Results

After forming the aforementioned regression models we found 12 coefficients that represent the best fit for each model. The resulting predictive function looks like:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \beta_{11} X_{11}$$

Table 1: Information about Model Coefficients

	Variables	OLS	Ridge	Lasso	PCR	PLSR
1	Intercept	-0.003	0.000	0.000	0.000	0.000
2	Income	-0.605	-0.574	-0.540	-0.605	-0.605
3	Limit	1.067	0.718	0.815	1.067	1.067
4	Rating	0.286	0.603	0.470	0.286	0.286
5	Cards	0.058	0.044	0.040	0.058	0.058
6	Age	-0.013	-0.016	-0.005	-0.013	-0.013
7	Education	-0.006	-0.005	0.000	-0.006	-0.006
8	GenderFemale	-0.015	-0.013	0.000	-0.015	-0.015
9	StudentYes	0.282	0.276	0.266	0.282	0.282
10	MarriedYes	-0.009	-0.012	0.000	-0.009	-0.009
11	EthnicityAsian	0.011	0.010	0.000	0.011	0.011
12	EthnicityCaucasian	0.011	0.011	0.000	0.011	0.011

This table presents the fit between the prediction variables and the response variable (`Balance`) determined from the `cv.glmnet()`, `pcr()`, and `plsr()` functions, for each of our 5 models (OLS, Ridge, Lasso, PCR, PLSR).

After finding the 5 predictive functions, we found the MSE's for each model:

Table 2: Information about Mean Squared Errors

	Model	MSE
1	OLS	0.047
2	Ridge	0.045
3	Lasso	0.047
4	PCR	0.415
5	PLSR	0.308

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