A Regression Analysis of College Tuition and Mean Income

William Johnson

Kristin Lax

Connor Mitnick

Abstract

Studies spanning the second half of the 20th century have indicated that the more an university costs to attend, the greater return on investment in the form of eventual earnings. However, the current educational market suggests that these previous assumptions may no longer apply, and that cost of a university is no longer a significant indicator of how successful its graduates will be. With several notable shifts in this market in the last decade - including an increase in online educational programs, a rising student population incurring student loans, and the suggestion that all public universities charge no tuition, the necessity for empirical testing of the effects of cost on eventual success became apparent. While cost of a university did prove to be an influential factor in determining wages when examined on its own, it fell flat when other education based variables, such as average school expenditures per student and average SAT scores, were added into the equation. It became even more insignificant when other underlying demographic and socioeconomic factors such as race and average family income were taken into consideration. What a series of regressions conclusively shows is that cost is in fact, insignificant compared to a myriad of other relevant variables from several different influential categories. The positive implications of these findings are hopefully that students expecting to enter college within the next several years can avoid making a painful investment, assuming that their investment will pay off just because of the amount of money they have spent suggests a certain outcome. That investment can rather be made based on considerations of more important factors, such as what a university can offer a student in terms of resources, whether that be money or man hours, and how the success of an institution's graduates reflect its quality.

I. Introduction

Both social and statistical evidence supports the notion that the necessity of college education in building a successful life is growing. However, there is little indication about whether the return on investment of a college education is growing alongside it. While this paper does not intend to deter anyone from obtaining a college degree, it does aim to urge prospective college students to consider their motivation when choosing where to earn that degree. Oftentimes the high price of an elite school indicates a certain prestige that creates the illusion of greater opportunity, but is there evidence to support that these greater opportunities actually exist? Especially in a time period where providing subsidized public college education is being seriously considered in government and political forums, the true motivation for attending college is now simply that it is a necessity. Popular public opinion tells us that a college degree is now absolutely necessary to obtain employment in most if not all professional fields, and it is becoming abundantly clear that what is now attractive to students is not prestige, but efficiency. If the same opportunity can be found at a lower cost, it is reasonable to take it. This paper aims to prove, in the current educational market, that those opportunities do exist.

The impact of college selectivity on post-graduate earnings is not a new topic. However, while past trends indicate that the return on investment of a more expensive education, though declining (Fox 1993), is highly positive compared to investment in lower cost college education, the market for education within this decade is different than it's ever been. A heavy focus is now being put on cost before rankings, and the ideal is a university that can produce successful graduates while still providing an education that can, in short, minimize the now massive student loans that every twenty-something seems to accrue. Take Georgia Tech, for example, who was recently named the number one school in an article titled, "9 Non-Ivys That Give Students a Better Bang for Their Buck Than Harvard or Yale" (2016). Past studies have shown that the more you spend, the more you make, but this paper aims to investigate if that is still true. Given the current social climate regarding educational costs and the rising spirit of competition between both ranked and unranked universities to produce highly capable graduates, this paper also expects to find that the old studies no longer apply.

II. Literature Review

The great issue with obtaining a college education is not simply that it is costly, but that its costs seems to continuously rise. Fox (1993) examined how this rise in tuition has affected the ability of elite private universities to remain a worthwhile investment. For decades the most selective universities have been private. Throughout this time it has also held true that private colleges, particularly the most elite, charge far more in tuition than both less elite private institutions and public universities of all levels of

selectivity, and the gap in those prices is also steadily growing. Despite this vast difference in upfront investment, many prospective students and their families have reasoned that these investments are still worth making, given the reputation of these highly selective schools and a number of other factors. When determining how to quantify these investments, deemed human capital investments by Fox, the author examined postgraduate net earnings from those who attended elite private colleges as well as those who attended less selective public universities, comparing the two to obtain a differential figure. While the findings indicated that investing in a private education offers at least a comparable rate of return to investing in higher education, period, it showed that the attractiveness of an investment in private education versus a cheaper public education may be dwindling. The differential between the "most high" college earnings and the earnings yielded by a graduate of an average public institution, while still positive, represents a gap that is shrinking as the years roll on and the tuition of colleges across the U.S. continue to rise, leading Fox to believe that it would not take more than a decade for the investment in an elite private college to no longer be attractive. This paper was written in 1993. While this analysis offers a great amount of insight into a trend that can continuously be observed today, it fails to account for those students who continue to pursue higher education in the form of a master's or doctorate degree, or in some other form of specialized schooling such as medical or law school. The author acknowledges that, because a great and growing percentage of students who choose to pursue higher education come from universities of higher ranking, there is potential for the data to be skewed in the direction of cheaper public colleges. The author also acknowledges that if other factors besides cost differential were considered in determining return on a college investment, the results of this analysis might display a more full and comprehensive representation of the market for college education as it stands today.

When discussing the effect college tuition cost has on future earnings, it is important to note, understand, and access the variety of other factors that come into play. Ernest T. Pascarella builds on this idea in a paper entitled "College tuition costs and early career socioeconomic achievement: do you get what you pay for?" He formed five dependent variables in his final analysis: educational attainment, occupational status, earnings, women's entrance into traditional male-dominated careers, and spouse's earnings (1992). Therefore, an OLS was performed to determine the direct or unmediated effect of college tuition on each dependent variable. Each analysis was separated based on race (black or white) and sex, which is not common across literature pertaining to college tuition. This might contribute to the desire to reduce impact of different styles of colleges could have on results. Pascarella concludes through statistical analysis that "attendance at a relatively high cost college may make a unique and significant, if modest, positive contribution to early socioeconomic achievement" (1992). What is important to note is that this conclusion still stands given their appreciation of controls for "student background characteristics and

aspirations," among others. This leads to the conclusion that "in terms of impact on early socioeconomic outcomes, college costs appear to function as more than a proxy for the characteristics of students recruited to schools varying in tuition" (1992). A main issue brought up during discussion over varying causal mechanisms is what actually happens to a college student during his or her time in school. The increased tuition cost may lead to a student earning more, but not necessarily because of any extraordinary cognitive development during college. That is, even though tuition may be higher, indicating a potential for better professors or facilities, the development of the student may not have a positive correlation with tuition. Then, the reasoning behind higher earnings may still come from how employers view degrees from different schools, and how those degrees differed in cost.

Additionally, while considering a multitude of factors in determining educational return certainly adds value to an analysis, it may also be beneficial to consider how differing methods of computation can affect the results of that analysis mathematically. In his 2008 paper entitled "Comparing two approaches to the rate of return to investment in education," Orhan Kara observes the results of a rate of return analysis using two different methods. The first method of using the internal rate of return, where benefits and a multitude of costs including tuition and opportunity cost are taken into account, showed that for the country of Turkey, the rate of return on investment is actually declining with higher education, probably due to Turkey's unique job market. Conversely, when using an earnings equation where variables such as years of schooling, expenditures and a multitude of social household factors are used in a multiple linear regression, the results show that the returns are actually increasing with education, a near exact opposite of the results of the previous analysis. While these results would likely differ in a country like the United States, and this analysis does not actually differentiate between colleges as it would if applied to the United States, they can offer insight into the sensitivity of the results of an analysis like this. While Kara believes that the internal rate of return analysis better reveals an estimation of return "from the actual amount invested in schooling", he also believes that the earnings equation method yields a more accurate result for estimating the earnings gaps associated with additional schooling (2008). Regardless of the numerical results, the most important piece of insight that this paper provides is that using several methods of analysis can yield different, but equally important data in terms of the implications of each analysis and its individual results.

While this paper examines a topic similar to all of the papers mentioned above, and uses a method particularly similar to Fox's 1993 paper on investment in elite public colleges, it intends to capitalize on the potential improvements mentioned in the above reviews. While it will show the relationship between the cost of a university and the average salary of its graduates, it will also show how other factors affect this relationship, much like in Pascarella's analysis, though the additional variables will differ to better

align with this paper's hypothesis. Ultimately this paper will tell whether a larger investment secures a larger return, but it does not ignore variables that develop over the period of time between when that money is invested and the return is received. Loans, length of college career, and a variety of other factors can conceivably have a large effect on the post-graduate success of a student, as well as university expenditures per student. Additionally, this paper will offer several different analyses to show the impact of both the addition and removal of pertinent variables. It is likely that these additional analyses will lead to a more comprehensive understanding of the educational market as it stands today, thus helping to either support or disprove the original hypothesis with a greater degree of certainty. Lastly, a key feature of this paper is its currency among other papers of similar content that attempt to apply conclusions obtained from a dated data set to similar data obtained today. This paper not only adds depth to previous studies, but also brings those studies into the present.

III. Data

Variables

This paper analyzes the mean income of students 7 years after enrollment across 1327 different universities and colleges in the United States. This includes different kinds of four year institutions such as community colleges, private universities, and public universities, but excludes one or two year secondary education programs.

Using this data, a simple regression of the average student's cost of attendance per year (that is considering tuition, food, room, and board) should reveal any correlation between a rising cost in education and rising earnings thereafter. However, this simple regression will not represent an accurate model as due to the lack of consideration of other factors which might affect the mean income of graduated students.

A series of separate multiple regressions against various considerations listed in the table below should create a more accurate model of the mean income, which in turn will provide more information about the impact of school cost on a student's eventual income. Variables such as the mean income of students, percentage of students graduating within six years of enrollment, percentage of graduates with unpaid student loans after seven years, average SAT scores of the student population, and whether the university in public or private were included because they can intuitively be expected to indicate, either separately or jointly, the quality of education provided by a university, thus indicating how successful its students will be. However, demographic information such as the average family income of students at the university, as well as the percentages of students or differing races was included to determine to what degree factors outside a student's education contribute to eventual wealth.

A final regression of the institution's instructional expenditures per student and the students' four year completion rate should demonstrate any impact the institution's focus on its students has.

Name	Description	Units	Type
l_earn	The mean income of students 7 years after enrollment in the respective institution	US Dollars	Dependent
private	Dummy variable representing whether or not a university is private or public	Boolean	Independent
lcost	The average total cost of attending the respective institution	Log of US Dollars	Independent
lexp	The instructional expenditures spent per student by the institution.	Log of US Dollars	Independent
c150_4	The percentage of students that earn a degree at the institution within 150% of the expected graduation time (4 to 6 years)	Percentage of Student Class	Independent
rpy_7yr_rt	The percentage of students that have not defaulted on, have payed off, or are still paying off their student loans 7 years after enrollment	Percentage of Student Class	Independent
sat_avg_all	The average SAT score of accepted students at the respective institution	SAT Score	Independent
lfaminc	The average family income of students at the university	Log of US Dollars	Independent
udgs_white	The percentage of students at the university who are white	Percentage of Student Class	Independent

udgs_black	The percentage of students at the university who are black	Percentage of Student Class	Independent
udgs_hisp	The percentage of students at the university who are hispanic	Percentage of Student Class	Independent
udgs_asian	The percentage of students at the university who are asian	Percentage of Student Class	Independent

Figure 3.1: A description of all the variables considered in this document

The data used in this document was collected by the U.S. Department of Education's College Scorecard program (https://collegescorecard.ed.gov/data/). The program holds data beginning in 1996, but this document only uses the most recent data collected in 2015.

Statistics

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
l_earn	1,316	10.58388	0.2204357	9.961757	11.71913
private	1,316	.6025836	0.4895495	0	1
lcost	1,316	10.28061	0.4112766	9.19766	11.04443
lexp	1,316	9.04798	0.4589147	7.543803	11.57321
c150_4	1,316	0.5350456	0.1711883	0	0.9779
rpy_7yr_rt	1,316	0.8646419	0.0867108	0.4051144	0.9917763
sat_avg_all	1,316	1057.924	123.3705	666	1504
lfaminc	1,316	11.09673	0.3076053	9.864335	11.80635
ugds_white	1,316	0.6209891	0.2192719	0	0.9611
ugds_black	1,316	0.1328103	0.18378	0.0007	0.9987

ugds_hisp	1,316	0.0880611	0.1034942	0	0.9454
ugds_asian	1,316	0.0409061	0.0590354	0	0.4371

Figure 3.2: Statistics on all the variables used in this document

Regression Models

Simple Regression:

$$l\_earn = \beta_0 + \beta_1 * (lcost) + u$$

This regression most directly examines the return on investment of college education.

Multiple Regression 1:

$$l\_earn = \beta_0 + \delta_1 private + \beta_1 * (lcost) + \beta_2 * (lexp) + \beta_3 * (c150\_4) + \beta_4 * (rpy\_7yr\_rt) + \beta_5 * (sat\_avg\_all) + \beta_6 * (lfaminc) + \beta_7 * (ugds\_white) + \beta_8 * (ugds\_black) + \beta_9 * (ugds\_hisp) + \beta_{10} * (ugds\_asian) + u$$

This regression tries to eliminate bias by considering all variables in the dataset that might influence the mean income.

Multiple Regression 2:

$$l\_earn = \beta_0 + \beta_1 * (lexp) + \beta_2 * (rpy\_7yr\_rt) + \beta_3 * (sat\_avg\_all) + \beta_4 * (lfaminc) + u$$

This regression removes the independently insignificant variables from the first regression to try to achieve a better model.

Multiple Regression 3:

$$l\_earn = \beta_0 + \beta_1 * (lexp) + \beta_4 * (rpy\_7yr\_rt) + \beta_5 * (sat\_avg\_all) + \beta_6 * (lfaminc) + \beta_7 * (ugds\_white) + \beta_8 * (ugds\_black) + \beta_9 * (ugds\_hisp) + \beta_{10} * (ugds\_asian) + u$$

This regression adds back in the independently insignificant variables that are jointly significant.

Model Assumptions

The Gauss Markov Assumptions must be met in order to ensure a degree of reliability of a linear regression.

1. Linearity of Model:

   Each model uses the standard form $Y = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \ldots$, meeting this assumption.

2. Random sampling:

As the U.S. Department of Education has records kept for nearly every university related to our hypothesis, the data does indeed create random variables that span the width of the domain we wish to examine. However, we have eliminated approximately half of the original universities in the dataset as many institutions withheld data from records for privacy reasons. That acknowledged, this should not create any bias as the institutions' decisions to withhold varies arbitrarily.

3. Zero Conditional Mean:

It is likely that this simple regression will not meet this assumption due to the lack of an account for the graduation rate in this model. As such, there will be a positive bias in this model due to the positive correlations listed in Figure 3.3 and the positive coefficients in Multiple Regression 1. In the multiple regressions, it is harder to know if other variables ought to be included. The use of the varied regressions should help to give insight into the unbiased model.

4. No Perfect Collinearity:

Listed below are the correlations between the different variables used in the regressions. Since no variables have perfect collinearity, this assumption is met. Yellow boxes indicate correlations greater than |0.5|, and pink boxes indicate correlations greater than |0.7|

	l_earn	**private**	**lcost**	**lexp**	**c150_4**	**rpy_7yr_rt**	**sat_avg_all**	**lfaminc**
l_earn	1.0000							
private	0.0809	1.0000						
lcost	0.3714	0.8063	1.0000					
lexp	0.5523	0.0717	0.4093	1.0000				
c150_4	0.5981	0.2254	0.5537	0.6421	1.0000			
rpy_7yr_rt	0.5742	0.1670	0.4157	0.4085	0.7217	1.0000		
sat_avg_all	0.6505	0.0627	0.3845	0.6463	0.7643	0.6833	1.0000	
lfaminc	0.4677	0.2944	0.5876	0.3832	0.7126	0.7370	0.6136	1.0000
ugds_white	0.0109	0.0276	0.0385	-0.0702	0.2328	0.4981	0.2871	0.4684

ugds_black	-0.3426	-0.0513	-0.1907	-0.1714	-0.4045	-0.7307	-0.4871	-0.4872
ugds_hisp	0.1539	-0.0810	-0.0373	0.0565	-0.0451	0.0107	-0.0673	-0.1977
ugds_asian	0.5001	-0.0810	0.1378	0.4500	0.3271	0.2668	0.3942	0.0283

	ugds_white	**ugds_black**	**ugds_hisp**	**ugds_asian**
ugds_white	1.0000			
ugds_black	-0.6652	1.0000		
ugds_hisp	-0.1977	-0.1276	1.0000	
ugds_asian	0.0283	-0.1638	0.3222	1.0000

Figure 3.3: Correlations between all the variables being considered in this document

5. Homoskedacity:

The variance of the residuals in the models used in this paper grows quickly, but then remains relatively constant, as seen in the residual graphs below. This behavior should closely enough approximate this assumption to minimize bias generated by the breaking of this assumption.

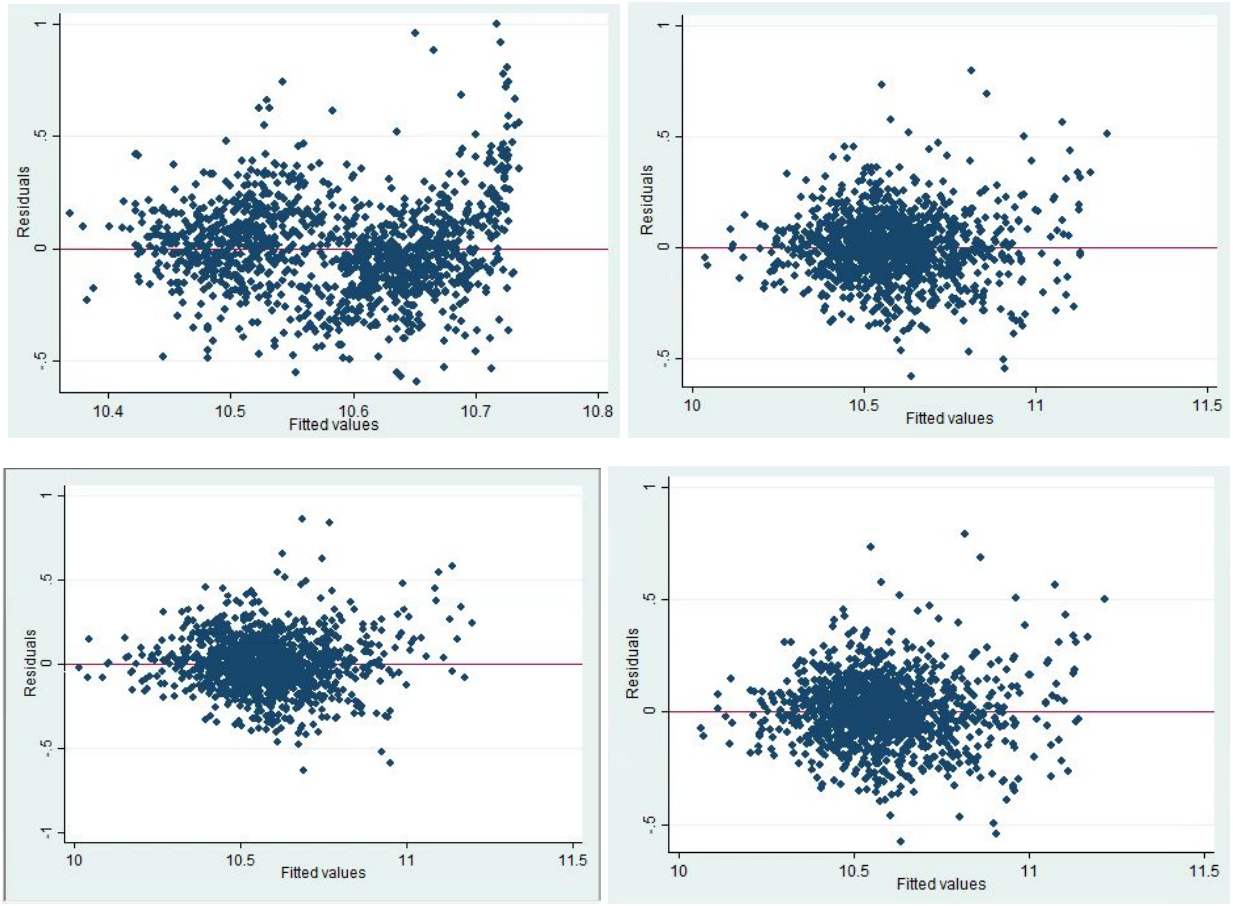

Figure 3.4: Residual graphs for the simple regression (top left), multiple regression 1 (top right), multiple regression 2 (bottom left), and multiple regression 3 (bottom right)

IV. Results

Simple Regression

The simple regression used in this document analyzes the log of the mean income of a student seven years after enrollment against the log of the average cost of attendance per student. Running this regression in STATA, we found the following model:

$$l\_earn = 8.53762 + 0.199041 * (lcost)$$
$$(0.1412514)\ (0.0137286)$$

Adjusted R^2: 0.1373

The R-squared value reveals that the cost on its own does not explain the earnings. The cost does have a positive coefficient however, suggesting a general correlation between the two variables. This is supported by the correlation graph seen in the assumptions section and the graph below. Note though,

that this is a fairly weak positive correlation. As it stands, this simple regression indicates that a 1% increase in the cost of an institution will cause a 0.19904% increase in the average eventual earnings of its students.

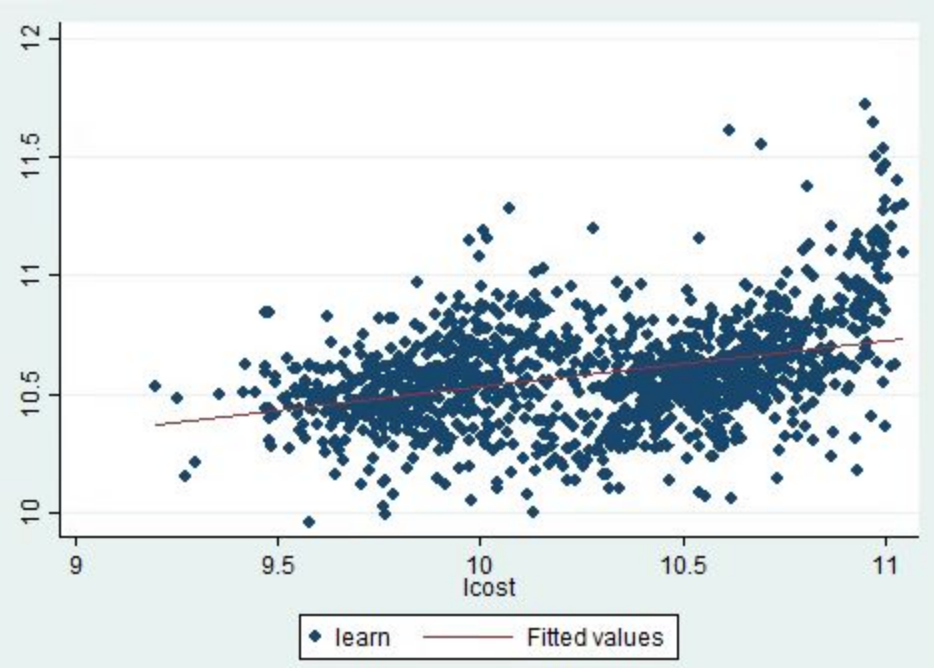

Figure 4.1: A graph of income versus annual education cost

This graph helps to highlight that there is a positive correlation. More importantly however, it helps to demonstrate the wide variance of income for each considered cost, especially at higher cost levels, and explain the extraordinarily high sum of squared error value in the model.

Multiple Regression 1

The first multiple regression in this document adds in several new variables to the regression to help eliminate bias and raise the meaningfulness of the calculated coefficients.

l_earn = 7.722512 + -0.0298125 * private + 0.049877 * lcost + 0.0424658 * lexp + -0.0219198 * c150_4
(0.2765967) (.0174395) (0.0259314) (.0136046) (0.0481947)
+ 0.6964007 * rpy_7yr_rt + 0.0005535 * sat_avg_all + 0.075109 * lfaminc +
(0.1028329) (0.0000607) (0.0262999)
-0.121927 * ugds_white + 0.0493894 * ugds_black + 0.1551053 * ugds_hisp + 0.6939077 * ugds_asian
(0.056542) (0.0590631) (0.0686017) (0.1188667)

Adjusted R^2: 0.5607

The higher Adjusted R-squared value of this model confirms that these added variables have indeed eliminated some bias. Interestingly, the constant coefficient no longer seems to represent the earnings without college education, at least not intuitively. This may be influenced in part by the nature of the variables that have been added. If the loan repayment rate were 0%, for example, then it would seem that that university has actually had an intense negative affect on its students, as well as if there was a low rate of graduation.

It is also interesting to note that the coefficient of the variable costt4_a has decreased exponentially from the simple regression, indicating that, while intuitively it was the best choice to determine and explicit return on investment, it is not as strong an indicator among other variables in determining graduate success as was previously assumed. Other results of note include the coefficient of the variable private, which was actually negative, indicating that going to a private institution actually slightly lowers eventual earnings. This directly contradicts the findings in the paper mentioned above written by Marc Fox (1993). Additionally, the negative coefficient for the variable udgs_white and the increasingly positive coefficients for ugds_black, ugds_hisp, and ugds_asian could indicate that the less diverse an institution is, the less likely it will be to produce higher earning graduates.

Multiple Regression 2

The second multiple regression in this document analyzes the additional impact of student's performance on their later income. The variables private, lcost, c150_4, ugds_white, ugds_black, ugds_hisp, and ugds_asian were removed because they lacked individual significance, which can be seen in Figure 4.2 below. A regression calculation in STATA generated the following model:

$$\begin{array}{l} \text{l\_earn} = 8.513634 + 0.1159641 * \text{lexp} + 0.6922984 * \text{rpy\_7yr\_rt} + 0.0005769 * \text{sat\_avg\_all} + \\ \qquad (0.2086717)\ (0.0124539) \qquad\quad (0.0825405) \qquad\qquad (0.0000589) \\ \qquad\qquad\qquad\qquad -0.0169367 * \text{lfaminc} \\ \qquad\qquad\qquad\qquad (0.0214811) \end{array}$$

Adjusted R^2: 0.4871

The R-squared value dropped significantly with the removal of so many variables. The coefficients all raised from the positive bias nonetheless except, peculiarly, for the SAT variable

coefficient. While the presence of bias is acknowledged, it seems that the change in the R-squared value may be exaggerated by the removal of the demographics data. Given that the F tests reveal the joint significance of the demographics data, the third regression will take these variables into account.

Multiple Regression 3

The third multiple regression analyzes the effect of all significant (joint and independent) variables on the log of eventual income based on a 5% significance level, again reported in Figure 4.2. All race variables were added back into the regressions based on their joint significance, while lcost, private, and c150_4 were all removed. The A regression calculation in STATA generated the following model:

l_earn = 7.958961 + 0.048674 * lexp + 0.6734149 * rpy_7yr_rt + 0.00055 * sat_avg_all +
(0.2095167) (0.0125185) (0.0976378) (0.0000573)
0.0954401 * lfaminc + -0.1380448 * ugds_white + 0.0379168 * ugds_black + 0.1509219 * ugds_hisp +
(0.0218838) (0.0529705) (0.0572998) (0.0672237)
0.7014964 * ugds_asian
(0.1161863)

Adjusted R^2: 0.5605

Not surprisingly, this final regression has an adjusted R^2 value only 0.0002 less than multiple regression one, while possessing 3 less variables. This is expected, as only the most statistically significant variables were included. In support of our hypothesis, lcost is left off of this final regression. A summary of these statistical findings is presented below.

F testing

Because both lcost and private were both only significant at a two-tailed value of 10% in Multiple Regression 1, though they were expected to be significant at all levels based on previous research, an F test was performed to test their joint significance level. The R^2 value of the unrestricted model was 0.5643, while the R^2 of the restricted model was 0.5631. With 2 degrees of freedom and an n-k-1 value of 1103, the F value for these two variables comes out to 1.6455, indicating that the variables are not significant at 5%.

An F test was also performed on ugds_white, ugds_black, ugds_hisp, and ugds_asian because the majority of these race variables were not significant at the highest confidence interval, if they were significant at all. With an unrestricted R^2 value of 0.5631 and a restricted R^2 value of 0.4887, the F value of the four variables was calculated to be 47.1279. This indicates that, while not every race variable is

individually significant, their joint significance is incredibly high. For reference, the critical value for 5% significance for the degrees of freedom in this calculation is 2.37.

Dependent Variable mn_earn_wne_p7				
Independent Variables	**Simple Model**	**Multiple (1)**	**Multiple (2)**	**Multiple (3)**
private	----------	-.0298125* (.0174395) (0.088)	----------	----------
lcost	0.199041*** (0.0137286) (0.00)	.049877* (.0259314) (0.055)	----------	----------
lexp	----------	.0424658*** (.0136046) (0.002)	.1159641*** (.0124539) (0.000)	.048674*** (.0125185) (0.000)
c150_4	----------	-.0219198 (.0481947) (0.649)	----------	----------
rpy_7yr_rt	----------	.6964007*** (.1028329) (0.000)	.6922984*** (.0825405) (0.000)	.6734149*** (.0976378) (0.000)
sat_avg_all	----------	.0005535*** (.0000607) (0.000)	.0005769*** (.0000589) (0.000)	.00055*** (.00055) (0.000)
lfaminc	----------	.075109*** (.0262999) (0.004)	-.0169367 (.0214811) (0.431)	.0954401*** (.0218838) (0.000)
ugds_white	----------	-.121927** (.056542) (0.031)	----------	-.1380448** (.0539705) (0.011)
udgs_black	----------	.0493894 (.0590631) (0.403)	----------	.0379168 (.0572998) (0.508)
ugds_hisp	----------	.1551053** (.0686017) (0.024)	----------	.1509219** (.0672237) (0.025)
ugds_asian	----------	.6939077*** (.1188667) (0.000)	----------	.7014964*** (.1161863) (0.000)
Observations	1316	1316	1316	1316
R-squared	0.1379	0.5607	.4887	.5631

Figure 4.2: Coefficients, standard errors, and p values of coefficients for each regression

The above table shows the significance of each variables in the respective regressions, as well as the standard error (in the first parentheses) and the p value (in the second parentheses). One asterisk represents a two-tailed, 10% significance level, while two represents a two-tailed 5% significance level, and three represents significance at a two-tailed 1%. As shown, Multiple Regression 3 includes all variables either jointly or individually significant at a 5% significance level or smaller, and thus forms the best fit equation.

V. Conclusion

The cost of an institution appears to have little impact on the eventual earnings that students receive after graduating from the institution. The far better heuristic is the measure of the university's instructional expenditure per student. The cost, while significant during the simple regression, becomes insignificant immediately upon adding more variables in the first multiple regression. Moreover, even the extra consideration of the public or private nature of the university, which might theoretically reduce the lack of significance of the cost due to the cost difference between public and private institutions, did not result in joint significance. Thus the first multiple regression validates the hypothesis.

That said, looking at graph 4.1, there remains a correlation between the cost of an institution and the eventual earnings achieved by the attending students, however weak that correlation may be. This likely comes from the higher instructional expenditures generally present at higher cost institution. Since the expenditure has such a high impact on the eventual earnings as seen in the regressions, even this low correlation between cost and expenditure could result in the small correlation observed between earnings and cost.

These observations all support the idea that the marginal returns received from college education are not highly significant. In many cases then, it may make sense to choose a cheaper institution over a more expensive one, despite the increased prestige achieved by attending a more elite institution.

References

Fox, M. (1993). Is it a good investment to attend an elite private college?*Economics of Education Review, 12*(2), 137-151. Retrieved February 18, 2016.

Pascarella, E., Smart, J., & Smylie, M. (1992). College tuition costs and early career socioeconomic achievement: do you get what you pay for?. *High Educ*, *24*(3), 275-290. http://dx.doi.org/10.1007/bf00128447

Kara, O. (2010). Comparing two approaches to the rate of return to investment in education. *Education* Economics, 18(2), 153-165. http://dx.doi.org/10.1080/09645290802416486

9 Non-Ivys That Give Students a Better Bang for Their Buck Than Harvard or Yale. (2015). Mic. Retrieved 14 April 2016, from http://mic.com/articles/116550/9-non-ivy-colleges-that-give-students-a-better-bang-for-their-buck-than-harvard-or-yale#.PNzpdgAxA