Senior Project
Department of Economics

"The Consequence of Institution Category on Post-Graduation Job Market Outcome"

Ryan Bayonnet
May, 2014

Advisor: *Dr. Francesco Renna*

Abstract

This study analyzes the effect on employment outcomes of attending a private university in comparison to a public university. Using a high school student's college choice, which looks at a student's probability of attending a private institution based on personal characteristics. The study is able to estimate selectivity-corrected labor market outcome probabilities. The data I am using for this study comes from two sources. The first from a survey conducted by the National Center for Education Statistics and is titled "The High School and Beyond: Sophomore and Senior Cohort.". The second data source was also conducted by National Center for Education Statistics and is titled "The National Education Longitudinal Study: Base Year through Fourth Follow-Up." After adjusting for the effects from college selection, the study outcome suggests an increase in employment opportunities that results from attending a private educational institution.

Table of Contents

I. Introduction, Hypothesis, Motivation ……4

II. Review of Literature……6

III. Data sources and Descriptions……9

IV. Theoretical Overview & Methodology……10

V. Results ……13

VI. Limitations and Conclusions……15

VII. Appendix ……16

VIII. References……18

IX. SAS Code……19

I. Introduction, Hypothesis and Motivation

According to USA Today (2005) the mean tuition of attending a public university was $3,113 in 2003. At the same point in time, the mean tuition of attending a private university cost $11,447. Over the four year period a student attends either a private or public university, tuition differentiates by $34,866 between the two school types. The gap between public and private universities has been increasing since the 1960's (USA 2005). Average tuition costs of private universities have grown by $20,089 from 1964 to 2006. During this same time, average tuition costs at public universities have increased by only $4,898. This large difference in tuition suggests that employment rates following a private education should increase compared to employment rates following a public university education. Thus far, little research has been done to quantify these changes, as well as compare employment rates from both public and private universities (Iberman 2009). Reviewing historical research, it appears that employment rates of institution types vary. There appears to be little confirmation on the correlation between college quality and employment rates.

This study will review post - baccalaureate employment in public versus private universities. The two have possible differences in education quality that private universities claim have an effect on post graduate earnings/employment (Black 2004). Data sources include the National Longitudinal Study of the High School Class of 1972 and High School and Beyond. Using this data the study looks at the effects on institution type on employment rates of college graduates. Initially a standard linear regression, not adjusted for a student's college selection bias was used, to see if private institutions have a positive effect on employment outcome. As outlined later, this method does not appropriately represent the hypothesis. To compensate for this a linear probability model was used in which the probability of a student selecting a college is determined. After I estimating the predicted probabilities of post-graduate employment for each institution type employment rate adjusted institution type. Comparing

the means of these two probabilities, the study finds the selectivity effected differences in both institutions probability of employment.

Evidence uncovered increased employment rates to attending a private institution, relative to a public school, this figure is found to be lower once controlled for college selection bias. The testable hypothesis is as follows: *The increased tuition of private universities increases the quality of education and allows for a higher employment rate of a student.*

II. Literature Review

In recent years there has been an increased amount of concern about the rising cost of sending children to college. Is the investment of sending children to college still worth the increasing cost? Some have even speculated that private colleges will be restricted to students from privileged backgrounds because of expensive tuition rates. Tuition at private and public universities have increased faster than the Consumer Price Index (CPI) over the last seven years. CPI has increased by 38% where average tuition rates between both private and public universities has increased between 98-102%, Marc Fox (1993). The importance of determining the return on investment from private and public universities has been previously acknowledged by researchers, such as: Marc Fox (1993), Dominic J. Brewer, Eric R. Eide, & Ronald G. Ehrenberg (1999), Dan Black & Jeffrey Smith (2004), Estelle James, Nabeel Alsalam, Joseph C. Conaty, & Duc-Le To (2005), and Scott Iberman (2009).

Ehrenberg (1999) found evidence of differences in labor market premiums of private vs. public classifications of institutions. He also sought to determine whether this higher return applies uniformly to attendees of all types of four-year colleges, or only to those at certain types of institutions. Ehrenberg looked at factors regarding the selectivity of the undergraduate body (SAT scores, GPA, class rank), as well as a universities resource measures (Student to teacher ratio, library size, cost per pupil) to determine a college's quality. This measure plus a school's tuition cost were the two main factors affecting a student's college choice. Once this choice was made, the study looked to see if an institution's public/private classification affected the future earnings of the pupil. The study combined the National Longitudinal Study of the High School Class of 1972, which looked at the characteristics of colleges and the survey "High School and Beyond" looking at individual characteristics of the pupils. This study found that different types of institutions had different future earnings. Elite Private Institutions had the largest labor market premium to attending. Middle-Rated Private Institutions were second with

a large, but smaller than the above premium to attending, and Elite Public Universities was third. It was shown that private colleges of equal level had a larger labor market premium, in all levels.

Kahn (2009) attempted to study labor market experiences of college graduates as a function of economic conditions at time of graduation. Specifically, studying the long term consequences of graduating from college in a bad economy. Kahn used the National Longitudinal Survey of Youth, a survey also conducted by the NLS that had a slightly different focus that the National Longitudinal Study: Fourth follow up. Kahn believed that initial unemployment occurs the effects on long term earnings can be substantial. Stating, "If respondent has initial unemployment, will be less productive and will have long term effects years after graduation." Kahn controls for the selection bias that occurs when a respondent's decision to enter college is affected by labor market conditions at high school graduation. Kahn results find that college unemployment rates have significant negative impact on wages, as a result periods of unemployment have significant effects on future earnings.

Black (2004) addresses two related weaknesses of Ehrenberg (1999): (1) the identification strategy, which looks at the linear outcome model's limits, such as sample selection; (2) While Ehrenberg (1999) shows that, in all six cases, attendance at an elite private institution increases future earnings. The increased wages also came with a sizable wage premium relative to bottom public universities. Black (2004) tries to match equal ability students by looking at their choice of university quality and their earnings after the students graduate. The study finds that there are more high-ability students in low-quality colleges than low-ability students in high-quality colleges. Conditioning on ability makes a large difference on the estimates, reducing the effects of college quality by about 25%.

Fox (1993) attempts to quantify the cost differential between the most elite private colleges and elite public colleges. College is viewed as a human capital investment in educational quality. Fox (1993) attempts to calculate the rate of return on this investment using a sample of recent college graduates and earnings, looking specifically at attending elite private colleges compared to a public college. The data used in this paper is supplied by the 1980 High School and Beyond (HSB) supplement of the NLSY72. This paper studies the rate of return, while comparing the net earning profiles of graduates of elite private colleges and less selective public colleges.

James (2005) to determine if it matters which college you attend and which college characteristics or other aspects of the college experience, lead to a higher value-added to future earnings. This study also uses data from the "National Longitudinal Study of the High School Class of 1972 and High School and Beyond." James (2005) used Ehrenberg's model on students' choice of college type to investigate a few of the determinants of college quality and their effects on undergraduate earnings. These quality determinants included not only whether an institution is public or private, but also at per-pupil expenditures, and the selectivity of acceptance. Private institutions had a positive effect on future earnings, while per-pupil expenditures had no impact. James's study begins to fill the gap on why there may be increased earnings and employment for private institutions.

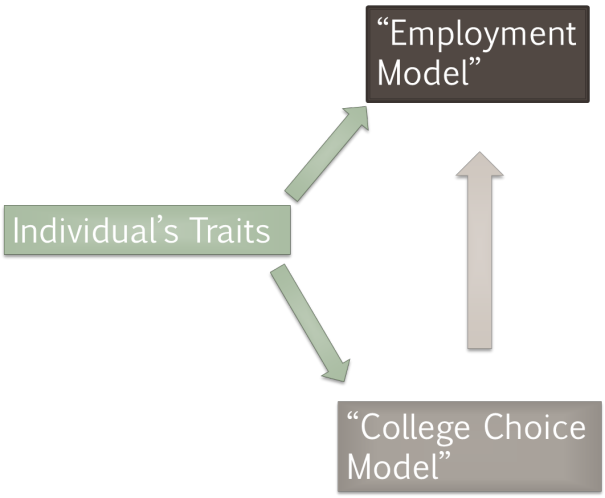
In my study I will try to determine which types of institutions will have the lowest chance of unemployment, this will allow for higher future wages for students who avoid unemployment. This study hopes to fill the gap between Ehrenberg (1999), & Kahn (2009).

III. Data Sources and Descriptions

The data used comes from two separate data sources. Both stem from the National Education Longitudinal Studies (NELS) program of the National Center for Education Statistics, which was created to study educational, vocational, and personal development of young students starting with their elementary or high school years and following them over time into adulthood. This took a look at over 22,000 public high school students and over 15,000 private schools. The NELS program has five major data surveys: the National Longitudinal Study of the High School Class of 1972, High School and Beyond, the National Education Longitudinal Study of 1988, the Education Longitudinal Study of 2002, and the High School Longitudinal Study of 2009. This study uses the High School and Beyond (HS&B) and the National Education Longitudinal Study of 1988. The National Education Longitudinal Study: Base Year through Fourth Follow-Up provides college characteristics, which allow us to determine what factors are specific to a university. This data was collected with a multipurpose study in mind, addressing issues that are relevant students in high school and post high school effects on life (NCES, 2012). The study was followed up four times, annually, until 1988. Results of the study branched into the High School and Beyond data set, a survey of current high school sophomores and seniors in the year of the study. The High School and Beyond: Sophomore and Senior Cohort Third follow up is a study, which supplies the data for the background characteristics of a pupil. The study was designed to elaborate on information obtained from the NLS-72 high school seniors, the sophomores were used to capture information about students prior to high school completion and to include those who might drop out prior to completion (NCES, 2012).

IV. Theoretical Overview & Methodology

Why do students choose to attend higher cost institutions? Is it worth the higher-costs for private institutions if public institutions offer the same quality education? Students hope that their attendance at private college will allow for greater opportunities. There has been prior analysis examining the relationship between college quality and wages or earnings employ the same basic methodology. This study will use this methodology to examine the relationship between college quality and employment rates.



Showing that a respondent's employment is effected by both their individual traits and by the institution type they attend. This also shows that an individual's traits may affect their college choice as well. This will make a traditional OLS regression biased because of selection issues.

Initially a T-Test was able to determine the differences in mean employment rate for students that attend private and public institutions[1], from this it was determined that, uncorrected for college selection, there was a 7.2% higher probability that a student with a private education would be

[1] Appendix – Table 3

employed. My study will determine if this 7.2% increase in employment rates is a result of increased education from private institutions or if that increased probability of employment is a result of the students who attend private institutions.

1. OLS – Linear Regression

The dependent variable "Employed" (Y_i) will be regressed on a set of the respondent's personal characteristics (X_i), with college characteristics (j) for the school that the respondent attended (C_{ij}), and a normally distributed error term(u_i):

$$Y_i = \beta_0 + \beta_1(X_i) + \beta_2(C_{ij}) + (u_i)$$

The measures used for college characteristics are represented by C, many of these use a single variable or a small set of dummy variables. This method could have possible weaknesses because it does not account for the choice of college category in terms of the expected labor market payoff. In the linear regression, there are ten independent variables used[2]. "Standardized Test' which is post-high school graduation test, looked at student's math, science, reading, and social studies abilities. The second variable is the institution classification of the four year degree granting institution that the student attended (Private/Public). The third variable is whether or not the respondent graduated from the four year degree granting university. The fourth, fifth, and sixth variable are about the respondent's gender and race. The next three variables are about the respondent's home life. These include a respondent's family income, which is the immediate number of family members of the respondent. The next variable is family income which is the annual income of the respondent at the time he/she graduated high school. The last two are parental education levels. This is scaled by degree, from not graduating high school, to receiving their doctorate. [3]

[2]Appendix – Table 1
[3]Appendix – Table 2

2. *College Choice*

The second regression used is a switching regression using the equations below. This study uses linear probability models to determine the probability that a student will attend a specific type of university. The first equation I used use is to determine the probability that a respondent will attend a private institution based on personal characteristics. The variables signified as X_i are the same as in the previous OLS, added is θ to represent the tuition of the institution j stating the individual institution.

$$Pr(Z = 1) = \beta_0 + \beta_1(X_i) + \beta_2(\theta_j) + (u_i)$$

The second equation determines the probability that a respondent will attend a public institution based on personal characteristics.

$$\Pr(Z = 0) = \beta_0 + \beta_1(X_i) + \beta_2(\theta_j) + (u_i)$$

3. Job Market Outcome

From the college choice model I am able to create two equations, the first predicting the probability that a respondent who attended a private institution would be employed. With the λ' controlling for the selectivity of a student's college institution type.

$$Y_1 = \beta_0 + \beta_1(X_i) + \gamma\lambda' + (u_i)$$

The second equation in this set will do much the same, predicting the probability that a respondent who attended a public institution would be employed. From this I am able to control for the selection bias of a student's college institution type. $\lambda^o = (1 - \lambda')$

$$Y_2 = \beta_0 + \beta_1(X_i) + \gamma\lambda^o + (u_i)$$

4. Difference

After the predicted probabilities of employment for both private and public institutions, a T-Test I run on the predicted means. This will determine the differences the in mean predicted employment between public and private institutions, controlling for college selection bias.

V. Results

1. OLS

This study will now look at the estimates on the effect of employment by institution type. This is done by using a linear "Employment Model" [4]. Here I try to determine if earnings increase from attending a private college. Looking at the OLS regression, private universities do have a small effect on employment, about a 1% increase in probability of being employed compared to public universities, but this is not statistically significant. The R-Squared for this regression is just above 17%, meaning the model explains 17% of the variation in the data. In this regression it is estimated that males have a 9% lower probability of employment compared to females. The results from the standardized test the students took in high school has a significant impact on employment. A 10 point increase in test score increased the probability of employment by 1.1%. Family size plays a role on income. For each increased member in an immediate family, it is estimated that the respondent will have a 3% lower chance of probability of being in the work force. This is statistically significant. An OLS does not control for the college selection choice that would bias the results. The next model will control for this bias.

2. Switching Regression - Linear Probability Model

From the switching regression used a linear probability model to determine the probability that a student would choose a private or public school based on personal characteristics. From these two predicted equations, I was able to use the equation's betas to determine the probability that students who attend public institutions would employed, and students that attend private institutions would be employed.

[4] Appendix – Table 2

3. Difference

Taking the mean employment rates from both of these institution types, a T-Test was used to determine mean differences between the two. The results show that attending a private college, corrected for college selection; increase the probability that a student would be employed by 3.1%. These results are statistically significant.[5] The 7.2% increase in probability of attending a private institution decreased by 4.1% after controlled for selectivity.

[5] Appendix – Table 4

VI. Limitations and Conclusions

This study presents a relationship between privatized education and post-graduation employment rates. The results of this study are unique from past research, taking a similar model to Ehrenberg (1999), editing his wage model to fit a linear probability model to determine employment rates. There is possible two-way causality (Endogeneity) that is not explained in the OLS model, but using a switching regression this study is able to show differences in employment rates that were controlled for selectivity. Employment rates increase when a student attends a private institution, but there is evidence that students with a higher probability of being employed increasingly attend private institutions. The study determines that students should take into account institution type when determining college choice.

This study could be improved by increased resources, more time, or more in-depth data. To further investigate U.S. regional colleges' differences and possible variations in domestic versus international universities. These future studies could also inquire into the benefits of attending different levels of private institutions, be it high school, graduate schools, or doctoral programs.

VII. Appendix

Table 1 - Descriptive Statistics	N	Mean	Std Dev	Minimum	Maximum
Present Income	8011	10.6399278	0.7749204	4.6051702	13.331408
Employed	9313	0.844003	0.4968811	0	1
Standardized Test	7919	59.3346344	12.0133132	32.52	108.99
Private University	9313	0.3250295	0.468411	0	1
Employed Part-Time	9313	0.1924192	0.3942217	0	1
Graduated Undergraduate	9313	0.4485128	0.4973687	0	1
Male	9313	0.4965498	0.4986559	0	1
Hispanic	9313	0.1625685	0.3689913	0	1
Black	9313	0.0431655	0.2032404	0	1
Family Size	7839	4.5806863	1.299239	2	10
Family Income	7191	68706.09	62632.45	1	300000

Table 2 - OLS	Parameter	Standard	t Value	Pr > |t|
	Estimate	Error		
Intercept	0.61265	0.04156	15.05	<.0001
Standardized Test	0.00111	0.00051200	-2.11	0.0286
Private University	0.01222	0.012221	0.99	0.3223
Male	-0.09104	0.00909	-4.77	<.0001
Hispanic	-0.22336	0.01654	29.65	<.0001
Black	-0.21655	0.02116	6.11	<.0001
Family Size	-0.03122	0.00365	-7.44	<.0001
Family Income	1.31E-07	1.12E-07	-1.58	0.1987

Root MSE	**0.44887**	**R-Square**	**0.1742**	**Error**	**7191**
Dependent Mean	**0.50065**	**Adj R-Sq**	**0.1695**	**Corrected Total**	**7191**
Coeff Var	**89.9954**				

Table 3 - Uncorrected T-Test Difference						
	N	Mean	Std Dev	Std Err	Minimum	Maximum
Public College	6154	**0.8894**	0.4994	0.00456	0	1
Private College	3159	**0.9612**	0.4912	0.00914	0	1
Diff (1-2)		**-0.0718**	0.4944	0.01112		

Method	Variances	DF	t Value	Pr > |t|
Pooled	**Equal**	**8009**	**-2.54**	**0.021**

Table 4 Corrected T-Test Difference						
Corrected	N	**Mean**	Std Dev	Std Err	Minimum	Maximum
Public College	4861	**0.9685**	0.2343	0.00752	0.3657	1.4522
Private College	2330	**0.9994**	0.2321	0.00236	0.4699	1.7654
Diff (1-2)		**-0.0309**	0.2763	0.00963		

Method	Variances	DF	t Value	Pr > |t|
Pooled	**Equal**	**7191**	**-48.32**	**<.0001**

Table 5 Private University Probit				
Parameter - Probit	Estimate	Standard Error	t Value	Approx Pr > |t|
PrivateCol.Intercept	-1.15682	0.122336	-9.46	<.0001
PrivateCol.BY2XCOMP	0.007226	0.001548	4.67	<.0001
PrivateCol.BYS14	0.32949	0.045958	7.17	<.0001
PrivateCol.Parttime	0.022617	0.04849	0.47	0.6409
PrivateCol.Tuitfees	0.069311	0.033975	2.04	0.0413
PrivateCol.F4S12	-0.137474	0.033669	-4.08	<.0001
PrivateCol.Hispanic	-0.116645	0.064051	-1.82	0.0686
PrivateCol.Black	0.051255	0.062776	0.82	0.4142
PrivateCol.BYFAMSIZ	0.021566	0.013296	1.62	0.1048
PrivateCol.FamilyIncome	0.000001035	0	.	.
PrivateCol.GradHSf	0.057078	0.06071	0.94	0.3471
PrivateCol.Juncollf	0.041113	0.073459	0.56	0.5757
PrivateCol.Collegef	-0.028597	0.07568	-0.38	0.7055
PrivateCol.GradCollegef	0.125368	0.062876	1.99	0.0462
PrivateCol.Mastersf	0.076319	0.071485	1.07	0.2857
PrivateCol.PHDf	0.26696	0.081376	3.28	0.001
PrivateCol.GradHSm	-0.11012	0.060084	-1.83	0.0668
PrivateCol.Juncollm	-0.079836	0.070699	-1.13	0.2588
PrivateCol.Collegem	-0.143613	0.07451	-1.93	0.0539
PrivateCol.GradCollegem	0.033297	0.06569	0.51	0.6122
PrivateCol.Mastersm	0.052034	0.075382	0.69	0.49
PrivateCol.PHDm	-0.004488	0.111355	-0.04	0.9679

Table 6 Private University OLS				
Parameter - OLS	Estimate	Standard Error	t Value	Approx Pr > |t|
lnincome.Intercept	10.95355	0.314966	34.78	<.0001
lnincome.BY2XCOMP	0.009911	0.001819	5.45	<.0001
lnincome.Parttime	-0.373303	0.049629	-7.52	<.0001
lnincome.F4S12	0.177618	0.039399	4.51	<.0001
lnincome.Hispanic	-0.125297	0.072002	-1.74	0.0818
lnincome.Black	-0.152893	0.065632	-2.33	0.0198
lnincome.BYFAMSIZ	-0.037866	0.014204	-2.67	0.0077
lnincome.FamilyIncome	-3.19E-07	0.00000034	-0.94	0.348
lnincome.GradHSf	0.009612	0.066597	0.14	0.8852
lnincome.Juncollf	0.122389	0.078882	1.55	0.1208
lnincome.Collegef	0.232959	0.082361	2.83	0.0047
lnincome.GradCollegef	-0.024051	0.06871	-0.35	0.7263
lnincome.Mastersf	0.078688	0.07536	1.04	0.2964
lnincome.PHDf	-0.212511	0.090291	-2.35	0.0186
lnincome.GradHSm	0.038442	0.065834	0.58	0.5593
lnincome.Juncollm	-0.12925	0.075232	-1.72	0.0858
lnincome.Collegem	-0.036775	0.080271	-0.46	0.6469
lnincome.GradCollegem	-0.036659	0.068345	-0.54	0.5917
lnincome.Mastersm	-0.0921	0.077211	-1.19	0.2329
lnincome.PHDm	-0.153668	0.111142	-1.38	0.1668
_lnincome.lambda	-0.534945	0.182865	-2.93	0.0034

Table 7 Public University Probit				
Parameter - Probit	Estimate	Standard Error	t Value	Approx Pr > |t|
PrivateCol.Intercept	-1.121967	0.122453	-9.16	<.0001
PrivateCol.BY2XCOMP	0.007234	0.00155	4.67	<.0001
PrivateCol.BYS14	0.32682	0.045937	7.11	<.0001
PrivateCol.Parttime	0.029122	0.048519	0.6	0.5484
PrivateCol.Tuitfees	0.0494	0.033975	1.45	0.1459
PrivateCol.BYS12	-0.147779	0.033669	-4.39	<.0001
PrivateCol.Hispanic	-0.093643	0.063906	-1.47	0.1428
PrivateCol.Black	0.082482	0.062496	1.32	0.1869
PrivateCol.BYFAMSIZ	0.020319	0.013306	1.53	0.1268
PrivateCol.FamilyIncome	0.000000978	0	.	.
PrivateCol.GradHSf	0.054222	0.060532	0.9	0.3704
PrivateCol.Juncollf	0.030148	0.073397	0.41	0.6813
PrivateCol.Collegef	-0.032325	0.075491	-0.43	0.6685
PrivateCol.GradCollegef	0.128291	0.062823	2.04	0.0411
PrivateCol.Mastersf	0.080599	0.071494	1.13	0.2596
PrivateCol.PHDf	0.27067	0.081437	3.32	0.0009
PrivateCol.GradHSm	-0.112693	0.059974	-1.88	0.0602
PrivateCol.Juncollm	-0.090133	0.070628	-1.28	0.2019
PrivateCol.Collegem	-0.158329	0.074474	-2.13	0.0335
PrivateCol.GradCollegem	0.032892	0.06563	0.5	0.6162
PrivateCol.Mastersm	0.039398	0.075399	0.52	0.6013
PrivateCol.PHDm	-0.011649	0.111682	-0.1	0.9169

Table 8 Public University OLS				
Parameter - OLS	Estimate	Standard Error	t Value	Approx Pr > |t|
lnincome.Intercept	10.930219	0.257906	42.38	<.0001
lnincome.BY2XCOMP	0.009781	0.001425	6.86	<.0001
lnincome.Parttime	-0.296046	0.036609	-8.09	<.0001
lnincome.BYS12	0.150505	0.029913	5.03	<.0001
lnincome.Hispanic	-0.184929	0.047691	-3.88	0.0001
lnincome.Black	-0.122355	0.046959	-2.61	0.0092
lnincome.BYFAMSIZ	-0.01712	0.010176	-1.68	0.0925
lnincome.FamilyIncome	-0.000000286	0.000000262	-1.09	0.2745
lnincome.GradHSf	0.014073	0.044082	0.32	0.7495
lnincome.Juncollf	0.09429	0.0536	1.76	0.0786
lnincome.Collegef	0.147356	0.054527	2.7	0.0069
lnincome.GradCollegef	0.033361	0.048932	0.68	0.4954
lnincome.Mastersf	0.065885	0.054096	1.22	0.2232
lnincome.PHDf	-0.331729	0.070423	-4.71	<.0001
lnincome.GradHSm	0.054251	0.045526	1.19	0.2334
lnincome.Juncollm	0.04416	0.053004	0.83	0.4048
lnincome.Collegem	0.065496	0.056813	1.15	0.249
lnincome.GradCollegem	0.036012	0.049284	0.73	0.465
lnincome.Mastersm	0.053998	0.057274	0.94	0.3458
lnincome.PHDm	0.059285	0.085356	0.69	0.4873
_lnincome.lambda	-0.651243	0.152423	-4.27	<.0001

VIII. References

Dominic J. Brewer, Eric R. Eide, Ronald G. Ehrenberg. "Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings." *The University of Wisconsin Press Journals Division,* December, 1999, *Vol. 34* (I), pp.104-123.

Estelle James; Nabeel Alsalam; Joseph C. Conaty, Duc-Le To. "College Quality and Future Earnings: Where Should You Send Your Child to College?" *The American Economic Review,* January, 2005, *Vol. 79* (II), pp. 247-252.

Scott Iberman. "Are there Returns to Attending a Private College or University?" *University of Houston,* February, 2009.

Dan Black, Jeffrey Smith. "How Robust is the Evidence on the Effects of College Quality? Evidence from Matching." *Journal of Economics*, March, 2004, *Vol 121*, pp. 99-124.

Marc Fox, "Is it a Good Investment to Attend an Elite Private College?" *Department of Economics, Brooklyn College of the City University of New York*, April, 1993, *Vol. 12*, (II) pp. 137-151.

Lisa B. Kahn. "The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy". *Yale School of Management*, October 2007, Vol. 17, pp. 303-316.

National Center for Educational Statistics, "Overview" US Department of Education, July, 2012. "https://nces.ed.gov/surveys/hsb/"

IX. SAS Code

```
data one;
        set SP1;
proc sort;
        by stu_id;
run;

data two;
        set SP2;
proc sort DATA = TWO;
        by incode;
run;

data three;
        set SP3;
proc sort DATA = THREE;
        by stu_id;
run;

data seniorproject1;
        merge one three;
                by Stu_id;
run;

data four;
        set SP4;
proc sort;
        by incode;
run;
data seniorproject2;
merge two four;
by incode;
run;
data seniorproject3;
set seniorproject2;
proc sort data = seniorproject1;
by Stu_id;
run;
proc sort data = seniorproject3;
by Stu_id;
run;
data seniorprojectall;
merge seniorproject1 seniorproject3;
by stu_id;
run;

data FINAL;
set seniorprojectall;
```

```
if f4hi99 < 1 then f4hi99=.;
lnincome=log(f4hi99);

if F4S12 >2 then F4S12=.;
if F4S12=2 then F4S12=0;

if bys14 >2 then bys14=.;
if bys14=1 then bys14=0;
if bys14=2 then bys14=1;

if byfamsiz=99 then byfamsiz =.;
if byfamsiz=98 then byfamsiz =.;

if f4sect=5 then PrivateCol = 1;
        if  f4sect=6 then PrivateCol = 0;

select(f4sect);
        when (-1, 2, 3, 4) Privatecol = .;
        otherwise;
end;

if bys14= 2 then PrivateHS= 1;
if bys14= 3 then PrivateHS= 1;
/*else PrivateHS =0;*/

if BYS31A =1 then API = 1; else API= 0;

if BYS31A = 2 then Hispanic = 1; else Hispanic=0;
if BYS31A = 3 then Black = 1; else Black=0;

if BYS34A = 1 then notfinf =1; else Notfinf=0;
if BYS34A = 2 then GradHSf =1; else GradHSf=0;
if BYS34A = 3 then Juncollf=1; else JunCollf=0;
if BYS34A=4 then Collegef = 1; else Collegef=0;
if BYS34A =5 then GradCollegef =1; else GradCollegef=0;
if BYS34A=6 then Mastersf = 1; else Mastersf=0;
if BYS34A=7 then PHDf = 1; else PHDf=0;

if BYS34B = 1 then notfinm =1; else Notfinm =0;
if BYS34B = 2 then GradHSm =1; else GradHSm =0;
if BYS34B= 3 then Juncollm=1; else JunCollm =0;
if BYS34B=4 then Collegem = 1; else Collegem=0;
if BYS34B =5 then GradCollegem =1; else GradCollegem =0;
if BYS34B=6 then Mastersm = 1; else Mastersm=0;
if BYS34B=7 then PHDm = 1; else PHDm=0;

if F2P11a = 1 then Parttime = 1; else parttime=0;
```

```
if F2p11a = 1 then In_Labor_Force = 1;
if F2p11a = 2 then In_Labor_Force = 1; else In_Labor_force = 0;
lnIn_Labor_Force=log(In_Labor_Force);

if F4TYPEDG=1 then NoDG =1; else NoDG=0;
if F4TYPEDG=2 then Cert =1; else Cert=0;
if F4TYPEDG=3 then Associate =1; else Associate=0;
if F4TYPEDG=4 then Bachelor =1; else Bachelor=0;
if byfaminc=1 then FamilyIncome=1;
if byfaminc=2 then FamilyIncome =1000;
if byfaminc=3 then FamilyIncome =3000;
if byfaminc=4 then FamilyIncome =5000;
if byfaminc=5 then FamilyIncome =7500;
if byfaminc=6 then FamilyIncome =10000;
if byfaminc=7 then FamilyIncome =15000;
if byfaminc=8 then FamilyIncome =20000;
if byfaminc=9 then FamilyIncome =25000;
if byfaminc=10 then FamilyIncome =35000;
if byfaminc=11 then FamilyIncome =50000;
if byfaminc=12 then FamilyIncome =75000;
if byfaminc=13 then FamilyIncome =100000;
if byfaminc=14 then FamilyIncome =200000;
if byfaminc=15 then FamilyIncome =300000;
if BYS34A=1 then notfinf=1; else notfinf=0;
if BYS34A=2 then Gradhsf=1; else Gradhsf=0;
if BYS34A=3 then Juncollf=1; else Juncollf=0;
if BYS34A=4 then Collegef=1; else Collegef=0;
if BYS34A=5 then GradCOllegef=1; else GradCOllegef=0;
if BYS34A=6 then Mastersf=1; else Mastersf=0;
if BYS34A=7 then PHDf=1; else PHDf=0;
if BYS34B=1 then notfinm=1; else notfinm=0;
if BYS34B=2 then Gradhsm=1; else Gradhsm=0;
if BYS34B=3 then Juncollm=1; else Juncollm=0;
if BYS34B=4 then Collegem=1; else Collegem=0;
if BYS34B=5 then GradCOllegem=1; else GradCOllegem=0;
if BYS34B=6 then Mastersm=1; else Mastersm=0;
if BYS34B=7 then PHDm=1; else PHDm=0;
if f4typedg= 4 then Undergraduate=1; else Undergraduate=0;
if f4typedg =9 then Graduate =1; else Graduate = 0;
if f4typedg=10 then Graduate =1; else Graduate = 0;

if by2xcomp <0 then by2xcomp= .;
if by2xcomp =-9 then by2xcomp= .;

label
by2xcomp= 'Standardized Test'
lnincome= 'Logged Income'
Undergraduate= 'Graduated Undergraduate'
```

```
lnincome= 'Present Income'
PrivateCol='Private University'
F4S12='Male'
Hispanic='Hispanic'
Black='Black'
byfamsiz= 'Family Size'
byfaminc='Family Income'
bys34a= 'Fathers Education'
bys34b= 'Mothers Education'
bachelor = 'Bachelors Degree'
tuitfees='College Tuition'
totattnd= 'Colleges Attendance'
bys14='Private High School'
In_Labor_Force = 'In Labor Force'
NoInc='Family Income of Zero'
famincLess1000 ='Family Income Less Than $1,000'
famincLess3000 ='Family Income $1,000-$2,999'
famincLess5000 ='Family Income $3,000-$4,999'
famincLess7500 ='Family Income $5,000-$7,499'
famincLess10000 ='Family Income $7,500-$9,999'
famincLess15000 ='Family Income $10,000-$14,999'
famincLess20000 ='Family Income $15,000-$19,999'
famincLess25000 ='Family Income $20,000-$24,999'
famincLess35000 ='Family Income $25,000-$34,999'
famincLess50000 ='Family Income $36,000-$49,999'
famincLess75000 ='Family Income $50,000-$74,999'
famincLess100000 ='Family Income $75,000-$99,999'
famincLess200000 ='Family Income $100,000-$199,999'
famincMore200000 ='Family Income $200,000 or More'
FamilyIncome='Family Income'
notfinf= 'Father Did not Finish High School'
Gradhsf= 'Father Graduated High School'
Juncollf= 'Father Junior College Graduate'
Collegef = 'Father College Attandee'
GradCOllegef='Father Graduated College'
Mastersf= 'Father Recieved Masters'
PHDf='Father Recieved PHD'
notfinm= 'Mother Did not Finish High School'
Gradhsm= 'Mother Graduated High School'
Juncollm= 'Mother Junior College Graduate'
Collegem = 'Mother College Attandee'
GradCOllegem='Mother Graduated College'
Mastersm= 'Mother Recieved Masters'
PHDm='Mother Recieved PHD'
parttime = 'Employed Part-Time';
run;

DATA FINAL;
```

```
        SET FINAL;
                IF PRIVATECOL = . THEN DELETE;
RUN;

proc means;
        var lnincome In_Labor_Force by2xcomp PrivateCol parttime bys14 Undergraduate F4S12
hispanic black BYFAMSIZ
                FamilyIncome Gradhsf Juncollf Collegef GradCOllegef Mastersf PHDf Gradhsm
                Juncollm Collegem GradCOllegem Mastersm PHDm;
run;

proc reg data = final;
        model In_Labor_Force = by2xcomp PrivateCol  Undergraduate F4S12 hispanic black
                BYFAMSIZ FamilyIncome Gradhsf Juncollf Collegef GradCOllegef Mastersf PHDf Gradhsm
                Juncollm Collegem GradCOllegem Mastersm PHDm;
run;


PROC QLIM DATA = Final heckit;
   PROB: MODEL PrivateCol = by2xcomp  tuitfees F4S12 hispanic black BYFAMSIZ FamilyIncome
        Gradhsf Juncollf Collegef GradCOllegef Mastersf PHDf Gradhsm
                Juncollm Collegem GradCOllegem Mastersm PHDm  / DISCRETE (D=PROBIT);
   PROB:  MODEL In_Labor_Force = by2xcomp   F4S12 hispanic black BYFAMSIZ FamilyIncome Gradhsf
                Juncollf Collegef GradCOllegef Mastersf PHDf Gradhsm
                Juncollm Collegem GradCOllegem Mastersm PHDm  / SELECT(PrivateCol=1);
                Output OUT=PRED1 predicted MILLS;
RUN;

PROC QLIM DATA = Final heckit;
   PROB: MODEL PrivateCol = by2xcomp  tuitfees F4S12 hispanic black BYFAMSIZ FamilyIncome G
                radhsf Juncollf Collegef GradCOllegef Mastersf PHDf Gradhsm
                Juncollm Collegem GradCOllegem Mastersm PHDm  / DISCRETE (D=PROBIT);
   PROB:  MODEL In_Labor_Force= by2xcomp   F4S12 hispanic black BYFAMSIZ FamilyIncome Gradhsf
                Juncollf Collegef GradCOllegef Mastersf PHDf Gradhsm
                Juncollm Collegem GradCOllegem Mastersm PHDm / SELECT(PrivateCol=0);
                Output out = PRED2  predicted MILLS;
RUN;

DATA PRED1 (RENAME=(P_In_Labor_Force=PS1 Mills_PrivateCol=IMR1));
        SET PRED1;
RUN;
PROC SORT DATA = PRED1;
        BY STU_ID;
RUN;
DATA PRED2 (RENAME=(P_In_Labor_Force=PS0 Mills_PrivateCol=IMR2));
        SET PRED2;
RUN;
PROC SORT DATA = PRED2;
```

```
        BY STU_ID;
RUN;
DATA PRED;
        MERGE PRED1 PRED2;
                BY STU_ID;
RUN;
DATA PRED;
        SET PRED;
                P_Employed=.;
                        IF PRIVATECOL = 1 THEN P_Employed = (PS1 + IMR1*(-0.123280));
                        IF PRIVATECOL = 0 THEN P_Employed = (PS0 + IMR2*(-0.686235));
RUN;

PROC MEANS DATA = PRED;
        VAR PS0 PS1 P_Employed;
RUN;

PROC TTEST DATA = PRED;
        CLASS PRIVATECOL;
        VAR P_Employed In_Labor_force;
RUN;
```