Ekonometrija 1

Enajste vaje:
Konstantnost variance slučajne spremenljivke in heteroskedastičnost. Odsotnost koreliranosti slučajne spremenljivke in avtokorelacija.

Na enajstih vajah bomo nadaljevali z obravnavo homoskedastičnosti kot ene temeljnih predpostavk klasičnega linearnega regresijskega modela. Spoznali bomo dva dodatna testa za preverjanje te predpostavke, tj. Breusch–Paganov in Whiteov test ter dodaten postopek ocenjevanja regresijskih modelov v primeru heteroskedastičnosti. Nato pa bomo obravnavali še eno temeljnih predpostavk klasičnega linearnega regresijskega modela, in sicer predpostavko o odsotnosti avtokorelacije.

♣-♣-♣

Primer 1: Na podlagi podatkov za 50 ameriških zveznih držav za leto 1985 ocenite naslednjo funkcijo porabe bencina (glej datoteko `bencin.dta`):

$$CON_i = \beta_1 + \beta_2 REGMV_i + \beta_3 TAX_i + u_i .$$

Pri tem imajo posamezne spremenljivke naslednji pomen:

- *CON*: poraba bencina v posamezni državi (v milijonih sodčkov);
- *REGMV*: število registriranih motornih vozil (v tisočih);
- *TAX*: stopnja trošarine za bencin (v centih na galono).

a) Opravite Breusch–Paganov test prisotnosti heteroskedastičnosti in pojasnite vaše ugotovitve. S pomočjo metode PNK–TNK zapišite model v obliki, ki bo zagotavljala, da je predpostavka o homoskedastičnosti ponovno izpolnjena in to tudi dokažite. Ocenite ta model, nato pa zapišite ustrezne (NENALICE) ocene prvotnega regresijskega modela.

b) Opravite Whiteov test prisotnosti heteroskedastičnosti in pojasnite vaše ugotovitve. S pomočjo metode PNK–TNK zapišite model v obliki, ki bo zagotavljala, da je predpostavka o homoskedastičnosti ponovno izpolnjena in to tudi dokažite. Ocenite ta model, nato pa zapišite ustrezne (NENALICE) ocene prvotnega regresijskega modela.

c) Poskušajte odpraviti ugotovljeno prisotnost heteroskedastičnosti še s pomočjo logaritemske transformacije spremenljivk. Ste bili pri tem uspešni?

d) Nazadnje s pomočjo Huber/White robustne cenilke variance izračunajte nepristranske standardne napake ocen parametrov. Katero predpostavko klasičnega linearnega regresijskega modela ste pri tem sprostili in kako?

Izpis rezultatov obdelav v programskem paketu Stata:

```
. regress con regmv tax

      Source |       SS       df       MS              Number of obs =      50
-------------+------------------------------           F(  2,    47) =  152.36
       Model |  19504553.9     2  9752276.93           Prob > F      =  0.0000
    Residual |  3008446.64    47  64009.5029           R-squared     =  0.8664
-------------+------------------------------           Adj R-squared =  0.8607
       Total |  22513000.5    49   459448.99           Root MSE      =     253

------------------------------------------------------------------------------
         con |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       regmv |   .1861319   .0117189    15.88   0.000     .1625565    .2097074
         tax |  -53.59101   16.85588    -3.18   0.003    -87.50067   -19.68134
       _cons |    551.688   186.2709     2.96   0.005     176.9592    926.4168
------------------------------------------------------------------------------

. predict econ, resid
. gen econ2=econ^2


. * Breusch-Paganov test

. sum econ2

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       econ2 |        50    60168.93    207477.4   60.18166    1365618

. regress econ2 regmv tax

      Source |       SS       df       MS              Number of obs =      50
-------------+------------------------------           F(  2,    47) =    9.17
       Model |  5.9199e+11     2  2.9600e+11           Prob > F      =  0.0004
    Residual |  1.5173e+12    47  3.2283e+10           R-squared     =  0.2807
-------------+------------------------------           Adj R-squared =  0.2500
       Total |  2.1093e+12    49  4.3047e+10           Root MSE      =  1.8e+05

------------------------------------------------------------------------------
       econ2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       regmv |   23.26149   8.322487     2.80   0.007     6.518812    40.00418
         tax |  -29576.46   11970.63    -2.47   0.017    -53658.25   -5494.665
       _cons |   284600.5   132284.9     2.15   0.037     18477.51    550723.5
------------------------------------------------------------------------------

. display invchi2tail(2,0.05)
5.9914645

. display chi2tail(2,81.774988)
1.749e-18


. qui regress con regmv tax
. estat hettest regmv tax

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: regmv tax

         chi2(2)      =    81.76
         Prob > chi2  =   0.0000
```

```
. qui regress econ2 regmv tax
. predict econ2fit, xb
. inspect econ2fit

econ2fit:  Linear prediction                      Number of Observations
----------------------------                 ------------------------------
                                             Total    Integers   Nonintegers
|      #                     Negative          14           1            13
|      #                     Zero               -           -             -
|      #                     Positive          36           -            36
|  #   #                                    -----       -----         -----
|  #   #                     Total             50           1            49
|  #   #   #    .   .        Missing            -
+----------------------                     -----
-111465.5      416881.8                        50
  (50 unique values)

. gen lecon2=log(econ2)
. qui regress lecon2 regmv tax
. predict lecon2fit, xb

. gen td=sqrt(exp(lecon2fit))

. gen cont=con/td
. gen regmvt=regmv/td
. gen taxt=tax/td
. gen kt=1/td

. regress cont regmvt taxt kt, nocons

      Source |       SS       df       MS              Number of obs =      50
-------------+------------------------------           F(  3,    47) =  204.85
       Model |  1590.87775     3  530.292582           Prob > F      =  0.0000
    Residual |   121.6675    47  2.58867021           R-squared     =  0.9290
-------------+------------------------------           Adj R-squared =  0.9244
       Total |  1712.54525    50  34.2509049           Root MSE      =  1.6089

------------------------------------------------------------------------------
        cont |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      regmvt |   .1628148   .0108746    14.97   0.000      .140938    .1846916
        taxt |  -20.32982   8.020062    -2.53   0.015     -36.4641   -4.195536
          kt |   253.5014   95.08184     2.67   0.010     62.22145    444.7814
------------------------------------------------------------------------------

. gen contfit=_b[regmvt]*regmv+_b[taxt]*tax+_b[kt]

. correlate con contfit
(obs=50)

             |      con  contfit
-------------+------------------
         con |   1.0000
     contfit |   0.9261   1.0000


. scalar r2p=r(rho)^2
. display r2p
.85758636

. drop econ2fit td cont regmvt taxt kt contfit lecon2 lecon2fit


. * Whiteov test

. gen regmv2=regmv^2
. gen tax2=tax^2
. gen regmvtax=regmv*tax
```

```
. regress econ2 regmv tax regmv2 tax2 regmvtax

      Source |       SS       df       MS              Number of obs =      50
-------------+------------------------------           F(  5,    44) =   17.43
       Model |  1.4017e+12     5  2.8033e+11           Prob > F      =  0.0000
    Residual |  7.0764e+11    44  1.6083e+10           R-squared     =  0.6645
-------------+------------------------------           Adj R-squared =  0.6264
       Total |  2.1093e+12    49  4.3047e+10           Root MSE      =  1.3e+05

------------------------------------------------------------------------------
       econ2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       regmv |   140.5628   39.92468     3.52   0.001     60.09989    221.0257
         tax |    -237873   103340.2    -2.30   0.026    -446141.4   -29604.56
      regmv2 |  -.0005147     .00106    -0.49   0.630    -.002651    .0016216
        tax2 |   12346.99   4688.527     2.63   0.012      2897.89     21796.1
    regmvtax |  -12.84508    3.60418    -3.56   0.001    -20.10883    -5.58133
       _cons |    1098291   556502.1     1.97   0.055    -23265.35     2219847
------------------------------------------------------------------------------

. scalar theta=e(N)*e(r2)

. display theta, invchi2tail(e(rank)-1,0.05), chi2tail(e(rank)-1,theta)
33.22564 11.070498 3.394e-06


. qui regress con regmv tax
. estat imtest, white

White's test for Ho: homoskedasticity
         against Ha: unrestricted heteroskedasticity

         chi2(5)      =     33.23
         Prob > chi2  =    0.0000

Cameron & Trivedi's decomposition of IM-test

---------------------------------------------------
              Source |       chi2     df      p
---------------------+-----------------------------
  Heteroskedasticity |      33.23      5    0.0000
            Skewness |      13.51      2    0.0012
            Kurtosis |       1.57      1    0.2095
---------------------+-----------------------------
               Total |      48.31      8    0.0000
---------------------------------------------------


. qui regress econ2 regmv tax regmv2 tax2 regmvtax
. predict econ2fit, xb
. inspect econ2fit

econ2fit:  Linear prediction                     Number of Observations
----------------------------                -----------------------------
                                            Total   Integers   Nonintegers
|  #                         Negative          23          -           23
|  #                         Zero               -          -            -
|  #                         Positive          27          -           27
|  #                                        -----      -----        -----
|  #                         Total             50          -           50
|  #   #   .   .   .         Missing            -
+----------------------                     -----
-98912.37       1020180                        50
  (50 unique values)

. gen lecon2=log(econ2)
. qui regress lecon2 regmv tax regmv2 tax2 regmvtax
. predict lecon2fit, xb
```

```
. gen td=sqrt(exp(lecon2fit))

. gen cont=con/td
. gen regmvt=regmv/td
. gen taxt=tax/td
. gen kt=1/td

. regress cont regmvt taxt kt, nocons

      Source |       SS       df       MS              Number of obs =      50
-------------+------------------------------           F(  3,    47) =  286.87
       Model |  2201.7405     3  733.913501           Prob > F      =  0.0000
    Residual | 120.240758    47  2.55831399           R-squared     =  0.9482
-------------+------------------------------           Adj R-squared =  0.9449
       Total |  2321.98126    50  46.4396252           Root MSE      =  1.5995

------------------------------------------------------------------------------
        cont |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      regmvt |   .1662324    .009169    18.13   0.000     .1477868     .184678
        taxt |  -24.70128   10.15112    -2.43   0.019    -45.12271   -4.279852
          kt |    286.102   110.0896     2.60   0.012     64.63024    507.5738
------------------------------------------------------------------------------

. gen contfit=_b[regmvt]*regmv+_b[taxt]*tax+_b[kt]

. correlate con contfit
(obs=50)

             |      con  contfit
-------------+------------------
         con |   1.0000
     contfit |   0.9274   1.0000


. scalar r2p=r(rho)^2
. display r2p
.86002384

. drop econ2fit td cont regmvt taxt kt contfit lecon2 lecon2fit


. * Logaritemska transformacija spremenljivk

. gen lcon=log(con)
. gen lregmv=log(regmv)
. gen ltax=log(tax)

. regress lcon lregmv ltax

      Source |       SS       df       MS              Number of obs =      50
-------------+------------------------------           F(  2,    47) =  253.25
       Model | 38.3992068     2  19.1996034           Prob > F      =  0.0000
    Residual | 3.56316785    47  .075812082           R-squared     =  0.9151
-------------+------------------------------           Adj R-squared =  0.9115
       Total | 41.9623747    49  .856374993           Root MSE      =  .27534

------------------------------------------------------------------------------
        lcon |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lregmv |   .8796717   .0415988    21.15   0.000     .7959857    .9633578
        ltax |   -.649004   .1747917    -3.71   0.001     -1.00064   -.2973685
       _cons |   .7274507    .555704     1.31   0.197    -.3904817    1.845383
------------------------------------------------------------------------------
```

```
. estat hettest lregmv ltax

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: lregmv ltax

         chi2(2)      =    21.25
         Prob > chi2  =   0.0000

. estat imtest, white

White's test for Ho: homoskedasticity
         against Ha: unrestricted heteroskedasticity

         chi2(5)      =     13.61
         Prob > chi2  =    0.0183

Cameron & Trivedi's decomposition of IM-test

---------------------------------------------------
              Source |       chi2     df      p
---------------------+-----------------------------
  Heteroskedasticity |      13.61      5    0.0183
            Skewness |       1.34      2    0.5109
            Kurtosis |       2.05      1    0.1525
---------------------+-----------------------------
               Total |      17.00      8    0.0301
---------------------------------------------------


. * Uporaba robustnih standardnih napak

. regress con regmv tax, robust

Linear regression                                      Number of obs =      50
                                                       F(  2,    47) =   41.10
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.8664
                                                       Root MSE      =     253

------------------------------------------------------------------------------
             |               Robust
         con |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       regmv |   .1861319    .021533     8.64   0.000     .1428132    .2294507
         tax |  -53.59101    23.8975    -2.24   0.030    -101.6666   -5.515436
       _cons |    551.688   237.1574     2.33   0.024     74.58876    1028.787
------------------------------------------------------------------------------
```

■

<u>Primer 2:</u> Na podlagi četrtletnih podatkov za ZDA za obdobje 1976q1–1990q4 ocenite naslednjo funkcijo povpraševanja po novih avtomobilih (glej datoteko `avtotrg.dta`):

$$AVTOP_t = \beta_1 + \beta_2 CENA_t + \beta_3 RRDP_t + \beta_4 OM_t + u_t \, .$$

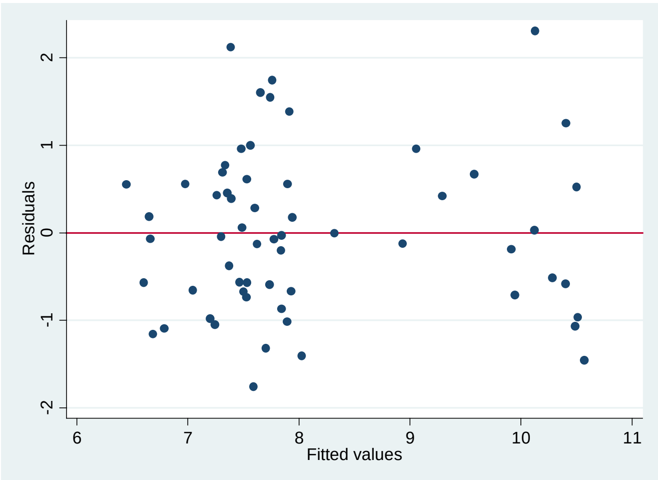
Pri tem imajo posamezne spremenljivke naslednji pomen:

- *AVTOP*: število prodanih novih avtomobilov na 1.000 prebivalcev;
- *CENA:* indeks cen novih avtomobilov (1982q1=100);
- *RRDP*: realni razpoložljivi dohodek na prebivalca (v 1.000 USD);
- *OM*: referenčna obrestna mera pri poslovnih bankah (v %).

Najprej s pomočjo grafične metode odkrivanja avtokorelacije preverite, ali je v zgornji funkciji povpraševanja po novih avtomobilih mogoče zaznati prisotnost avtokorelacije prvega in četrtega reda. Grafično preverite tudi morebitne znake slabe specifikacije modela. Nato opravite še Durbin–Watsonov *d*–test. Pri tem ocenite tudi koeficient avtokorelacije prvega in četrtega reda (na različne načine). Pojasnite vaše ugotovitve.

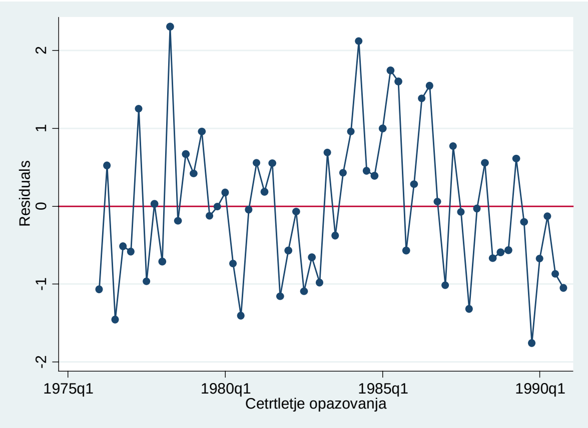
Izpis rezultatov obdelav v programskem paketu Stata:

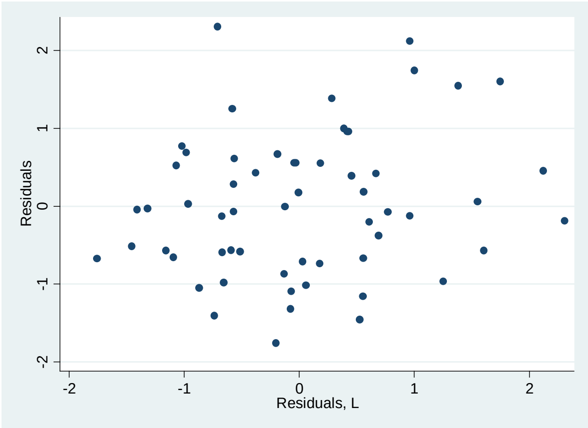
```
. tsset kvartal
        time variable:  kvartal, 1976q1 to 1990q4
                delta:  1 quarter

. regress avtop cena rrdp om

      Source |       SS       df       MS              Number of obs =      60
-------------+------------------------------           F(  3,    56) =   31.69
       Model |  85.0940388     3  28.3646796           Prob > F      =  0.0000
    Residual |  50.1162514    56  .894933061           R-squared     =  0.6293
-------------+------------------------------           Adj R-squared =  0.6095
       Total |   135.21029    59  2.29169983           Root MSE      =  .94601

------------------------------------------------------------------------------
       avtop |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        cena |  -.1062054   .0200516    -5.30   0.000    -.1463737   -.0660371
        rrdp |   1.241812   .4282211     2.90   0.005     .3839823    2.099642
          om |  -.1344411    .042009    -3.20   0.002    -.2185954   -.0502869
       _cons |   7.012139   3.027984     2.32   0.024     .9463581    13.07792
------------------------------------------------------------------------------

. predict eavto, resid

. * Analiza prisotnosti avtokorelacije - Graficna metoda

. rvfplot, yline(0)
```



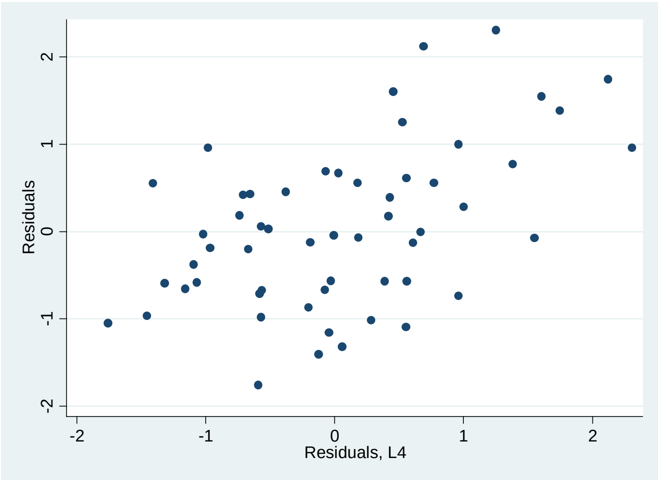

```
. gen eavto1=eavto[_n-1]
(1 missing value generated)

. gen eavto4=eavto[_n-4]
(4 missing values generated)
```

. twoway connected eavto kvartal, yline(0) **. scatter eavto l.eavto**





. scatter eavto l4.eavto



. * Analiza prisotnosti avtokorelacije - Ocenjevanje koeficienta AK 1. reda

. gen eeavtott1=eavto*eavto1
(1 missing value generated)

. qui sum eeavtott1, detail
. scalar rhostevec=r(N)*r(mean)

. gen eavto2=eavto^2
. qui sum eavto2, detail
. scalar rhoimenovalec1=r(N)*r(mean)

. scalar rhoapprox1=rhostevec/rhoimenovalec1

. display rhostevec, rhoimenovalec1, rhoapprox1
9.4984437 50.116252 .18952821

. corrgram eavto

LAG	AC	PAC	Q	Prob>Q	-1 0 1 [Autocorrelation]	-1 0 1 [Partial Autocor]
1	0.1895	0.1935	2.2648	0.1323	|-	|-
2	0.1233	0.0937	3.2397	0.1979	|	|
3	0.1002	0.0692	3.8949	0.2730	|	|

```
4        0.4777   0.4853   19.056  0.0008          |---              |---
5        0.0258  -0.1839   19.102  0.0018          |                -|
6       -0.0894  -0.2036   19.653  0.0032          |                -|
7       -0.0759  -0.0910   20.057  0.0054          |                 |
8        0.1829   0.0772    22.45  0.0041          |-                |
9       -0.1178  -0.1764   23.462  0.0052          |                -|
10      -0.1922  -0.0379   26.209  0.0035         -|                 |
```

```
. gen eavto12=eavto1^2
(1 missing value generated)

. qui sum eavto12, detail
. scalar rhoimenovalec2=r(N)*r(mean)

. scalar rhoapprox2=rhostevec/rhoimenovalec2

. display rhostevec, rhoimenovalec2, rhoapprox2
9.4984437 49.011816 .19379906


. regress eavto eavto1, nocons

      Source |       SS       df       MS              Number of obs =      59
-------------+------------------------------           F(  1,    58) =    2.27
       Model |  1.84078945     1  1.84078945           Prob > F      =  0.1378
    Residual |  47.1365277    58  .812698753           R-squared     =  0.0376
-------------+------------------------------           Adj R-squared =  0.0210
       Total |  48.9773171    59  .830124019           Root MSE      =   .9015

------------------------------------------------------------------------------
       eavto |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      eavto1 |   .1937991   .1287699     1.51   0.138    -.0639619      .45156
------------------------------------------------------------------------------

. scalar rho=_b[eavto1]
. display rho
.19379906


. * Analiza prisotnosti avtokorelacije - Durbin-Watsonov d-test

. gen eavtott12=(eavto-eavto1)^2
(1 missing value generated)

. qui sum eavtott12, detail
. scalar dwstevec=r(N)*r(mean)

. qui sum eavto2, detail
. scalar dwimenovalec=r(N)*r(mean)

. scalar dw=dwstevec/dwimenovalec
. display dwstevec, dwimenovalec, dw
78.992245 50.116252 1.5761802


. qui regress avtop cena rrdp om
. estat dwatson

Durbin-Watson d-statistic(  4,    60) =   1.57618


. * Analiza prisotnosti avtokorelacije - Ocenjevanje koeficienta AK 4. reda

. gen eeavtott4=eavto*eavto4
(4 missing values generated)
```

```
. qui sum eeavtott4, detail
. scalar rhostevec=r(N)*r(mean)

. qui sum eavto2, detail
. scalar rhoimenovalec1=r(N)*r(mean)

. scalar rhoapprox1=rhostevec/rhoimenovalec1

. display rhostevec, rhoimenovalec1, rhoapprox1
23.94273 50.116252 .47774381


. corrgram eavto

                                          -1       0       1 -1       0       1
 LAG       AC       PAC      Q     Prob>Q  [Autocorrelation]  [Partial Autocor]
-------------------------------------------------------------------------------
1        0.1895   0.1935   2.2648  0.1323          |-                 |-
2        0.1233   0.0937   3.2397  0.1979          |                  |
3        0.1002   0.0692   3.8949  0.2730          |                  |
4        0.4777   0.4853   19.056  0.0008          |---               |---
5        0.0258  -0.1839   19.102  0.0018          |                 -|
6       -0.0894  -0.2036   19.653  0.0032          |                 -|
7       -0.0759  -0.0910   20.057  0.0054          |                  |
8        0.1829   0.0772    22.45  0.0041          |-                 |
9       -0.1178  -0.1764   23.462  0.0052          |                 -|
10      -0.1922  -0.0379   26.209  0.0035         -|                  |


. gen eavto42=eavto4^2
(4 missing values generated)

. qui sum eavto42, detail
. scalar rhoimenovalec2=r(N)*r(mean)

. scalar rhoapprox2=rhostevec/rhoimenovalec2

. display rhostevec, rhoimenovalec2, rhoapprox2
23.94273 47.781433 .50108856


. regress eavto eavto4, nocons

      Source |       SS       df       MS              Number of obs =      56
-------------+------------------------------           F(  1,    55) =   19.23
       Model |  11.9974278     1  11.9974278           Prob > F      =  0.0001
    Residual |  34.3161111    55  .623929294           R-squared     =  0.2590
-------------+------------------------------           Adj R-squared =  0.2456
       Total |   46.313539    56  .827027482           Root MSE      =  .78989

------------------------------------------------------------------------------
       eavto |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      eavto4 |   .5010886   .1142715     4.39   0.000     .2720833    .7300939
------------------------------------------------------------------------------

. scalar rho=_b[eavto4]
. display rho
.50108856
```

■