

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 $)"
. summarize educ wage
```

Variable	Obs	Mean	Std. Dev.	Min	Max
educ	2000	13.633	2.0877	9	18
wage	2000	18.34701	11.49495	.7	82.42857

```
. reg wage educ
```

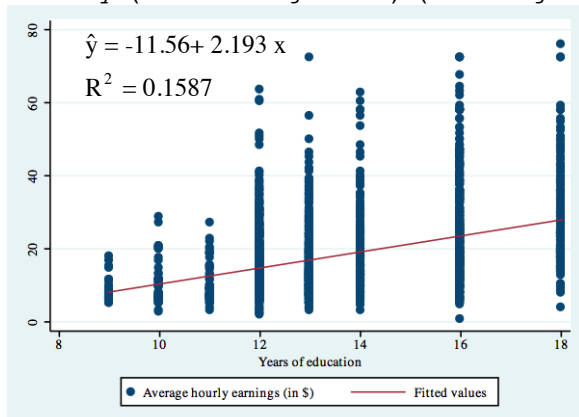
Source	SS	df	MS	Number of obs =	2000
Model	41922.0349	1	41922.0349	F(1, 1998) =	376.94
Residual	222213.443	1998	111.217939	Prob > F =	0.0000
				R-squared =	0.1587
				Adj R-squared =	0.1583
Total	264135.478	1999	132.133806	Root MSE =	10.546

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

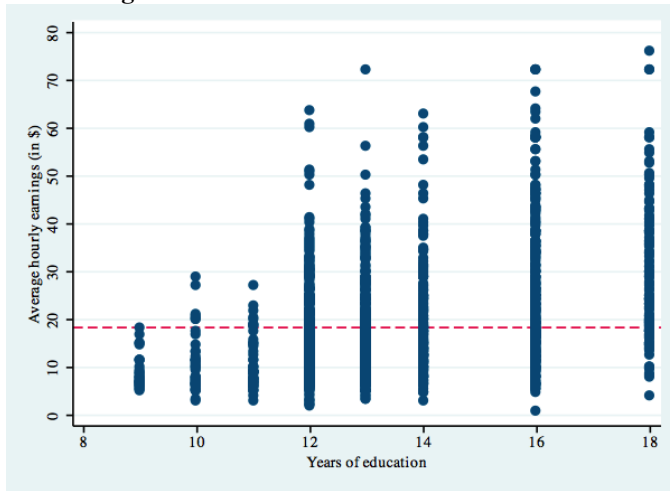
```
. predict wagehat
. list educ wage wagehat
```

	educ	wage	wagehat
1.	9	16.825	8.184312
2.	9	11.225	8.184312
3.	9	9	8.184312
4.	13	36	16.9585
5.	12	8	14.76495
6.	12	6.25	14.76495
7.	18	24.05	27.92623
8.	13	12	16.9585
9.	16	37.5	23.53913
10.	12	19.225	14.76495

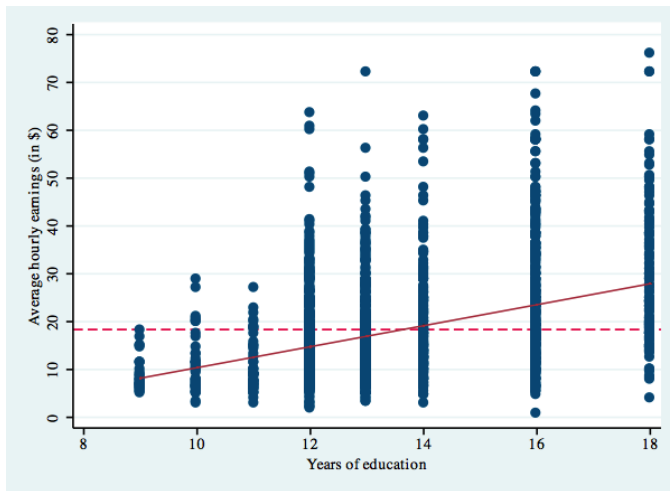
```
. twoway (scatter wage educ) (line wagehat educ)
```



Illustrating R²



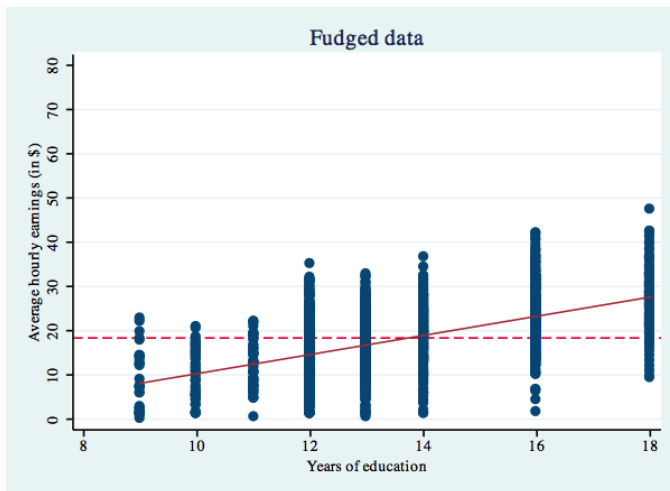
Sample mean earnings = \$18.35
Sample variance (earnings) = 132.1



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

$$R^2 = 0.1587$$

FUDGED DATA:



Sample mean earnings = \$18.15
Sample variance (earnings) = 63.4

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

$$R^2 = 0.322$$