Unemployment and its Educational Predictors

Alcorn, Bryan
Shane, Joelle
Vogel, Abby
Vogel, Todd

December 5, 2016

1 Results

1.1 Modeling

After forming the aforementioned regression models we found 10 coefficients that represent the best fit for each model. The resulting predictive function looks like:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} \tag{1}$$

Table 1: Information about Model Coefficients

	Variables	Ridge	Lasso	PCR
1	UNEMP_RATE	0.000	0.000	0.000
2	INEXPFTE	0.000	-0.000	-0.008
3	TUITIONFEE_IN	-0.963	-0.156	-0.182
4	AVGFACSAL	-0.786	-0.067	-0.089
5	C150_4	-0.695	-0.119	-0.123
6	C150_L4	0.594	0.000	0.002
7	RET_FT4	0.978	0.000	0.027
8	PELL_COMP_ORIG_YR2_RT	-2.752	0.000	-0.111
9	PELL_COMP_ORIG_YR3_RT	0.491	0.042	0.125
10	PELL_COMP_ORIG_YR4_RT	1.372	0.000	0.020
11	CDR2	0.498	0.120	0.166

This table presents the fit between the prediction variables and the response variable ($Unemployment$) determined from the $cv.glmnet()$ and $pcr()$ functions, for each of our 3 models (Ridge, Lasso, PCR).

After finding the 5 predictive functions, we found the MSE's for each model:
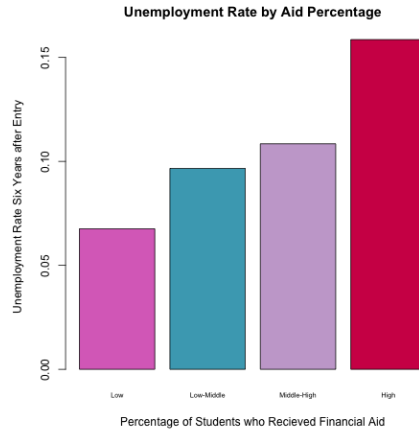
Figure 1: Effect of spending on unemployment rate

Table 2: Information about Mean Squared Errors

	Model	MSE
1	Ridge	0.186
2	Short Ridge	12.797
3	Lasso	0.195
4	Short Lasso	17.114
5	PCR	0.277
6	Short PCR	1.446

Analysis of these numbers will reveal which model has the most predictive power.

1.2 Hypothesis Testing

For hypothesis testing method unemployment rate was set as our response variable and school funding was set as the independent variable (percent of students recieving financial aid).

If the amount of funding received by the school has no effect on quality of education rate, we should see little difference in the response variable (unemployment rate), and that difference is solely due to chance.

We looked for estimate (of effect of spending on unemployment). This value will reveal the relationship between our two variables. A higher estimate means that spending has a significant effect on unemployment and a negative estimate means that spending has an inverse effect on unemployment.

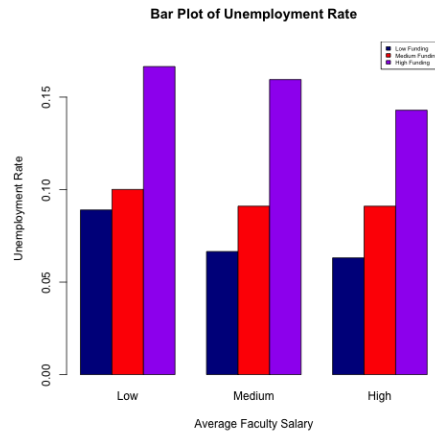Estimate of effect of spending on unemployment rate:

Figure 2: Effect of spending on unemployment rate, controlling for quality of teachers

```
## [1] -0.05141263
```

To control for a difference in unemployment rate generated from the difference in faculty quality, the data was split into three groups, corresponding to schools that had low, medium, and high average faculty salary.

While the results still show that more spending is associated with more unemployment, we can see that average faculty salaray did influence unemployment rate, as schools that have higher average faculty salary also had a lower unemployment rate, in general.

Estimate of the effect of spending on unemployment rate (low average faulty salary):

```
## [1] -0.07196604
```

Estimate of the effect of spending on unemployment rate (middle average faulty salary):

```
## [1] -0.08068517
```

Estimate of the effect of spending on unemployment rate (high average faulty salary):

```
## [1] -0.06582947
```