

SEMINAR 1 $\bar{x} = \frac{1}{n} \sum x_i$

$$\textcircled{A} 1. \sum (x_i - \bar{x}) = \sum x_i - n\bar{x}$$

$$\Rightarrow \sum x_i - \sum x_i = \underline{\underline{0}}$$

$$2. \sum (x_i - \bar{x})(y_i - \bar{y})$$

$$= \sum x_i y_i - \sum x_i \bar{y} - \sum y_i \bar{x} + \sum \bar{x} \bar{y}$$

$$= \sum x_i y_i - n\bar{y}\bar{x} - n\bar{y}\bar{x} + n\bar{x}\bar{y}$$

$$= \sum x_i y_i - \bar{x} \sum y_i$$

$$= \underline{\underline{\sum (x_i - \bar{x}) y_i}}$$

$$3. \sum x_i y_i - \bar{y} \sum x_i$$

$$= \underline{\underline{\sum (y_i - \bar{y}) x_i}}$$

Useful summation rules

$$\text{I) } \sum a = na$$

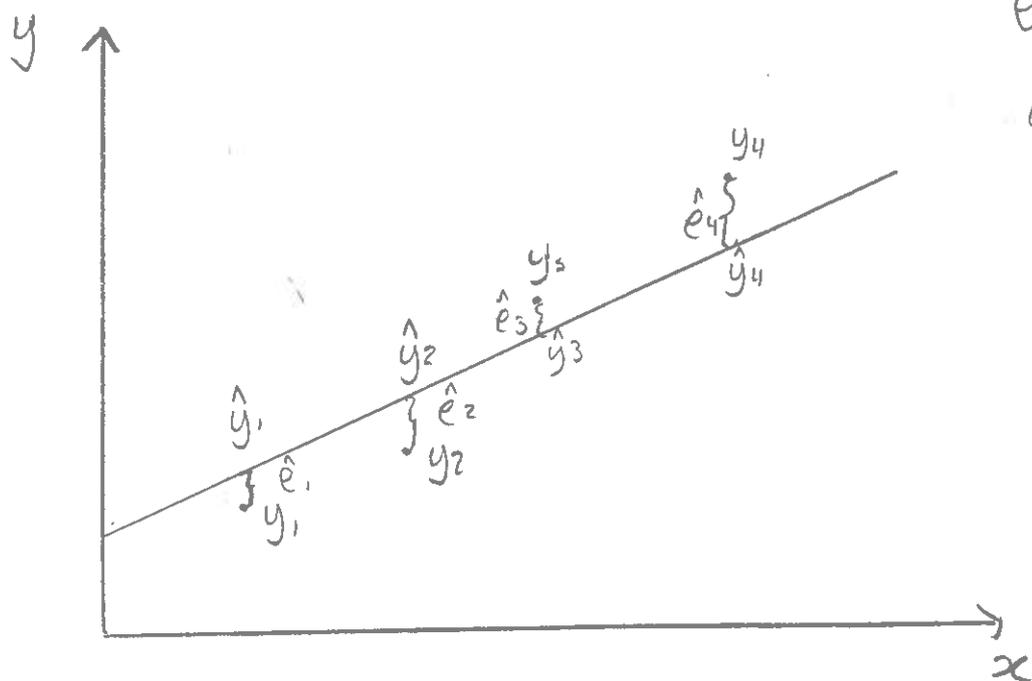
$$\text{II) } \sum aX = a \sum X$$

$$\text{III) } \sum (X + Y) = \sum X + \sum Y$$

We can also combine these.

② By the name "least squares", we get the idea. We want the squared distances to some fitted line to be as small as possible. As we saw in ①, we must square these distances (else their sum is zero).

$\hat{e}_1, \dots, \hat{e}_4$
are residuals



The residuals are the vertical distance between the real and the fitted value, so:

$$\hat{e}_i = y_i - \hat{y}_i = y_i - b_1 - b_2 x_i$$

Note that $\hat{y}_i = b_1 + b_2 x_i$ is the mean of y_i

$$\Rightarrow \sum_{i=1}^n (y_i - \hat{y}_i) = 0 \quad \text{So we must square it.}$$

$$\Rightarrow \text{Sum of only positive numbers} = > 0$$

(Ref: Appendix 2A)

○ To find these, do the minimization

$$\min_{\beta_0, \beta_1} S(\beta_0, \beta_1) = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2$$

FOC

○ I)
$$-2 \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = 0$$

$$= -2 \sum Y_i + 2n \hat{\beta}_0 + 2 \hat{\beta}_1 \sum X_i = 0 \quad | \times \left(-\frac{1}{2n}\right)$$

$$= \bar{Y} - \hat{\beta}_0 - \hat{\beta}_1 \bar{X} = 0 \quad (5)$$

○ II)
$$-2 \sum (X_i (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)) = 0$$

$$= \sum X_i Y_i - \hat{\beta}_0 \sum X_i - \hat{\beta}_1 \sum X_i^2 = 0 \quad (6)$$

By using the following rule backwards:

$$\sum_{i=1}^n a X_i = a \sum_{i=1}^n X_i$$

(3)

$$(5) \Rightarrow \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

We now use the trick to multiply (5) by $n \sum x_i$ and (6)' by n and then subtract (5)' from (6)'

$$(5)' \quad n \hat{\beta}_0 \sum x_i = \sum x_i \sum y_i - \hat{\beta}_1 (\sum x_i)^2$$

$$(6)' \quad n \sum x_i y_i - n \hat{\beta}_0 \sum x_i - n \hat{\beta}_1 \sum x_i^2 = 0$$
$$- n \hat{\beta}_0 \sum x_i = n \hat{\beta}_1 \sum x_i^2 - n \sum x_i y_i$$

$$\Rightarrow n \hat{\beta}_1 \sum x_i^2 - \hat{\beta}_1 (\sum x_i)^2 = n \sum x_i y_i - \sum x_i \sum y_i$$

$$\Rightarrow \hat{\beta}_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

- We now need to make this pretty by introducing the "deviation from mean" form. To do this, divide by n to get

$$\hat{\beta}_1 = \frac{\sum X_i Y_i - \frac{1}{n} \sum X_i \sum Y_i}{\sum X_i^2 - \frac{1}{n} (\sum X_i)^2}$$

Then remember that:

$$\begin{aligned} \sum (X_i - \bar{X})^2 &= \sum X_i^2 - n\bar{X}^2 = \\ &= \sum X_i^2 - \bar{X} \sum X_i = \sum X_i^2 - \frac{(\sum X_i)^2}{n} \end{aligned}$$

○ And similarly:

$$\begin{aligned} \sum (X_i - \bar{X})(Y_i - \bar{Y}) &= \sum X_i Y_i - n\bar{X}\bar{Y} \\ &= \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n} \end{aligned}$$

$$\Rightarrow \hat{\beta}_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

QB 1.

$$(7) \min_{\alpha, \beta_1} S(\alpha, \beta_1) = \sum_{i=1}^n (Y_i - \alpha - \beta_1 (X_i - \bar{X}))^2$$

$$\frac{\partial S}{\partial \alpha} = -2 \left(\sum (Y_i - \hat{\alpha} - \hat{\beta}_1 (X_i - \bar{X})) \right) = 0$$

$$\frac{\partial S}{\partial \beta_1} = -2 \left(\sum [(Y_i - \hat{\alpha} - \hat{\beta}_1 (X_i - \bar{X})) (X_i - \bar{X})] \right) = 0$$

We find FOC giving us:

$$\sum Y_i (X_i - \bar{X}) - \hat{\alpha} \underbrace{\sum (X_i - \bar{X})}_{=0} - \hat{\beta}_1 \sum (X_i - \bar{X})^2 = 0$$

$$\Rightarrow \hat{\beta}_1 = \frac{\sum (X_i - \bar{X}) Y_i}{\sum (X_i - \bar{X})^2}$$

And also:

$$\sum [Y_i - \hat{\alpha} - \hat{\beta}_1 \sum (X_i - \bar{X})] = 0$$

$$\Rightarrow \hat{\alpha} = \frac{1}{n} \sum Y_i - \frac{1}{n} \frac{\sum (X_i - \bar{X}) Y_i}{\sum (X_i - \bar{X})^2} \times \underbrace{\sum (X_i - \bar{X})}_{=0}$$

$$\Rightarrow \underline{\underline{\hat{\alpha} = \bar{Y}}}$$

2. We see that :

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$\hat{\alpha} = \bar{Y}$$

So we easily spot that $\hat{\alpha} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X}$

3. The sample correlation measures the strength of linear association. Since the correlation is positive, this indicates a positive relation between X and Y , hence $\hat{\beta}_1 > 0$. Since we know that:

$$r_{X,Y} = \frac{S_{xy}}{S_x S_y} \quad \text{and} \quad \hat{\beta}_1 = \frac{S_{xy}}{S_x^2}$$

We easily see that since $S_x S_y > 0 \Rightarrow S_{xy} > 0$ when covariance is positive. But this in turn must imply $\hat{\beta}_1 > 0$.

④ We now have:

$$X_i = \beta_0' + \beta_1' Y_i + \epsilon$$

So by analogy, we get

$$\hat{\beta}_1' = \frac{\sum (X_i - \bar{X}) Y_i}{\sum (Y_i - \bar{Y})^2} = \frac{S_{xy}}{S_y^2}$$

meaning $\hat{\beta}_1' \neq \frac{1}{\hat{\beta}_1}$ as one could imagine.

They are not the same.

⑤

In general, we know that

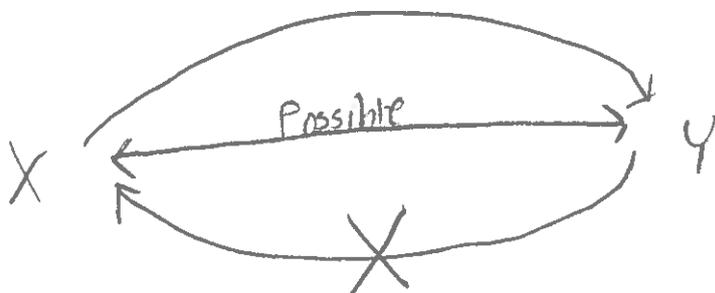
$$R_{X,Y}^2 = \hat{\beta}_1 \times \hat{\beta}_1'$$

$$\Rightarrow \text{Meaning } \hat{\beta}_1' = \frac{R^2}{\hat{\beta}_1}$$

$$\text{First sample: } \frac{0,5^2}{0,5} = \underline{\underline{0,5}}$$

$$\text{Second sample: } \frac{0,9^2}{0,5} = \underline{\underline{1,62}}$$

⑥ Since $\hat{\beta}_1$ is constant despite the break in $R_{X,Y}$, we have evidence of one-way causality (From X to Y).



QC STATA - handout!

Q So what really happened here?

Let's start by looking at $\hat{\beta}_0$:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$\hat{\beta}_0 = \frac{1}{n} \sum Y_i - \hat{\beta}_1 \frac{1}{n} \sum X_i$$

$$\Rightarrow \hat{\beta}'_0 = \frac{1}{n} \sum (Y_i - \bar{Y}) - \hat{\beta}_1 \frac{1}{n} \sum (X_i - \bar{X}) = 0$$

While for the coefficient:

$$\hat{\beta}_1 = \frac{\sum (X_i - \bar{X}) Y_i}{\sum (X_i - \bar{X})^2}$$

$$\Rightarrow \hat{\beta}'_1 = \frac{\sum \left[(X_i - \bar{X}) \left(\frac{1}{n} \sum (X_i - \bar{X}) \right) (Y_i - \bar{Y}) \right]}{\sum \left[(X_i - \bar{X}) \left(\frac{1}{n} \sum (X_i - \bar{X}) \right)^2 \right]} = 0$$

$$\Rightarrow \hat{\beta}_1 = \beta_1 = \frac{\sum (X_i - \bar{X}) Y_i}{\sum (X_i - \bar{X})^2}$$

Seminar 2

(QA) (1) $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, $i = 1, 2, \dots, n$

$$\text{Var}(\varepsilon_i) = \sigma^2 \quad \forall i$$

① Find $\text{var}(\hat{\alpha})$ and $\text{var}(\hat{\beta}_1)$

When $Y_i = \alpha + \beta_1 (X_i - \bar{X}) + \varepsilon_i$

We remember that

$$\hat{\beta}_1 = \sum \left[\frac{(X_i - \bar{X})}{\sum (X_i - \bar{X})^2} \right] Y_i$$

Note: $\sum (X_i - \bar{X})^2$ is a fixed number, hence it can be taken outside the exterior summation operator, and we're back at the "usual" form.

Now let $w_i \equiv \frac{(X_i - \bar{X})}{\sum (X_i - \bar{X})^2}$ for simplicity

$$\Rightarrow \sum w_i Y_i = \hat{\beta}_1$$

$$\Rightarrow \hat{\beta}_1 = \sum w_i (\beta_0 + \beta_1 X_i + \varepsilon_i)$$

$$\text{Now: } \hat{\beta}_1 = \sum [w_i \beta_0 + \beta_1 w_i x_i + w_i \varepsilon_i]$$

$$\text{Then: } \sum w_i = \sum \left[\frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right]$$

$$= \frac{1}{\sum (x_i - \bar{x})^2} \underbrace{\sum (x_i - \bar{x})}_{=0} = 0$$

$$\text{White } \sum w_i x_i = \sum \left[\frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right] x_i = 1$$

$$\Rightarrow \hat{\beta}_1 = \beta_1 + \sum w_i \varepsilon_i$$

$$\text{Now } \text{var}(\hat{\beta}_1) = \text{var}(\beta_1 + \sum w_i \varepsilon_i)$$

$$\Rightarrow \text{var}(\hat{\beta}_1) = \text{var}(\sum w_i \varepsilon_i)$$

$$\Rightarrow \text{var}(\hat{\beta}_1) = \sum w_i^2 \text{var}(\varepsilon_i) + 2 \sum_{i < j}^n w_i w_j \underbrace{\text{cov}(\varepsilon_i, \varepsilon_j)}_{=0}$$

$$\Rightarrow \text{var}(\hat{\beta}_1) = \sum w_i^2 \text{var}(\varepsilon_i) = \sum w_i^2 \sigma^2$$

And then:

$$\sum w_i^2 = \left[\frac{\sum (X_i - \bar{X})^2}{(\sum (X_i - \bar{X})^2)^2} \right] = \frac{1}{\sum (X_i - \bar{X})^2}$$

Hence:
$$\text{var}(\hat{\beta}_1) = \frac{\sigma^2}{\sum (X_i - \bar{X})^2}$$

For $\hat{\alpha} = \bar{Y}$, we have $\bar{Y} = \frac{1}{n} \sum Y_i$

$$\Rightarrow \text{var}(\hat{\alpha}) = \text{var}\left(\frac{1}{n} \sum Y_i\right)$$

$$= \frac{1}{n^2} \text{var}\left(\sum Y_i\right) = \frac{1}{n^2} \text{var}\left(\sum \varepsilon_i\right)$$

$$= \frac{1}{n^2} \sum \text{var}(\varepsilon_i) = \frac{1}{n^2} \sum \sigma^2 = \frac{1}{n^2} n \sigma^2 = \underline{\underline{\frac{\sigma^2}{n}}}$$

Remember here that:

$$\text{var}(\varepsilon_i + \varepsilon_j) = 1^2 \text{var} \varepsilon_i + 1^2 \text{var} \varepsilon_j + 2 \text{cov}(\varepsilon_i, \varepsilon_j) = 0$$

$$\Rightarrow \sum_i \text{var}(\varepsilon_i) = \text{var} \sum \varepsilon_i$$

②

RULE

$$\text{var}(aX + bY) = a^2 \text{var}X + b^2 \text{var}Y + 2abcov(X, Y)$$

$$\hat{\beta}_0 = \hat{\alpha} - \hat{\beta}_1 \bar{X} \quad (\text{from last seminar}).$$

$$\text{var}(\hat{\beta}_0) = \text{var}(\hat{\alpha}) + \bar{X}^2 \text{var}(\hat{\beta}_1) - 2\bar{X} \text{cov}(\hat{\alpha}, \hat{\beta}_1)$$

$$\text{var}(\hat{\beta}_0) = \frac{\sigma^2}{n} + \bar{X}^2 \left(\frac{\sigma^2}{\sum (X_i - \bar{X})^2} \right)$$

$$\text{var}(\hat{\beta}_0) = \frac{\sigma^2}{n} \left(1 + \frac{n\bar{X}^2}{\sum (X_i - \bar{X})^2} \right)$$

Let us make it like 2.14 in HGL:

Work with parenthesis:

$$\frac{\sum (X_i - \bar{X})^2 + n\bar{X}^2}{\sum (X_i - \bar{X})^2} = \frac{\sum X_i^2 - 2\sum X_i \bar{X} + n\bar{X}^2 + n\bar{X}^2}{\sum (X_i - \bar{X})^2}$$

$$\frac{\sum x_i^2 - 2n\bar{x}^2 + 2n\bar{x}^2}{\sum (x_i - \bar{x})^2}$$

$$\Rightarrow \text{var}(\hat{\beta}_0) = \sigma^2 \left[\frac{\sum x_i^2}{n \sum (x_i - \bar{x})^2} \right]$$

For the covariance $\text{cov}(\hat{\beta}_0, \hat{\beta}_1)$:

Rule: $\text{cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$

Then $\text{cov}(\hat{\beta}_0, \hat{\beta}_1) = E[(\hat{\beta}_0 - E(\hat{\beta}_0))(\hat{\beta}_1 - E(\hat{\beta}_1))]$

$$= E[(\hat{\beta}_0 - \beta_0)(\hat{\beta}_1 - \beta_1)]$$

Now $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ $\hat{\beta}_0 - E(\hat{\beta}_0) = \bar{y} - \hat{\beta}_1 \bar{x} - \bar{y} + \beta_1 \bar{x} = -\bar{x}(\hat{\beta}_1 - \beta_1)$

$E(\hat{\beta}_0) = \bar{y} - \beta_1 \bar{x}$ $= -\bar{x}(\hat{\beta}_1 - \beta_1)$

Note: $E[(\hat{\beta}_1 - \beta_1)^2] = 0$

giving $(\hat{\beta}_0 - E(\hat{\beta}_0)) = -\bar{x}(\hat{\beta}_1 - \beta_1)$

$$\Rightarrow E[(\hat{\beta}_0 - \beta_0)(\hat{\beta}_1 - \beta_1)] = -\bar{x} E[(\hat{\beta}_1 - \beta_1)^2]$$

$$\Rightarrow \text{cov}(\hat{\beta}_0, \hat{\beta}_1) = -\bar{x} \text{var}(\hat{\beta}_1) = \left(\frac{-\bar{x}}{\sum (x_i - \bar{x})} \right) \sigma^2$$

QB

① So we still have that

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \varepsilon_i$$

So as we understand, the intercept is unaffected. Hence

$$E(Y_i) = \begin{cases} \beta_1 + \beta_2 & \text{if } X_i = 1 \\ \beta_1 & \text{if } X_i = 0 \end{cases}$$

So we estimate $\hat{\beta}_0$ as usual and:

$$\hat{\beta}_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \bar{Y}_1 - \bar{Y}_0$$

(Difference estimator)

PROOF

② So we have:

Y : Number of damages on New Year's Eve

$$X: (\text{dummy}) = \begin{cases} 1 & \text{if illegal} \\ 0 & \text{if legal} \end{cases}$$

\bar{Y}_1 : Average when illegal: 14

\bar{Y}_0 : Average when legal: 20

$$\Rightarrow \hat{\beta}_1 = \bar{Y}_1 - \bar{Y}_0 = 14 - 20 = \underline{-6}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 17 + 3 = \underline{20}$$

$$\underline{\underline{Y_i = 20 - 6X_i}} \quad (+ \epsilon_i)$$

Probably significant
in this model, due to few obs!

③ If illegal \dots , $X_i = 1$ so $E(Y_i) = \hat{\beta}_0 + \hat{\beta}_1 = \underline{\underline{14}}$

BUT \rightarrow

The standard deviation on the coefficient in this model should be large, since there are so few observations.

Hence, a prediction interval would be wide.

There might also be other variables holding some explanatory power (like the weather).

Let $E(\epsilon_j) = 0$, $\text{var}(\epsilon_j) = \sigma^2$, $\text{cov}(\epsilon_j, \epsilon_i) = 0 \forall i$

$$\text{Now: } \hat{Y}_j = E(\hat{Y}_j) = \hat{\mu}_j = \hat{\beta}_0 + \hat{\beta}_1 X_j$$

We now have the prediction error as

$$f_j = Y_j - \hat{\mu}_j \quad \text{where } E(f_j) = 0$$

$$\text{var}(f_j) = \sigma^2 + \text{var}(\hat{\mu}_j) - 2\text{cov}(\hat{\mu}_j, \epsilon_j) = 0$$

$$\text{So we get: } \hat{\mu}_i \pm \widehat{\text{se}}(\hat{f}_j) \times t_{\frac{1-\alpha}{2}, (n-2)} = [\hat{\mu}_L, \hat{\mu}_U]$$

$$\text{where } \widehat{\text{se}}(\hat{f}_j) = \sqrt{\hat{\sigma}^2 + \text{var}(\hat{\mu}_i)}$$

$$= \sqrt{\hat{\sigma}^2 \left(1 + \frac{1}{n} + \frac{(X_j - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right)}$$

○ We can conclude that since :

1) $\hat{\sigma}^2$ is large

2) n is small

○ \Rightarrow Our prediction interval would be wide.

Proof

Note that the numerator is equal:

$$\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y}) = \sum (Y_i - \bar{Y}) X_i - \bar{X} \sum (Y_i - \bar{Y})$$

$$= \sum (Y_i - \bar{Y}) X_i$$

$$= \sum Y_i X_i - \bar{Y} \sum X_i$$

$$= N_1 \bar{Y}_1 - N_1 \bar{Y}$$

$$= N_1 \bar{Y}_1 - N_1 (N_1 \bar{Y}_1 + N_0 \bar{Y}_0) / N$$

$$= \frac{N_0 N_1}{N} (\bar{Y}_1 - \bar{Y}_0)$$

The denominator:

$$\sum (X_i - \bar{X})^2 = \sum X_i^2 - 2\bar{X} \sum X_i + \sum \bar{X}^2$$

$$= \sum X_i^2 - 2\bar{X} N_1 + N \bar{X}^2$$

$$= N_1 - 2 \frac{N_1}{N} N_1 + N \left(\frac{N_1}{N} \right)^2 = \frac{N_0 N_1}{N}$$

$$\Rightarrow \frac{\cancel{N_0 N_1}}{N} (\bar{Y}_1 - \bar{Y}_0) / \frac{\cancel{N_0 N_1}}{N} = (\bar{Y}_1 - \bar{Y}_0) \quad \text{End of proof.}$$

Seminar 3

QA Two models

$$(1) Y_i = \beta_0 + \epsilon_i$$

$$i = 1, 2, \dots, n$$

$$(2) Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2)$$

Now,

$$(1) \Rightarrow \frac{\sum_{i=1}^n \hat{\epsilon}_i^2}{\sigma^2} \sim \chi^2 (n-1)$$

Proof

Write (1) as vectors:

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = \hat{\beta}_0 \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + \begin{bmatrix} Y_1 - \hat{\beta}_0 \\ \vdots \\ Y_n - \hat{\beta}_0 \end{bmatrix}$$

Rotationally symmetric

\Rightarrow can be

turned upside/down.

Restricted to lie in the space where the sum of the components $= 0$.

(Recall that the sum of a demeaned variable $= 0$)

Since its distribution is rotationally symmetric

and it's an $(n-1)$ -dimensional subspace

Reason: $\hat{\beta}_0$ "removes" one dimension

and has expected value 0, the distribution

of the sum of squares of its entries

is the same as the distribution of the

squares of the norms of $(n-1)$ independent

random variables ^(length of vectors) each $\sim N(0, \sigma^2)$

Hence $\sum_{i=1}^n \frac{\hat{\xi}_i^2}{\sigma^2} \sim \chi^2(n-1)$ End of proof.
by the Hahn-Banach - theorem.

The main point: Number of degrees of freedom goes down by 1 for every parameter to be estimated (and now we know why \rightarrow due to one less dimension in the subspace).

②

We know that

$$\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma^2}{\sum (x_i - \bar{x})^2}\right)$$

with reference to
last seminar

Now use Slutsky's lemma to get

$$\hat{\beta}_1 - \beta_1 \sim N\left(0, \frac{\sigma^2}{\sum (x_i - \bar{x})^2}\right)$$

i.l.v.

$$\Rightarrow Z = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\sigma^2}{\sum (x_i - \bar{x})^2}}} \sim N(0, 1)$$

%

Now we know that:

$$P(-c \leq Z \leq c) = 1 - \alpha$$

$$\Rightarrow P\left(-c \leq \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\sigma^2 / \sum (x_i - \bar{x})^2}} \leq c\right) = 1 - \alpha$$

$$\Rightarrow \left[\hat{\beta}_1 \pm c \sqrt{\sigma^2 / \sum (x_i - \bar{x})^2} \right] = 1 - \alpha$$

And although σ^2 is unknown, we can estimate it.

$$\hat{\epsilon}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i \quad \text{and}$$

$$\hat{\sigma}^2 = \frac{\sum \hat{\epsilon}_i^2}{(n-2)} \quad \text{Now substituting}$$

into $Z \sim N(0,1)$ yields a t -distribution with $(n-2)$ degrees of freedom:

$$t = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\hat{\sigma}^2 / \sum (X_i - \bar{X})^2}} = \frac{\hat{\beta}_1 - \beta_1}{se(\hat{\beta}_1)} \sim t_{(n-2)}$$

○ (QB)

$$3) \text{BNPcap}_t = A \exp(\gamma \text{Trend}_t + \varepsilon_t)$$

$$\Leftrightarrow B = A \exp(\gamma T + \varepsilon_t)$$

$$\ln B = \ln A \exp(\gamma T + \varepsilon_t)$$

$$\ln B = \underbrace{\ln A}_{\beta_0} + \underbrace{(\gamma T + \varepsilon_t)}_{\beta_1 X_i}$$

So if we use $\ln \text{bnpcap}$ as regressand and trend as regressor, we get γ as the

○ Coefficient!

$$5) \text{ Let } \text{BNPcap} = B$$

$$L \text{BNPcap} = \ln B$$

$$L \text{BNPcap}_{-1} = \ln B_{-1}$$

$$D L \text{BNPcap} = D$$

$$\text{Trend} = T$$

, then:

$$\text{Trend}_{-1} = T_{-1}$$

$$D = \ln B - \ln B_{-1}$$

$$\Rightarrow D = \ln A + (\gamma T + \epsilon_i) - \ln A - (\gamma T_{-1} + \epsilon_j)$$

$$\underline{\underline{D = \gamma(T - T_{-1}) + u_t}}$$

○ 6) A model which suits the description is

$$Y_t = \beta_0 + \gamma X_t + \epsilon_t$$

AR(1)-process

where $\epsilon_t = \rho \epsilon_{t-1} + v_t$ $-1 < \rho < 1$

By assumption, v_t is "white noise" with

full classical assumptions, and $\text{var}(\epsilon_t) = \text{var}(\epsilon_{t-1})$

Also $\text{cov}(v_t, \epsilon_{t-1}) = 0$ | Hence $\text{var}(\epsilon_t) = \rho^2 \text{var}(\epsilon_{t-1}) + \sigma_v^2$
 $\Rightarrow \text{var}(\epsilon_t) = \frac{\sigma_v^2}{1-\rho^2}$

Now: $\text{cov}(\epsilon_t, \epsilon_{t-1}) = E[(\epsilon_t - 0)(\epsilon_{t-1} - 0)] = E[\epsilon_t \epsilon_{t-1}]$

$$= E[(\rho \epsilon_{t-1} + v_t) \epsilon_{t-1}]$$

$$= \rho E[\epsilon_{t-1} \epsilon_{t-1}] = \rho \text{var} \epsilon_{t-1}$$

Hence $\text{corr}(\epsilon_t, \epsilon_{t-1}) = \rho$

over-estimate var is better than underestimating! We want P(type I error) as small as possible.

Now, since we want classical disturbance, we get?

$$\Delta Y_t = \gamma \Delta X_t + \Delta \epsilon_t$$

and $\Delta \epsilon_t = u_t$ when $\rho=1$, but

if $\rho < 1 \Rightarrow$ neg. corr \Rightarrow over-estimate var

So we have all classical properties, but lose one observation (as we have seen in Stata).

So using this, we can construct a 95% CI the following way:

Recall from QA2 that:

$$\left[\hat{\beta}_1 \pm c \sqrt{\hat{\sigma}^2 / \sum (x_i - \bar{x})^2} \right] = 1 - \alpha$$

which turned out to be:

$$\left[\hat{y} \pm t_{0,975,63} \text{se}(\hat{y}) \right] = 0,95$$

$$= \left[0,0296377 \pm 1,96 \times 0,002658 \right] = \left[0,024, 0,035 \right]$$

Interpretation

When constructing many such intervals using the same method, we would expect that 95% of the time, we would get an interval which contains the true value. We trust the theory, and not the single confidence interval.

The probability of the true value being included in a single CI is either 1 or 0! (Not 0,95!!!)

7

Recall from last seminar that:

$$\left[\hat{Y}_0 \pm t_c \text{se}(f) \right] \approx (1-\alpha)$$

$$\text{Where } \text{se}(f) = \sqrt{\hat{\sigma}^2 \left[1 + \frac{1}{N} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right]}$$

$$= \sqrt{\hat{\sigma}^2 + \frac{\hat{\sigma}^2}{N} + (x_0 - \bar{x})^2 \text{var}(\hat{\beta}_1)}$$

In our case we get:

$$\text{se}(f) = \sqrt{0,004265084 + \frac{0,004265084}{64} + (182 - 149,5)^2 \times 0,004419^2}$$

$$\text{se}(f) = 0,07047248875$$

While:

$$\hat{Y}_0 = 7,471852 + 0,0305388 \times 182 = 13,03$$

$$\text{Hence: } \left[\hat{Y}_0 \pm t_{\frac{1-\alpha}{2}} \text{se}(f) \right] = \left[13,03 \pm 1,96 \times 0,07047248 \right]$$

$$\approx \underline{\underline{[12,89, 13,17]}}$$

8) Predicted BNP cap is

$$e^{13,03} \approx 455\,886,89$$

SSB: 2012 = 581 000

SEMINAR 4

QA

$$Y = a + bX + Z$$

$$E(Z) = 0, \text{ cov}(X, Z) = 0$$

①

$$\text{var}(Y) = \text{var}(a + bX + Z)$$

= 0

$$\text{var}(Y) = b^2 \text{var}(X) + \text{var}(Z) + 2b \text{cov}(X, Z)$$

$$E(Y - E(Y))^2 = b^2 E(X - E(X))^2 + E(Z - E(Z))^2$$

$$\Rightarrow b = \frac{E(Y - E(Y))}{E(X - E(X))} \quad \Bigg| \quad \times \frac{E(X - E(X))}{E(X - E(X))}$$

$$\Rightarrow b = \frac{\text{cov}(X, Y)}{\text{var}(X)}$$

②

$$E[E(Y|X)] = E(Y)$$

The law of iterated expectations ("double expectations")

Given knowledge of the outcome of X , and

hence conditional expectation of Y on X , does

not change our belief about Y .

One cannot use limited information to predict

the forecast error one would make if

one had superior information.

Proof: Recall that in the discrete case:

$$f(y) = P(Y=y_i) \quad f(x,y) = P(X=x_i \cap Y=y_i)$$

$$E(Y) = \sum_y y_i f(y_i) \quad f(y_i) = \sum_x f(x,y)$$

$$f(x_i, y_i) = f(y_i|x) f(x)$$

Hence;

$$E(Y) = \sum_y y_i \left[\sum_x f(y|x) f(x) \right]$$

p. 27 in HGL

$$= \sum_x \sum_y \left[y_i f(y|x) \right] f(x)$$

$$= \sum_x E(Y|X=x) f(x)$$

By assuming X is random

$$\Rightarrow \underline{\underline{E(Y) = E(E(Y|X))}}$$

RULE: Bayes' rule $\frac{P(Y=y_i \cap X=x_i)}{P(X=x_i)} = P(Y=y_i | X=x_i)$

③

a) $E(Y|X) = E(Y=y | X=x)$

$$\left. \begin{aligned} E(Y = -2 | X = -8) &= \frac{0,1}{0,1} = \underline{\underline{1}} \\ E(Y = 6 | X = -8) &= \frac{0}{0,1} = \underline{\underline{0}} \end{aligned} \right\} = 1$$

$$\left. \begin{aligned} E(Y = -2 | X = 0) &= \frac{0,5}{0,7} = \frac{\underline{\underline{5}}}{\underline{\underline{7}}} \\ E(Y = 6 | X = 0) &= \frac{0,2}{0,7} = \frac{\underline{\underline{2}}}{\underline{\underline{7}}} \end{aligned} \right\} = 1$$

$$\left. \begin{aligned} E(Y = -2 | X = 8) &= \frac{0,1}{0,2} = \frac{\underline{\underline{1}}}{\underline{\underline{2}}} \\ E(Y = 6 | X = 8) &= \frac{0,1}{0,2} = \frac{\underline{\underline{1}}}{\underline{\underline{2}}} \end{aligned} \right\} = 1$$

$$b) E(Y|X=x)$$

$$\Rightarrow E(Y|X=-8) = -2 \times 1 + 6 \times 0 = \underline{\underline{-2}}$$

$$E(Y|X=0) = -2 \times \frac{5}{7} + 6 \times \frac{2}{7} = \underline{\underline{\frac{2}{7}}}$$

$$E(Y|X=8) = -2 \times \frac{1}{2} + 6 \times \frac{1}{2} = \underline{\underline{2}}$$

$$c) E(E(Y|X)) = E(Y)$$

$$\Rightarrow -2 \times 0,1 + \frac{2}{7} \times 0,7 + 2 \times 0,2 = \underline{\underline{0,4}}$$

$$d) -2 \times 0,7 + 6 \times 0,3 = \underline{\underline{0,4}}$$

④

$$Y = E(Y|X) + \varepsilon$$

a) $\varepsilon = Y - E(Y|X)$

$$\varepsilon|X = [Y - E(Y|X)]|X$$

$$\varepsilon|X = Y|X - E(Y|X)|X$$

$$E(\varepsilon|X) = E(Y|X) - E(E(Y|X)|X) = \underline{\underline{0}}$$

b) From lectures: $\text{cov}(X, \varepsilon) = E(X\varepsilon) = 0$

$$E(Y|X) = \beta_0 + \beta_1 X$$

Let now;

$$Y = \beta_0 + \beta_1 X + \varepsilon, \text{ then:}$$

$$Y|X = (\beta_0 + \beta_1 X + \varepsilon) | X$$

$$Y|X = E(Y|X)|X + \varepsilon|X$$

$$E(Y|X) = E(E(Y|X)|X) + E(\varepsilon|X)$$

$$\Rightarrow \underline{\underline{E(\varepsilon|X) = 0}}$$

Now recall $\text{cov}(X, \varepsilon) = E((X - E(X))(E(\varepsilon - E(\varepsilon))))$

$$\begin{aligned} &= 0 \\ &= E(X\varepsilon) - E(X)E(\varepsilon) \end{aligned}$$

$$\text{And } E(Y) = E(E(Y|X) + E(\varepsilon))$$

$$\Rightarrow E(\varepsilon) = 0$$

$$\Rightarrow \underline{\underline{\text{cov}(X, \varepsilon) = 0}}$$

⑤ Recall: $\hat{\beta}_1 = \frac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2}$

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x}) (\beta_0 + \beta_1 x_i + \varepsilon_i)}{\sum (x_i - \bar{x})^2}$$

$$\hat{\beta}_1 = \frac{\beta_0 \sum (x_i - \bar{x}) + \beta_1 \sum (x_i - \bar{x})^2 + \sum (x_i - \bar{x}) \varepsilon_i}{\sum (x_i - \bar{x})^2}$$

$$\hat{\beta}_1 = \beta_1 + \frac{\sum (x_i - \bar{x}) \varepsilon_i}{\sum (x_i - \bar{x})^2}$$

$$\hat{\beta}_1 - \beta_1 = \frac{\sum (x_i - \bar{x}) \varepsilon_i}{\sum (x_i - \bar{x})^2}$$

$= W$

$$\Rightarrow E(\hat{\beta}_1 - \beta_1) = EW.$$

$$E(W) = E(E(W|X))$$

$$= E \left[\frac{\sum (x_i - \bar{x}) E(\varepsilon_i | X)}{\sum (x_i - \bar{x})^2} \right] = 0$$

$$\Rightarrow \underline{\underline{E(\hat{\beta}_1 - \beta_1) = 0}}$$

SEMINAR 5

$$QA) \quad r_{Mat} = r_t + \Phi_t \sigma_a \rho_{am}$$

r_t = riskless interest rate in time t .

Φ_t = aggregated market parameter

σ_a and ρ_{am} = parameters of the pdf

$$f(r_{at}, r_{mt})$$

r_{at} = individual asset return

r_{mt} = stock market return

ρ_{am} = correlation between r_{at} and r_{mt}

$$\Phi_t = \frac{E(r_{mt}) - r_t}{\sigma_m}$$

\Rightarrow By inserting we get

$$\mu_{at} = \frac{E(r_{mt}) - r_t}{\sigma_m} \sigma_a \rho_{am} + r_t$$

Now recall formula for correlation as;

$$\rho_{am} = \frac{\sigma_{am}}{\sigma_a \sigma_m}, \quad \text{thus:}$$

$$\mu_{at} = \frac{E(r_{mt}) - r_t}{\sigma_m} \sigma_a \frac{\sigma_{am}}{\sigma_a \sigma_m} + r_t$$

$$\Rightarrow \mu_{at} = r_t + \left[E(r_{mt}) - r_t \right] \frac{\sigma_{am}}{\sigma_m^2}$$

2)

The BETA of an asset describes the correlated volatility of an asset.

So if $\beta = 0$, movement of the asset is uncorrelated with movement of the benchmark. If $\beta > 0$, the asset moves in the same direction as the benchmark.

3)

β_a^{CAPM} a parameter in $E(r_{at} | r_{mt})$

To show this, use the well-known fact that the bivariate normal distribution has the pdf:

$$f(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{\left[-\frac{1}{2}\left(\frac{z}{1-\rho^2}\right)\right]}$$

$$z = \frac{(x_1 - \mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2}$$

Now if the conditional expectation function

is linear, that is, if $E(r_{at} | r_{mt}) = \mu_{r_{at}|r_{mt}} = \beta_0 + \beta_1 X$

where $\beta_1 = \beta_a^{\text{CAPM}}$, then we can assume

the normal distribution. Then;

$$f(r_{at}, r_{mt}) = \frac{1}{\sigma_a \sigma_m 2\pi \sqrt{1-\rho_{am}^2}} e^{\left[-\frac{1}{2}\left(\frac{z_a^2 - 2\rho_{am}z_a z_m + z_m^2}{1-\rho_{am}^2}\right)\right]}$$

○ Now

$$f(r_{at}) = \int_{-\infty}^{\infty} f(r_{at}, r_{mt}) dr_{mt}$$

which can be shown to take the form

$$f(r_{at}) = \frac{1}{\sigma_a \sqrt{2\pi}} e^{\left[-\frac{1}{2} \left(\frac{r_{at} - \mu_a}{\sigma_a}\right)^2\right]}$$

Insert into $f(r_{at}|r_{mt}) = \frac{f(r_{at}, r_{mt})}{f(r_{at})}$

gives:

$$f(r_{at}|r_{mt}) = \frac{1}{\sqrt{2\pi\sigma_m^2(1-\rho_{am}^2)}} e^{\left\{-\frac{\left(\frac{r_{mt} - \mu_m}{\sigma_m} - \frac{\sigma_{am}}{\sigma_a\sigma_m} \frac{r_{at} - \mu_a}{\sigma_a}\right)^2}{2(1-\rho_{am}^2)}\right\}}$$

$$= \frac{1}{\sqrt{2\pi\left(\sigma_m^2 - \sigma_m^2 \left(\frac{\sigma_{am}^2}{\sigma_a\sigma_m}\right)\right)}} e^{\left\{-\frac{1}{2} \left[\frac{r_{mt} - \left(\mu_m - \frac{\sigma_{am}}{\sigma_m^2} \mu_a + \frac{\sigma_{am}}{\sigma_a^2} r_a\right)}{\sigma_m^2 - \sigma_m^2 \left(\frac{\sigma_{am}^2}{\sigma_a\sigma_m}\right)}\right]^2\right\}}$$

Now write:

$$r_{mt} = \left(\mu_m - \frac{\sigma_{am}}{\sigma_m^2} \mu_a + \frac{\sigma_{am}}{\sigma_m^2} r_{at} \right)$$

as $r_m = \mu_{mla}$

where

$$\mu_{mla} = \underbrace{\mu_m - \frac{\sigma_{am}}{\sigma_m^2} \mu_a}_{=\beta_0} + \underbrace{\frac{\sigma_{am}}{\sigma_m^2} r_{at}}_{=\beta_1}$$

We also define

$$\sigma_{mla}^2 = \sigma_m^2 \left(1 - \frac{\sigma_{am}^2}{\sigma_a^2 \sigma_m^2} \right)$$

Hence;

$$f(r_{at} | r_{mt}) = \frac{1}{\sqrt{2\pi\sigma_{mla}^2}} e^{\left\{ -\frac{1}{2} \frac{[r_{mt} - \mu_{mla}]^2}{\sigma_{mla}^2} \right\}}$$

Which is a normal pdf with expectation μ_{mla} and variance σ_{mla}^2

Now $\mu_{alm} = E(r_{at} | r_{mt}) = \beta_0 + \beta_1 X$

where $\beta_0 = \mu_m - \frac{\sigma_{am}}{\sigma_m^2} \mu_a$

$$\beta_1 = \frac{\sigma_{am}}{\sigma_m^2} = \beta_a^{\text{CAPM}}$$

4) Specification of model:

So we have the model

$$Y = \beta_0 + \beta_a^{\text{CAPM}} X + \varepsilon_t \text{ where;}$$

$$Y = r_{at} \quad \text{and} \quad \varepsilon_t = r_{at} - E(r_{at} | r_{m0}, m_1, \dots, m_T)$$

(strict exogeneity)

$$X = r_{mt}$$

5) $H_0: \beta_a^{\text{CAPM}} = 0$ vs. $H_1: \beta_a^{\text{CAPM}} \neq 0$

Two-sided test. Hence reject H_0 if

$$|\theta_{\text{obs}}| > t_{\frac{1-\alpha}{2}, (n-2)} = t_{0,975, 179} \approx 1,96$$

$$|\theta_{\text{obs}}| = \left| \frac{\hat{\theta} - \theta}{\text{se}(\hat{\theta})} \right| = \left| \frac{-0,1731595 - 0}{0,1985225} \right| = \underline{0,87}$$

\Rightarrow Do not reject H_0 at 5% significance level.

$$p\text{-value: } (1 - 0,8078) \times 2 = \underline{0,3844}$$

○ QB)

2) Say Z_a and Y_a uncorrelated. Then information on Z_a has no influence on Y_a , hence $\partial \tilde{Y}_a / \partial Z_b = 0,58$ would be a

○ good estimate when $\tilde{Y}_a = E(Y_{ai} | Z_{ai}, Z_{bi})$

Let us instead assume Z_a is an omitted variable in the regression. Then the magnitude of the coefficient can be incorrectly estimated. Let us assume;

○
$$Y_a = \beta_0 + \beta_1 Z_b + \beta_2 Z_a + \epsilon$$

Now omitting Z_a from the regression is equivalent to imposing $Z_a = 0$. Let β_1^* be the estimate from Q1, then:

○
$$\text{bias}(\beta_1^*) = E(\beta_1^*) - \beta_1 = \beta_2 \frac{\widehat{\text{COV}}(Z_b, Z_a)}{\widehat{\text{Var}}(Z_b)}$$

So if Z_a and Z_b are uncorrelated, that is, if $\text{reg } Z_a \text{ } Z_b$ returns a coefficient equal to 0, or if Z_a and Y_a are uncorrelated (or both), then 0,58 would be a relevant estimate.

3) Frisch-Waugh - theorem.

$$\text{Let } E(Y|Z_{bi}) = \beta_0^* + \beta_1^* E(Z_{ai}|Z_{bi}) + \varepsilon_i^*$$

and define $E(Y|Z_{bi})$ and $E(Z_a|Z_{bi})$ as the residuals:

$$E(Y|Z_{bi}) = Y_i - \left[\bar{Y} + \frac{\hat{\sigma}_{Y Z_b}}{\hat{\sigma}_{Z_b}^2} (Z_{bi} - \bar{Z}_b) \right]$$

$$E(Z_a|Z_b) = Z_{ai} - \left[\bar{Z}_a + \frac{\hat{\sigma}_{Z_a Z_b}}{\hat{\sigma}_{Z_b}^2} (Z_{bi} - \bar{Z}_b) \right]$$

$$\hat{\beta}_1^* = \frac{\frac{1}{n} \sum E(Y|Z_{bi}) E(Z_a|Z_{bi})}{\frac{1}{n} \sum E(Z_a|Z_b)^2}$$

Now expand the numerator to get:

$$E(Y|Z_{bi}) E(Z_a|Z_{bi}) =$$

$$\left[(Y_i - \bar{Y}) - \frac{\hat{\sigma}_{Y, Z_b}}{\hat{\sigma}_{Z_b}^2} (Z_{bi} - \bar{Z}_b) \right] \left[(Z_{ai} - \bar{Z}_a) - \frac{\hat{\sigma}_{Z_a, Z_b}}{\hat{\sigma}_{Z_b}^2} (Z_{bi} - \bar{Z}_b) \right]$$

By collecting terms, we get:

$$= (Y_i - \bar{Y})(Z_{ai} - \bar{Z}_a) - \frac{\hat{\sigma}_{Z_a, Z_b}}{\hat{\sigma}_{Z_b}^2} (Y_i - \bar{Y})(Z_{bi} - \bar{Z}_b)$$

$$- \frac{\hat{\sigma}_{Y, Z_b}}{\hat{\sigma}_{Z_b}^2} (Z_{bi} - \bar{Z}_b)(Z_{ai} - \bar{Z}_a) + \frac{\hat{\sigma}_{Y, Z_b}}{\hat{\sigma}_{Z_b}^2} \frac{\hat{\sigma}_{Z_a, Z_b}}{\hat{\sigma}_{Z_b}^2} (Z_{bi} - \bar{Z}_b)^2$$

Hence:

$$\frac{1}{n} \sum E(Y|Z_{bi}) E(Z_a|Z_{bi}) =$$

$$\hat{\sigma}_{Y, Z_a} - \frac{\hat{\sigma}_{Z_a, Z_b}}{\hat{\sigma}_{Z_b}^2} \hat{\sigma}_{Y, Z_b} - \frac{\hat{\sigma}_{Y, Z_b}}{\hat{\sigma}_{Z_b}^2} \hat{\sigma}_{Z_a, Z_b} + \frac{\hat{\sigma}_{Y, Z_b}}{\hat{\sigma}_{Z_b}^2} \frac{\hat{\sigma}_{Z_a, Z_b}}{\hat{\sigma}_{Z_b}^2} \hat{\sigma}_{Z_b}^2$$

From remembering that: $\frac{1}{n} \sum (Y_i - \bar{Y})(Z_{ai} - \bar{Z}_a) = \hat{\sigma}_{Y, Z_a}$

$$= \hat{\sigma}_{y,za} - 2 \frac{\hat{\sigma}_{y,zb} \hat{\sigma}_{za,zb}}{\hat{\sigma}_{zb}^2} + \frac{\hat{\sigma}_{za,zb} \hat{\sigma}_{y,zb}}{\hat{\sigma}_{zb}^2}$$

$$= \frac{\hat{\sigma}_{y,za} \hat{\sigma}_{zb}^2 - \hat{\sigma}_{y,zb} \hat{\sigma}_{za,zb}}{\hat{\sigma}_{zb}^2}$$

Now the denominator:

$$\frac{1}{n} \sum E(z_a | z_b)^2 = \frac{1}{n} \sum \left((z_{ai} - \bar{z}_a) - \frac{\hat{\sigma}_{za,zb}}{\hat{\sigma}_{zb}^2} (z_{bi} - \bar{z}_b) \right)^2$$

$$= \hat{\sigma}_{za}^2 - 2 \frac{\hat{\sigma}_{za,zb}^2}{\hat{\sigma}_{zb}^2} - \left(\frac{\hat{\sigma}_{za,zb}}{\hat{\sigma}_{zb}^2} \right)^2 \hat{\sigma}_{zb}^2$$

$$= \frac{\hat{\sigma}_{za}^2 \hat{\sigma}_{zb}^2 - \hat{\sigma}_{za,zb}^2}{\hat{\sigma}_{zb}^2}$$

Hence;

$$\hat{\beta}_1^* = \frac{\hat{\sigma}_{y,za} \hat{\sigma}_{zb}^2 - \hat{\sigma}_{y,zb} \hat{\sigma}_{za,zb}}{\hat{\sigma}_{zb}^2} \Bigg/ \frac{\hat{\sigma}_{za}^2 \hat{\sigma}_{zb}^2 - \hat{\sigma}_{za,zb}^2}{\hat{\sigma}_{zb}^2}$$

$$\hat{\beta}_1^* = \frac{\hat{\sigma}_{y,za} \hat{\sigma}_{zb}^2 - \hat{\sigma}_{y,zb} \hat{\sigma}_{za,zb}}{\hat{\sigma}_{za}^2 \hat{\sigma}_{zb}^2 - \hat{\sigma}_{za,zb}^2} = \hat{\beta}_1$$

Now this is the theory behind the theorem, i.e. "the whole story". But we can do this step-wise by first regressing y_a on z_b , then z_a on z_b , and then finally regress the residuals from the first regression on the residuals from the second regression. This returns

$$\underline{\underline{\hat{\beta}_1 = 0,8418998}}$$

5) We need the disturbances
to have all classical properties
We also need the disturbances to
be normally distributed in order for
the statistical inference to be logically valid.

a) Test $H_0: \alpha = 0$ $H_1: \alpha \neq 0$

Two-sided test: Reject H_0 if

$$|\theta_{obs}| > t_{\frac{1-\alpha}{2}, (n-2)} = t_{0.975, 98} \approx 1,984$$

[In stata: `di invttail(98, 0, 025)`]

$$|\theta_{obs}| = \left| \frac{\hat{\theta} - \theta}{se(\hat{\theta})} \right| = \left| \frac{0,593733 - 0}{0,0462636} \right| = 12,83$$

\Rightarrow Reject H_0

b) p-value = 0,0000 \Rightarrow Reject H_0 .

6) From the previous question, we see that both coefficients are significantly different from 0

By investigating the Stata output

and the 95% CI from the univariate regressions, we see that the coefficients from the multivariate case are not included. Since also the explanatory power is significantly improved and the p-values still 0,0000, we should be left to believe that omitting either Z_a or Z_b would give an omitted variable bias.

To actually test this, write:

$$Y_i = \beta_0 + \beta_1 Z_a + \beta_2 Z_b + u_i$$

$$Z_a = \alpha_0 + \alpha_1 Z_b + v_i$$

$$Y_i = \psi_0 + \psi_1 Z_b + w_i$$

Then it's easy to see that;

$$\Psi_1 = \beta_2 \quad \text{iff} \quad \beta_1 = 0 \quad \text{and/or} \quad \alpha_1 = 0$$

So the test is;

$$H_0: \Psi_1 = \beta_1 \quad \text{vs.} \quad H_1: \Psi_1 \neq \beta_1$$

$$\Leftrightarrow H_0: \beta_1 = 0 \quad \text{and/or} \quad \alpha_1 = 0$$

$$H_1: \beta_1 \quad \text{and} \quad \alpha_1 \neq 0$$

$$7) \quad \text{corr}(z_a, z_b) = -0,5081$$

Multi collinearity not a problem, as

only exact collinearity violates OLS principle.

To further investigate this, we use the regression between z_a and z_b and check R^2

$R^2 = 0,2582$ which is not too high. A

typical threshold is 0,8, where we would

consider multicollinearity a problem.

SEMINAR 6

QA) $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$, $i = 1, 2, \dots, n$

1) We need the following specification of the model:

I) X_{ji} ($j = 1, 2$) ($i = 1, 2, \dots, n$) Stochastic means

$$\text{var}(X_{ji}) = \sigma_{x_j}^2 > 0 \quad \text{and} \quad \rho_{x_1 x_2}^2 < 1$$

II) $E(\varepsilon_i) = 0 \quad \forall i$ => Truly separate variables

III) $\text{var}(\varepsilon_i) = \sigma^2 \quad \forall i$

IV) $\text{cov}(\varepsilon_i, \varepsilon_j) = 0 \quad \forall i \neq j$

V) $\beta_0, \beta_1, \beta_2, \sigma^2$ constant parameters

VI) $\varepsilon_i \sim N(0, \sigma^2)$

2)

From the Frisch - Waugh - theorem,
we now know that finding the
partial effect of X_2 on Y and X_1 ,
we can by using the residuals
from these regressions find the partial effect
of X_1 on Y . Finding the explanatory
power of $e_{X_1|X_2}$ on $e_{Y|X_2}$ is the
same as finding the partial effect of
 X_1 on Y directly through the bivariate
case. Compared to the bivariate, we would
get exactly the same estimate $\hat{\beta}_1$ and
same standard error $se(\hat{\beta}_1)$.

○ 3) Let us start by showing what we're supposed to "know" already

$$\min_{\beta_0, \beta_1, \beta_2} S(\beta_0, \beta_1, \beta_2) = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2$$

○ FOC

$$(I) -2 \sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i}) = 0$$

$$(II) -2 \sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i}) X_{1i} = 0$$

$$(III) -2 \sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i}) X_{2i} = 0$$

○ Rearrange to get:

$$(I)' \quad \sum Y_i = n \hat{\beta}_0 + \hat{\beta}_1 \sum X_{1i} + \hat{\beta}_2 \sum X_{2i}$$

$$(II)' \quad \sum Y_i X_{1i} = \hat{\beta}_0 \sum X_{1i} + \hat{\beta}_1 \sum X_{1i}^2 + \hat{\beta}_2 \sum X_{2i} X_{1i}$$

$$(III)' \quad \sum Y_i X_{2i} = \hat{\beta}_0 \sum X_{2i} + \hat{\beta}_1 \sum X_{1i} X_{2i} + \hat{\beta}_2 \sum X_{2i}^2$$

○

Now rewrite in terms of normal equations

$$X_1^* = X_{1i} - \bar{X}_1$$

$$X_2^* = X_{2i} - \bar{X}_2$$

$$Y^* = Y_i - \bar{Y}$$

$$\stackrel{(I)'}{\Rightarrow} \left(\sum Y^* + \bar{Y} \right) = n \hat{\beta}_0 + \hat{\beta}_1 \sum (X_1^* + \bar{X}_1) + \hat{\beta}_2 \sum (X_2^* + \bar{X}_2)$$

$= 0 \quad = n\bar{Y}$

$$= n \bar{Y} = n \hat{\beta}_0 + n \hat{\beta}_1 \bar{X}_1 + n \hat{\beta}_2 \bar{X}_2$$

$$\Rightarrow \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}_1 - \hat{\beta}_2 \bar{X}_2$$

It can also be shown that

$$\hat{\beta}_1 = \frac{\sum (Y_i^* X_{1i}^*) \sum X_{2i}^{*2} - \sum (Y_i^* X_{2i}^*) \sum (X_{1i}^* X_{2i}^*)}{\sum X_{1i}^{*2} \sum X_{2i}^{*2} - \left(\sum X_{1i}^* X_{2i}^* \right)^2}$$

$$\hat{\beta}_2 = \frac{\sum (Y_i^* X_{2i}^*) \sum X_{1i}^{*2} - \sum (Y_i^* X_{1i}^*) \sum (X_{2i}^* X_{1i}^*)}{\sum X_{1i}^{*2} \sum X_{2i}^{*2} - \left(\sum X_{1i}^* X_{2i}^* \right)^2}$$

○ Put denominator = M , then

$$\hat{\beta}_2 = \frac{\hat{\sigma}_{x_1}^2 \hat{\sigma}_{y, x_2} - \hat{\sigma}_{y, x_1} \hat{\sigma}_{x_2, x_1}}{M}$$

and $E(\hat{\beta}_2 | X_1, X_2) = \frac{\hat{\sigma}_{x_1}^2}{M} E(\hat{\sigma}_{y, x_2} | X_1, X_2)$

○ $-\frac{\hat{\sigma}_{x_1, x_2}}{M} E(\hat{\sigma}_{y, x_1} | X_1, X_2)$

Now $E(\hat{\beta}_2) = E(E(\hat{\beta}_2 | X_1, X_2))$

Find first $\hat{\sigma}_{y, x_1} = \frac{1}{n} \sum Y_i (X_{1i} - \bar{X}_1)$

○ $= \frac{1}{n} \sum (\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i)(X_{1i} - \bar{X}_1)$

$$= \frac{1}{n} \sum (\beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i)(X_{1i} - \bar{X}_1)$$

$$= \beta_1 \hat{\sigma}_{x_1}^2 + \beta_2 \hat{\sigma}_{x_1, x_2} + \frac{1}{n} \sum \epsilon_i (X_{1i} - \bar{X}_1), \text{ and symmetry dictates:}$$

○ $\hat{\sigma}_{y, x_2} = \beta_2 \hat{\sigma}_{x_2}^2 + \beta_1 \hat{\sigma}_{x_1, x_2} + \frac{1}{n} \sum \epsilon_i (X_{2i} - \bar{X}_2)$

Hence;

$$E(\hat{\sigma}_{y, x_1} | X_1, X_2) = \beta_1 \hat{\sigma}_{x_1}^2 + \beta_2 \hat{\sigma}_{x_1, x_2}$$

$$E(\hat{\sigma}_{y, x_2} | X_1, X_2) = \beta_1 \hat{\sigma}_{x_1, x_2} + \beta_2 \hat{\sigma}_{x_2}^2$$

Thus;

$$E(\hat{\beta}_2 | X_1, X_2) = \frac{\hat{\sigma}_{x_1}^2}{M} (\beta_1 \hat{\sigma}_{x_1, x_2} + \beta_2 \hat{\sigma}_{x_2}^2) - \frac{\hat{\sigma}_{x_1, x_2}}{M} (\beta_1 \hat{\sigma}_{x_1}^2 + \beta_2 \hat{\sigma}_{x_1, x_2})$$

$$E(\hat{\beta}_2 | X_1, X_2) = \frac{\hat{\sigma}_{x_1}^2}{M} \beta_1 \hat{\sigma}_{x_1, x_2} + \beta_2 \frac{\hat{\sigma}_{x_2}^2 \hat{\sigma}_{x_1}^2}{M} - \frac{\hat{\sigma}_{x_1}^2}{M} \beta_1 \hat{\sigma}_{x_1, x_2} - \beta_2 \frac{\hat{\sigma}_{x_1, x_2}^2}{M}$$

$$E(\hat{\beta}_2 | X_1, X_2) = \beta_2 \left(\frac{\hat{\sigma}_{x_2}^2 \hat{\sigma}_{x_1}^2 - \hat{\sigma}_{x_1, x_2}^2}{\hat{\sigma}_{x_2}^2 \hat{\sigma}_{x_1}^2 - \hat{\sigma}_{x_1, x_2}^2} \right) = \beta_2$$

$$\Rightarrow E(E(\hat{\beta}_2 | X_1, X_2)) = E(\hat{\beta}_2) = E(\beta_2) = \beta_2$$

○ So what if we have

$e_{Y_1 | X_{2i}}$ as regressand and

$e_{X_1 | X_{2i}}$ as regressor?

○ By FW-theorem, no change.

4)

The partial empirical correlation coefficient between X_1, X_2 and Y, X_2 measures how much X_1 explains of the variation in Y after we've considered the contribution from X_2 .

Squared partial correlation coefficients cannot be less than the regular correlation coefficients, and no larger than the determinant coefficient, R^2 .

This means (from sem 5) that :

- R^2 in reg $\hat{y}_a z_a$ cannot be less than R^2 in reg $y_a z_a$ or reg $\hat{z}_b z_a$
- Cannot exceed R^2 in reg $y_a z_a z_b$

4.13

$$a) \ln P = \alpha + \beta S$$

$$\Rightarrow \frac{dP}{dS} = \beta e^{\alpha + \beta S}$$

$$\text{At sample mean: } \frac{dP}{dS} = \beta e^{\alpha + \beta \bar{S}} = \beta e^{\bar{P}}$$

$$\Rightarrow \frac{dP}{dS} = 0,000596 e^{112810,8} = \underline{\underline{67,23}}$$

$$\begin{aligned} EL_S(P) &= \hat{\beta}_1 \times \bar{S} = 0,000596 \times 1611,9682 \\ &= \underline{\underline{0,9607}} \end{aligned}$$

The coefficient 0,000596 suggests an increase of one square foot gives $\approx 0,06\%$ price increase.

b) The coefficient 1,0066 is the elasticity

$$P = e^{4,1707 + 1,0066 \ln S}$$

$$\Rightarrow \frac{dP}{dS} = \frac{1,0066}{S} e^{4,1707 + 1,0066 \ln \bar{S}} = \frac{1,0066}{\bar{S}} e^{\ln \bar{P}}$$

At sample mean: $\frac{dP}{dS} = \underline{\underline{70,44}}$

e) Residuals seem to increase in magnitude

as SQFT increases.

Most evident in linear model, where

all residuals are positive below 1000

f) In this problem, we could have estimated this by writing:

$$\hat{p} = e^{\{\alpha + \beta \widehat{\ln S}\}}$$

but since we know from Jensen's inequality that since e is a strictly convex function and $\ln P$ has a strictly positive variance, then

$$E(P) = E(e^{\ln P}) > e^{E(\ln P)}, \text{ or similarly}$$

$$\text{mean}(P) > e^{\text{mean}(\ln P)} \quad \text{Hence we will}$$

systematically underestimate if we take the

antilog. It can be shown that a better

estimate is:

$$\hat{p} = \widehat{E(P)} = e^{\{\alpha + \beta \ln \bar{S} + \hat{\sigma}^2 / 2\}}$$

g)

$$se(f) = \sqrt{\widehat{\text{var}}(f)}$$

$$\widehat{\text{var}}(f) = \hat{\sigma}^2 + \frac{\hat{\sigma}^2}{n} + (X_0 - \bar{X})^2 \widehat{\text{var}}(\hat{\beta}_1)$$

log-lin
I)

$$\Rightarrow \widehat{\text{var}}(f) = 0,041222602 + \frac{0,041222602}{880} + (2700 - 1611,968)^2 \times 0,0000129^2$$

$$= 0,041466$$

$$\Rightarrow se(f) = \sqrt{0,041466} = 0,2036$$

$$\text{Hence: } [12,20299 \pm 1,9626634 \times 0,2036]$$

$$= [11,8033, 12,6026] \text{ for } \ln P$$

$$\Rightarrow \underline{\underline{[1133683, 297316] \text{ for } P}}$$

log-log
II)

$$\widehat{\text{var}}(f) = 0,043368435 + \frac{0,043368435}{880} + (\ln 2700 - \ln 1611,968)^2 \times 0,0225423^2$$

$$= 0,04358074$$

$$\Rightarrow \text{se}(f) = 0,20876$$

Hence []

[]

$$\Rightarrow [127\ 267, \ 277\ 454] \text{ for } P$$

linear
(III)

$$\widehat{\text{var}}(f) = 915618929 + \frac{915618929}{880} + (2700 - 1611,968)^2 \times 1,918489^2$$

$$= 921016549,6$$

$$\Rightarrow \text{se}(f) = 30348,26$$

$$\Rightarrow \underline{\underline{[141801, \ 260928]}}$$

h) Because of its skewness to the right, and hence far from being normally distributed, the linear model is not good. Between log-lin and log-log, it is more difficult to choose.

Log-lin might be preferred due to higher R^2 and smaller standard deviation of the error terms.

SEMINAR 7

$$WALC = \beta_1 + \beta_2 \ln(TOTEXP) + \beta_3 AGE + \beta_4 NK + e$$

$$NK \in \{1, 2\}$$

a) $n = 1519$

Variable	Coefficient	Std. error	t-stat	p-value
C	0,0091	0,0190	(0,4789)	0,6347
$\ln(TOTEXP)$	0,0276	(0,004176)	6,6086	0,0000
AGE	(-0,00139)	0,0002	-6,924	0,0000
NK	-0,0133	0,0033	-4,0750	0,0000

R^2	(0,05419)	Mean dep. var.	0,0606
S.E. of regression	(0,06162)	S.D. dep. var.	0,0633
Sum of sq. res.	5,752896		

Bonus-question: Which test is the t-stat for?

Answer: $H_0: \beta_1 = 0$ vs. $H_1: \beta_1 \neq 0$

$$\Rightarrow t\text{-stat} = \frac{\hat{\beta}_1 - 0}{se(\hat{\beta}_1)}$$

Hence;

$$i) \quad t\text{-stat} = \frac{0,0091 - 0}{0,0190} = \underline{\underline{0,4789}}$$

$$ii) \quad \text{std. error}(\hat{\beta}_2) = \frac{\hat{\beta}_2}{t\text{-stat}(\hat{\beta}_2)} = \frac{0,0276}{6,6086} = \underline{\underline{0,004176}}$$

$$iii) \quad \hat{\beta}_3 = t\text{-stat}(\hat{\beta}_3) \times \text{St. err}(\hat{\beta}_3) = 0,0002 \times (-6,9624) \\ = \underline{\underline{-0,00139}}$$

iv) R^2 is the explained variance in the model,

$$\text{hence } R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

And we know SSE already, so all we need is SST.

$$SST = (n-1) s_y^2 = (n-1) \hat{\sigma}_{walc}^2$$

$$SST = 1518 \times 0,0633^2 = \underline{\underline{6,0825}}$$

$$\Rightarrow R^2 = 1 - \frac{5,752896}{6,0825} = \underline{\underline{0,05419}}$$

v)

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n \hat{e}_i^2}{N - K}$$

$$\Rightarrow \hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n \hat{e}_i^2}{N - K}} = \sqrt{\frac{5,752896}{1515}} = \underline{\underline{0,00162}}$$

i.e. Root MSE

b) WALC is proportion of households budget spent on alcohol.

NK is number of kids in household

AGE is age of household head.

Interpretations:

$\hat{\beta}_2$: When total expenditures increase by 1%, the share of alcohol in the budget increases by 0,000276

Why:

other necessary goods, such

as medicines, diapers, food, health care etc.

$\hat{\beta}_3$: When the age of the oldest in the household increases, share of alcohol goes down.

$\hat{\beta}_4$: When number of kids increase, less is spent on alcohol.

c) 95% CI for β_3

$$t_c = 1,96 \quad (\text{since } n \text{ large})$$

Hence; $(\hat{\beta}_3 \pm se(\hat{\beta}_3) \times t_c)$ is a 95% CI

$$\Rightarrow [-0,00139 \pm 1,96 \times 0,0002] = \underline{\underline{[-0,001782, -0,001]}}$$

Interpretation:

When constructing many such intervals, we would expect that 95% of the time, we will get an interval which contains the true value.

d) $H_0: \beta_4 = 0$ vs. $H_1: \beta_4 \neq 0$

$$t_{\text{stat}} = \frac{t_{\text{obs}} - 0}{se(\hat{\beta}_4)} = \frac{-0,0133}{0,0033} = \underline{\underline{-4,0750}}$$

Reject H_0 if $|t_{\text{obs}}| > t_{\alpha/2, n-4} = 1,96$

\Rightarrow Reject H_0 at 5% significance level.

Why this outcome?

- We are measuring how many standard errors we are from 0. At 1.96 standard errors, we say it is a 5% chance of committing a type I (rejection) error. Hence when we are this far away, we need a higher significance level to claim insignificance of NK (in fact p-value is 0.0000).
- So even though the coefficient itself is not that far from 0, the standard error is also small, making $NK^{\text{coeff.}}$ significantly different from 0.

5,4

$$a) WTRANS = \beta_1 + \beta_2 \ln(TOTEXP) + \beta_3 AGE + \beta_4 NK + \epsilon$$

$$WTRANS = -0,0315 + 0,0414 \ln(TOTEXP) - 0,0001 AGE - 0,0130 NK$$

$(0,0322) \quad (0,0071) \quad (0,0004) \quad (0,0055)$

b) $\hat{\beta}_2$: When total expenditures increase by 1%, transports share of the budget is increased positively. Makes sense, perhaps going from public transport to car etc.

$\hat{\beta}_3$: Age has a minorly negative effect on transports share of budget. Seems false, older people use more expensive transport, taxis etc. Unreasonable.

$\hat{\beta}_4$: Seems a little unreasonable that the effect is negative. More kids \Rightarrow more driving to trainings/events etc., picking up kids at places?

Alternatively, the negative sign can imply that when number of kids increase, expenditures on other items increase more than transportation.

c) The p-value for significance of AGE ($H_0: \beta_3 = 0$ vs. $H_1: \beta_3 \neq 0$) is 0,8690 suggesting AGE is a candidate for exclusion.

d) The explained variation is simply found by recalling that $R^2 = 0,0247$, indicating that 2,47% of the variation in WTRANS is explained by these explanatory variables.

e) One-children households:

$$\begin{aligned} WTRANS &= -0,0315 + 0,0414 \ln 98,7 - 0,0001 \times 36 - 0,0130 \times 1 \\ &= \underline{\underline{0,1420}} \end{aligned}$$

Two-children households:

$$WTRANS = -0,0315 + 0,0414 \ln 98,7 - 0,0001 \times 36 - 0,0130 \times 2 = \underline{\underline{0,1290}}$$

5.14

b) When the number of square feet increases by a hundred, the price of the house increases by 3,876%.

$$c) \frac{\partial \ln \text{price}}{\partial \text{AGE}} = \alpha_3 + 2\alpha_4 \text{AGE}$$

$$\text{AGE} = 2 \Rightarrow -0,0175549 + 2 \times 0,0001734 \times 2 = -0,0169$$

$$\text{AGE} = 15 \Rightarrow -0,0175549 + 2 \times 0,0001734 \times 15 = -0,01235$$

So if a house is two years old, a one year increase in age reduces the price by 1,7% while it's only 1,235% decrease if the house is 15 years old. Hence younger houses fall faster in price than older, perhaps since older houses are looked upon as "hip and trendy". (see for instance housing prices in Teitthusbakken in Oslo!)

d)

$$\text{Rule: } \frac{d \ln x}{dY} = \frac{1}{x} \frac{dx}{dY}$$

by remembering rules for computing differentials of logarithms:

$$d f(x) = \frac{1}{f(x)} f'(x)$$

$$\text{Hence; } \frac{\partial \text{PRICE}}{\partial \text{AGE}} = \text{PRICE} \left[\frac{\partial \ln \text{PRICE}}{\partial \text{AGE}} \right]$$

$$= \text{PRICE} [\alpha_3 + 2\alpha_4 \text{AGE}]$$

$$\text{And; } \frac{\partial \ln \text{PRICE}}{\partial \text{SQFT100}} = \alpha_2$$

$$\Rightarrow \frac{\partial \text{PRICE}}{\partial \text{SQFT100}} = \alpha_2 \text{ PRICE}$$

e) First, we have to find estimated price.

$$\text{AGE} = 15, \quad \text{SQFT100} = 20$$

$$\Rightarrow \widehat{\ln \text{PRICE}} = 11,11959 + 0,0387624 \times 20 - 0,0175549 \times 15 + 0,0001734 \times 15^2$$

$$\widehat{\ln \text{PRICE}} = 11,6705295$$

Jensens inequality
from sem 6.

$$\Rightarrow \widehat{\text{PRICE}} = e^{11,6705295 + \frac{1}{2} \sigma^2} = \underline{\underline{121898,998}}$$

$$\text{Thus; } \frac{\partial \text{PRICE}}{\partial \text{AGE}} = 121898,998 \times [-0,0175549 + 2 \times 0,0001734 \times 15] = \underline{\underline{-1505,81}}$$

$$\frac{\partial \text{PRICE}}{\partial \text{SQFT100}} = 0,0387624 \times 121898,998 = \underline{\underline{4725,1}}$$

When AGE goes from 15 to 16 years, nominal house price falls by -1505 dollars, and when number of sqft increases from 2000, price increase by 4725

f) Assume PRICE = 120 000 (CONSTANT)

$$\text{var} \left[\frac{\partial \text{PRICE}}{\partial \text{AGE}} \right] = \text{var}(\alpha_3 \text{PRICE} + 2\alpha_4 \text{AGE} \cdot \text{PRICE})$$

$$= \text{PRICE}^2 \text{var} \alpha_3 + 4 \text{AGE}^2 \text{PRICE}^2 \text{var} \alpha_4$$

$$+ 4 \text{PRICE}^2 \times \text{AGE} \text{cov}(\alpha_3, \alpha_4)$$

$$= 120\,000^2 \times 1,84 \times 10^{-6} + 4 \times 15^2 \times 120\,000^2 \times (5,133 \times 10^{-10})$$
$$+ 4 \times 120\,000^2 \times 15 \times (-2,849 \times 10^{-8})$$

$$= \underline{8533,008}$$

$$\Rightarrow \text{se} \left(\frac{\partial \text{PRICE}}{\partial \text{AGE}} \right) = \sqrt{8533,008} = \underline{\underline{92,37}}$$

$$\text{var} \left[\frac{\partial \text{PRICE}}{\partial \text{SQFT100}} \right] = \text{PRICE}^2 \text{var}(\alpha_2)$$

$$= 120\,000^2 \times 7,557 \times 10^{-7} = 10882,08$$

$$\Rightarrow \text{se} \left[\frac{\partial \text{PRICE}}{\partial \text{SQFT100}} \right] = \sqrt{10882,08} = \underline{\underline{104,32}}$$

g) 95% interval estimate:

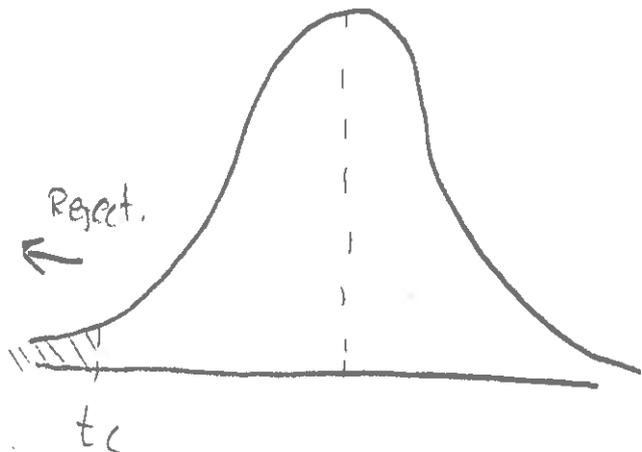
(For $\frac{\partial \text{PRICE}}{\partial \text{SQFT100}}$)

$$\left[4651,488 \pm 1,96 \times 104,32 \right] = \underline{\underline{[4447, 4856]}}$$

h) $H_0: \frac{\partial \text{PRICE}}{\partial \text{AGE}} \geq -1000$ vs. $H_1: \frac{\partial \text{PRICE}}{\partial \text{AGE}} < -1000$

One-sided test. So we reject if:

$$\theta_{\text{obs}} < \frac{\hat{\theta} - \theta}{\text{se}(\hat{\theta})}$$



$$\hat{\theta} = 120\,000 \times [-0,0175549 + 30 \times 0,0001734] = -1482,348$$

$$\Rightarrow \theta_{\text{obs}} = \frac{-1482,348 - (-1000)}{92,37} = \underline{\underline{-5,22}}$$

$-5,22 < -1,645 \Rightarrow$ Reject at 5% sign. level.

SEMINAR 8

6.1

$$Y = \beta_1 + \beta_2 X + \beta_3 Z + \varepsilon$$

$$SSE = 979,830 \quad s_y = 13,45222$$

a) From last time, recall:

$$R^2 = 1 - \frac{SSE}{SST} \quad \text{and} \quad SST = (N-1) s_y^2$$

$$\Rightarrow SST = 39 \times 13,45222^2 = 7057,526694$$

$$\Rightarrow R^2 = 1 - \frac{979,830}{7057,526694} = \underline{\underline{0,861}}$$

b). First of all some general stuff on
F-testing

These tests test the significance of the model. We can test significance jointly for all explanatory variables.

So in general, when we have a multiple regression model on the form:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

We can test the overall significance of our model by setting up:

$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$ vs. At least one nonzero.

Now if we impose a restriction to all but the intercept, we can write the F-statistic as:

$$F = \frac{(SST - SSE) / (k-1)}{SSE / (N-k)}$$

And we compare this value to a critical value $F_{(k-1, N-k)}$

Now, in general, we can use F-tests to test significance of some of the variables (instead of all). Then our

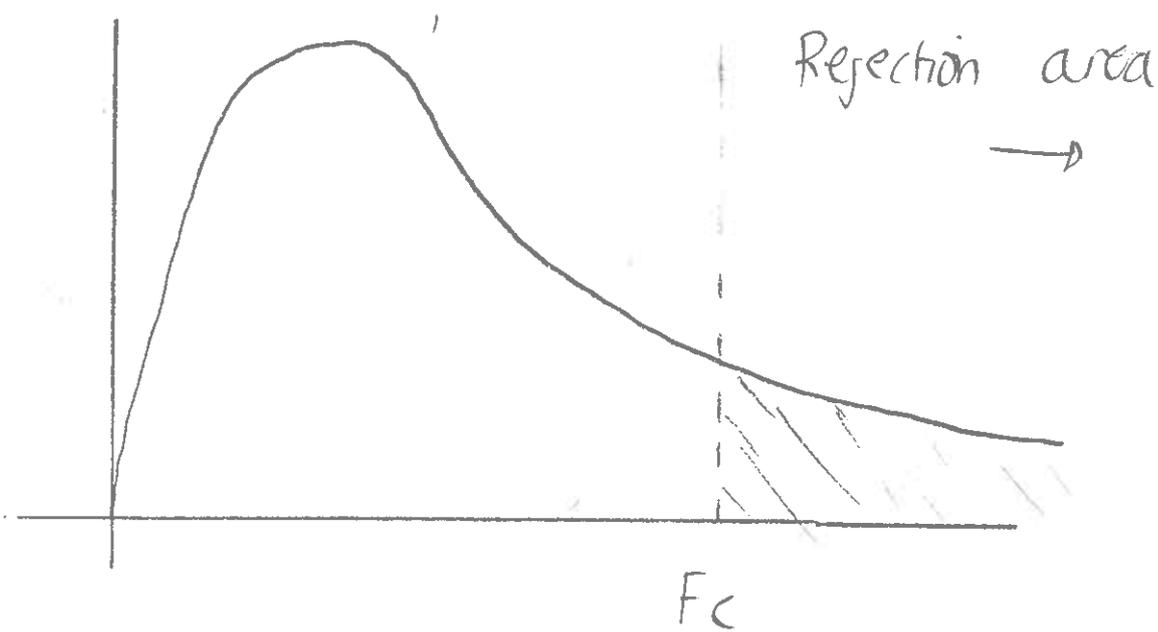
F-statistic looks slightly different

$$F = \frac{(SSE_R - SSE_U) / J}{SSE_U / (N-k)}$$

Where SSE_R is SSE in restricted model and J is number of linear restrictions.

F is now $F(J, N-k)$

In F-testing, we always reject if $F_{obs} > F_c$ and there are no \geq, \leq etc.



Now to the question:

We are testing overall significance, so

$$F_{obs} = \frac{(7057,526694 - 979,830) / 3 - 1}{979,830 / 40 - 3}$$

$$F_{obs} = \underline{\underline{114,7519}}$$

$$\text{And our } F_{stat} = F_{(0,05, 2, 37)}$$

$$\Rightarrow F_{stat} \approx 3,25$$

\Rightarrow Reject H_0 at 5% sign. level.

6.2 RESET - test for misspecification

New model is:

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 Z_i + \gamma_1 \hat{Y}_i^2 + \gamma_2 \hat{Y}_i^3 + \epsilon_i$$

Hence we want to test if $\gamma_1 = \gamma_2 = 0$

$\Rightarrow H_0: \gamma_1 = \gamma_2 = 0$ vs. $H_1: \text{At least one different from zero.}$

$$\Rightarrow F_{obs} = \frac{(979,830 - 696,5357) / 2}{696,5357 / (40 - 5)} = \underline{7,12}$$

$$\text{And } F_c = F(0,95, 2, 35) = 3,27$$

Hence reject H_0 at 5% sign. level.

(6.3)

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

a) We can start out by remembering

$$SSE = \sum_{i=1}^n \hat{e}_i^2 \quad \text{and} \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n e_i^2}{N-K}$$

$$\text{Hence } SSE = \hat{\sigma}^2 \times (N-K) = 2,5193 \times 17 = \underline{\underline{42,8281}}$$

$$\text{And then } R^2 = 1 - \frac{SSE}{SST}, \text{ giving}$$

$$SST = \frac{SSE}{(1-R^2)} = \frac{42,8281}{1-0,9466} = \underline{\underline{802,0243}}$$

$$\text{And } R^2 = \frac{SSR}{SST} \Rightarrow SSR = R^2 \times SST = \underline{\underline{759,1962}}$$

b) 95% interval estimates

$$\beta_2: [\hat{\beta}_2 \pm se(\hat{\beta}_2) \times t_c] = (1 - \alpha)$$

$$t_c = t_{0,025,17} \approx 2,11$$

$$\text{Hence; } [0,69914 \pm 2,11 \times \sqrt{0,048526}]$$

$$= \underline{\underline{[0,2343, 1,1639]}}$$

$$\beta_3: [\hat{\beta}_3 \pm se(\hat{\beta}_3) t_c] = (1 - \alpha)$$

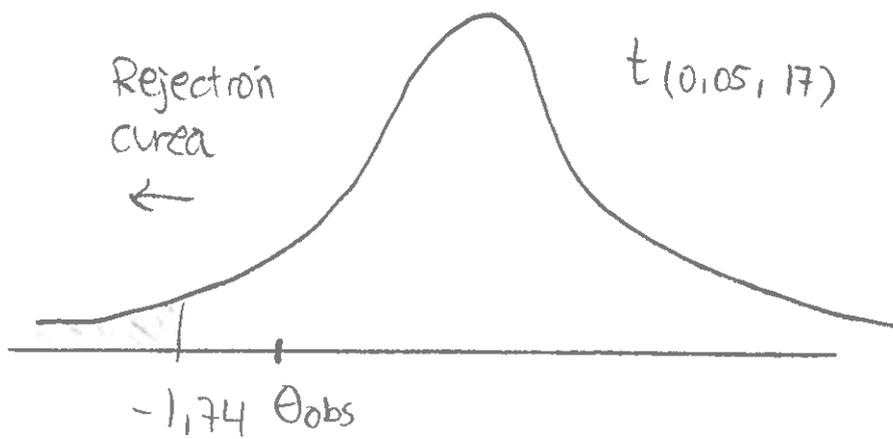
$$\text{Hence; } [1,7769 \pm 2,11 \times \sqrt{0,037120}]$$

$$= \underline{\underline{[1,3704, 2,1834]}}$$

$$c) H_0: \beta_2 \geq 1 \quad \text{vs.} \quad H_1: \beta_2 < 1$$

$$\theta_{\text{obs}} = \frac{\hat{\beta}_2 - 1}{\text{se}(\hat{\beta}_2)} = \frac{0,69914 - 1}{\sqrt{0,048526}} = \underline{\underline{-1,3658}}$$

Reject H_0 if $\theta_{\text{obs}} < t_c = -1,74$



Hence do not reject H_0 at 5% significance level!
There is not enough evidence in data to suggest $\beta_2 < 1$

d) $H_0: \beta_2 = \beta_3 = 0$ vs. $H_1: \beta_2 \neq 0$ or $\beta_3 \neq 0$
or $\beta_2, \beta_3 \neq 0$

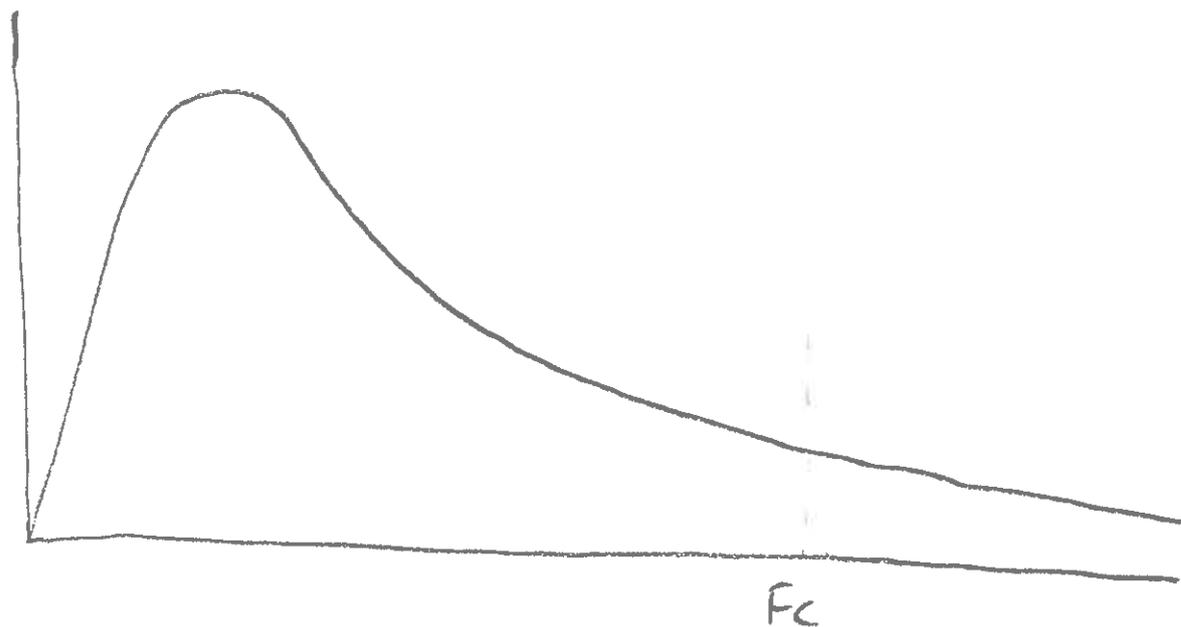
Using F-test for overall significance:

$$F_{obs} = \frac{(SST - SSE) / (K - 1)}{SSE / (N - K)}$$

$$F_{obs} = \frac{(802,0243 - 42,8281) / (3 - 1)}{42,8281 / (20 - 3)} = \underline{150,676}$$

Reject H_0 if $F_{obs} > F_{(0,95, 2, 17)} \approx 3,6$

Hence reject H_0 at 5% sign. level



$$e) \quad H_0: 2\beta_2 = \beta_3 \quad \text{vs.} \quad H_1: \text{Not } H_0.$$

Now rewrite the test into:

$$H_0: 2\beta_2 - \beta_3 = 0 \quad \text{vs.} \quad H_1: \text{Not } H_0.$$

Then a regular t -test gives us:

$$\theta_{\text{obs}} = \frac{(2\hat{\beta}_2 - \hat{\beta}_3) - 0}{\text{se}(2\hat{\beta}_2 - \hat{\beta}_3)}$$

So we need $\text{se}(2\hat{\beta}_2 - \hat{\beta}_3)$

$$\text{Find then} \quad \text{var}(2\hat{\beta}_2 - \hat{\beta}_3) = 4 \text{var}\hat{\beta}_2 + \text{var}\hat{\beta}_3 \\ - 4 \text{cov}(\hat{\beta}_2, \hat{\beta}_3)$$

$$\Rightarrow \text{var}(2\hat{\beta}_2 - \hat{\beta}_3) = 4 \times 0,048526 + 0,037120 - 4 \times (-0,031223) \\ = 0,356116 \Rightarrow \text{se}(\cdot) = 0,59675$$

$$\Rightarrow \theta_{\text{obs}} = \frac{(2 \times 0,69914 - 1,7769)}{0,59675} = \underline{\underline{-0,6345}}$$

And we have a two-sided test,

so we reject H_0 if:

$$|\theta_{\text{obs}}| > t_c = t_{0,025,17} = 2,11$$

Hence do not reject H_0 . There is not enough evidence in data to suggest

$$2\beta_2 \neq \beta_3.$$

$$\textcircled{6.5} \quad \ln(\text{WAGE}) = \beta_1 + \beta_2 \text{ EDUC} + \beta_3 \text{ EDUC}^2 \\ + \beta_4 \text{ EXPER} + \beta_5 \text{ EXPER}^2 + \beta_6 \text{ HRSWK} + \varepsilon$$

a) We would need to jointly test if the linear effect from education and experience, and the squared effect from the two both are equal simultaneously.

Hence; $H_0: \beta_2 = \beta_4$ and $\beta_3 = \beta_5$

vs. $H_1: \beta_2 \neq \beta_4$ or $\beta_3 \neq \beta_5$ or both

b) Assuming H_0 true, the restricted model is now:

$$\ln(\text{WAGE}) = \beta_1 + \beta_2[\text{EDUC} + \text{EXPER}] + \beta_3[\text{EDUC}^2 + \text{EXPER}^2] + \beta_6 \text{HRSWK} + \varepsilon.$$

$$c) F_{\text{obs}} = \frac{(SSE_R - SSE_U) / J}{SSE_U / (N - K)}$$

$$F_{\text{obs}} = \frac{(254,1726 - 222,6674) / 2}{222,6674 / (1000 - 6)} = \underline{70,32}$$

$$\text{And } F_c = F(0,95, 2, 994) \approx 3$$

\Rightarrow Reject H_0 at 5% sign. level.

SEMINAR 9

7.1

$$\widehat{SAL} = 24200 + 1643 \text{ GPA} + 5033 \text{ METRICS}$$

(1078) (352) (456)

$$R^2 = 0,74$$

a) 24200 (intercept) has the interpretation "minimum salary", or the amount of which a student with GPA = 0 would have had in starting salary.

1643 (coefficient) says that a one unit increase in grade point average gives a 1643 dollar increase in salary.

5033 (coefficient) says that taking econometrics gives a 5033 dollar increase in starting salary.

b)

So using the hint: Create a dummy such that:

$$\text{FEMALE} = \begin{cases} 1 & \text{if female} \\ 0 & \text{if male} \end{cases}$$

Now respecify into:

$$\text{SAL} = \beta_1 + \beta_2 \text{GPA} + \beta_3 \text{METRICS} + \beta_4 \text{FEMALE} + \epsilon$$

β_4 will then just add to the intercept

c)' Let's keep FEMALE as a part of the model. To check if males and females have the same effect from econometrics, add:

$$\begin{aligned} \text{SAL} = & \beta_1 + \beta_2 \text{GPA} + \beta_3 \text{METRICS} + \beta_4 \text{FEMALE} \\ & + \beta_5 \text{FEMALE} \times \text{METRICS} + \epsilon \end{aligned}$$

β_5 will affect the slope if there are differences between females relative to males.

Now we have also:

$$E(\widehat{SAL}) = \begin{cases} \beta_1 + \beta_2 \text{GPA} + \beta_3 \text{METRICS} & \text{FEM}=0 \\ (\beta_1 + \beta_4) + \beta_2 \text{GPA} + (\beta_3 + \beta_5) \text{METRICS} & \text{FEM}=1 \end{cases}$$

β_6

$$(8.1) \frac{\sum_{i=1}^n (X_i - \bar{X})^2 \sigma_i^2}{\left[\sum_{i=1}^n (X_i - \bar{X})^2 \right]^2}$$

Now assume $\sigma_i^2 = \sigma^2$ (i.e. not dependant on i)

Then;

$$\frac{\sigma^2 \sum_{i=1}^n (X_i - \bar{X})^2}{\left[\sum_{i=1}^n (X_i - \bar{X})^2 \right]^2} = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

8.2.

$$Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i$$

$$\text{var}(\varepsilon_i) = \sigma_i^2$$

$$Y_i^* = \beta_1 \sigma_i^{-1} + \beta_2 X_i^* + \varepsilon_i^*$$

$$Y_i^* = \sigma_i^{-1} y_i, \quad X_i^* = \sigma_i^{-1} X_i$$

$$\varepsilon_i^* = \sigma_i^{-1} \varepsilon_i$$

$$(\sum \sigma_i^{-2}) \hat{\beta}_1 + (\sum \sigma_i^{-1} X_i^*) \hat{\beta}_2 = \sum \sigma_i^{-1} Y_i^* \quad (\text{I})$$

$$(\sum \sigma_i^{-1} X_i^*) \hat{\beta}_1 + (\sum X_i^{*2}) \hat{\beta}_2 = \sum X_i^* Y_i^* \quad (\text{II})$$

a) Start by showing $\hat{\beta}_1$ (easiest):

Solve (I) for $\hat{\beta}_1$:

$$\hat{\beta}_1 = \frac{-\sum \sigma_i^{-1} Y_i^* - (\sum \sigma_i^{-1} X_i^*) \hat{\beta}_2}{\sum \sigma_i^{-2}}$$

Now we do realise that since

$$Y_i^* = \sigma_i^{-1} Y_i \quad \text{etc, we get:}$$

$$\hat{\beta}_1 = \frac{\sum \sigma_i^{-2} Y_i}{\sum \sigma_i^{-2}} - \left(\frac{\sum (\sigma_i^{-2} X_i)}{\sum \sigma_i^{-2}} \right) \hat{\beta}_2$$

And now for $\hat{\beta}_2$:

Insert result for $\hat{\beta}_1$ into (II) to get:

$$\left(\sum \sigma_i^{-2} X_i \right) \left(\frac{\sum \sigma_i^{-2} Y_i}{\sum \sigma_i^{-2}} - \frac{\sum (\sigma_i^{-2} X_i)}{\sum \sigma_i^{-2}} \hat{\beta}_2 \right)$$

$$+ \sum X_i^* \hat{\beta}_2 = \sum X_i^* Y_i^*$$

Now divide by $\sum \sigma_i^{-2}$ in all terms and get:

$$\frac{\sum \sigma_i^{-2} Y_i}{\sum \sigma_i^{-2}} \frac{\sum \sigma_i^{-2} X_i}{\sum \sigma_i^{-2}} - \left(\frac{\sum \sigma_i^{-2} X_i}{\sum \sigma_i^{-2}} \right)^2 + \frac{\sum X_i^* \hat{\beta}_2}{\sum \sigma_i^{-2}} = \frac{\sum X_i^* Y_i^*}{\sum \sigma_i^{-2}}$$

Rearrange to get:

$$\frac{\sum X_i^* Y_i^*}{\sum \sigma_i^{-2}} - \left(\frac{\sum \sigma_i^{-2} Y_i}{\sum \sigma_i^{-2}} \right) \left(\frac{\sum \sigma_i^{-2} X_i}{\sum \sigma_i^{-2}} \right)$$

$$= \left(\frac{\sum X_i^2}{\sum \sigma_i^{-2}} - \left(\frac{\sum \sigma_i^{-2} X_i}{\sum \sigma_i^{-2}} \right)^2 \right) \hat{\beta}_2$$

$$\hat{\beta}_2 = \frac{\frac{\sum \sigma_i^{-2} X_i Y_i}{\sum \sigma_i^{-2}} - \left(\frac{\sum \sigma_i^{-2} Y_i}{\sum \sigma_i^{-2}} \right) \left(\frac{\sum \sigma_i^{-2} X_i}{\sum \sigma_i^{-2}} \right)}{\frac{\sum \sigma_i^{-2} X_i^2}{\sum \sigma_i^{-2}} - \left(\frac{\sum \sigma_i^{-2} X_i}{\sum \sigma_i^{-2}} \right)^2}$$

Note that $\sum Y_i^* X_i^* = \sum \sigma_i^{-2} Y_i X_i$

$\sum X_i^{*2} = \sum \sigma_i^{-2} X_i^2$



for last equality
HUSK bog c!!!

② Investigate no autocorrelation in Q4, Q5

So we are testing whether:

$$e_t = \rho e_{t-1} + v_t$$

We use then a Lagrange multiplier test, and we can choose freely which one to use. Then lets use:

$$H_0: \rho = 0 \quad \text{vs.} \quad H_1: \rho \neq 0$$

$$LM = (T-1) \times R^2 = 63 \times 0,9360 \approx \underline{58,968}$$

Which is a lot higher than 3,84

And it can be proven that LM has a $\sim \chi_{(1)}$ distribution, meaning the critical value is 3,84. Hence we reject H_0

and conclude there's sufficient evidence in data to suggest autocorrelation.

Now let's test the other model. Note that we get 2 missing values for \hat{e}_{t-1} here, since \hat{e}_t itself contains a missing value. Thus we estimate

$$LM = (T-1) \times R^2 = 63 \times 0,2136 \approx 13,4568$$

And we reject H_0 ($H_0: \rho=0$) and conclude autoregressive error terms are present.

$$\textcircled{3} \quad H_0: \gamma_1 = \gamma_2 \quad \text{vs} \quad H_1: \gamma_1 \neq \gamma_2$$

DUMMY-test' in STATA. Or:

This means we have. (it test for equivalence)

$$E(LBNPCAP) = \begin{cases} \beta_0 + \gamma_1 \text{TREND} \\ \beta_0 + \underbrace{(\gamma_1 + \kappa_1)}_{=\gamma_2} \text{TREND} \end{cases}$$

Then we can test using normal F-test by testing: (Chow test)

$H_0: \beta_1 = 0$ vs $H_1: \text{Not } H_0$

Now using Stata - printout, we find:

$$F_{\text{obs}} = \frac{(SSE_R - SSE_U) / J}{SSE_U / (N - K)}$$

Note: SSEU is now the sum of SSE from the two partial sample periods.

$$F_{\text{obs}} = \frac{(0,143478791 - 0,141068877) / 1}{0,141068877 / 105} = \underline{\underline{2,08}}$$

$$F_c = F_{(0,95,1,105)} \approx \underline{\underline{3,95}}$$

Hence we ^{do not} reject H_0 and conclude there's

not sufficient evidence in data to suggest a structural breakdown after WWII.

Other breaks: The great recession, WWI

④ All assumptions, but possibly one (MR4)

a) are still in place.

They are:

$$1) Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i$$

$$2) E(Y_i) = \beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik} \quad i=1, 2, \dots, N$$

$$\Leftrightarrow E(\varepsilon_i) = 0$$

$$3) \text{Var}(Y_i) = \text{Var}(\varepsilon_i) = \sigma^2$$

$$4) \text{Cov}(Y_i, Y_j) = \text{Cov}(\varepsilon_i, \varepsilon_j) = 0 \quad (i \neq j)$$

5) Stochastic X's and they are not exact linear functions of each other (multicollinearity).

$$b) Y_i \sim N(\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}, \sigma^2) \Leftrightarrow \varepsilon_i \sim N(0, \sigma^2)$$

We now use t and s -subscripts to indicate time-series (instead of i and j).

Note, however, that 4) might be violated.

OBS !! \Rightarrow NOTE

Flytt i handocet.

b)

$$LM = (T-1) R^2 = 63 \times 0,0001 = 0,0063$$

\Rightarrow NOT Evidence of autoregressive error terms.

c) di $\text{invtail}(38, 0,4/2) \approx 0,85118276$

$$\Rightarrow \left[\hat{\beta}_1 \pm t_{(38, \frac{0,4}{2})} \text{se}(\hat{\beta}_1) \right]$$

$$= \left[0,9599643 \pm 0,85118276 \times 0,0526567 \right]$$

$$= \underline{\underline{[0,915, 1,005]}}$$

d) Here, we see that (1) in the problem set is a special case of the two regressions from seminar 3 if

$$i) \beta_1 = \beta_3 = \beta_4 = \dots = \beta_8 = 0$$

$$ii) \beta_1 = 1 \text{ and } \beta_2 = \beta_3 = \dots = \beta_8 = 0$$

Hence we can conduct F-tests to test by using (1) as unrestricted and respectively the two models from seminar 3 as restricted models.

$$H_0: \beta_1 = \beta_3 = \beta_4 = \dots = \beta_8 \quad \text{vs. } H_1: \text{Not } H_0.$$

$$F_{obs} = \frac{(SSE_R - SSE_U) / 7}{SSE_U / (N - k)} = \frac{(0,264435223 - 0,008954321) / 7}{0,008954321 / ((64 - 9))}$$

$$F_{obs} = \underline{\underline{224,176679}}$$

Reject H_0 and conclude our first specification was wrong.

Test for second model:

$$\text{Recall } DLBNPCAP = \sum_{i=1}^8 \beta_i DTREND + u_t$$

and since we assume u_t has classical properties,

(1) is a special case of this if $\beta_1 = 1$ and $\beta_2, \beta_3, \dots, \beta_8 = 0$.

$$\text{So } H_0: \beta_1 = 1 \text{ and } \beta_2, \beta_3, \dots, \beta_8 = 0$$

vs. H_1 : Not H_0 .

$$F_{\text{obs}} = \frac{(0,028486275 - 0,008954321) / 8}{0,008954321 / (64 - 9)} = \underline{\underline{15,0}}$$

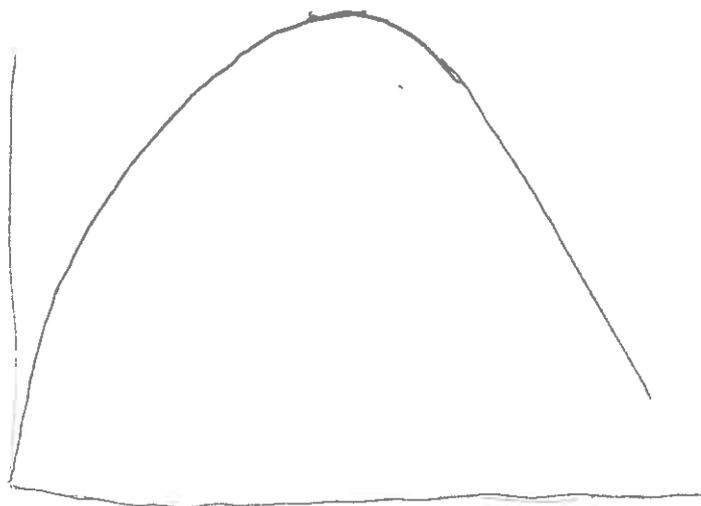
$$F_c = F(0,95, 8, 57) \approx 2,11$$

\Rightarrow Reject H_0 at 5% sign. level.

e) Since coefficient of trend is positive in the linear case, but negative in the squared, $\ln pcap$ will be concave (strictly) in trend. Hence at some point, it turns and has a more and more negative impact.

This might seem unrealistic, and is a clear suggestion that this model might have a limited relevant time frame.

(Since it is not at all obvious that GDP over time has a negative trend!)



8.2

b) So $\sigma_i^2 = \sigma^2 \quad \forall i$

This implies:

$$\sum \sigma_i^2 Y_i X_i = \sigma^{-2} \sum Y_i X_i$$

$$\sum \sigma_i^{-2} Y_i = \sigma^{-2} \sum Y_i$$

$$\sum \sigma_i^{-2} X_i = \sigma^{-2} \sum X_i$$

$$\sum \sigma_i^{-2} = n \sigma^{-2}$$

And thus:

$$\hat{\beta}_2 = \frac{\frac{\sigma^{-2} \sum X_i Y_i}{N \sigma^{-2}} - \left(\frac{\sigma^{-2} \sum Y_i}{N \sigma^{-2}} \right) \left(\frac{\sigma^{-2} \sum X_i}{N \sigma^{-2}} \right)}{\frac{\sigma^{-2} \sum X_i^2}{N \sigma^{-2}} - \left(\frac{\sigma^{-2} \sum X_i}{N \sigma^{-2}} \right)^2}$$

$$\hat{\beta}_2 = \frac{\frac{\sum Y_i X_i}{n} - \bar{Y} \bar{X}}{\frac{\sum X_i^2}{n} - \bar{X}^2}$$

And;

$$\hat{\beta}_1 = \frac{\sigma^{-2} \sum Y_i}{N \sigma^{-2}} - \left(\frac{\sigma^{-2} \sum X_i}{N \sigma^{-2}} \right) \hat{\beta}_2 = \underline{\underline{\bar{Y} - \hat{\beta}_2 \bar{X}}}$$

And these are our usual least squares estimators.

c) In the weighted averages in generalized least squares, each observation is weighted by the inverse of the error variance. Reliable observations with small error variances are weighted more heavily than those with higher error variances, making them more unreliable.

NOTE (specification)

Sequential exogeneity, meaning

$\ln p_{cap,t-1}$ can (most likely is) correlated with ε_{t-k} $k \in (1, \dots)$. And thus we

might have (most likely do have) bias in the samples. However, we have

consistent estimators, meaning

$$\text{plim}_{n \rightarrow \infty} (\hat{\beta}_k) = \beta_k.$$

SEMINAR 10

$$\textcircled{9.7} \quad e_t = \rho e_{t-1} + v_t$$

Rewrite now such that:

$$e_t = \rho(\rho e_{t-2} + v_{t-1}) + v_t \quad \text{by inserting for the period before.}$$

Continue substituting like this (iterating) k times;

$$e_t = \rho^k e_{t-k} + v_t + \rho v_{t-1} + \dots + \rho^{k-1} v_{t-k+1}$$

Let $k \rightarrow \infty$, then the first & last terms go to 0 as $-1 < \rho < 1$

$$\textcircled{\text{Hence;}} \quad e_t = v_t + \rho v_{t-1} + \rho^2 v_{t-2} + \dots$$

$$\text{And } E(e_t) = E(v_t) + \rho E(v_{t-1}) + \rho^2 E(v_{t-2}) \dots$$

$$\Rightarrow E(e_t) = 0$$

$$\text{var}(e_t) = \text{var}(v_t) + \rho^2 \text{var}(v_{t-1}) + \rho^4 \text{var}(v_{t-2})$$

$$\Rightarrow \text{var}(e_t) = \frac{\sigma_v^2}{1 - \rho^2} \quad \left| \begin{array}{l} \text{From;} \\ \sigma_v^2 \sum_{i=0}^{\infty} \rho^{2i} = \sigma_v^2 \frac{1}{1 - \rho^2} \\ \text{from infinite sum rule} \end{array} \right.$$

White

$$\text{cov}(e_t, e_{t-1}) = E(e_t e_{t-1})$$

$$= E[(v_t + \rho v_{t-1} + \dots)(v_{t-1} + \rho v_{t-2} + \dots)]$$

$$= \rho E(v_{t-1}^2) + \rho^3 E(v_{t-2}^2) + \rho^5 E(v_{t-3}^2) + \dots$$

(Since $\text{cov}(v_{t-k}, v_{t-j}) = E(v_{t-k} v_{t-j}) = 0$ $k \neq j$)

$$\sum_{k=0}^{\infty} \rho^{2k} = \frac{1}{1-\rho^2}$$

$$\Rightarrow \rho \sigma_v^2 (1 + \rho^2 + \rho^4 + \dots)$$

$$\Rightarrow \text{cov}(e_t, e_{t-1}) = \frac{\rho \sigma_v^2}{1-\rho^2}$$

It can then be shown that in general

$$\text{cov}(e_t, e_{t-k}) = \frac{\rho^k \sigma_v^2}{1-\rho^2} \quad k > 0$$

$$\text{Thus } \text{corr}(e_t, e_{t-k}) = \frac{\text{cov}(e_t, e_{t-k})}{\sqrt{\text{var}(e_t)} \sqrt{\text{var}(e_{t-k})}}$$

$$\text{i) } \sigma_v^2 = 1 \quad \rho = 0,9$$

$$\Rightarrow \text{corr}(e_t, e_{t-1}) = \frac{0,9 \times 1}{1 - 0,9^2} \cdot \frac{\sqrt{\frac{1}{1-0,9^2}} \sqrt{\frac{1}{1-0,9^2}}}{\sqrt{\frac{1}{1-0,9^2}} \sqrt{\frac{1}{1-0,9^2}}}$$

$$\Rightarrow \text{corr}(e_t, e_{t-1}) = \underline{\underline{0,9}}$$

$$\text{ii) } \text{corr}(e_t, e_{t-4}) = \frac{0,9^4 \times 1}{1 - 0,9^2} \cdot \frac{1}{1 - 0,9^2} = \underline{\underline{0,6561}}$$

Pattern shows clearly:

$$\boxed{\text{corr}(e_t, e_{t-k}) = \rho^k}$$

$$\forall \sigma_v^2 \in \mathbb{R}^+$$

We showed that

$$\text{var}(e_t) = \sigma_e^2 = \frac{\sigma_v^2}{1-\rho^2}$$

$$\Rightarrow \text{iii)} \quad \sigma_e^2 = \frac{1}{1-0,9^2} = \underline{\underline{5,263}}$$

$$\text{b) i)} \quad \text{corr}(e_t, e_{t-1}) = \underline{\underline{0,4}}$$

$$\text{ii)} \quad \text{corr}(e_t, e_{t-4}) = 0,4^4 = \underline{\underline{0,0256}}$$

$$\text{iii)} \quad \sigma_e^2 = \frac{1}{1-0,4^2} = \underline{\underline{1,1905}}$$

We now have weaker correlation between current and previous period errors. Thus the correlation between very distant errors dies out more quickly. More correlated errors also make them more volatile.

②

$$(1) E(Y) = e^{\mu + \frac{1}{2}\sigma^2}$$

$$(2) E[\ln(Y)] = \mu$$

$$a) \ln E(Y) = \ln e^{\mu + \frac{1}{2}\sigma^2} = \mu + \frac{1}{2}\sigma^2$$

$$\sigma^2 > 0 \Rightarrow \ln E(Y) = \mu + \frac{1}{2}\sigma^2 > E[\ln(Y)] = \mu$$

b) We start by recalling that the expected value is an average over an infinite number of observations.

Hence;

$$E(Y) = \mu = \frac{1}{n} \sum_{i=1}^n Y_i = \overline{Y}$$

$$E(\ln Y) = \frac{1}{n} \sum_{i=1}^n \ln Y_i = \overline{\ln Y}$$

Now this means we immediately from (2)

get ;

$$E(\ln Y) = \hat{\mu} = \underline{\underline{\overline{\ln Y}}}$$

While (1) gives ; $\ln E(Y) = \ln \bar{Y} = \mu + \frac{1}{2} \sigma^2$

$$\Rightarrow \underline{\underline{\hat{\sigma}^2}} = 2 (\ln \bar{Y} - \overline{\ln Y})$$

c) Now $\text{var}(\ln Y) = \sigma^2$

This immediately indicates an overdetermined

System with three equations and two unknowns. Thus there might be a problem

if they lead to multiple σ^2 , making it

a subjective choice choosing between them.

$$\text{var}(\ln Y) = E((\ln Y)^2) - [E(\ln Y)]^2 = \sigma^2$$

$$\Rightarrow \sigma^2 = E((\ln Y)^2) - [E(\ln Y)]^2$$

$$\hat{\sigma}^2 = \frac{\overline{\ln Y^2} - \overline{\ln Y}^2}{n}$$

$$(3) Y = \alpha + \beta_x X + \beta_z Z + u$$

a) In order to find the variance, we could use rule for sums of stochastic variables inside a variance operator, but let's do it the "hard" way by using expectation terms. Then recall

$$\text{var}(Y) = E(Y^2) - [E(Y)]^2 \quad \text{Thus we need } Y^2$$

$$Y^2 = (\alpha + \beta_x X + \beta_z Z + u)^2$$

$$Y^2 = 2\alpha\beta_x X + 2\alpha\beta_z Z + 2\alpha u + 2\beta_x\beta_z XZ + 2\beta_x Xu + 2\beta_z Zu + \alpha^2 + \beta_x^2 X^2 + \beta_z^2 Z^2 + u^2$$

$$\text{And then } E(Y^2) = E(2\alpha\beta_x X + \dots)$$

$$\Rightarrow \text{var}(Y)$$

$$= 2\alpha\beta_x E(X) + 2\alpha\beta_z E(Z) + 2\alpha E(u) + 2\beta_x\beta_z E(XZ) \\ + 2\beta_x E(Xu) + 2\beta_z E(Zu) + \alpha^2 + \beta_x^2 E(X^2) + \beta_z^2 E(Z^2) \\ + E(u^2) - \left[\alpha + \beta_x E(X) + \beta_z E(Z) + E(u) \right]^2$$

$$= \left[\alpha^2 + 2\alpha\beta_x E(X) + 2\alpha\beta_z E(Z) + 2\alpha E(u) + 2\beta_x\beta_z E(X)E(Z) \\ + 2\beta_x E(X)E(u) + 2\beta_z E(Z)E(u) + \beta_x^2 E(X)^2 + \beta_z^2 E(Z)^2 \\ + E(u)^2 \right]$$

$$\Rightarrow \text{var}(Y) = \beta_x^2 [E(X^2) - E(X)^2] + \beta_z^2 [E(Z^2) - E(Z)^2]$$

$$+ [E(u^2) - E(u)^2] + 2\beta_x\beta_z [E(XZ) - E(X)E(Z)]$$

$$+ 2\beta_x [E(Xu) - E(X)E(u)] + 2\beta_z [E(Zu) - E(Z)E(u)]$$

$$\Rightarrow \text{var}(Z) = \beta_x^2 \text{var}(X) + \beta_z^2 \text{var}(Z) + \text{var}(u) \\ + 2\beta_x\beta_z \text{cov}(X, Z) + 2\beta_x \text{cov}(X, u) + 2\beta_z \text{cov}(Z, u)$$

$$\text{cov}(Y, X) = E(Y - E(Y))(X - E(X))$$

$$\Rightarrow \text{cov}(Y, X) = E\left[\left(\alpha + \beta_x X + \beta_z Z + u - (\alpha + \beta_x E(X) + \beta_z E(Z) + Eu)\right)\right. \\ \left.(X - E(X))\right]$$

$$\text{cov}(X, Y) = E\left[\left(\beta_x (X - E(X)) + \beta_z (Z - E(Z)) + (u - E(u))\right)\right. \\ \left.(X - E(X))\right]$$

$$\text{cov}(X, Y) = \beta_x E\left[(X - E(X))^2\right] + \beta_z E\left[(Z - E(Z))(X - E(X))\right] \\ + E\left[(u - E(u))(X - E(X))\right]$$

$$\text{cov}(X, Y) = \beta_x \text{var}(X) + \beta_z \text{cov}(Z, X) + \text{cov}(u, X)$$

Similar result (try yourself!) for $\text{cov}(Y, Z)$

Assume now:

$$\sigma_y^2 = \text{var}(Y), \quad \sigma_x^2 = \text{var}(X), \quad \sigma_z^2 = \text{var}(Z)$$

$$\sigma_{yx} = \text{cov}(Y, X), \quad \sigma_{yz} = \text{cov}(Y, Z), \quad \sigma_{xz} = \text{cov}(X, Z)$$

with corresponding empirical moments

$$s_y^2, \quad s_x^2, \quad s_z^2, \quad s_{yx}, \quad s_{yz}, \quad s_{xz}$$

that are consistent (convergence).

$$b) \quad \text{cov}(u, X) = \text{cov}(u, Z) = 0$$

$$\text{Hence; } \beta_x \text{var}(X) + \beta_z \text{cov}(Z, X) = \text{cov}(Y, X)$$

$$\beta_z \text{var}(Z) + \beta_x \text{cov}(Z, X) = \text{cov}(Y, Z)$$

$$\hat{\beta}_x = \frac{s_{xy} - \hat{\beta}_z s_{xz}}{s_x^2}$$

$$\Rightarrow \hat{\beta}_z s_z^2 + \left(\frac{s_{xy} - \hat{\beta}_z s_{xz}}{s_x^2} \right) s_{zx} = s_{zy}$$

$$\hat{\beta}_z \left(\frac{S_z^2}{S_{2x}} - \frac{S_{x2}}{S_x^2} \right) = \frac{S_{2y}}{S_{2x}} - \frac{S_{xy}}{S_x^2}$$

$$\Rightarrow \hat{\beta}_z = \frac{\frac{S_{2y}}{S_{2x}} - \frac{S_{xy}}{S_x^2}}{\frac{S_z^2}{S_{2x}} - \frac{S_{x2}}{S_x^2}}$$

$$\Rightarrow \hat{\beta}_z = \frac{S_{2y} S_x^2 - S_{xy} S_{2x}}{S_z^2 S_x^2 - S_{x2}^2}$$

$$\text{And } \hat{\beta}_x = \frac{S_{xy}}{S_x^2} - \frac{S_{x2}}{S_x^2} \hat{\beta}_z$$

$$\Rightarrow \hat{\beta}_x = \frac{S_{xy}}{S_x^2} - \frac{S_{x2}}{S_x^2} \left(\frac{S_{2y} S_x^2 - S_{xy} S_{2x}}{S_z^2 S_x^2 - S_{x2}^2} \right)$$

c) Assuming n observations of (x, y, z)

then $\lim_{n \rightarrow \infty} S_{xy} = \sigma_{xy}$

$$\lim_{n \rightarrow \infty} S_x^2 = \sigma_x^2 \quad \text{etc.}$$

Hence

$$\lim_{n \rightarrow \infty} \hat{\beta}_2 = \lim_{n \rightarrow \infty} \left(\frac{S_{zy} S_x^2 - S_{xy} S_{zy}}{S_z^2 S_x^2 - S_{xz}^2} \right)$$

$$\lim_{n \rightarrow \infty} \hat{\beta}_2 \stackrel{\text{Slutsky}}{=} \frac{\lim_{n \rightarrow \infty} S_{zy} S_x^2 - \lim_{n \rightarrow \infty} S_{xy} S_{zy}}{\lim_{n \rightarrow \infty} S_z^2 S_x^2 - \lim_{n \rightarrow \infty} S_{xz}^2}$$

$$\Rightarrow \lim_{n \rightarrow \infty} \hat{\beta}_2 = \frac{\sigma_{zy} \sigma_x^2 - \sigma_{xy} \sigma_{zy}}{\sigma_z^2 \sigma_x^2 - \sigma_{xz}^2} = \beta_2$$

(Refer seminar 5 Q3 where we showed this form of $\hat{\beta}_1$ using FW-theorem).

$$\text{plim}_{n \rightarrow \infty} \hat{\beta}_x = \frac{\text{plim}_{n \rightarrow \infty} S_{xy}}{\text{plim}_{n \rightarrow \infty} S_x^2} - \frac{\text{plim}_{n \rightarrow \infty} S_{xz}}{\text{plim}_{n \rightarrow \infty} S_x^2} \hat{\beta}_z$$

$$\Rightarrow \text{plim}_{n \rightarrow \infty} \hat{\beta}_x = \frac{\sigma_{xy}}{\sigma_x^2} - \frac{\sigma_{xz}}{\sigma_x^2} \hat{\beta}_z = \beta_x$$

d) $\beta_z = 0$, $\text{cov}(u, z) = 0$ $\text{cov}(u, x) \neq 0$
 $\text{cov}(x, z) \neq 0$ Because in denominator of β_x

(I) $\text{cov}(Y, X) = \beta_x \text{var}(X) + \text{cov}(u, X)$

(II) $\text{cov}(Y, Z) = \beta_x \text{cov}(X, Z)$

(III) Implies $\beta_x = \frac{\text{cov}(Y, Z)}{\text{cov}(X, Z)}$ immediate.

(I) then implies, inserting for β_x that

$$\text{cov}(u, X) = \text{cov}(Y, X) - \text{var}(X) \frac{\text{cov}(Y, Z)}{\text{cov}(X, Z)}$$

Instrument variable estimator

Since X is correlated with u (disturbance)
we use another variable Z with properties

- 1) Z does not have direct effect on Y
(which is true since $\beta_2 = 0$)
- 2) Z is uncorrelated with disturbance
(true since $\text{cov}(u, Z) = 0$)
- 3) Z is strongly correlated with X
(or at least not weakly)

Which is the real reason why $\text{cov}(X, Z) \neq 0$.

And we can then get our instrument variable estimators (IV estimators) as

$$E(Zu) = 0 \Rightarrow E(Z(Y - \beta_1 - \beta_2 X)) = 0$$

And our sample moment condition for β_2

is thus

$$\frac{1}{n} \sum Z_i (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)$$

And we hence get

$$\hat{\beta}_2 = \frac{\sum (z_i - \bar{z})(y_i - \bar{y})}{\sum (z_i - \bar{z})(x_i - \bar{x})} = \frac{E[(z - \bar{z})(y - \bar{y})]}{E[(z - \bar{z})(x - \bar{x})]} = \beta_2$$

$E(z)$
↓

e) We now have the equation set

$$\text{cov}(Y, X) = \beta_x \text{var}(X) + \beta_z \text{cov}(Z, X) + \text{cov}(u, X)$$

$$\text{cov}(Y, Z) = \beta_x \text{cov}(X, Z) + \beta_z \text{var}(Z)$$

We can no longer estimate β_x because Z is no longer an IV for X since it now has direct effect on Y . We have two equations in three unknowns ($\beta_x, \text{cov}(u, X), \beta_z$) which is an underdetermined system.

