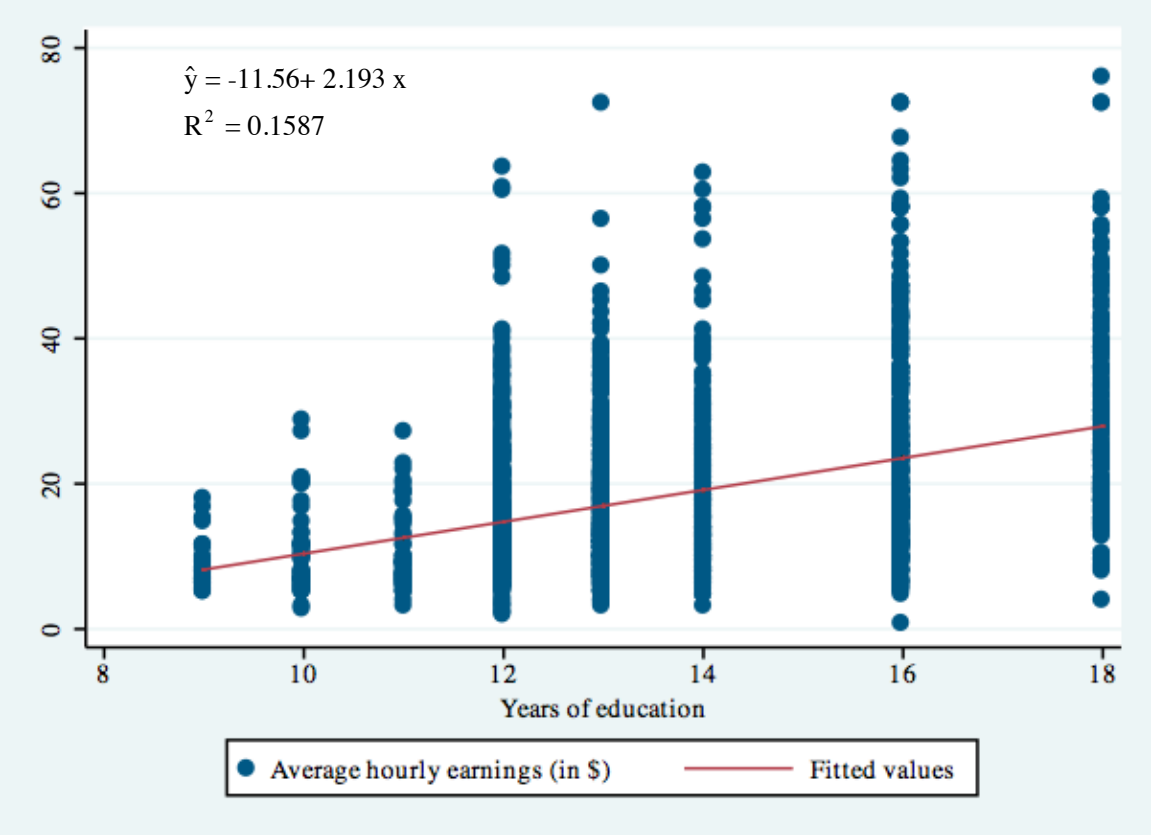
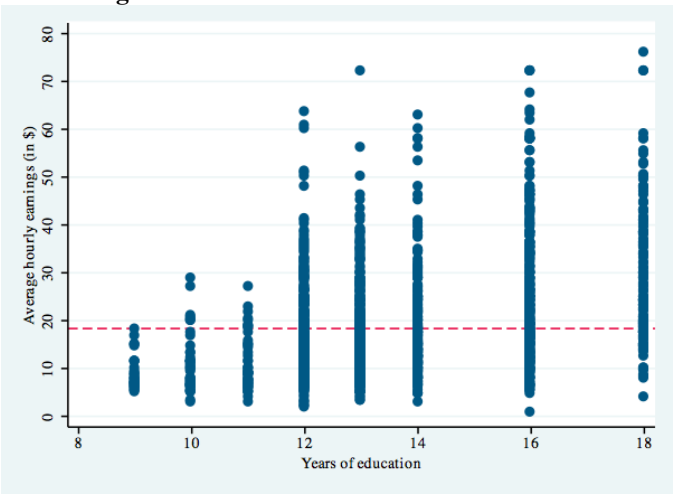
Data source: Current Population Survey 2006. Data available from the Inter-University Consortium for Political and Social Research at the University of Michigan.http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/4559?archive=ICPSR&q=current+population+survey+asec, from which I extracted a random sample of 2000 observations

```
. insheet using "/Users/sadoulet/…/data/CPS_06_2000.csv";
. label variable educ "Years of education";
. label variable wage "Average hourly earnings (in $)";
. reg wage educ;
. predict wagehat;
. twoway (scatter wage educ) (line wagehat educ);
```

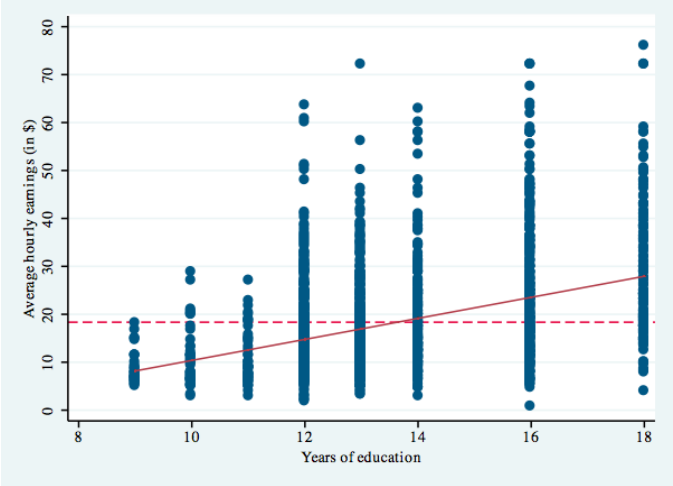



```
. reg wage educ;

      Source |       SS       df       MS              Number of obs =    2000
-------------+------------------------------           F(  1,  1998) =  376.94
       Model |  41922.0349     1  41922.0349           Prob > F      =  0.0000
    Residual |  222213.443  1998  111.217939           R-squared     =  0.1587
-------------+------------------------------           Adj R-squared =  0.1583
       Total |  264135.478  1999  132.133806           Root MSE      =  10.546

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |    2.193546    .112983    19.41   0.000     1.971969    2.415123
       _cons |   -11.5576   1.558244    -7.42   0.000    -14.61355   -8.501649
------------------------------------------------------------------------------
```

Illustrating R^2



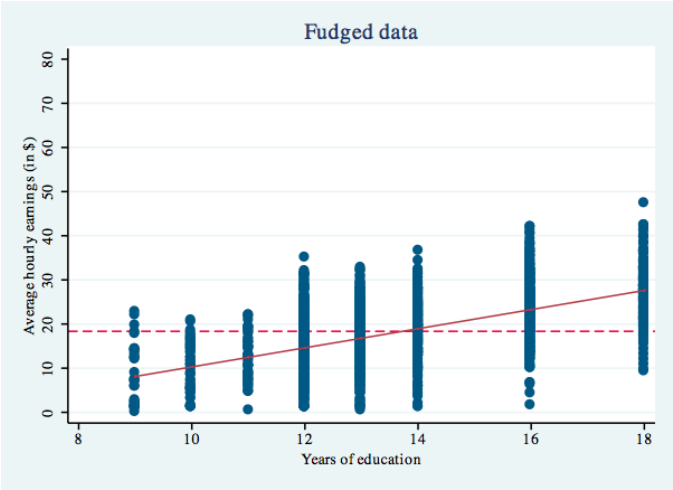
Average earnings = $18.37
variance(earnings) = 132.4



$\hat{y} = -11.56 + 2.193\ x$

$R^2 = 0.1587$

FUDGED DATA:



Average earnings = $18.15
var(earnings) = 63.4
(Std. Dev.= $7.96)

$\hat{y} = -11.36 + 2.165\ x$

$R^2 = 0.322$