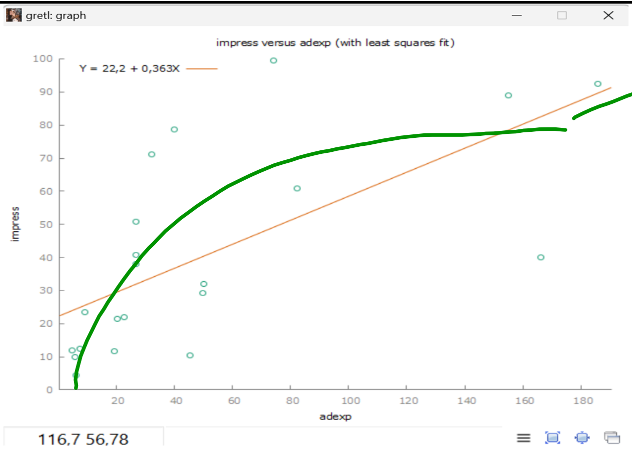
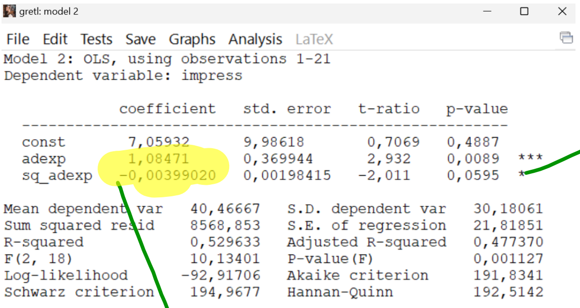
PS5



non-linear model

b) Estimate Model 2: $Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + u_i$.



```
Model 2: OLS, using observations 1-21
Dependent variable: impress

             coefficient   std. error   t-ratio   p-value
  --------------------------------------------------------
  const        7,05932     9,98618       0,7069   0,4887
  adexp        1,08471     0,369944      2,932    0,0089   ***
  sq_adexp    -0,00399020  0,00198415   -2,011    0,0595   *

Mean dependent var   40,46667   S.D. dependent var   30,18061
Sum squared resid    8568,853   S.E. of regression   21,81851
R-squared            0,529633   Adjusted R-squared   0,477370
F(2, 18)             10,13401   P-value(F)           0,001127
Log-likelihood      -92,91706   Akaike criterion     191,8341
Schwarz criterion    194,9677   Hannan-Quinn         192,5142
```

significant at a 10% sign level

negative

adexp↑ → impress↑ but effect is dampened when X is large (X^2 is very large)

Interpret: effect of X on Y?

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X + \hat{\beta}_2 X^2$$

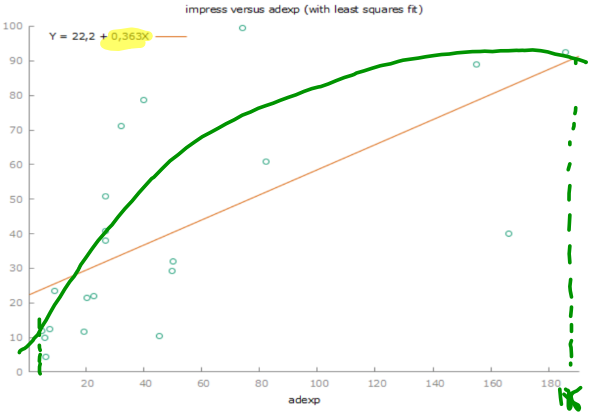
$$\frac{\partial \hat{y}}{\partial x} = \hat{\beta}_1 + 2\hat{\beta}_2 X$$

c) $$\left.\frac{\partial \hat{y}}{\partial x}\right|_{X=5,0} = \hat{\beta}_1 + 2\hat{\beta}_2 \cdot X = 1{,}08 + 2\cdot(-0{,}004)\cdot 5{,}0$$

$$= \underline{1{,}04}$$

d) $$\left.\frac{\partial \hat{y}}{\partial x}\right|_{X=185} = 1{,}08 + 2\cdot(-0{,}04)\cdot 185 = \underline{-0{,}4}$$

Slope of the regression function



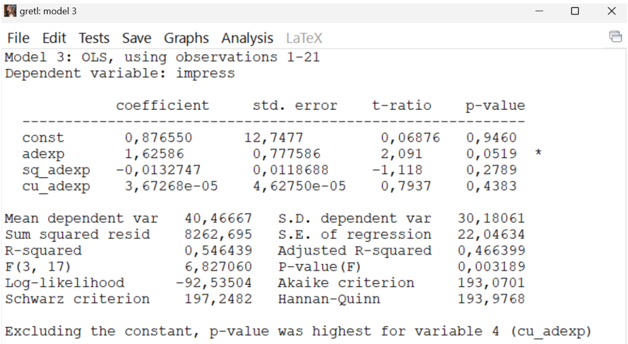
Gretl: $X^3 = X^2 \cdot X$
$= X \cdot X \cdot X$

e) Estimate Model 3: $Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \beta_3 X_i^3 + u_i$.

Add > Define new variable
cu_adexp= adexp^3

$$\frac{\partial \hat{Y}}{\partial X} = \hat{\beta}_1 + 2\hat{\beta}_2 X + 3\cdot \hat{\beta}_3 X^2$$

$$\boxed{(X^a)' = a\cdot X^{a-1}}$$



```
Model 3: OLS, using observations 1-21
Dependent variable: impress

             coefficient     std. error    t-ratio    p-value
  ----------------------------------------------------------
  const       0,876550      12,7477         0,06876   0,9460
  adexp       1,62586        0,777586       2,091     0,0519  *
  sq_adexp   -0,0132747      0,0118688     -1,118     0,2789
  cu_adexp    3,67268e-05    4,62750e-05    0,7937    0,4383

Mean dependent var   40,46667   S.D. dependent var   30,18061
Sum squared resid    8262,695   S.E. of regression   22,04634
R-squared            0,546439   Adjusted R-squared   0,466399
F(3, 17)             6,827060   P-value(F)           0,003189
Log-likelihood      -92,53504   Akaike criterion     193,0701
Schwarz criterion    197,2482   Hannan-Quinn         193,9768

Excluding the constant, p-value was highest for variable 4 (cu_adexp)
```

e-05: $\cdot 10^{-5}$

1e-05 ≑ $1\cdot 10^{-5} = 0{,}00001$

f) What is the effect of increasing expenditures with one million dollars on impressions if expenditures are 5,0 millions of dollars in Model 3?

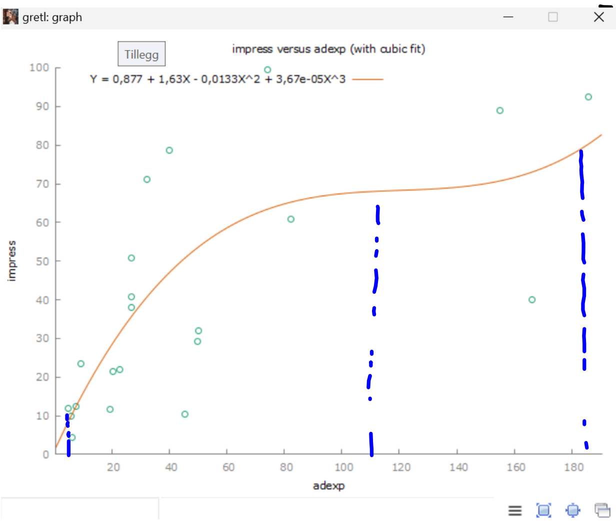
g) What is the effect of increasing expenditures with one million dollars on impressions if expenditures are 185 millions of dollars in Model 3?

f) $$\left.\frac{\partial \hat{Y}}{\partial X}\right|_{X=5,0} = \hat{\beta}_1 + 2\hat{\beta}_2 5{,}0 + 3\cdot\hat{\beta}_3\cdot 5{,}0^2 = 1{,}63 + 2\cdot(-0{,}013)\cdot 5 + 3\cdot 0{,}000037\cdot 5^2$$

$$\underline{=1{,}5}$$

g) $$\left.\frac{\partial \hat{Y}}{\partial X}\right|_{X=185} = 1{,}63 + 2\cdot(-0{,}013)\cdot 185 + 3\cdot 0{,}000037\cdot 185^2$$

$$= \underline{0{,}58}$$

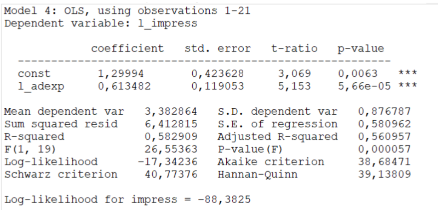


Non-linear models:
Effect of advertising on impressions is larger if X is small.

Effect of increasing advertising by 1 mill. dollars

h) Estimate Model 4: $\ln Y_i = \beta_0 + \beta_1 \ln X_i + u_i$

i) Interpret the estimated slope coefficient in Model 4?



```
Model 4: OLS, using observations 1-21
Dependent variable: l_impress

             coefficient   std. error   t-ratio   p-value
  -------------------------------------------------------
  const       1,29994      0,423628     3,069     0,0063   ***
  l_adexp     0,613482     0,119053     5,153     5,66e-05 ***

Mean dependent var   3,382864   S.D. dependent var   0,876787
Sum squared resid    6,412815   S.E. of regression   0,580962
R-squared            0,582909   Adjusted R-squared   0,560957
F(1, 19)             26,55363   P-value(F)           0,000057
Log-likelihood      -17,34236   Akaike criterion     38,68471
Schwarz criterion    40,77376   Hannan-Quinn         39,13809

Log-likelihood for impress = -88,3825
```

log-log function

$\hat{\beta}_1$: If X increase by 1%, Y is expected to increase by β_1%.

i) If ad exp. increase by 1%, impressions is expected to increase by 0,61%

If X=5: 1% increase in X is

5 → 5,05 (increase by 0,05)

X=185 : 1% increase in X is

185 → 186,85 (increase by 1.85)

non-linearity

Forecast

j) Predict the number of impressions if expenditures are 150 million dollars in ~~i) Model 1~~, ii) Model 2, iii) Model 3 and iv) Model 4.

ii) $\hat{Y} = 7{,}06 + 1{,}08 \cdot X - 0{,}004 \cdot X^2 = 7.06 + 1{,}08 \cdot 150 - 0{,}004 \cdot 150^2 = \boxed{79.06}$

iii) $\hat{Y} = 0{,}88 + 1{,}63 \cdot X - 0{,}013\, X^2 + 0{,}00004 \cdot X^3 \boxed{= 87{,}88}$

iv) $\widehat{\ln Y} = 1{,}30 + 0{,}61 \cdot \ln X$ $\hat{Y} = \dots$ not $\widehat{\ln Y} = \dots$

$e^{\ln X} = X$

$e^{\widehat{\ln Y}} = e^{1{,}30 + 0{,}61 \cdot \ln X}$

$\hat{Y} = e^{1{,}30 + 0{,}61 \cdot \ln X} = e^{1{,}30 + 0{,}61 \cdot \ln 150} = e^{4{,}36} \boxed{= 78{,}26}$

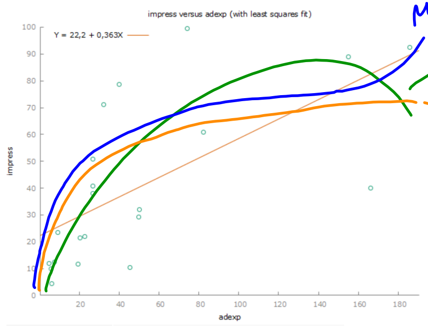
k) Model 2, Model 3 or Model 4?

1) Residuals when X=154,9 $\hat{u} = Y - \hat{Y}$. But also need to look at other values of X

2) R^2 adjusted because different numbers of regressors in Models 2 and 3

Cannot use R^2 adj. for Model 4 since dep. variable is not the same (ln Y)

3) Consider the data (XY scatterplot)



Model 3

Model 2

Model 3

More in line with theory

↓

effect of X on Y decrease when X becomes large.

$\widehat{sales} = -3001.06 + 109.97 \cdot price$

If price increase by 1,
we expect sales to increase
by 109.97.

$\boxed{\text{SPURIUS REGRESSION}}$

startup. More customers as you get more known → Sales increase from 2001-2006

inflation → price increase from 2001-2006

$Sales_t = \beta_0 + \beta_1 \cdot price_t + \beta_2 t + u_t$

$\widehat{Sales} = 7811,37 - 68,81 \cdot price_t + 991,37 \cdot t$

↑ insignificant ↑ significant

spurious regression taken into account by adding a trend

Autoregressive (AR) model:

$X_t = \beta_0 + \beta_1 X_{t-1} + u_t$ — AR(1)

$\boxed{\text{lags...}}$

AR(2): $X_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-2} + u_t$

AR-model with drift:

AR(1) with drift: $X_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 t + u_t$

Seasonal dummies:

Gretl: Add > Periodic dummies

(do not center)