Regression Analysis of Higher Education Outcomes

Gilbert Watson

Monday, November 25th, 2013

Contents

1 Introduction

Higher education is big business blah, blah, blah, blah.

2 The Data

2.1 Raw Data

Data on five year graduate rates, tuition and fees, and other post-secondary institutional characteristics are available through the US Department of Education's National Center for Education Statistics (NCES). For this study, we chose to only consider data for those institutions that possess the following characteristics:

a) Title IV participating

b) US only

c) Carnegie Foundation ranking in a research, doctoral, or baccalaureate category

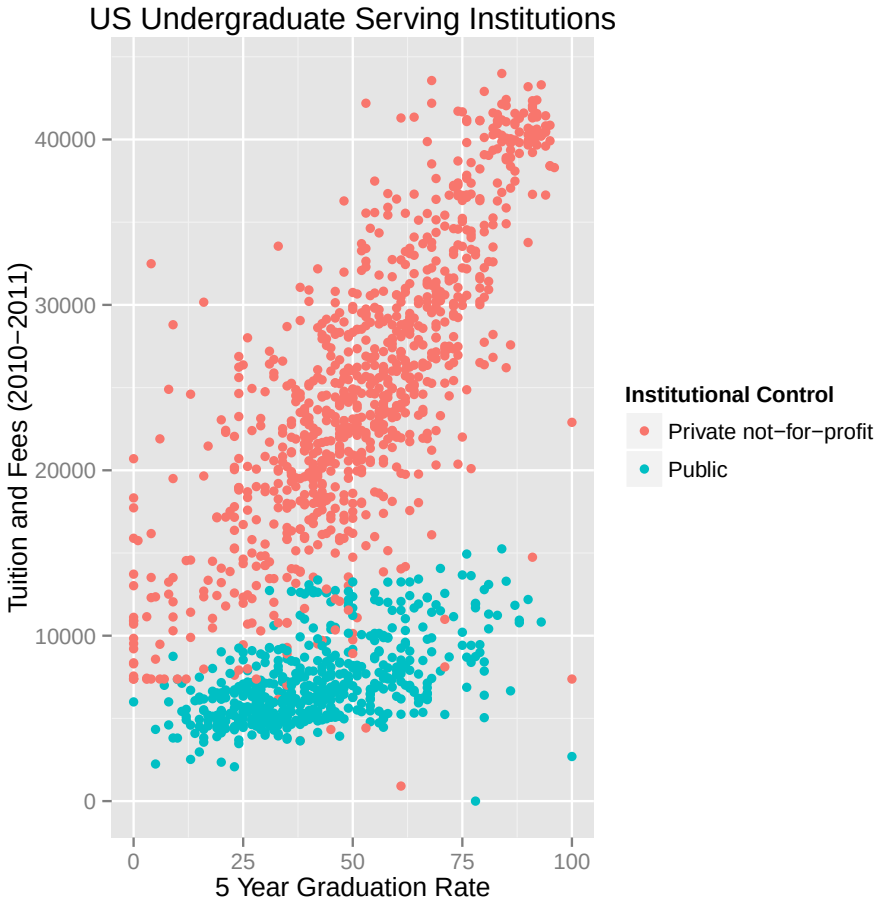
d) The institution has at least some full-time undergraduates

e) The institution has at least one program that is not offered through distance education

f) No for-profit institutions, only public and not-for-profit

The survey that collects this data is administered yearly, but for this study, we chose to only consider the cross section available from the 2010 survey. We do not intend to examine the time dynamics of institutional characteristics.

The original dataset contains 1499 observational units and 320 variables. Variables contain data on characteristics such as admission rates, admission yeilds, tuition and other costs, student body demographics, graduation rates, averages and percentages of students recieving various types of aid, pay rates for different types of university faculty, and various institutional financials on a per faculty basis. In particular, this paper seeks to examine two of these variables relationships to the others - five year undergraduate graduation rates and 2010-2011 educational year tuition and fees. What institutional characteristics are associated with strong educational outcomes (a high graduation rate) and does this have any relationship to the institutional determinants of high cost? A quick examination would show that there is in fact a relationship. A scatter plot of five year graduation rates and tuition and fees shows the existence of a strong relationship, and beyond that a relationship that is bifurcated between private and public institutions. This isn't surprising. Elite educational institutions often cost more and often have higher graduation rates if for no other reason than they attract top notch students. There is a bit of self selction bias going on. This paper is not concerned with this relationship alone. We are concerned about the individual predictors associated with tuition and fees (of which graduate rates may be one) and the individual predictors of graduation rates. Are the predictors similar, or are they determined by wholy different factors when accounting for things like selectivity?



3 ANOVA Analysis

3.1 One-Way ANOVA

```
> library(alr3)
> library(xtable)
> simpleanovalm <- lm(DRVGR2011_RV.Graduation.rate...Bachelor.degree.within.5.years..total~DRVIC2011_R
+                     data=data)
> simpleanova <- anova(simpleanovalm)
> print(xtable(pureErrorAnova(simpleanovalm)))
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
DRVIC2011_RV.Tuition.and.fees..2010.11	1	3423654.98	3423654.98	12287.08	0.0000
Residuals	1464	689073.02	470.68		
Lack of fit	1349	657029.58	487.05	1.75	0.0001
Pure Error	115	32043.44	278.64		

3.2 Two-Way ANOVA

```
> anovapublicprivatelm <- lm(DRVGR2011_RV.Graduation.rate...Bachelor.degree.within.5.years..total~DRVI
> anovapublicprivate <- anova(anovapublicprivatelm)
> print(xtable(pureErrorAnova(anovapublicprivatelm)))
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
DRVIC2011_RV.Tuition.and.fees..2010.11	1	3423654.98	3423654.98	13738.77	0.0000
HD2011.Control.of.institution	2	376291.62	188145.81	755.01	0.0000
Residuals	1462	312781.41	213.94		
Lack of fit	1353	285618.97	211.10	0.85	0.8935
Pure Error	109	27162.44	249.20		

```
> print(xtable(anova(anovapublicprivatelm,simpleanovalm)))
```

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	1462	312781.41				
2	1464	689073.02	-2	-376291.62	879.43	0.0000

```
> anovasectorlm <- lm(DRVGR2011_RV.Graduation.rate...Bachelor.degree.within.5.years..total~DRVIC2011_R
> anovasector <- anova(anovasectorlm)
> print(xtable(pureErrorAnova(anovasectorlm)))
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
DRVIC2011_RV.Tuition.and.fees..2010.11	1	246149.95	246149.95	924.11	0.0000
HD2011.Carnegie.Classification.2010..Undergraduate.Profile	8	176170.44	22021.30	82.67	0.0000
Residuals	1455	220506.73	151.55		
Lack of fit	1413	209319.48	148.14	0.56	0.9986
Pure Error	42	11187.25	266.36		

```
> print(xtable(anova(anovasectorlm,simpleanovalm)))
```

3.3 Data Cleaning

Several steps were taken to clean the data prior to analysis. A full accounting of the steps taken is in the source code for this document, which will be made available on Github. A list of the high points which do concern the analysis follows:

a) Institutional identification variables were removed.

b) A small collection of variables concerning a partitular cohort of students measured by NCES was removed.

c) The small differences between FASB and GASB financial reporting methods for educational institutions are ignored and those variables are merged. These largely comprise per full-time equivalent institutional finance variables.

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	1455	220506.73				
2	1464	689073.02	-9	-468566.29	343.53	0.0000

d) In-state and out-of-state cost variables are merged for private institutions where they have no applicability

e) Admission rates for institutions that only serve one gender were adjusted to zero for the appropriate gender. Additionally, a categorical variable indicating whether an institution serves males, females, or both was generated.

f) One pair of variables that partitioned the data were removed to eliminate the possibility of linear dependence between columns. For example, percentage of students enrolled in non-degree programs was removed as a variable in the presence of full-time and part-time enrollment because all three would perfectly partition the total enrollment variable.

g) Several variable transformations were preformed before analysis to eliminate obvious sources of colinearity. For example, the percentage of enrollment comprised of undergrads was generated to account for the fact that undergraduate enrollment and total enrollment are highly correlated. None of the original variables were removed and are still taken into consideration as predictors, but the transformations made seemed like obvious normalizations.

h) Variables that are missing for over 100 observations are removed from the dataset. This eliminates variables like the revenue from state appropriations since most private schools do not recieve state appropriations.

4 Dimensionality Reduction

319 potential predictors is a lot of variables to deal with. In the face of this, and without any qualitative understanding of the drivers behind graduation rates and tuition and fees, an automated variable selection method was used to reduce the dataset to a manageable number of predictors. For each response variable, the following steps were taken to reduce the dataset.

1) Variables closely related to the response were removed from the dataset. For example, in the case of graduation rates, 4 and 6 year graduation rates were removed when using the 5 year graduation rate as a response variable.

2) The full model with all remaining candidate predictors is estimated.

3) Using the full model, forward, backward, and sequential automated variable selection algorithms are used to search for the best full model. Each method is was allowed to pick a set of up to 50 predictors which optimized model BIC. The union of these predictors is passed to the next step.

4) All but those predictors proposed by the previous step are removed from the dataset. During this step, missing values are retained. A full model with those predictors remaining is estimated. Again, the three methods choose the best model availiable to the algorithm using BIC as a criterion. Each is allowed to consider models of up to 20 predictors. The union of these predictors is passed to the next step.

5) Only those predictors selected by the previous step are retained in the dataset. An efficient branch-and-bound algorithm is used to search the remaining predictors for the best model. The criterion doesn't matter since the search is exhaustive. The predictors selected by this step are used as the initial step in modeling.

The reason for using multiple stages of automated variable selection is that exhaustive search for the best model using just first order terms would be very computationally intensive. This method gives us a manageable number of predictors to begin model refinement. Even graphical search for potential predictors would have been difficult with so many potential variables. There would have been over fifty thousand paired scatter plots to consider.

5 Regression Analysis: 5 Year Graduation Rates

5.1 Initial Model

After the initial data reduction effort described above, an initial model for five year undergraduate graduation rates results. The table below gives the summary statistics for the dataset underlying the model:

	Min	Median	Mean	Max
Five.Year.Bachelors.Graduation.Rate	0.00	49.00	49.69	100.00
Tuition.Fees.2009.2010	876.00	18200.00	17877.57	42335.00
Tuition.Fees.2010.2011	910.00	19015.00	18741.26	43990.00
Mid.East	0.00	0.00	0.19	1.00
Distant.Degree.of.Urbanization	0.00	0.00	0.03	1.00
UgradProfile.FullTime.MoreSelective.HighTransfer	0.00	0.00	0.05	1.00
UgradProfile.FullTime.MoreSelective.LowTransfer	0.00	0.00	0.18	1.00
UgradProfile.FullTime.Selective.HighTransfer	0.00	0.00	0.20	1.00
UgradProfile.FullTime.Selective.LowTransfer	0.00	0.00	0.19	1.00
UgradProfile.Medium.FullTime.Selective.Inclusive	0.00	0.00	0.08	1.00
SizeSetting.Large.FourYear.NotResidential	0.00	0.00	0.08	1.00
SizeSetting.Medium.FourYear.NotResidential	0.00	0.00	0.08	1.00
SizeSetting.Small.FourYear.NotResidential	0.00	0.00	0.06	1.00
SizeSetting.Small.FourYear.Residential	0.00	0.00	0.11	1.00
SizeSetting.VSmall.FourYear.NotResidential	0.00	0.00	0.01	1.00
Undergraduate.Enrollment	71.00	2752.00	5880.41	58404.00
First.Time.Certificate.Seeking.Ugrad.Enrollment	10.00	559.00	1060.16	9254.00
First.Time.Full.Time.Certificate.Seeking.Ugrad.Enrollment	6.00	548.00	1029.94	9082.00
Percent.Total.Enrollment.White	0.00	68.00	62.05	97.00
Full.Time.Retention.Rate.2011	21.00	75.00	74.63	99.00
Bachelors.Degrees.Awarded	8.00	511.00	1103.17	12194.00
Percent.First.Time.Full.Time.Ugrad.Pell.Grant.Recipients	6.00	39.00	40.73	100.00
Investment.Return.As.Percent.Core.Revenue	-5.00	6.00	12.30	89.00
Endowment.Per.FTE.Enrollment	11.00	9974.00	44175.56	2398707.00
Male.Female.Full.Time.Admission.Yeild.Difference	-48.00	2.00	1.95	40.00
Percent.First.Time.Transfer.Ugrad.Enrollment	0.00	0.06	0.07	0.33
Assistant.Percentage.Full.Prof.Salary	-0.37	-0.13	-0.13	0.18
Percent.Average.Student.Loan.Of.Tuition.Ugrad	0.06	0.40	0.60	2.59

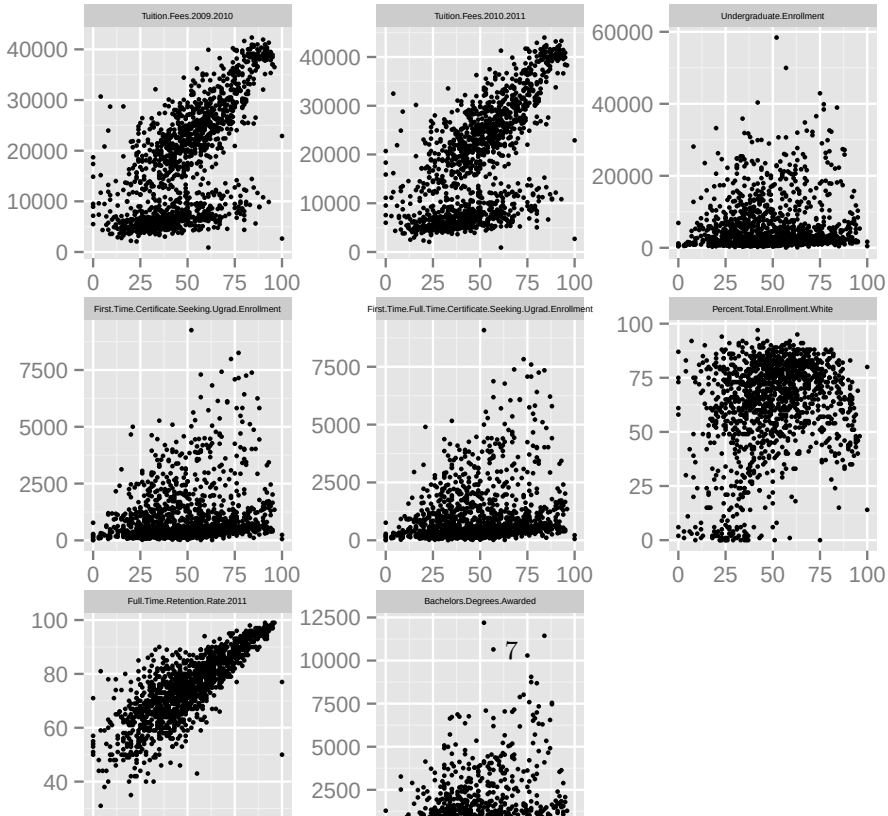
Table 1: Summary Statistics For Dataset (n = 1351)

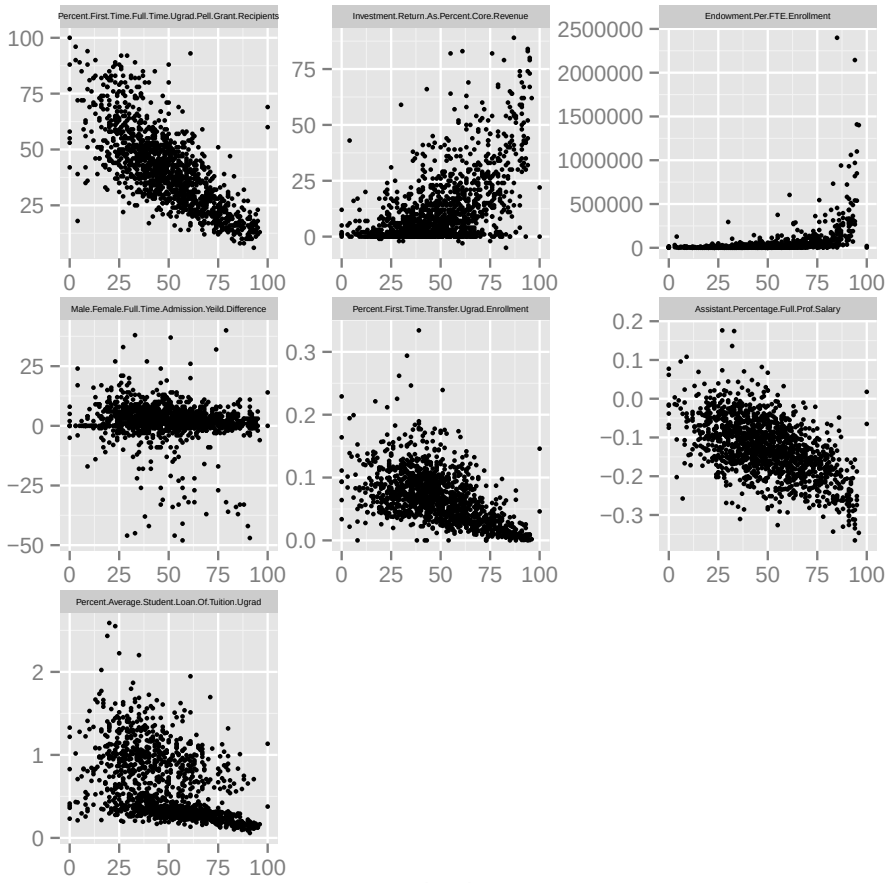
A table outlining the full model associated with this dataset follows:

The first thing we should do is examine scatter plots of the five year graduation rate with other, non-binary variables in the reduced dataset, just to get a sense of what is going on.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-1.7389	3.4773	-0.50	0.6171
Tuition.Fees.2009.2010	0.0009	0.0005	1.94	0.0532
Tuition.Fees.2010.2011	-0.0006	0.0005	-1.36	0.1730
Mid.East	1.9509	0.5710	3.42	0.0007
Distant.Degree.of.Urbanization	-5.8324	1.3631	-4.28	0.0000
UgradProfile.FullTime.MoreSelective.HighTransfer	4.7225	1.2972	3.64	0.0003
UgradProfile.FullTime.MoreSelective.LowTransfer	5.9283	1.0681	5.55	0.0000
UgradProfile.FullTime.Selective.HighTransfer	2.1294	0.6975	3.05	0.0023
UgradProfile.FullTime.Selective.LowTransfer	2.4710	0.7805	3.17	0.0016
UgradProfile.Medium.FullTime.Selective.Inclusive	-2.5668	0.8899	-2.88	0.0040
SizeSetting.Large.FourYear.NotResidential	-6.1480	1.1380	-5.40	0.0000
SizeSetting.Medium.FourYear.NotResidential	-6.4913	0.9431	-6.88	0.0000
SizeSetting.Small.FourYear.NotResidential	-4.9056	0.9843	-4.98	0.0000
SizeSetting.Small.FourYear.Residential	-2.8344	0.7578	-3.74	0.0002
SizeSetting.VSmall.FourYear.NotResidential	-6.0641	1.8418	-3.29	0.0010
Undergraduate.Enrollment	-0.0016	0.0002	-8.19	0.0000
First.Time.Certificate.Seeking.Ugrad.Enrollment	0.0105	0.0031	3.41	0.0007
First.Time.Full.Time.Certificate.Seeking.Ugrad.Enrollment	-0.0087	0.0029	-2.98	0.0029
Percent.Total.Enrollment.White	0.0713	0.0127	5.59	0.0000
Full.Time.Retention.Rate.2011	0.6332	0.0315	20.07	0.0000
Bachelors.Degrees.Awarded	0.0069	0.0007	9.99	0.0000
Percent.First.Time.Full.Time.Ugrad.Pell.Grant.Recipients	-0.1298	0.0235	-5.52	0.0000
Investment.Return.As.Percent.Core.Revenue	0.0562	0.0224	2.51	0.0121
Endowment.Per.FTE.Enrollment	0.0000	0.0000	3.22	0.0013
Male.Female.Full.Time.Admission.Yeild.Difference	-0.1236	0.0282	-4.37	0.0000
Percent.First.Time.Transfer.Ugrad.Enrollment	-19.6618	7.2701	-2.70	0.0069
Assistant.Percentage.Full.Prof.Salary	-11.7602	3.6215	-3.25	0.0012
Percent.Average.Student.Loan.Of.Tuition.Ugrad	-2.9308	1.0626	-2.76	0.0059

Table 2: F = 289.7726 (n = 1351 on 1323 degrees of freedom)





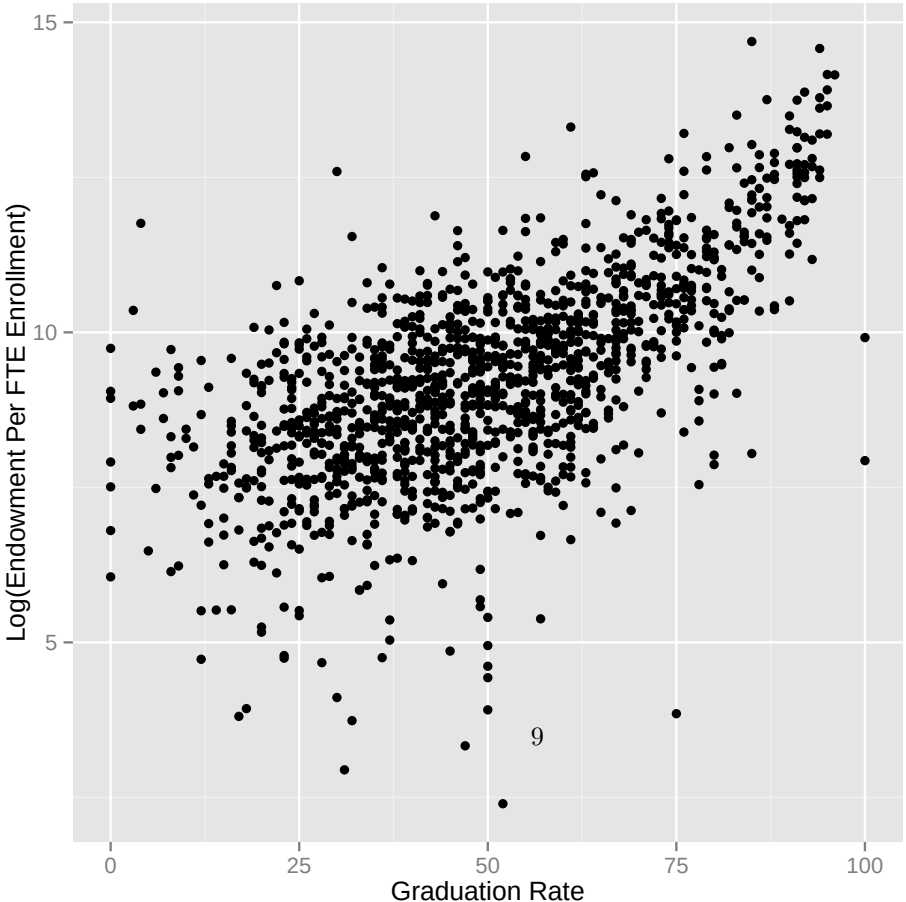
It is clear from this excersize that our full model has several serious problems. There are two pairs of variables that are clearly very collinear. "Tuition.Fees.2009.2010" and "Tuition.Fees.2010.2011" as well as "First.Time.Certificate.Seeking.Ugrad.Enrollment" and "First.Time.Full.Time.Certificate.Seeking.Ugrad.Enrollment" are providing almost the same information to the model. This can be confirmed by computing the variance inflation factors for all the variables in the full model. Clearly one variable of each pair should be removed from the model.

It is also clear from this table that "Undergraduate.Enrollment" and "Bachelor.Degrees.Awarded" are also causing collinearity issues. One would suspect that an intstitution with a large undergraduate enrollment would also have a large number of bachelor degrees awarded. One of these pair should likely be removed as well.

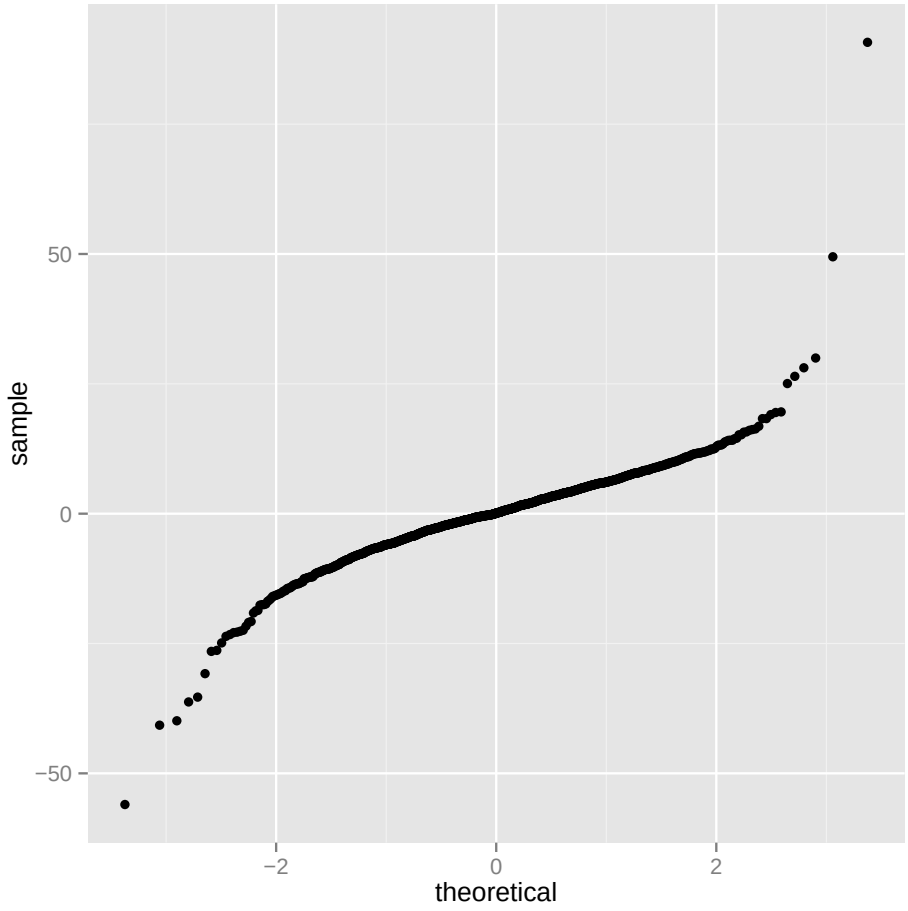
Another obvious correction would be to include an indicator variable for public institutions back into the model. The scatter plot panel with tuition on the y axis and graduation rate on the x axis is just the same as our first scatter plot comparing the two. The is a very obvious interaction between tuition and an institution's identity as public or private.

One more obvious correction would be to add a variable transformations for endowment per FTE. Earlier scatter plots indicate a diminishing relationship between this variable and graduation rates. An appropriate transform for endowment per FTE appears to be a log transformation.

	Variable	VIF
1	Tuition.Fees.2009.2010	604.98
2	Tuition.Fees.2010.2011	605.49
3	Mid.East	1.13
4	Distant.Degree.of.Urbanization	1.04
5	UgradProfile.FullTime.MoreSelective.HighTransfer	1.66
6	UgradProfile.FullTime.MoreSelective.LowTransfer	3.77
7	UgradProfile.FullTime.Selective.HighTransfer	1.73
8	UgradProfile.FullTime.Selective.LowTransfer	2.09
9	UgradProfile.Medium.FullTime.Selective.Inclusive	1.25
10	SizeSetting.Large.FourYear.NotResidential	2.04
11	SizeSetting.Medium.FourYear.NotResidential	1.39
12	SizeSetting.Small.FourYear.NotResidential	1.22
13	SizeSetting.Small.FourYear.Residential	1.23
14	SizeSetting.VSmall.FourYear.NotResidential	1.10
15	Undergraduate.Enrollment	43.08
16	First.Time.Certificate.Seeking.Ugrad.Enrollment	337.65
17	First.Time.Full.Time.Certificate.Seeking.Ugrad.Enrollment	292.15
18	Percent.Total.Enrollment.White	1.75
19	Full.Time.Retention.Rate.2011	3.26
20	Bachelors.Degrees.Awarded	23.94
21	Percent.First.Time.Full.Time.Ugrad.Pell.Grant.Recipients	3.66
22	Investment.Return.As.Percent.Core.Revenue	2.74
23	Endowment.Per.FTE.Enrollment	1.77
24	Male.Female.Full.Time.Admission.Yeild.Difference	1.06
25	Percent.First.Time.Transfer.Ugrad.Enrollment	2.03
26	Assistant.Percentage.Full.Prof.Salary	1.61
27	Percent.Average.Student.Loan.Of.Tuition.Ugrad	4.15



Before making these corrections though, we have one larger problem. There are several indicators that the residuals are not normally distributed, likely due to outliers. A qq plot of the residuals suggests that they are distributed with heavy tails.

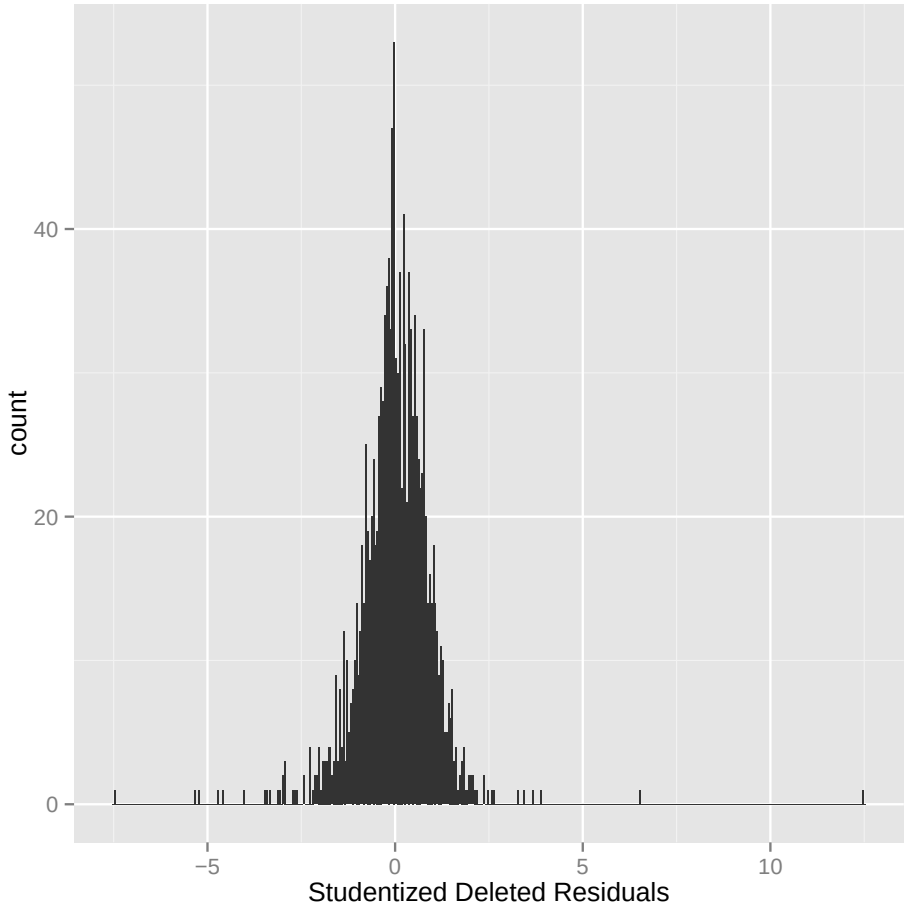


5.2 Outlier Removal

A histogram of the studentized deleted residuals would suggest that we have several outliers. After calculating the bonferroni critical value for outliers, we can assume that the following observations for five year graduation rate should be removed from the dataset. Note that I have set the α level at $\alpha = 0.10$ for the Bonferroni test.

```
> library(MASS)
> stud.del.res <- studres(grfull)
> qplot(stud.del.res,xlab="Studentized Deleted Residuals",binwidth=0.05)
> alpha <- 0.1
> ct <- qt(1-alpha/(2*length(stud.del.res)),length(stud.del.res) - length(grfull$coefficients))
> groutliers <- which(abs(stud.del.res) > ct)
> print(xtable(data.frame(Institutions = rownames(GradRate)[which(abs(stud.del.res) > ct)])))
```

	Institutions
1	Arizona Christian University : Arizona
2	Holy Cross College : Indiana
3	Franklin Pierce University : New Hampshire
4	Rivier University : New Hampshire
5	Northern New Mexico College : New Mexico
6	Northwest Christian University : Oregon
7	Gwynedd Mercy College : Pennsylvania
8	Hampden-Sydney College : Virginia

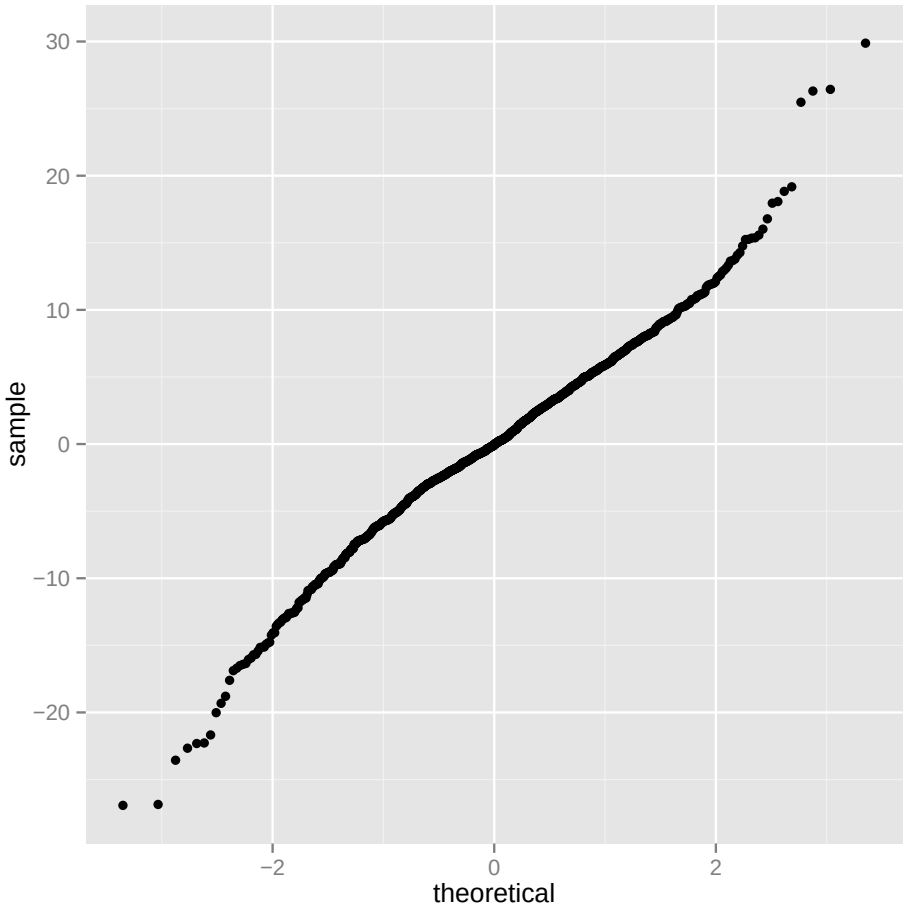


The predictor variables also have outliers. We use the hat matrix to find them:

```
> diagonal <- lm.influence(grfull)$hat
> threshold <- 2*(length(grfull$coefficients) - 1)/length(grfull$model$Five.Year.Bachelors.Graduation.
> xoutliers <- which(diagonal > threshold)
> print(xtable(data.frame(Institutions = rownames(GradRate)[xoutliers]),))
```

Now let's remove these from our full model and see if our residuals are more normally distributed.

```
> grfull_nouts <- lm(Five.Year.Bachelors.Graduation.Rate~.,data=GradRate[-c(groutliers,xoutliers),setd
> grfull_noutssumm <- summary(grfull_nouts)
> print(xtable(grfull_noutssumm,caption=paste0("F = ",round(grfull_noutssumm$fstatistic,4)[1]," (n = "
```



It appears that our residuals are now normally distributed. Now we can make the apply corrections we noticed necessary in our initial model.

5.3 Modeling

Now that we have identified outliers in the data, we can estimate the model, incorporating the transformation to endowment per FTE and the interaction of tuition and fees with an institution's status as public or private.

Some coefficients have lost individual significance of thier coefficients. Scaling of the number of first time certificate seeking undergrads aids it's significance, but others are not aided. It may be necessary to remove them from the model. Additionally, computing the VIF statistic again shows that the tuition and fees interaction with public or private status, as well as undegraduate enrollment and first time certificate seeking undergrad enrollment are causing multicollinearity issues. Centering only appears to make the interaction term's contribution to the model's collinearity issues greater according to the VIF statistic, so we will live with that. We will now test if we can remove the first time certificate seeking undegrad enrollment variable.

```
> grred.r <- lm(Five.Year.Bachelors.Graduation.Rate~. + Tuition.Fees.2010.2011*Private + I(log(Endowme
> grred.rsum <- summary(grred.r)
> F_star <- ((deviance(grred.r) - deviance(grred)) /
+              ((length(grred$model$Five.Year.Bachelors.Graduation.Rate) - grred.rsum$df[1]) -
```

```
+                 (length(grred$model$Five.Year.Bachelors.Graduation.Rate) - grredsum$df[1]))) /
+   (deviance(grred)/length(grred$model$Five.Year.Bachelors.Graduation.Rate) - grredsum$df[1])
> cv <- qf(0.99,1,length(grred$model$Five.Year.Bachelors.Graduation.Rate)-grredsum$df[1])
> F_star < cv

[1] FALSE
```

F^* is greater than the critical value of 6.65580753872202 so we cannot remove the first time certificate seeking undergrad variable, the next test shows that we can remove all three of the variables endowment per FTE, the percentage that the average student loan is of tuition, and the percent of first time undgrad transfers enrolled.

```
> grred.r <- lm(Five.Year.Bachelors.Graduation.Rate~. + Tuition.Fees.2010.2011*Private - Endowment.Per
> grred.rsum <- summary(grred.r)
> F_star <- ((deviance(grred.r) - deviance(grred)) /
+              ((length(grred$model$Five.Year.Bachelors.Graduation.Rate) - grred.rsum$df[1]) -
+                 (length(grred$model$Five.Year.Bachelors.Graduation.Rate) - grredsum$df[1]))) /
+   (deviance(grred)/length(grred$model$Five.Year.Bachelors.Graduation.Rate) - grredsum$df[1])
> cv <- qf(0.99,3,length(grred$model$Five.Year.Bachelors.Graduation.Rate)-grredsum$df[1])
> F_star < cv

[1] TRUE
```

We still have one more model problem to deal with if we are concerned about using our model for inference - heterskedasticity. The model exhibits heterskedasticity according the Breusch-Pagan test. We can quickly correct for this by using robust regression estimation.

```
> library(lmtest)
> bptest(grred.r)

        studentized Breusch-Pagan test

data:  grred.r
BP = 68.5881, df = 22, p-value = 1.101e-06

> test <- rlm(Five.Year.Bachelors.Graduation.Rate~. + Tuition.Fees.2010.2011*Private - Endowment.Per.F
```

5.4 Intepretation

	Institutions
1	The University of Alabama : Alabama
2	Troy University : Alabama
3	University of Alaska Anchorage : Alaska
4	Arizona State University : Arizona
5	University of Arizona : Arizona
6	Northern Arizona University : Arizona
7	Central Baptist College : Arkansas
8	California Institute of Technology : California
9	University of California-Berkeley : California
10	University of California-Irvine : California
11	University of California-Los Angeles : California
12	University of California-San Diego : California
13	National University : California
14	Pomona College : California
15	San Diego State University : California
16	Thomas Aquinas College : California
17	Metropolitan State University of Denver : Colorado
18	Naropa University : Colorado
19	Yale University : Connecticut
20	Trinity Washington University : District of Columbia
21	University of Central Florida : Florida
22	Florida Atlantic University : Florida
23	Florida State University : Florida
24	University of Florida : Florida
25	University of South Florida-Main Campus : Florida
26	Brewton-Parker College : Georgia
27	Macon State College : Georgia
28	Thomas University : Georgia
29	Calumet College of Saint Joseph : Indiana
30	Saint Mary-of-the-Woods College : Indiana
31	Southwestern College : Kansas
32	Berea College : Kentucky
33	Mid-Continent University : Kentucky
34	Midway College : Kentucky
35	Saint Catharine College : Kentucky
36	Grambling State University : Louisiana
37	University of Maine at Fort Kent : Maine
38	University of Maine at Presque Isle : Maine
39	Saint Joseph's College of Maine : Maine
40	University of Maryland-University College : Maryland
41	University of Maryland-College Park : Maryland
42	St Mary's College of Maryland : Maryland
43	Cambridge College : Massachusetts
44	Harvard University : Massachusetts
45	Cornerstone University : Michigan
46	Rochester College : Michigan
47	University of Michigan-Ann Arbor : Michigan
48	Michigan State University : Michigan
49	Finlandia University : Michigan
50	Alcorn State University : Mississippi
51	Blue Mountain College : Mississippi
52	Columbia College : Missouri
53	University of Missouri-St Louis : Missouri
54	William Jewell College : Missouri
55	Bellevue University : Nebraska
56	Peru State College : Nebraska
57	College of Saint Mary : Nebraska
58	University of New Hampshire at Manchester : New Hampshire

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.6623	3.3150	0.20	0.8417
Tuition.Fees.2009.2010	-0.0000	0.0006	-0.07	0.9481
Tuition.Fees.2010.2011	0.0003	0.0005	0.55	0.5811
Mid.East	2.4110	0.4997	4.82	0.0000
Distant.Degree.of.Urbanization	-5.4021	1.8528	-2.92	0.0036
UgradProfile.FullTime.MoreSelective.HighTransfer	2.6543	1.1422	2.32	0.0203
UgradProfile.FullTime.MoreSelective.LowTransfer	4.2649	0.9401	4.54	0.0000
UgradProfile.FullTime.Selective.HighTransfer	1.1345	0.6062	1.87	0.0615
UgradProfile.FullTime.Selective.LowTransfer	2.0683	0.6822	3.03	0.0025
UgradProfile.Medium.FullTime.Selective.Inclusive	-2.5974	0.7920	-3.28	0.0011
SizeSetting.Large.FourYear.NotResidential	-5.3214	1.0896	-4.88	0.0000
SizeSetting.Medium.FourYear.NotResidential	-6.0833	0.8555	-7.11	0.0000
SizeSetting.Small.FourYear.NotResidential	-6.1999	0.8790	-7.05	0.0000
SizeSetting.Small.FourYear.Residential	-2.7491	0.6495	-4.23	0.0000
Undergraduate.Enrollment	-0.0018	0.0002	-8.13	0.0000
First.Time.Certificate.Seeking.Ugrad.Enrollment	-0.0018	0.0057	-0.32	0.7481
First.Time.Full.Time.Certificate.Seeking.Ugrad.Enrollment	0.0041	0.0056	0.74	0.4614
Percent.Total.Enrollment.White	0.0837	0.0114	7.33	0.0000
Full.Time.Retention.Rate.2011	0.6201	0.0289	21.43	0.0000
Bachelors.Degrees.Awarded	0.0077	0.0008	9.42	0.0000
Percent.First.Time.Full.Time.Ugrad.Pell.Grant.Recipients	-0.1549	0.0216	-7.16	0.0000
Investment.Return.As.Percent.Core.Revenue	0.0302	0.0234	1.29	0.1965
Endowment.Per.FTE.Enrollment	0.0000	0.0000	3.92	0.0001
Male.Female.Full.Time.Admission.Yeild.Difference	-0.1292	0.0264	-4.89	0.0000
Percent.First.Time.Transfer.Ugrad.Enrollment	-19.2938	6.9698	-2.77	0.0057
Assistant.Percentage.Full.Prof.Salary	-8.6323	3.2296	-2.67	0.0076
Percent.Average.Student.Loan.Of.Tuition.Ugrad	-3.3480	1.0810	-3.10	0.0020

Table 3: F = 374.5623 (n = 1240 on 1213 degrees of freedom)

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-7.7089	4.0720	-1.89	0.0586
Tuition.Fees.2010.2011	0.0003	0.0001	5.96	0.0000
Mid.East	2.1592	0.5212	4.14	0.0000
Distant.Degree.of.Urbanization	-5.2771	1.9233	-2.74	0.0062
UgradProfile.FullTime.MoreSelective.HighTransfer	3.6546	1.1782	3.10	0.0020
UgradProfile.FullTime.MoreSelective.LowTransfer	5.2423	0.9752	5.38	0.0000
UgradProfile.FullTime.Selective.HighTransfer	1.1431	0.6291	1.82	0.0695
UgradProfile.FullTime.Selective.LowTransfer	2.2058	0.7046	3.13	0.0018
UgradProfile.Medium.FullTime.Selective.Inclusive	-3.0169	0.8246	-3.66	0.0003
SizeSetting.Large.FourYear.NotResidential	-6.4665	1.1246	-5.75	0.0000
SizeSetting.Medium.FourYear.NotResidential	-6.7317	0.8767	-7.68	0.0000
SizeSetting.Small.FourYear.NotResidential	-5.7885	0.9151	-6.33	0.0000
SizeSetting.Small.FourYear.Residential	-2.3676	0.6737	-3.51	0.0005
Undergraduate.Enrollment	-0.0004	0.0001	-2.63	0.0086
First.Time.Full.Time.Certificate.Seeking.Ugrad.Enrollment	0.0028	0.0008	3.65	0.0003
Percent.Total.Enrollment.White	0.0665	0.0117	5.70	0.0000
Full.Time.Retention.Rate.2011	0.6633	0.0297	22.36	0.0000
Percent.First.Time.Full.Time.Ugrad.Pell.Grant.Recipients	-0.1642	0.0227	-7.23	0.0000
Investment.Return.As.Percent.Core.Revenue	0.0712	0.0227	3.14	0.0018
Male.Female.Full.Time.Admission.Yeild.Difference	-0.1360	0.0274	-4.97	0.0000
Percent.First.Time.Transfer.Ugrad.Enrollment	-2.1561	7.1037	-0.30	0.7616
Assistant.Percentage.Full.Prof.Salary	-11.7636	3.3780	-3.48	0.0005
Percent.Average.Student.Loan.Of.Tuition.Ugrad	-0.9725	1.5585	-0.62	0.5328
Private	-6.2954	2.0761	-3.03	0.0025
I(log(Endowment.Per.FTE.Enrollment))	0.2719	0.2257	1.20	0.2286
Tuition.Fees.2010.2011:Private	0.0008	0.0002	3.88	0.0001

Table 4: F = 357.8953 (n = 1351 on 1214 degrees of freedom)