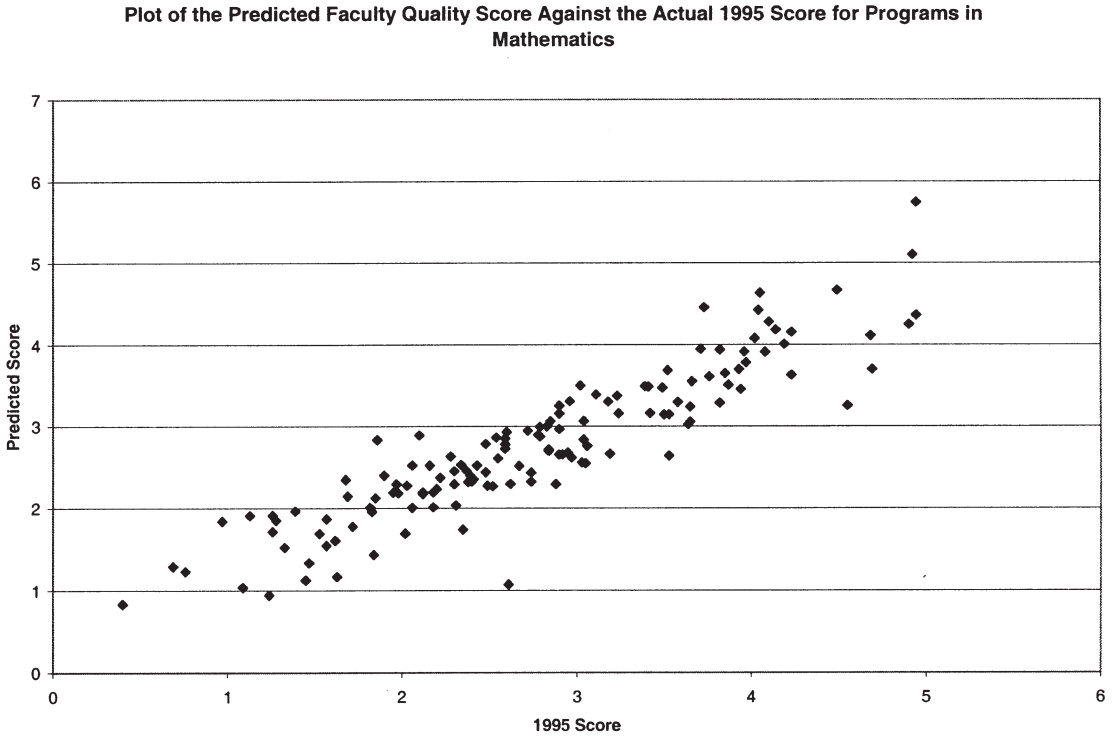
quality	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
phds	.3489197	.0544665	6.41	0.000	.2411721	.4566674
perfull	.008572	.0027864	3.08	0.003	.0030598	.0140842
persupp	.0183162	.0025146	7.28	0.000	.0133418	.0232906
perfpub	-.0150464	.0035235	-4.27	0.000	-.0220167	-.0080762
ratiocit	.0258671	.0077198	3.35	0.001	.0105955	.0411387
myd	-.7737551	.1995707	-3.88	0.000	-1.168553	-.3789567
ginipub	-.0294944	.0044222	-6.67	0.000	-.0382425	-.0207462
_cons	3.070145	.3625634	8.47	0.000	2.352908	3.787382

The resulting predictor equation is:

$$\mathbf{f(x)} = 3.07 + 0.349(\text{phds}) + 0.009(\text{perfull}) + 0.018(\text{persupp}) - 0.015(\text{perfpub}) + 0.026(\text{ratiocit}) - .774(\text{myd}) - 0.029(\text{ginipub})$$

It is noted that the Root Mean Square Error (RMSE) from the regression is 0.4186, and the variation in scores from the 1995 confidence interval calculation has an RMSE of 0.2277.

The following is scatter plot of the actual 1995 ratings and the predicted ratings.



The 95%-confidence interval for each of the variables used in the regression can now be used to find a new estimate for the quality score. As described above, values for the

coefficients in the regression equation are randomly selected in the intervals and tested to see if that set of coefficients satisfies the relation $\| \hat{\mathbf{r}} - \mathbf{r} \|^2 < p\, s^2 F$. For Mathematics data the bound $p\, s^2 F = (7)(.4186)^2(2.12) = 2.563556$. For this example 3,000 random selections were made in the coefficient intervals and 220 coefficients sets satisfied the inequality. The corresponding maximum and minimum interval are:

	phds coefficient	persupp coefficient	ginipub coefficient	myd coefficient	perfpub coefficient	ratiocit coefficient	perfull coefficient	constant
Max	0.35469	0.018583	-0.029026	-0.7526	-0.014673	0.026686	0.0088674	3.10858
Min	0.34314	0.018049	-0.029964	-0.79495	-0.015421	0.025047	0.0082761	3.03164

Using the values in the above table, the maximum and minimum predicted quality scores can be calculated, and the scores for Mathematics programs are displayed in the table below.

As described earlier, these maximum and minimum coefficient values could be used to construct new quality scores, by randomly selecting the coefficients in the regression equation between the corresponding maximum and minimum values. If this is done repeatedly a collection of quality scores is obtained for each program and the interquartile range of this collection could be generated. This was done 100 times and the results are given as the Predicted Ranks in the table with the Bootstrap rankings.

	Quality Score		Predicted Ranks		Bootstrap Ranks	
	Maximum	Minimum	1st Quartile	3rd Quartile	1st Quartile	3rd Quartile
Institution						
Dartmouth College	2.73	2.51	73	76	53	62
Boston University	2.70	2.42	77	80	48	52
Brandeis University	3.17	2.88	49	51	32	36
Harvard University	4.41	4.09	8	9	2	4
Massachusetts Inst of Technology	5.27	4.93	2	2	3	4
U of Massachusetts at Amherst	3.40	3.11	38	40	54	60
Northeastern University	2.41	2.13	99	103	70	80
Brown University	4.60	4.31	5	6	26	29
Brown University-Applied Math	4.59	4.26	6	6	14	17
University of Rhode Island	1.69	1.40	128	129	122	125
University of Connecticut	2.66	2.39	79	83	98	102
Wesleyan University	2.31	2.09	104	107	101	110
Yale University	3.38	3.13	38	40	7	8
Adelphi University	1.07	0.82	138	138	130	133
CUNY - Grad Sch & Univ Center	3.38	3.10	40	41	30	32
Clarkson University	2.49	2.21	90	94	109	118
Columbia University	4.32	3.99	11	11	10	12
Cornell University	4.81	4.46	3	4	14	16
New York University	4.83	4.50	3	4	7	8
Polytechnic University	2.15	1.88	112	114	98	105
Rensselaer Polytechnic Inst	3.64	3.36	27	30	48	52
University of Rochester	3.10	2.83	52	54	56	62

State Univ of New York-Albany	2.55	2.33	85	88	82	90
State Univ of New York-Binghamton	2.55	2.33	85	87	65	75
State Univ of New York-Buffalo	3.00	2.76	57	59	61	70
State Univ of New York-Stony Brook	3.60	3.31	30	32	19	22
Syracuse University	2.42	2.18	95	100	76	84
Princeton University	4.52	4.21	7	7	2	3
Rutgers State Univ-New Brunswick	4.06	3.77	16	18	17	20
Stevens Inst of Technology	1.73	1.48	127	127	121	128
Carnegie Mellon University	3.63	3.33	28	31	34	40

English Language and Literature

Applying the same method to the 1995 programs in English Language and Literature, a slightly different result is obtained, since programs in this field do not have the same productivity characteristics as those in Mathematics. Again, forward stepwise least squares linear regression was applied to a large number of quantitative variables, and the following were identified as being the most significant:

(nopubs2) Number of Publications During the Period 1985-1992

(perfawd) Percentage of Program Faculty with at Least One Honor or Award for the Period 1986-1992

(acadplan) Total Number of Doctorates FY 1986-1992 with academic employment plans at the 4-year college or university level.

(ginicit) Gini Coefficienticient for Program Citations, 1988-1992: The Gini coefficienticient is an indicator of the concentration of citations on a small number of the program faculty during the period 1988-1992.

(nocits1) Number of Citations During the Period 1981-1992

(fullprof) Percentage of Full Professors Participating in the Program

(empplan) Total Number of Doctorates FY 1986-1992 with Employment Plans.

None of the variables identified in the Mathematics regression are present in this regression analysis.

Results of this regression analysis are shown below. About 95% of the variation is explained by these variables, where $R^2 = 0.8106$.

```
      Source |       SS       df       MS              Number of obs =     117
-------------+------------------------------           F(  7,   109) =   66.65
       Model |  83.985691      7  11.9979559           Prob > F      =  0.0000
    Residual |  19.6227839   109    .18002554           R-squared     =  0.8106
-------------+------------------------------           Adj R-squared =  0.7984
       Total | 103.608475    116    .893176507          Root MSE      =   .42429

------------------------------------------------------------------------------
       q93a |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------+-----------------------------------------------------------------
    nopubs2 |   .1202936   .1017753     1.18   0.240    -.0814218     .322009
    perfawd |   .0326877   .0041423     7.89   0.000     .0244777    .0408977
   acadplan |   .7961931   .2416467     3.29   0.001     .3172573    1.275129
    ginicit |  -.0007486   .0001839    -4.07   0.000    -.001113    -.0003842
    nocits1 |   .0827859   .0234272     3.53   0.001     .036354     .1292178
   fullprof |   .2942413   .1096454     2.68   0.008     .0769276    .511555
    empplan |  -.599897    .2698761    -2.22   0.028    -1.134783   -.0650113
      _cons |  1.955276    .1533968    12.75   0.000     1.651249    2.259304
```

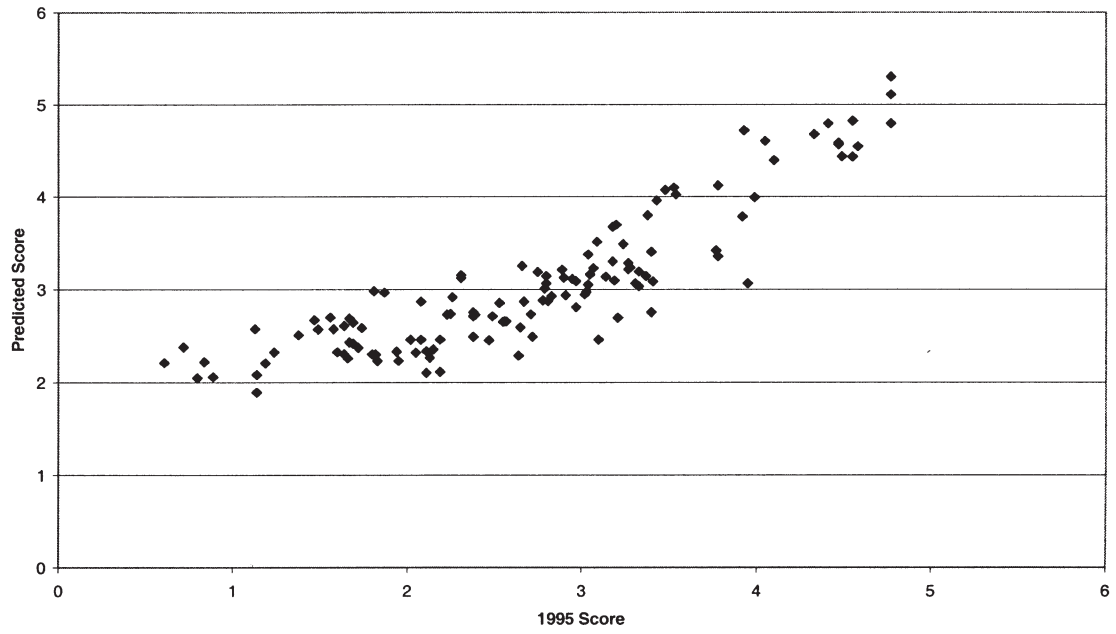
The resulting predictor equation is:

$$\mathbf{f(x)} = 1.955 + 0.12(\text{nopubs2}) + 0.033(\text{perfawd}) + 0.796(\text{acadplan}) - 0.001(\text{ginicit}) + 0.083(\text{nocits1}) + 0.294(\text{fullprof}) - 0.6(\text{empplan}).$$

The following is a scatter plot of the Random Halves draw from the 1995 rankings and the predicted ranking for that draw.

For programs in English Language and Literature, the Root Mean Square Error (RMSE) from the regression is 0.42429, and the variation in scores from the 1995 confidence interval calculation has an RMSE of 0.2544.

Plot of the Predicted Faculty Quality Score Against the Actual 1995 Score for Programs in English Language and Literature



In Mathematics the 95%-confidence interval for each of the variables used in the regression can be used to determine a new estimate for the quality score. In this case, the bound $p\,s^2\,F = (7)(.42869)^2(2.18) = 2.747136$. For this example 3,000 random selections were also made in the coefficient intervals and 242 coefficients sets satisfied the inequality. The corresponding maximum and minimum intervals are:

	nopubs2 coefficient	perfawd coefficient	acadplan coefficient	ginicit coefficient	nocits coefficient	fullprof coefficient	empplan coefficient	constant
Max	0.13384	0.033239	0.82835	-0.00072	0.085903	0.30883	-0.56399	1.97569
Min	0.10684	0.03214	0.76425	-0.00077	0.079689	0.27975	-0.63557	1.935

For the example used with Mathematics programs, the maximum and minimum values for the coefficients can be used to calculate the maximum and minimum predicted quality scores for the programs in English Language and Literature. These scores are displayed in the table below.

Repeating the exercise, described for Mathematics, of randomly selecting coefficient values in the maximum-minimum intervals a large number of times, an interquartile range can be generated for programs in English Language and Literature. This was again done 100 times and the results are given as the Predicted Ranks in the table with the Random Halves rankings.

	Quality Score		Predicted Ranking		Random Halves Ranks	
Institution	Maximum	Maximum	1st Quartile	3rd Quartile	1st Quartile	3rd Quartile
University of New Hampshire	2.74	2.56	91	93	70	77
Boston College	2.57	2.42	96	98	59	64
Boston University	3.80	3.59	20	21	38	42
Brandeis University	3.63	3.40	19	21	44	55
Harvard University	5.55	5.05	1	1	2	3
U of Massachusetts at Amherst	3.84	3.51	30	34	38	43
Tufts University	2.35	2.22	108	110	67	74
Brown University	4.21	3.78	15	16	13	15
University of Rhode Island	2.39	2.22	113	115	94	113
University of Connecticut	3.26	3.05	53	57	79	87
Yale University	5.07	4.52	5	6	2	3
CUNY - Grad Sch & Univ Center	3.50	3.21	42	48	18	19
Columbia University	4.90	4.24	9	10	7	9
Cornell University	4.71	4.16	13	13	6	8
St John's University	1.93	1.86	127	127	119	122
Fordham University	2.38	2.23	103	106	104	112
New York University	3.59	3.25	26	28	18	20
Drew University	2.30	2.15	116	119	123	126
University of Rochester	3.30	3.02	30	33	44	48
State Univ of New York-Binghamton	3.01	2.72	62	64	65	69

State Univ of New York-Buffalo	3.65	3.16	30	37	25	27
State U of New York-Stony Brook	3.17	2.77	48	55	46	52
Syracuse University	2.53	2.38	95	98	71	76
Indiana Univ of Pennsylvania	2.19	1.93	124	126	122	124
Princeton University	4.82	4.39	5	6	12	14
Rutgers State Univ-New Brunswick	3.96	3.62	22	23	16	18
Carnegie Mellon University	3.17	3.01	33	35	52	54