ECON4150 - Introductory Econometrics

Lecture 11: Nonlinear Regression Functions

Monique de Haan
(moniqued@econ.uio.no)

Stock and Watson Chapter 8

Lecture outline

- What are nonlinear regression functions?
- Data set used during lecture.
- The effect of change in X_1 on Y depends on X_1
- The effect of change in X_1 on Y depends on another variable X_2

What are nonlinear regression functions?

So far you have seen the **linear** multiple regression model

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki} + u_i$$

- The effect of a change in X_j by 1 is constant and equals β_j.

There are 2 types of **nonlinear** regression models

1. Regression model that is a nonlinear function of the independent variables $X_{1i}, ..., X_{ki}$
   - Version of multiple regression model, can be estimated by OLS.
2. Regression model that is a nonlinear function of the unknown coefficients $\beta_0, \beta_1, ..., \beta_k$
   - Can't be estimated by OLS, requires different estimation method.

This lecture we will only consider first type of nonlinear regression models.

What are nonlinear regression functions?

General formula for a nonlinear population regression model:

$$Y_i = f(X_{1i}, X_{2i}, \ldots.., X_{ki}) + u_i$$

Assumptions:

1. $E(ui|X_{1i}, X_{2i}, \ldots, X_{ki}) = 0$ (same); implies that f is the conditional expectation of Y given the X's.
2. $(X_{1i}, \ldots, X_{ki}, Y_i)$ are i.i.d. (same).
3. Big outliers are rare (same idea; the precise mathematical condition depends on the specific f).
4. No perfect multicollinearity (same idea; the precise statement depends on the specific f).

What are nonlinear regression functions?

Two cases:

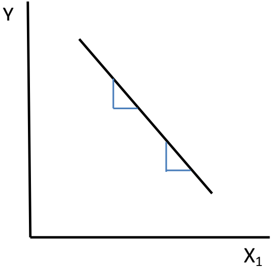
1. The effect of change in X_1 on Y depends on X_1
   - for example: the effect of a change in class size is bigger when initial class size is small
2. The effect of change in X_1 on Y depends on another variable X_2
   - For example: the effect of class size depends on the percentage of disadvantaged pupils in the class

We start with case 1 using a regression model with only 1 independent variable

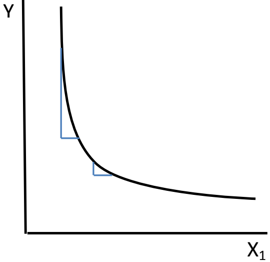
$$Y_i = f(X_{1i}) + u_i$$

What are nonlinear regression functions?

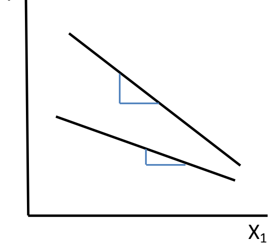
1. Linear model: constant slope



2. Nonlinear model: slope depends on X_1



3. Nonlinear model: slope depends on X_2



Data

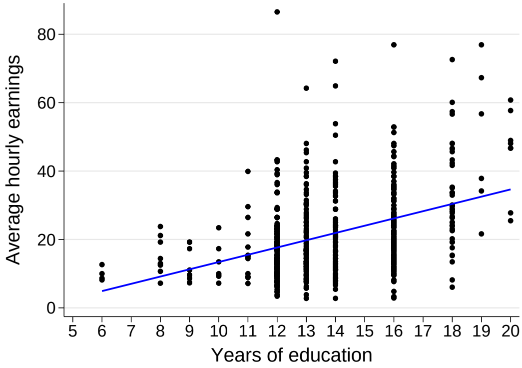
Examples in this lecture are based on data from the CPS March 2009.

- Current Population Survey" (CPS) collects information on (among others) education, employment and earnings.
- Approximately 65,000 households are surveyed each month.
- We use a 1% sample which gives a data set with 602 observations .

Summary Statistics

	Mean	SD	Min	Max	Nobs
Average hourly earnings	21.65	12.63	2.77	86.54	602
Years of education	13.88	2.43	6.00	20.00	602
Age	42.91	11.19	21.00	64.00	602
Gender (female=1)	0.39	0.49	0.00	1.00	602

We will investigate the association between years of education and hourly earnings.




```
. regress hourlyearnings education, robust

Linear regression                               Number of obs =        602
                                                   F(  1,   600) =   108.34
                                                   Prob > F      =   0.0000
                                                   R-squared     =   0.1674
                                                   Root MSE      =    11.53
```

hourlyearn~s	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
education	**2.12359**	**.2040197**	**10.41**	**0.000**	**1.722911**	**2.52427**
_cons	**-7.834347**	**2.728805**	**-2.87**	**0.004**	**-13.19352**	**-2.475178**

Linear model: interpretation

What is the effect of a change in education on average hourly earnings?

- When $E\left[u_i|X_{1i}\right] = 0 \longrightarrow E\left[Y_i|X_{1i}\right] = \beta_0 + \beta_1 X_{1i}$
- Taking the derivative of the conditional expectation w.r.t X_{1i} gives

$$\frac{\partial E\left[Y_i|X_{1i}\right]}{\partial X_{1i}} = \beta_1$$

- $$\begin{aligned} \triangle\widehat{Y} &= \left(\widehat{\beta}_0 + \widehat{\beta}_1(X_1 + \triangle X_1)\right) - \left(\widehat{\beta}_0 + \widehat{\beta}_1 X_1\right) \\ &= \widehat{\beta}_1 \cdot \triangle X_1 \end{aligned}$$

- An increase in years of education by 1 is expected to increase average hourly earnings by 2.12 dollars.

Polynomials

- If actual relationship is nonlinear with $f(X_{1i}) \neq \beta_0 + \beta_1 X_{1i}$ the linear model is misspecified and $E(u_i|X_{1i}) \neq 0$.

- One way to specify a nonlinear regression is to use a polynomial in X.

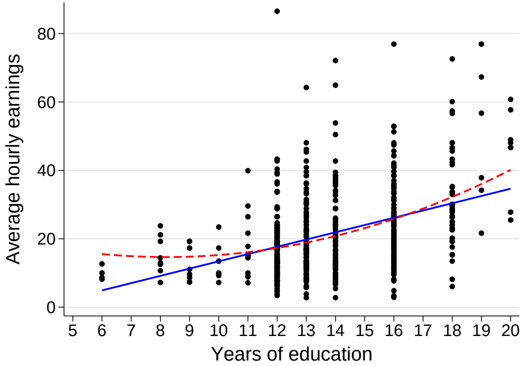
- The polynomial regression model of degree r is

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{1i}^2 + ... + \beta_r X_{1i}^r + u_i$$

- A quadratic regression is a polynomial regression with $r = 2$

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{1i}^2 + u_i$$

  - This is a multiple regression model with two regressors: X_{1i} and X_{1i}^2




```
Linear regression                               Number of obs =        602
                                                   F(  2,   599) =      62.56
                                                   Prob > F      =     0.0000
                                                   R-squared     =     0.1837
                                                   Root MSE      =     11.426
```

hourlyearn~s	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
education	**-3.004498**	**1.26951**	**-2.37**	**0.018**	**-5.49773**	**-.5112657**
education2	**.1831323**	**.0485472**	**3.77**	**0.000**	**.0877889**	**.2784757**
_cons	**26.98042**	**8.128804**	**3.32**	**0.001**	**11.01599**	**42.94484**

Polynomials: interpretation

- When $E\left[u_i|X_{1i}\right]=0 \longrightarrow E\left[Y_i|X_{1i}\right]=\beta_0+\beta_1X_{1i}+\beta_2X_{1i}^2+...+\beta_rX_{1i}^r$
- Taking the derivative of the conditional expectation w.r.t X_{1i} gives

$$\frac{\partial E\left[Y_i|X_{1i}\right]}{\partial X_{1i}}=\beta_1+2\beta_2X_{1i}+...+r\beta_rX_{1i}^{r-1}$$

- The predicted change in Y that is associated with a change in X_1:

$$\begin{aligned}\triangle\hat{Y} &= \widehat{f}(X_1+\triangle X_1)-\widehat{f}(X_1)\\ &= \left(\widehat{\beta}_1\left(X_1+\triangle X_1\right)+...+\widehat{\beta}_r\left(X_1+\triangle X_1\right)^r\right)-\left(\widehat{\beta}_1X_1+...+\widehat{\beta}_rX_1{}^r\right)\end{aligned}$$

Polynomials: interpretation

```
Linear regression                                      Number of obs =        602
                                                          F(  2,   599) =      62.56
                                                          Prob > F      =     0.0000
                                                          R-squared     =     0.1837
                                                          Root MSE      =     11.426
```

hourlyearn~s	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
education	-3.004498	1.26951	-2.37	0.018	-5.49773	-.5112657
education2	.1831323	.0485472	3.77	0.000	.0877889	.2784757
_cons	26.98042	8.128804	3.32	0.001	11.01599	42.94484

In the quadratic model the predicted change in hourly earnings when education increase from

10 to 11:

$$\widehat{\triangle Y} = \left(26.98 - 3.00 \cdot 11 + 0.18 \cdot 11^2\right) - \left(26.98 - 3.00 \cdot 10 + 0.18 \cdot 10^2\right) = 0.78$$

15 to 16:

$$\widehat{\triangle Y} = \left(26.98 - 3.00 \cdot 16 + 0.18 \cdot 16^2\right) - \left(26.98 - 3.00 \cdot 15 + 0.18 \cdot 15^2\right) = 2.58$$

Polynomials

- Is the quadratic model better than the linear model?
- We can test the null hypothesis that the regression function is linear against the alternative hypothesis that it is quadratic:

$$H_o: \ \beta_2 = 0 \quad vs \quad H_1: \beta_2 \neq 0$$

- Obtain the t-statistic:

$$t = \frac{\widehat{\beta}_2 - 0}{\widehat{SE}(\widehat{\beta}_2)} = \frac{0.183}{0.049} = 3.77$$

- Since $t = 3.77 > 2.58$ we reject the null hypothesis (the linear model) at a 1% significance level
- We can include higher powers of X_{1i} in the regression model
    - should we estimate a cubic regression model?

Polynomials

```
Linear regression                               Number of obs =        602
                                                   F(  3,   598) =      55.01
                                                   Prob > F      =     0.0000
                                                   R-squared     =     0.1933
                                                   Root MSE      =     11.368
```

hourlyearn~s	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
education	14.20664	5.252381	2.70	0.007	3.89128	24.52199
education2	-1.165764	.437365	-2.67	0.008	-2.024722	-.3068056
education3	.0338681	.0115973	2.92	0.004	.0110918	.0566444
_cons	-43.01427	19.90841	-2.16	0.031	-82.11317	-3.915365

Cubic versus quadratic model: $H_o: \beta_3 = 0 \quad vs \quad H_1: \beta_3 \neq 0$

- $t = 2.92 > 2.58 \longrightarrow H_0$ rejected at 1% significance level

Polynomials

Cubic versus linear model:

$$H_o: \ \beta_2 = 0, \beta_3 = 0 \quad vs \quad H_1: \beta_2 \neq 0 \ and/or \ \beta_2 \neq 0$$

```
. test education2=education3=0

 ( 1)  education2 - education3 = 0
 ( 2)  education2 = 0

       F(  2,   598) =      8.39
            Prob > F =      0.0003
```

- $F = 8.39 > 4.61(F_{2,\infty}) \longrightarrow H_0$ rejected at 1% significance level

Logarithms

- Another way to specify a nonlinear regression model is to use the natural logarithm of Y and/or X.
- Using logarithms allows changes in variables to be interpreted in terms of percentages

$$ln(x + \triangle x) - ln(x) \approx \frac{\triangle x}{x} \quad \left(when \ \frac{\triangle x}{x} \ is \ small \right)$$

- We will consider 3 types of logarithmic regression models:

1. The linear-log model

$$Y_i = \beta_0 + \beta_1 ln(X_{1i}) + u_i$$

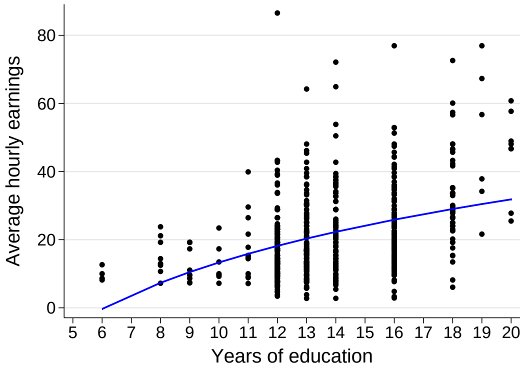
2. The log-linear model

$$ln(Y_i) = \beta_0 + \beta_1 X_{1i} + u_i$$

3. The log-log model

$$ln(Y_i) = \beta_0 + \beta_1 ln(X_{1i}) + u_i$$

The linear-log model




```
Linear regression                               Number of obs =        602
                                                   F(  1,   600) =      97.80
                                                   Prob > F      =     0.0000
                                                   R-squared     =     0.1499
                                                   Root MSE      =     11.651
```

hourlyearn~s	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
ln_education	26.72023	2.701844	9.89	0.000	21.41401	32.02645
_cons	-48.2151	6.942683	-6.94	0.000	-61.85002	-34.58019

The linear-log model: interpretation

- When $E\left[u_i|X_{1i}\right]=0 \longrightarrow E\left[Y_i|X_{1i}\right]=\beta_0+\beta_1 ln(X_{1i})$

- Taking the derivative of the conditional expectation w.r.t X_{1i} gives

$$\frac{\partial E\left[Y_i|X_{1i}\right]}{\partial X_{1i}}=\beta_1\cdot\frac{1}{X_{1i}}$$

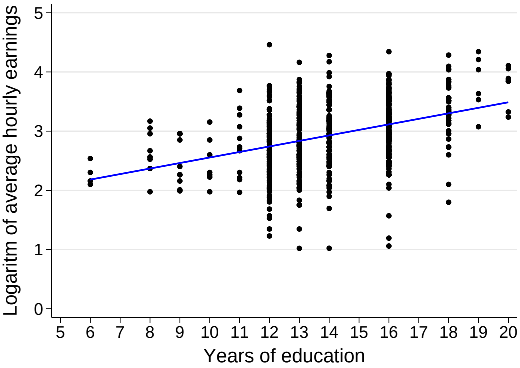
- Using that $\frac{\partial E[Y_i|X_{1i}]}{\partial X_{1i}}\approx\frac{\triangle E[Y_i|X_{1i}]}{\triangle X_{1i}}$ for small changes in X_1 and rewriting gives

$$\triangle E\left[Y_i|X_{1i}\right]\approx\beta_1\cdot\frac{\triangle X_{1i}}{X_{1i}}$$

- **Interpretation of** β_1: A 1% change in X_1 ($\frac{\triangle X_{1i}}{X_{1i}}=0.01$) is associated with a change in Y of $0.01\beta_1$

- A 1 % increase in years of education is expected to increase average hourly earnings by 0.27 dollars

The log-linear model




```
Linear regression                                   Number of obs =        602
                                                       F(  1,   600) =     139.52
                                                       Prob > F      =     0.0000
                                                       R-squared     =     0.1571
                                                       Root MSE      =     .52602
```

ln_hourlye~s	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
education	.0932827	.0078974	11.81	0.000	.0777728	.1087927
_cons	1.622094	.1112224	14.58	0.000	1.403662	1.840527

The log-linear model: interpretation

$$ln(Y_i) = \beta_0 + \beta_1 X_{1i} + u_i$$

- Suppose we have the following equation

$$ln(y) = a + b \cdot x$$

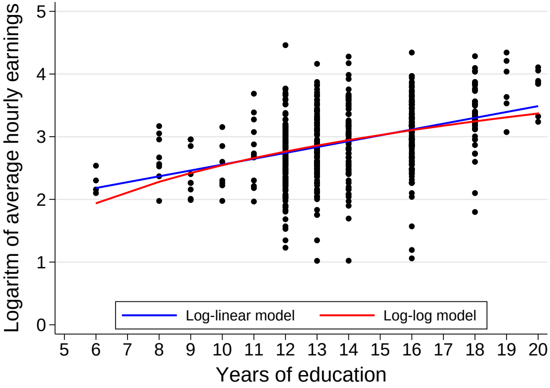
- Taking the derivative of both sides of the equation (using the chain rule) gives

$$\frac{1}{y}dy = b \cdot dx \quad \longrightarrow \quad 100 \cdot \frac{\triangle y}{y} \approx 100 \cdot b \cdot \triangle x$$

- **Interpretation of β_1:** A change in X_1 by one unit is associated with a $100 \cdot \beta_1$ percent change in Y

- An increase in years of education by 1 is expected to increase average hourly earnings by 9.3 percent.

The log-log model




```
Linear regression                                      Number of obs =        602
                                                          F(  1,   600) =     120.63
                                                          Prob > F      =     0.0000
                                                          R-squared     =     0.1447
                                                          Root MSE      =     .52989
```

ln_hourlye~s	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
ln_education	1.190072	.1083532	10.98	0.000	.9772749	1.40287
_cons	-.194417	.2832781	-0.69	0.493	-.7507542	.3619202

The log-log model: interpretation

$$ln(Y_i) = \beta_0 + \beta_1 ln(X_{1i}) + u_i$$

- Suppose we have the following equation

$$ln(y) = a + b \cdot ln(x)$$

- Taking the derivative of both sides of the equation (using the chain rule) gives

$$\frac{1}{y}dy = b \cdot \frac{1}{x}dx \quad \longrightarrow \quad 100 \cdot \frac{\triangle y}{y} \approx 100 \cdot b \cdot \frac{\triangle x}{x}$$

- **Interpretation of β_1:** A change in X_1 by one percent is associated with a β_1 percent change in Y

- An increase in years of education by 1 percent is expected to increase average hourly earnings by 1.2 percent.

Logarithms: which model fits the data best?

Difficult to decide which model fits data best.

- Sometimes you can compare the R^2 (don't rely too much on this!)
  - Linear-log model vs linear model:

$$R^2_{linear-log} = 0.1499 < 0.1674 = R^2_{linear}$$

  - Log-linear model vs log-log model:

$$R^2_{log-linear} = 0.1571 > 0.1477 = R^2_{log-log}$$

  - R^2 can never be compared when dependent variables differ
- Look at scatter plots and compare graphs
- Use economic theory or expert knowledge
  - Labor economist typically model earnings in logarithms and education in years
  - Wage comparisons most often discussed in percentage terms.

Interactions

- So far we discussed nonlinear models with 1 independent variable X_{1i}
- We now turn to models whereby the effect of X_{1i} depends on another variable X_{2i}
- We discuss 3 cases:

1. Interactions between two binary variables
2. Interactions between a binary and a continuous variable
3. Interactions between two continuous variables

Interpretation of a coefficient on a binary variable

$$Y_i = \beta_0 + \beta_1 D_{1i} + u_i$$

- Let D_{1i} equal 1 if an individual has more than a high school degree (*years of education* > 12) and zero otherwise.

```
Linear regression                               Number of obs =        602
                                                  F(  1,    600) =    58.09
                                                  Prob > F       =   0.0000
                                                  R-squared      =   0.0723
                                                  Root MSE       =   12.171
```

hourlyearnings	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
more_highschool	7.172748	.941093	7.62	0.000	5.324511	9.020984
_cons	16.89143	.6626943	25.49	0.000	15.58995	18.19291

- $\widehat{\beta}_0 = 16.89$ is the average hourly earnings for individuals with a high school degree or less.
- $\widehat{\beta}_0 + \widehat{\beta}_1 = 16.89 + 7.17 = 24.06$ is the average hourly earnings for individuals with more than a high school degree.

Interactions between two binary variables

- Effect of having more than a high school degree on earnings might differ between men and women

```
. regress hourlyearnings more_highschool if female==1, robust

Linear regression                               Number of obs =        237
                                                   F(  1,   235) =       9.81
                                                   Prob > F      =     0.0020
                                                   R-squared     =     0.0400
                                                   Root MSE      =     11.173
```

hourlyearnings	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
more_highschool	5.194752	1.658509	3.13	0.002	1.927306	8.462198
_cons	14.28346	1.428513	10.00	0.000	11.46913	17.09779

```
.
. regress hourlyearnings more_highschool if female==0, robust

Linear regression                               Number of obs =        365
                                                   F(  1,   363) =      69.19
                                                   Prob > F      =     0.0000
                                                   R-squared     =     0.1343
                                                   Root MSE      =     12.007
```

hourlyearnings	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
more_highschool	9.671839	1.162783	8.32	0.000	7.385202	11.95848
_cons	18.01175	.7031579	25.62	0.000	16.62898	19.39453

Interactions between two binary variables

- We can extend the model by including gender as an additional explanatory variable
- Let D_{2i} equal 1 for women and zero for men

$$Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + u_i$$

- This model allows the intercept to depend on gender
  - intercept for men: β_0
  - intercept for women: $\beta_0 + \beta_2$

Interactions between two binary variables

```
Linear regression                                      Number of obs =       602
                                                          F(  2,   599) =     44.33
                                                          Prob > F      =    0.0000
                                                          R-squared     =    0.1413
                                                          Root MSE      =    11.719
```

hourlyearnings	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
more_highschool	8.136047	.9585592	8.49	0.000	6.253501	10.01859
female	-6.85085	1.001335	-6.84	0.000	-8.817405	-4.884296
_cons	18.95006	.6887376	27.51	0.000	17.59742	20.30269

- The above regression model assumes that the effect of D_{1i} is the same for men and women
- We can extend the model by allowing the effect D_{1i} to depend on gender by including the interaction between D_{1i} and D_{2i}

$$Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_3 (D_{1i} \times D_{2i}) + u_i$$

Interactions between two binary variables

$$Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_3 (D_{1i} \times D_{2i}) + u_i$$

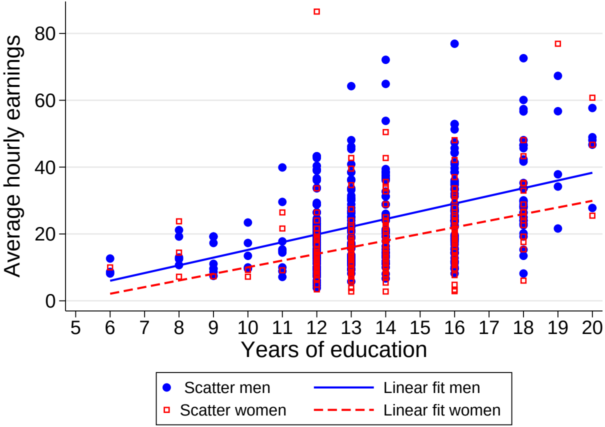
```
Linear regression                               Number of obs =        602
                                                  F(  3,    598) =   30.93
                                                  Prob > F      =   0.0000
                                                  R-squared     =   0.1476
                                                  Root MSE      =   11.686
```

hourlyearnings	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
more_highschool	9.671839	1.163464	8.31	0.000	7.386866	11.95681
female	-3.728292	1.591217	-2.34	0.019	-6.853346	-.603238
interaction	-4.477087	2.024681	-2.21	0.027	-8.453438	-.5007365
_cons	18.01175	.7035701	25.60	0.000	16.62998	19.39352

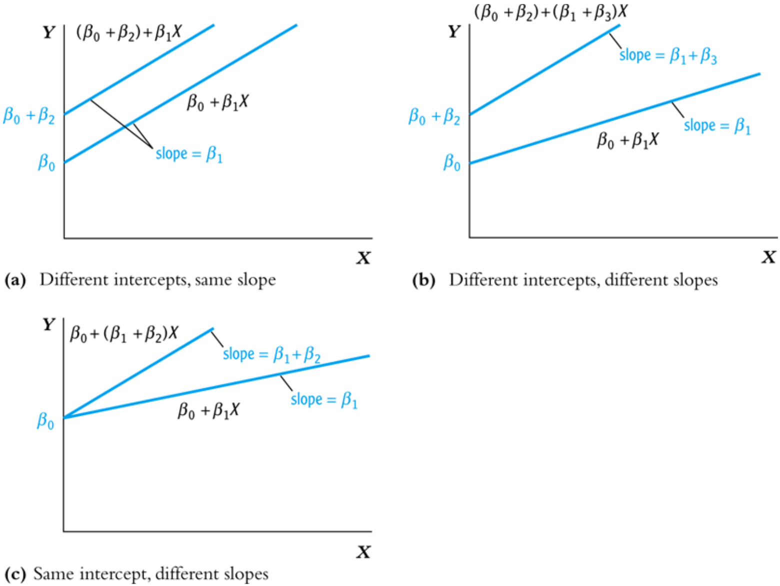
- $\widehat{\beta}_0 = 18.01$ is average hourly earnings for men with a high school degree or less
- $\widehat{\beta}_0 + \widehat{\beta}_1 = 18.01 + 9.67 = 27.68$ is average hourly earnings for men with more than a high school degree
- $\widehat{\beta}_0 + \widehat{\beta}_2 = 18.01 - 3.72 = 14.29$ is average hourly earnings for women with a high school degree or less
- $\widehat{\beta}_0 + \widehat{\beta}_1 + \widehat{\beta}_2 + \widehat{\beta}_3 = 18.01 + 9.67 - 3.72 - 4.48 = 19.48$ is average hourly earnings for women with more than a high school degree

Interaction between a continuous and a binary variable

- Consider the model $Y_i = \beta_0 + \beta_1 X_{1i} + u_i$ with X_{1i} the continuous variable years of education.
- The association between years of education and earnings might differ between men and women



Interaction between a continuous and a binary variable



Interaction between a continuous and a binary variable

- Consider the following regression model with

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 D_{2i} + \beta_3 (X_{1i} \times D_{2i}) + u_i$$

with X_{1i} years of education and D_{2i} the binary variable that equals 1 for women and 0 for men.

```
Linear regression                                      Number of obs =       602
                                                          F(  3,   598) =     49.24
                                                          Prob > F      =    0.0000
                                                          R-squared     =    0.2305
                                                          Root MSE      =    11.103
```

hourlyearn~s	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
education	2.307982	.232958	9.91	0.000	1.850467	2.765498
female	-1.961744	6.225225	-0.32	0.753	-14.18771	10.26422
interaction	-.3215831	.45654	-0.70	0.481	-1.2182	.5750335
_cons	-7.840784	3.038343	-2.58	0.010	-13.8079	-1.873664

Interaction between a continuous and a binary variable

- Is the effect of education on earnings significantly different between men and women?

$$H_0 : \beta_3 = 0 \quad vs \quad H_1 : \beta_3 \neq 0$$

- Compute the t-statistic:

$$t = \frac{-0.322}{0.457} = -0.70$$

- $|t| = 0.70 < 1.96 \longrightarrow H_0$ not rejected at 5% significance level

- Does gender matter?

```
. test female=interaction=0

 ( 1)   female - interaction = 0
 ( 2)   female = 0

       F(  2,   598) =    25.23
            Prob > F =     0.0000
```

Interaction between 2 continuous variables

- Multiple regression model with two continuous variables:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$$

with X_{1i} years of education and X_{2i} age (in years).

```
Linear regression                                      Number of obs =       602
                                                          F(  2,    599) =    56.78
                                                          Prob > F      =   0.0000
                                                          R-squared     =   0.1757
                                                          Root MSE      =   11.483
```

hourlyearn~s	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
education	2.1041	.2036148	10.33	0.000	1.704214	2.503986
age	.1024648	.040181	2.55	0.011	.0235521	.1813776
_cons	-11.96041	3.22028	-3.71	0.000	-18.28482	-5.636

- Earnings increase with age, estimated coefficient on age is significantly different from zero at 5% level
- Does the effect of education on earnings depend on age?

Interaction between 2 continuous variables

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 (X_{1i} \times X_{2i}) + u_i$$

```
Linear regression                               Number of obs =        602
                                                  F(  3,    598) =    38.28
                                                  Prob > F       =   0.0000
                                                  R-squared      =   0.1777
                                                  Root MSE       =   11.478
```

hourlyearn~s	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
education	1.195204	.7259149	1.65	0.100	-.2304487	2.620856
age	-.1857963	.2091314	-0.89	0.375	-.5965175	.2249249
interaction	.0210578	.0161605	1.30	0.193	-.0106804	.052796
_cons	.4588621	9.413582	0.05	0.961	-18.02884	18.94656

- Does the effect of education on earnings depend on age?
  - $\widehat{\beta}_3 = 0.021$
  - Compute the t-statistic:

$$t = \frac{0.021}{0.016} = 1.30$$

  - The coefficient on the interaction term between education and age is not significantly different from zero (at a 1%, 5% and 10% significance level)

Concluding remarks

- We discussed nonlinear regression models

$$Y_i = f(X_{1i}, X_{2i}, ..., X_{ki}) + u_i$$

- Models that are nonlinear in the independent variables are variants of the multiple regression model
  - and can therefore be estimated by OLS,
  - t- and F-tests can be used to test hypothesis about the values of the coefficients,
  - provided that the OLS assumptions hold (topic of next week)
- Often difficult to decide which (non)linear model best fits the data
  - Make a scatter plot
  - Use t- and F-tests
  - Use economic knowledge and intuition.