

EDU A206 Problem Set 2

March 6, 2020

0.0.1 Problem Set 2

Made By: Eric Englin, Eric Berube, Gabriella Pierre

0.0.2 Problem #1

A) Arteaga sought to find out if the return to college education lies within the human capital method, where going to more school and more classes would have impact on income and skill, or if it was part of the signaling method which would show that the people that got into Los Andes would have made that income whether they had extra classes or not.

B) In this study Arteaga uses the difference-in-difference approach to compare the wages earned by Los Andes graduates in the years before and after the policy change with other similar top 10 school graduates and their earned wages. Since the graduates from all schools were applying for the same jobs in the same sectors the only change in hiring trends and wages should be a result of the Los Andes curriculum change. She found that after the change the Los Andes graduates did worse on recruitment exams and were hired up to 17% less than before the change.

Equation: $wage_{it} = \beta_0 + \beta_1 Andes_i * Post_t + \beta_2 Andes_i + \beta_3 Post_t + \beta_4 experience_{i,t} + \epsilon_{it}$ **C)**

Variable meanings

1. $wage_{it}$: wages earned
2. $Andes$: flag to designate if person went to Universidad de los Andes
3. $Post$: lag that this is in the post-curriculum change period
4. $experience$: years since graduation
5. ϵ : error term

D) Beta 1 on Andes in the Post period because we want to see how the curriculum change affects wages. This change is only seen in post-andes beta1.

E) The human capital model is strongly supported. She found that after the change the Los Andes graduates did worse on recruitment exams and were hired up to 17% less than before the change.

F) Regression discontinuity uses arbitrary cutoffs by looking at individuals or samples right above and below this cutoff to find the average treatment effect. This has been used to find the effect of certain colleges that have a cutoff for test scores or GPAs. Thistlethwaite used this design to understand the effect of the National Merit Scholarship program by looking at students who were around the cutoff for receiving and not receiving the award.

0.0.3 Problem 2

A) Read the paper!

B) This is measured in 3 ways:

1. **SAT average for admissions:** This can be a fairly problematic indicator because it is going to favor richer and less diverse students, which biases our quality indicator. It would be beneficial if they had some measure to indicate the output, rather than inputs like test scores.

2. **Barron's index:** This may be a better indicator for signaling, but it seems fairly good for determining quality of colleges at a national level.

3. **Net tuition cost:** Not a good measure because many expensive schools cater to a different kind of experience that may be based around learning, but can also be based around having a quality "life experience" or lifestyle while on campus.

C) Beta = Holding quality measure 2 constant, every 1 point increase in quality measure 1 is related to a beta increase in log earnings. Theta is the same, but holding beta constant and using quality measure 2.

D) The major omitted variable is based around parental income. These incomes were missing in many cases, so they built a regression to engineer features for the parental income using occupation & education data available for student families. You definitely want to account for this, so its a good solution to do this procedure to minimize error rate. However, more accurate and consistent data around income would always be helpful. If the data is not available in any capacity, I would think they could predict family income a bit better if they also factored in location and demographic data along with occupation and education levels.

```
In [3]: #import library
import pandas as pd

#import data
df = pd.read_excel(r"./data/PS2_tables.xlsx",sheet_name="Problem2")

#create empty datafields for confidence intervals
df['CI']=" "
df['CI_noHBCU']=" "

#loop through each measure (SAT, tuition, Barron) and method (Basic, Self-revelation)
for x in df.index:
    change = round(2*df['StDev'][x],2) #twice the standard deviation
    coef = df['Coefficient'][x] # field coefficient
    df['CI'][x]='['+str(round(coef-change,3))+', '+(str(round(coef+change,3)))+']'

    change2 = round(2*df['StDev_noHBCU'][x],2) #twice the standard deviation
    coef2 = df['Coefficient_noHBCU'][x] # field coefficient
    df['CI_noHBCU'][x]='['+str(round(coef2-change2,3))+', '
                        +(str(round(coef2+change2,3)))+']'
```

Table for all black & hispanic students:

Field	Coefficient	StDev	CI
Basic* (SAT Score)	0.067	0.019	[0.027, 0.107]
Self-Revelation* (SAT Score)	0.076	0.032	[0.016, 0.136]
Basic* (Net Tuition)	0.173	0.056	[0.063, 0.283]
Self-Revelation (Net Tuition)	0.138	0.071	[-0.002, 0.278]
Basic* (Barron's Index)	0.063	0.022	[0.023, 0.103]
Self-Revelation (Barron's Index)	0.049	0.036	[-0.021, 0.119]

Table for all black & hispanic students, excluding HBCUs:

Field	Coefficient	StDev	CI
Basic* (SAT Score)	0.122	0.030	[0.062, 0.182]
Self-Revelation* (SAT Score)	0.120	0.042	[0.04, 0.2]
Basic* (Net Tuition)	0.187	0.064	[0.057, 0.317]
Self-Revelation (Net Tuition)	0.166	0.079	[0.006, 0.326]
Basic* (Barron's Index)	0.158	0.040	[0.078, 0.238]
Self-Revelation (Barron's Index)	0.143	0.053	[0.033, 0.253]

F) Once you look at the self-revelation model, the college effect on student wages becomes statistically similar to 0. However, this trend isn't seen among Black and Hispanic students excluding HBCUs. Among these students, the wages increase by all metrics used.

G) This paper shows that a higher quality college has a higher impact on black and hispanic students future wages, so these colleges, including Harvard, should take this as an opportunity to help increase future wages for these groups.

0.0.4 Problem #3

A) College educated workers are affected because there are spillover effects and complementary jobs needed to help the higher skill workers (example: assistant makes the business person more effective/productive). College educated workers help other high-skill workers when the spillover effect is greater than the decrease in private returns due to more competition in the labor market.

B) Our spillovers:

- **Eric Berube:** When I was in the Army there was a noticeable spillover in terms of jobs created when I was assigned to different units. I went to certain schools that resulted in certifications and authorization to perform specific roles as an officer. By holding those certifications my unit was allowed to assign more soldiers (high school educated) to work underneath me in key positions. By getting those extra soldiers that freed up work hours from other people who had been covering those responsibilities. There was a compound effect from my arrival to work efficiency and positions available. Having another manager also meant that my boss (college educated) could focus more on their specific tasks instead of using their work week to supervise the other people in the unit.

- **Gabriella Pierre:** As a soon to be worker in an educational non-profit committed to advancing equity for low-income, minoritized students, the spillovers I create on workers with less than a college degree are increased access to opportunities regardless of their educational background as

I advocate for their inclusion in the work that we do. For college-educated workers, I am causing those individuals who may work with me, to be more productive and perform better, increasing overall team success.

- **Eric Englin:** Before coming to grad school, I worked with homeless service providers to help get people into housing. One big part of my job was communicating to legislators that we were spending money effectively, so I think that a potential spillover would be if I did a good job in communicating that we needed more money to end homelessness, this could pave the way for more funding to all positions of people working in the same area. It would effectively increase the size of the pot for everyone.

C)

This is the table showing the current working population and the future working population assuming that this program increases the college educated group by 1%, from 291 at 31% to 298 at 32%.

Education Level	Total (thousands)	% of total	Total Under New Program (thousands)	% of total
Less than HS	81	9%	81	9%
HS Grad	309	33%	309	33%
Some College	250	27%	250	27%
College	291	31%	298	32%
Total	931		938	

NPV Calculation:

- Increased Income: \$ 7,000\$ people x55,640 (Annual Wages) = 389,480,000
- Program Cost: -\$70,000,000
- $NPV = -\$70,000,000 + 389,480,000 + \frac{389,480,000}{(1+0.03)^1} + \frac{389,480,000}{(1+0.03)^2} + \frac{389,480,000}{(1+0.03)^3}$
- $NPV = -70,000,000 + 389,480,000 + 378,135,922 + 367,122,255$
- $NPV = \$1,064,738,177$

D) An instrumental variable is needed to determine the causal effect of this policy. This means that we need to show that the social returns are increased solely due to the 93 additional college educated workers coming into Maine. There are many unobservable factors that could skew our effect and instrumental variables help with this.

E) We need to find an instrumental variable that is relevant (i.e. correlated with the outcome of social return to having more college educated workers) and exclusionary (i.e. uncorrelated with omitted variables).

F) To control for sources of potential bias, Moretti used Census data to produce two instrumental variables: (1) the lagged city demographic structure and (2) the presence of a land-grant college. We would use the land-grant cities as an instrumental variable because Maine has 2 land-grant universities. This would provide a fairly interpretable outcome that could be used to explain the gains from the policy to a wide audience group.

G) For our analysis, we are looking at the land grant analyses and will base our predictions off of the 1990 findings. With the increases in years 2 & 3, we found a total NPV increase in salaries to be **\$579,314,072.60**. These are shown in the table & calculations below:

Education Level	Percent Increase	Starting Salary	Year 2 & 3 Salary	Salary Difference	Total NPV per Person	Total People	Total NPV
Less than HS	0.77%	\$26,350	\$26,552	\$202.90	\$376.93	81,000	\$30,530,978
HS Grad	0.84%	\$37,120	\$37,431	\$311.81	\$579.26	309,000	\$178,990,547
Some College	0.94%	\$46,810	\$47,250	\$440.01	\$817.43	250,000	\$204,357,634
College	0.55%	\$55,640	\$55,946	\$306.03	\$568.50	291,000	\$165,434,911
Spillover Total	-	-	-	-	\$2,342.12	931,000	\$579,314,072

Calculations for NPV

- $Wages_{starting} * PercentIncrease = Wages_{Year2} = Wages_{Year3}$
- $Wages_{Year2or3} - Wages_{starting} = Wages_{difference}$
- $NPV_{Person} = \frac{Wages_{difference}}{(1+r)^2} + \frac{Wages_{difference}}{(1+r)^3}$

– Note: $Wages_{difference} = 0$ in year 1 because effect of program isn't seen until year 2

- $NPV_{total} = NPV_{perperson} * Workers$

Calculations shown for Less than HS:

- Percent increase taken from Moretti (2002 paper) as 0.77 for every 1 percentage point increase in college educated workers
 - $\$26,350 * 1.0077 = \$26,552$
 - $\$26,552 - \$26,350 = \$202.90$
 - $NPV = \frac{\$202.90}{(1+0.03)^2} + \frac{\$202.90}{(1+0.03)^3}$
 - $NPV = \$191.25 + \185.68
 - $NPV = \$376.93$ per person
 - $NPV_{total} = \$376.93 \text{ per person} * 81,000 \text{ people} = \$30,530,978.78$

H) Shown in the table below, the tax revenue from the social spillover among all existing groups is \$40,551,985.08. However, an additional \$104,892.07 will be gained through tax revenue from the 7,000 new college educated workers (for simplicity, our assumption is that these workers would not see their wages don't increase over time). This would bring the total to \$91,949,100.74. From all of this income, it would likely pay for itself as a program.

Education Level	Year 2 & 3 Salary Difference	Total NPV per Person	Total People	Total NPV	Tax Revenue
Less than HS	\$202.90	\$376.93	81,000	\$30,530,978	\$2,137,168
HS Grad	\$311.81	\$579.26	309,000	\$178,990,547	\$12,529,338
Some College	\$440.01	\$817.43	250,000	\$204,357,634	\$14,305,034
College	\$306.03	\$568.50	291,000	\$165,434,911	\$11,580,443
Spillover Total	-	\$2,342.12	931,000	\$579,314,072	\$40,551,985
-	-	-	-	-	-
New College Workers	\$55,640	\$104,892.07	7,000	\$734,244,509	\$51,397,115
Spillover Plus New Total	-	\$107,234.19	938,000	\$1,313,558,581	\$91,949,100

I) The tax revenue from the social spillover among all existing groups is \$40,551,985.08. However, an additional \$51,397,115.66 will be gained through tax revenue from the 9,300 new college educated workers (assuming their wages don't increase over time). This would bring the total to \$91,949,100.74. If the total program cost is \$10,000 per person, this would be a \$70 million program that would have a net benefit of \$21,949,100.74.

J) Wages will increase the most for less educated workers, which would decrease wage inequality.