

Business Statistics 41000

Wage Data

Mladen Kolar

Wage data

Data:

- ▶ Greenberg and Kusters, 1970, Rand Corporation.
- ▶ 39 demographic groupings of 6000 households with the male head earning less than \$15,000 annually in 1966.

Problem:

- ▶ Estimate the relationship between pay and labor supply.
- ▶ Use this information to influence social policy decisions and the debate on a guaranteed national wage.

Possible solution:

- ▶ Fit a linear model for the effect of pay on labor supply for the working poor.

Read in the data, and change to X and Y for convenience.

```
D = read.csv("wages.csv")  
Y = D$HRS  
X = D$RATE
```

Use correlation to slope the line

```
(b1 = cor(X,Y)*sd(Y)/sd(X))
```

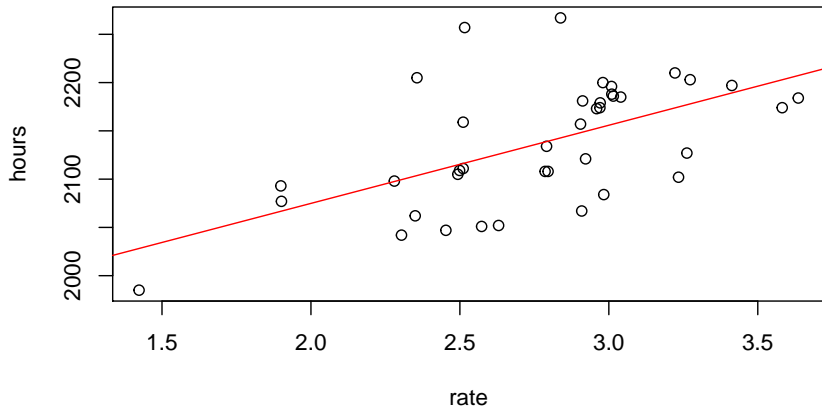
```
## [1] 80.937
```

Make sure that \bar{X} and \bar{Y} are on the line.

```
(b0 = mean(Y) - mean(X)*b1)
```

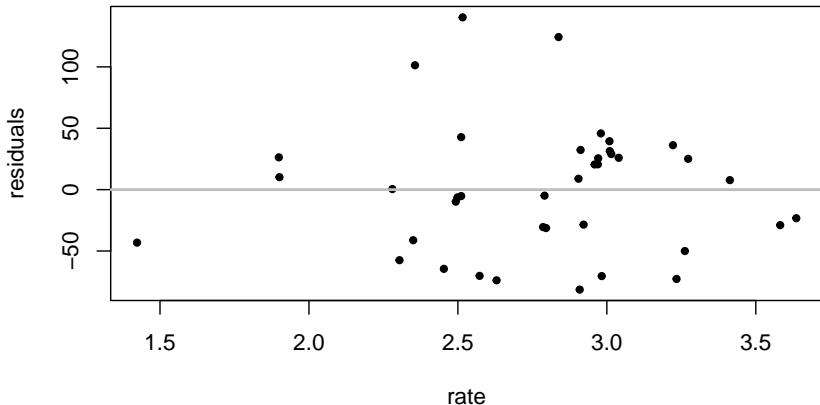
```
## [1] 1913
```

```
plot(X,Y, xlab="rate", ylab="hours")  
abline(b0, b1, col="red")
```



The **residual errors** are defined as $e_i = Y_i - (b_0 + b_1 X_i)$.

```
e = Y - (b0 + b1*X)
plot(X, e, xlab="rate", ylab="residuals", pch=20)
abline(h=0, col=8, lwd=2)
```



What does this plot of e_i versus X_i say about the fit?

How does labor supply change with pay?

We've estimated the linear model:

$$\text{hours} = 1913 + 81 \times \text{rate} + e$$

It increases: Every dollar extra per hour leads to an increase of 81 expected annual hours worked.

What would this mean for the debate (at the time of the study) on national wage/negative income tax?

Questions remain . . .

1. Does this apply today?
2. Does the same hold for those who make more money?
3. How good is this estimate?
4. Is pay rate the only variable of importance?
5. Do higher wages *cause* more work?