Estimation of Wage Equations: An Application of MLR

David Denteh [†]

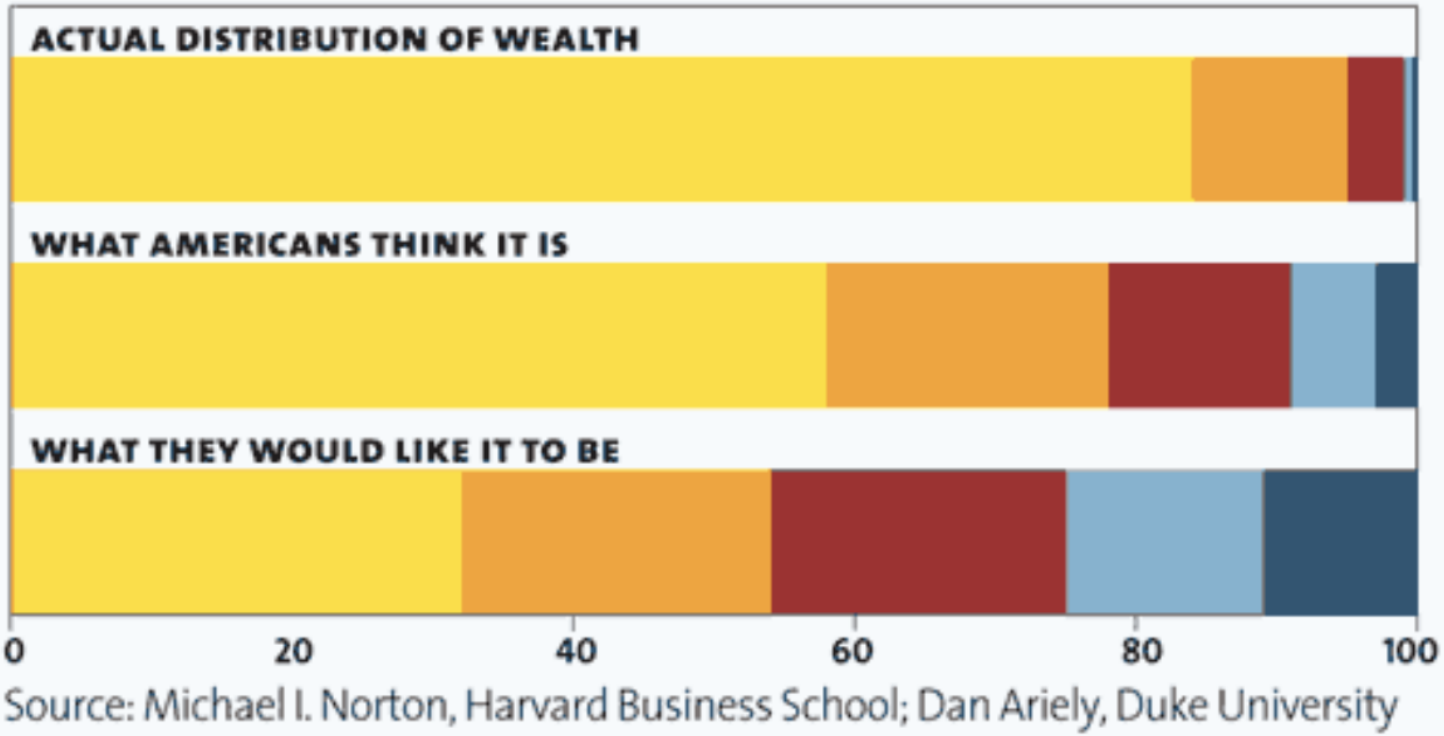
[†]University of South Carolina

Statistics Department(STAT 705)

April 26, 2019

Outline of Presentation

Income Equality or Inequality????



Source: Michael I. Norton, Harvard Business School; Dan Ariely, Duke University

What's Multiple Linear Regression (MLR)?

Suppose a researcher wants to answer the following questions;

What is the effect of gender,race,education and experience on hourly earning?Do these variables interact?

What is the un/adjusted average wage gap between males and females?

Does wage gap suggest gender wage discrimination?

What is the rate of return to schooling?

What is the estimated effect of a unit increase in experience for an individual with 20 years of experience?

What's Multiple Linear Regression (MLR)?

We can answer the above questions using Multiple Linear Regression!

MLR is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. MLR is an extension of linear (OLS) regression that uses just one explanatory variable.

In this presentation , we employ this simple but yet powerful statistical tool to answer the questions posed above.

2014 monthly Current Population Survey Outgoing Rotation Group (CPS-ORG)

The data set contains 137,374 non-student wage and salary workers, ages 18 to 65, in 2014, extracted from the Census Bureau's Current Population Survey Outgoing Rotation Group files (i.e., the 14 sample asked earnings and union questions each month).Define new variables;

- hrearn = hourly earnings
- femalemarried = female*married
- exper = yrs of potential experience.MIN([age-school-6]or[age-18])
- expersquare = the square of years of potential experience

Variable Description

school	Years of schooling completed (0-20)
age	18 to 65
female	1=females, 0=males
black	1=black, 0=other race
hispanic	1=Hispanic; 0=other ethnicity
forborn	1=foreign born (includes citizens and non-citizens); 0=native born
married	1=married and spouse present; 0=ever married, spouse not present or never
union	1=union member; 0=not a member
wkearn	Usual weekly earnings ($1 to estimated mean above the topcode)
wkhours	Usual hours worked per week (1-99)
parttime	1=less than 35 usual hours worked per week; 0=35+ usual hours

Models 1 and 2

Consider the Models:

log(hrearn)=B0+B1school+B2exper+B3exper2+B4female
+B5married+e

AND

log(hrearn)=B0+B1school+B2exper+B3exper2+B4female
+B5married+B6femalemarried+e

school	age	female	black	hispanic	forborn
Min. : 0.0	Min. :18	Min. :0.00	Min. :0.00	Min. :0.00	Min. :0.00
1st Qu.:12.0	1st Qu.:31	1st Qu.:0.00	1st Qu.:0.00	1st Qu.:0.00	1st Qu.:0.00
Median :13.3	Median :42	Median :0.00	Median :0.00	Median :0.00	Median :0.00
Mean :14.0	Mean :42	Mean :0.49	Mean :0.09	Mean :0.13	Mean :0.14
3rd Qu.:16.0	3rd Qu.:52	3rd Qu.:1.00	3rd Qu.:0.00	3rd Qu.:0.00	3rd Qu.:0.00
Max. :20.0	Max. :65	Max. :1.00	Max. :1.00	Max. :1.00	Max. :1.00

married	union	wkearn
Min. :0.00	Min. :0.00	Min. : 3
1st Qu.:0.00	1st Qu.:0.00	1st Qu.: 464
Median :1.00	Median :0.00	Median : 750
Mean :0.56	Mean :0.12	Mean :1008
3rd Qu.:1.00	3rd Qu.:0.00	3rd Qu.:1191
Max. :1.00	Max. :1.00	Max. :5940

wkhours	parttime	hrearn	exper	expersquare	femalemarried
Min. : 1	Min. :0.00	Min. : 3	Min. : 0	Min. : 0	Min. :0.00
1st Qu.:40	1st Qu.:0.00	1st Qu.: 12	1st Qu.:11	1st Qu.: 121	1st Qu.:0.00
Median :40	Median :0.00	Median : 19	Median :22	Median : 471	Median :0.00
Mean :40	Mean :0.14	Mean : 24	Mean :22	Mean : 624	Mean :0.26
3rd Qu.:40	3rd Qu.:0.00	3rd Qu.: 29	3rd Qu.:32	3rd Qu.:1024	3rd Qu.:1.00
Max. :99	Max. :1.00	Max. :297	Max. :47	Max. :2209	Max. :1.00

```
Call:
lm(formula = log(hrearn) ~ school + exper + expersquare + female +
    married)

Residuals:
   Min     1Q Median     3Q    Max
-2.795 -0.336 -0.027  0.301  3.445

Coefficients:
               Estimate  Std. Error t value             Pr(>|t|)
(Intercept)  1.05825796  0.00883214   119.8 <0.0000000000000002 ***
school       0.11397895  0.00055287   206.2 <0.0000000000000002 ***
exper        0.03138694  0.00045851    68.5 <0.0000000000000002 ***
expersquare -0.00050662  0.00000998   -50.8 <0.0000000000000002 ***
female      -0.22631766  0.00284662   -79.5 <0.0000000000000002 ***
married      0.10458052  0.00306233    34.1 <0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.53 on 137368 degrees of freedom
Multiple R-squared:  0.308,     Adjusted R-squared:  0.308
F-statistic: 1.22e+04 on 5 and 137368 DF,  p-value: <0.0000000000000002
```

```
Call:
lm(formula = log(hrearn) ~ school + exper + expersquare + female +
    married + femalemarried)

Residuals:
   Min     1Q Median     3Q    Max
-2.814 -0.335 -0.029  0.299  3.471

Coefficients:
                 Estimate  Std. Error t value           Pr(>|t|)
(Intercept)    1.03405525  0.00891359   116.0 <0.0000000000000002 ***
school         0.11373890  0.00055230   205.9 <0.0000000000000002 ***
exper          0.03114227  0.00045810    68.0 <0.0000000000000002 ***
expersquare   -0.00050450  0.00000997   -50.6 <0.0000000000000002 ***
female        -0.16618417  0.00427148   -38.9 <0.0000000000000002 ***
married        0.15887739  0.00419997    37.8 <0.0000000000000002 ***
femalemarried -0.10774463  0.00571207   -18.9 <0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.52 on 137367 degrees of freedom
Multiple R-squared:  0.31,      Adjusted R-squared:  0.31
F-statistic: 1.03e+04 on 6 and 137367 DF,  p-value: <0.0000000000000002
```

Unadjusted Wage differential??

One may say to find gender wage differential ,

Simply calculate the mean wage differential between men and women:

The unadjusted wage gap is 27.195 - 21.545 = 5.6

Some findings

What is the rate of return to sch.?The rate of return to schooling is 11.398 pct

What is the approx. pct difference between women and men? The approx. pct difference between women and men is 22.632 pct, on average.

What is the exact pct difference between women and men? The exact pct difference between women and men is [exp(-0.22631766)-1]*100 =20.254 pct, on average.

Does the coef. est. on female suggest gender wage discrimination? If the model is correctly specified model, the coef. on female is suggestive of gender wage discrimination. However??

Some findings

What is the est. effect of a unit increase in exper. for an individual with 20yrs of exper.? The est. increase is [0.0313869 – (0.0005066*20*2)] *100 = 1.11229pct.

Does exper. have an increasing or decreasing effect on hourly earnings? Exper. has a decreasing effect on hourly wages because of the neg. sign on the coef. on exper. squared.

Compute the turnaround point for the effect of exper. on hourly earnings? The turnaround point is given by 0.03114227/ (2*0.00050450) = 30.97 yrs.

Are hourly earnings minimized or maximized at this point? Hourly wages are maximized at this point.

Some findings

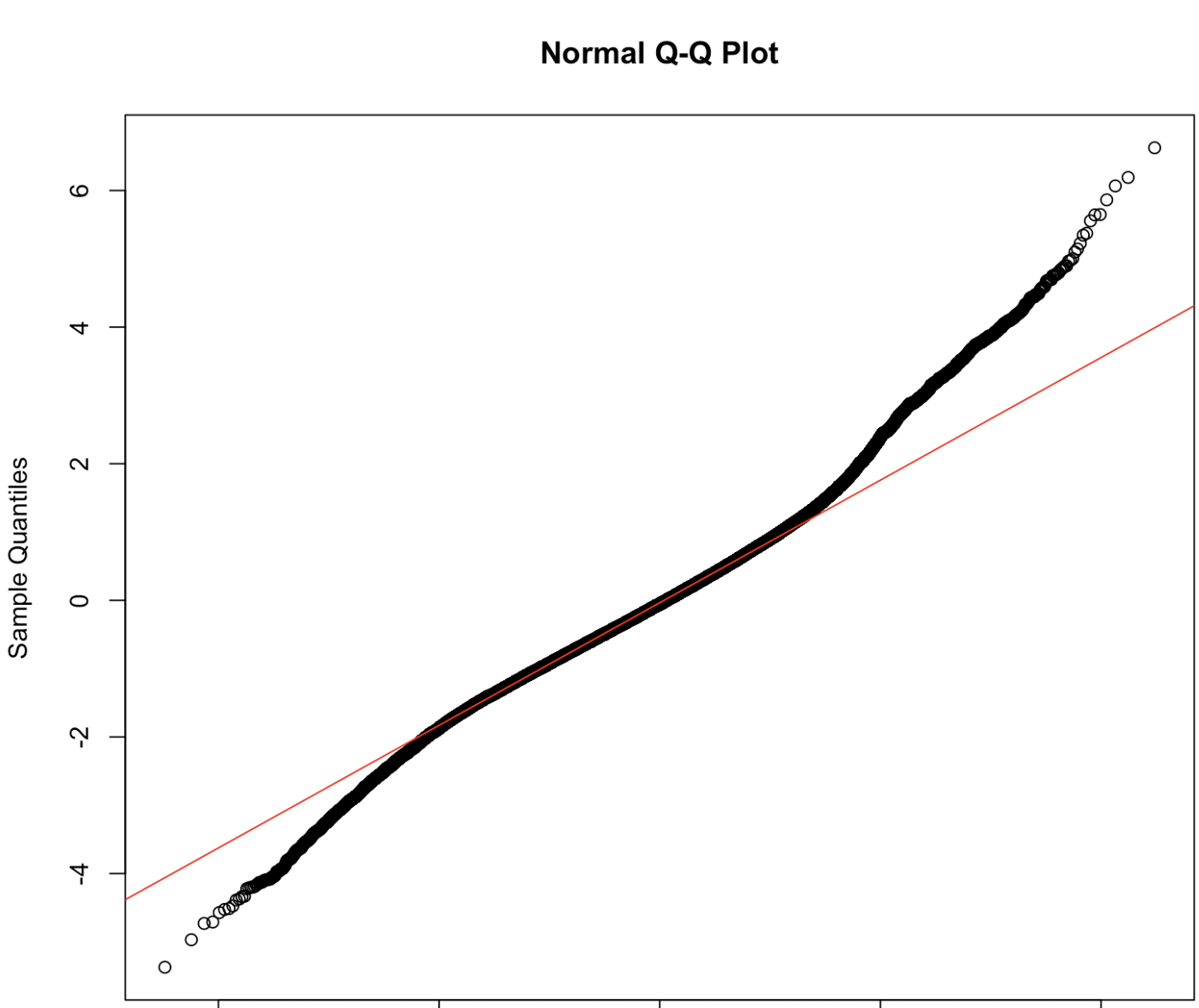
What is the omitted category in the 2 Model? Single males

What is the approx. pct wage difference b/tn married females and the single male? The approx. pct difference is -0.1661842+0.1588774-0.1077446 = - 0.1150514 (11.505pct lower).

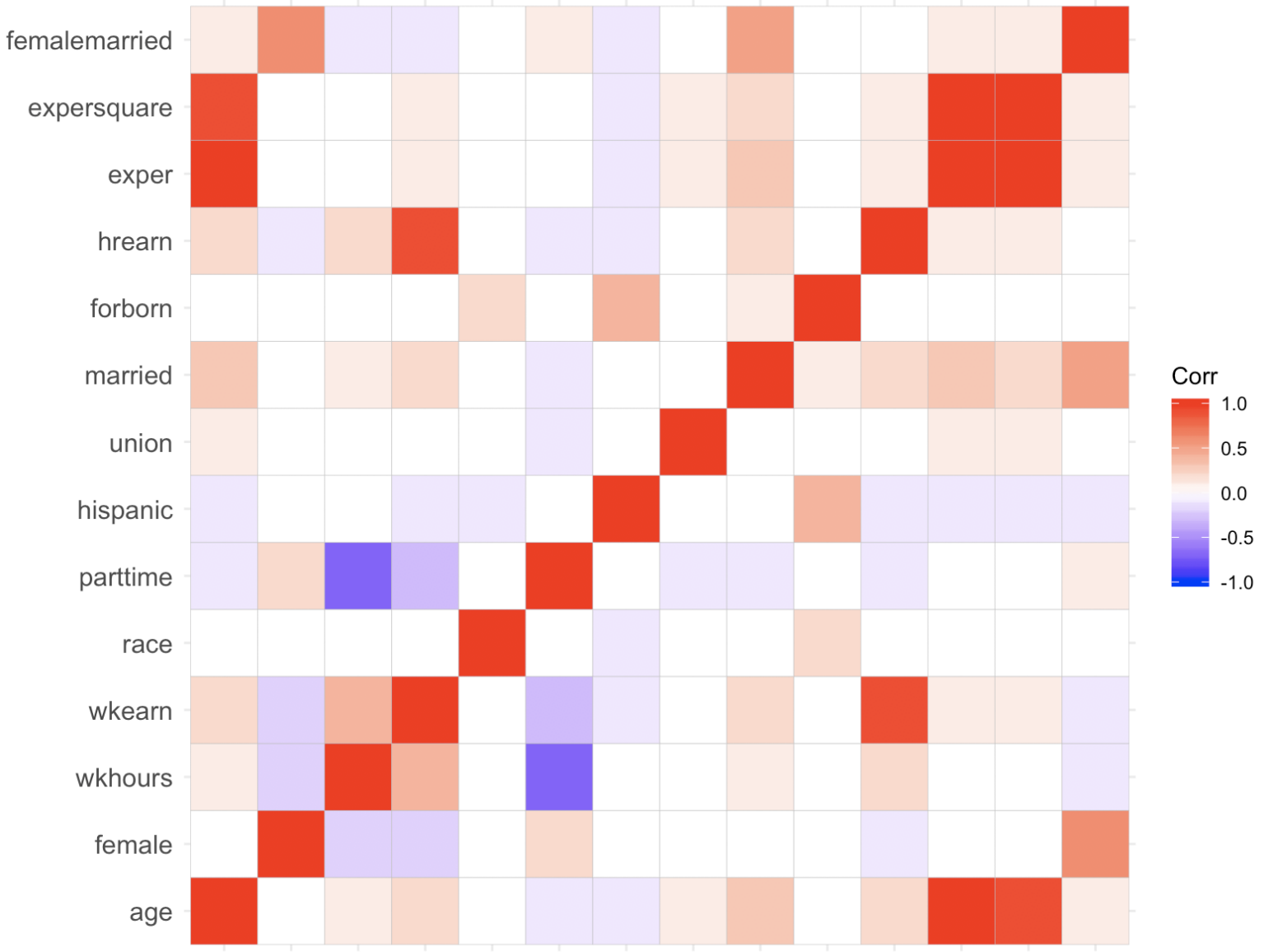
What is the interpretation of the coef. est. on femalemarried? The coefficient estimate on females suggests that marriage premium differs by gender by approx. 10.774pct. That is, the percentage hourly wage differential between the married and single is lower for women relative to men by approximately 10.774pct.

Construct a 99pct. CI for the coef. est. on femalemarried. -0.1077446 +/- (2.576*0.0057121) = (-0.12245897, -0.09303023)

QQ Plot



Correlation Matrix



reference

http://ceprdata.org/cps-uniform-data-extracts/cps-outgoing-rotation-group/.

R Code available upon request.(Email:davident85@gmail.com)

Thank You!