11/05/2018

**BAHCESEHIR UNIVERSITY ECONOMETRICS-II**

**ECO3062 – TUBA TORU DELIBASI**

**Relationship Between University Graduates And Their Wages in UK**

* Since Oxford and Cambridge UK were among the most prestigious schools in the world, we wanted to investigate this nation first.
* As the change in inflation and unemployment rates do not change much from year to year, we can compare data more easily with the ceteris paribus idea.
* Like the case for the other countries; when you have more education, your wages will increase relatively. Thats why we analysed the different education level and their effects on wages.
* Also we wanted to see the difference between in the same group but divided for male and female.
* For this reasons, this case and the country became our first choice.

**Review and Data**

* **We used in our surveys time range from 2010 to 2015 for UK.**
* **Although we cant see the survey data purely because we could found as a report, we used those Wage reports from ;**
* [http://stats.oecd.org/viewhtml.aspx?datasetcode=EAG\_EARNINGS&lang=en#](http://stats.oecd.org/viewhtml.aspx?datasetcode=EAG_EARNINGS&lang=en)
* **Also for Macro Datas we get from;**
* [http://hdr.undp.org/en/data#](http://hdr.undp.org/en/data)
* **In studies, ISCED standard was followed in education levels**
* **We got Unemployment, human development index, export rate data from UNITED NATIONS DEVELOPMENT PROGRAM**.
* **The subgroups: We separated our relative wage data by; age, year, sex and education level. With these subgroups, we have increased our number of observations. We wanted to create panel data in terms of comparison and we generated identifiers from age sex and education levels.**
* **With Excel, we corrected the format, combined the data and created our analysis data.**

**Definitions**

* L2 Under high school
* L3 High School
* L5 Associate degree
* L6 License
* L7T8 Master and Doctorate

**Model Reasoning**

* **Cross Section Data is more appropriate. Because we can put more variables into the regression by evaluating each row of 72 rows of data independently.**
* **We can not do the time series because we only have 6 years observations.**
* **We can not use panel data because if the number of periods is not too high, and if we eliminate the effects of the identifiers that we generate from the fractions in our hand, we will essentially destroy the values we are looking for.**
* **In the first examples we will show this. If we could translate the panel data, we could get meaningful results from the macro variables. Since our macro values are copied to all rows, R can not obtain sufficient variance from 6 values and can not return meaningful reliable coefficients to us.**

**Problems**

* **We can’t see the survey data purely and just the reports that publishes the average values were enabled and open to the public.**
* **The other problem is our time range. England had just 6 comparable census on our topic. Having only 6 years is not useful neither in panel nor in time data.**
* **So that our number of observations are;72.**

**Theory**

* It is a general opinion that university education brings a better future with higher salaries.
* That is one of the reason, one may pay many pounds to educational institutions in Britain.
* At the same time, gender disparity is a global problem apperas everywhere with different levels and kinds.
* In our research we wilt try to check their validity for England

**Results**

* First 3 regression we made with the panel data is not meaningful because it was insignificant.
* So thats why we used cross section in next steps

library(readxl)

library(plm)

dat = read\_excel(file.choose())

pdat = pdata.frame(dat)

**R Results Reg 1**

reg1 = plm(re ~ unem, data = pdat, model = "fd")

summary(reg)

# Coefficients:

# Estimate Std. Error t-value Pr(>|t|)

# unem 0.42182 1.63885 0.2574 0.7979

# Total Sum of Squares: 6479

# Residual Sum of Squares: 6470.8

# R-Squared: 0.0051996

# Adj. R-Squared: 0.0051996

# F-statistic: Inf on 0 and 51 DF, p-value: NA

**R Results Reg 2**

reg2 = plm(re ~ exp + hdi + sex, data = pdat, model = "fd")

summary(reg2)

# Coefficients:

# Estimate Std. Error t-value Pr(>|t|)

# exp -0.50083 1.31195 -0.3817 0.7043

# hdi -40.62177 578.71452 -0.0702 0.9443

# Total Sum of Squares: 6479

# Residual Sum of Squares: 6456.3

# R-Squared: 0.0037795

# Adj. R-Squared: -0.016145

# F-statistic: 0.175998 on 1 and 50 DF, p-value: 0.67663

**R Results Reg 3**

reg3 = plm(re ~ year + sex + age, data = pdat, model = "fd")

summary(reg3)

# Coefficients:

# Estimate Std. Error t-value Pr(>|t|)

# year2011 0.29775 5.68389 0.0524 0.9584

# year2012 -4.75300 8.03824 -0.5913 0.5572

# year2013 2.83675 9.84479 0.2881 0.7745

# year2014 1.29260 10.16765 0.1271 0.8994

# year2015 2.41130 10.48058 0.2301 0.8190

# Total Sum of Squares: 6479

# Residual Sum of Squares: 6073.6

# R-Squared: 0.062567

# Adj. R-Squared: -0.017214

# F-statistic: 0.784234 on 4 and 47 DF, p-value: 0.54119

**R Results Reg 4**

reg4 = lm(re ~ educ\*age, data = dat)

summary(reg12)

# Coefficients:

# Estimate Std. Error t value Pr(>|t|)

# (Intercept) 75.839 3.455 21.952 < 2e-16 \*\*\*

# educL3 24.161 5.984 4.038 0.000151 \*\*\*

# educL5 41.953 5.984 7.011 2.06e-09 \*\*\*

# educL6 70.442 5.984 11.772 < 2e-16 \*\*\*

# educL7T8 85.713 5.984 14.324 < 2e-16 \*\*\*

# ageY55T64 2.771 4.886 0.567 0.572678

# educL3:ageY55T64 -2.771 8.462 -0.327 0.744447

# educL5:ageY55T64 11.577 8.462 1.368 0.176234

# educL6:ageY55T64 18.320 8.462 2.165 0.034254 \*

# educL7T8:ageY55T64 43.647 8.462 5.158 2.78e-06 \*\*\*

# ---

# Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

#

# Residual standard error: 11.97 on 62 degrees of freedom

# Multiple R-squared: 0.9332, Adjusted R-squared: 0.9235

# F-statistic: 96.25 on 9 and 62 DF, p-value: < 2.2e-16

* **All the education levels are significant.**
* **Differences in all levels of education are meaningful.**
* **This ratio for educ 3 (primary high school) is 24% for educ 5 (pre-bachelor) is 41% and 70% for educ 6 (licence) and lastly for educ 7T8 (master and doctorate) 85% more salaries. Old age group for licence , doctorates and masters graduates salaries is higher than young age group of the same group.**

**R Results Reg 5**

reg5 = lm(re ~ educ\*sex, data = dat)

summary(reg13)

# Coefficients:

# Estimate Std. Error t value Pr(>|t|)

# (Intercept) 81.646 4.340 18.812 < 2e-16 \*\*\*

# educL3 18.354 7.517 2.441 0.0175 \*

# educL5 44.770 7.517 5.955 1.32e-07 \*\*\*

# educL6 76.858 7.517 10.224 6.30e-15 \*\*\*

# educL7T8 118.386 7.517 15.748 < 2e-16 \*\*\*

# sexM -8.844 6.138 -1.441 0.1547

# educL3:sexM 8.844 10.631 0.832 0.4087

# educL5:sexM 5.943 10.631 0.559 0.5781

# educL6:sexM 5.487 10.631 0.516 0.6076

# educL7T8:sexM -21.699 10.631 -2.041 0.0455 \*

# ---

# Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

# Residual standard error: 15.03 on 62 degrees of freedom

# Multiple R-squared: 0.8946, Adjusted R-squared: 0.8793

# F-statistic: 58.46 on 9 and 62 DF, p-value: < 2.2e-16

* **Now the ratios are also gender-free**
* **group educ 3: the salary they get is %18 higher than educ 2.**
* **This ratio for educ 5 is %44 and for educ 6 %76 and finally for educ 7t8 this ratio is %118.**
* **Sexes of individuals, influence their wages especially in academy and in management departments. We can clearly see such discrimination between male and female where females are earning more than males with a ratio of 21. This was a suprising result, and we checked a snapshot from our data base.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| id | year | sex | age | educ | re |
| F\_Y25T34\_L7T8 | 2013 | F | Y25T34 | L7T8 | 155.272 |
| M\_Y25T34\_L7T8 | 2013 | M | Y25T34 | L7T8 | 150.924 |
| F\_Y55T64\_L7T8 | 2013 | F | Y55T64 | L7T8 | 245.682 |
| M\_Y55T64\_L7T8 | 2013 | M | Y55T64 | L7T8 | 184.619 |
| F\_Y25T34\_L7T8 | 2014 | F | Y25T34 | L7T8 | 161.17 |
| M\_Y25T34\_L7T8 | 2014 | M | Y25T34 | L7T8 | 162.896 |
| F\_Y55T64\_L7T8 | 2014 | F | Y55T64 | L7T8 | 241.763 |
| M\_Y55T64\_L7T8 | 2014 | M | Y55T64 | L7T8 | 181.694 |
| F\_Y25T34\_L7T8 | 2015 | F | Y25T34 | L7T8 | 172.308 |
| M\_Y25T34\_L7T8 | 2015 | M | Y25T34 | L7T8 | 166.741 |
| F\_Y55T64\_L7T8 | 2015 | F | Y55T64 | L7T8 | 223.997 |
| M\_Y55T64\_L7T8 | 2015 | M | Y55T64 | L7T8 | 170.061 |

* Clearly shows that, in given education level, females were earning more than males, except 2013, with a really small difference.
* So this suprising outcome of our regression, highlighted a particular situation in England.
* Thus, can be due to some governmental support programs against gender disparity in UK.

**R Results Reg 6**

reg6 = lm(re ~ educ\*sex\*age, data = dat)

summary(reg14)

# Coefficients:

# Estimate Std. Error t value Pr(>|t|)

# (Intercept) 80.25350 3.11485 25.765 < 2e-16 \*\*\*

# educL3 19.74650 5.39507 3.660 0.00059 \*\*\*

# educL5 34.99517 5.39507 6.487 3.26e-08 \*\*\*

# educL6 66.63317 5.39507 12.351 < 2e-16 \*\*\*

# educL7T8 82.66317 5.39507 15.322 < 2e-16 \*\*\*

# sexM -8.82917 4.40506 -2.004 0.05026 .

# ageY55T64 2.78550 4.40506 0.632 0.52993

# educL3:sexM 8.82917 7.62978 1.157 0.25248

# educL5:sexM 13.91583 7.62978 1.824 0.07392 .

# educL6:sexM 7.61750 7.62978 0.998 0.32271

# educL7T8:sexM 6.09950 7.62978 0.799 0.42768

# educL3:ageY55T64 -2.78550 7.62978 -0.365 0.71653

# educL5:ageY55T64 19.54950 7.62978 2.562 0.01334 \*

# educL6:ageY55T64 20.45017 7.62978 2.680 0.00983 \*\*

# educL7T8:ageY55T64 71.44517 7.62978 9.364 9.56e-13 \*\*\*

# sexM:ageY55T64 -0.02933 6.22969 -0.005 0.99626

# educL3:sexM:ageY55T64 0.02933 10.79015 0.003 0.99784

# educL5:sexM:ageY55T64 -15.94500 10.79015 -1.478 0.14551

# educL6:sexM:ageY55T64 -4.26000 10.79015 -0.395 0.69460

# educL7T8:sexM:ageY55T64 -55.59700 10.79015 -5.153 4.03e-06 \*\*\*

# ---

# Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

#

# Residual standard error: 7.63 on 52 degrees of freedom

# Multiple R-squared: 0.9772, Adjusted R-squared: 0.9689

# F-statistic: 117.5 on 19 and 52 DF, p-value: < 2.2e-16

* **For group educ5 (pre-graduate); younger ones receive 19% less salary than the same education group’s of elders.**
* **In the same way, in educ6 (graduate); younger ones receive 20% less salary than the elders.**
* **Unequally educ7T8 (master and doctorate) young graduates are earning much more less then tehir leder collegues. The difference is 71% of a high-school gradeuated average salary.**
* **educL7T8:sexM:ageY55T64 shows that young male master and doctoral graduates of educ7T8 earned 55% more than elder male masters and doctoral graduates.**

**R Results Reg 7**

reg7 = lm(re ~ educ + unem + exp + hdi, data = dat)

summary(reg15)

# Coefficients:

# Estimate Std. Error t value Pr(>|t|)

# (Intercept) -979.38812 1896.84922 -0.516 0.60741

# educL3 21.19415 6.59221 3.215 0.00205 \*\*

# educL5 46.16007 6.59221 7.002 1.84e-09 \*\*\*

# educL6 78.02049 6.59221 11.835 < 2e-16 \*\*\*

# educL7T8 105.95474 6.59221 16.073 < 2e-16 \*\*\*

# unem 3.67745 7.31209 0.503 0.61674

# exp -0.08137 2.53408 -0.032 0.97448

# hdi 1143.10647 2028.26116 0.564 0.57500

# ---

# Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

#

# Residual standard error: 16.41 on 64 degrees of freedom

# Multiple R-squared: 0.8703, Adjusted R-squared: 0.8561

# F-statistic: 61.36 on 7 and 64 DF, p-value: < 2.2e-16

* **We did upper regression to see general differences of wages just with education level**
* **With hiding sex and age impacts on wages;**
* **educ3 graduates earn more than educ 2 graduates with a percentage of 21%.**
* **This percentage is %46 for educ5**
* **%78 for educ6**
* **%105 for educ7T8.**

**IN CONCLUSION**

* **With or without age and sex effects, we still get the same results:**
* **With an increase in education, wages are increasing too.**
* **Elders earns more when we compare with the youngs for higher education level.**
* **So in qualified jobs, experience really matters and return as higher payments.**
* **Women are positively dicriminated. It runs in qualified jobs, but not for unskilled workers.**
* **In UK completing a University program is sound for better jobs.**

**AYÇA EPİR – 1395239**

**RIFAT CAN GÜPGÜPOĞLU – 1335919**

**ENES AHMET ORHAN - 1310083**