Heteroscedasticity Problem

True Relationship

True Model: $y_i = 1 + 2x_{2i} + 3x_{3i} + 4x_{4i} + u_i$

where: $u_i \sim N(0, \sigma_i^2)$ and $\sigma_i^2 = x_{4i}^2 + e_i$ and $e_i \sim N(0,5)$

```
set obs 100
g e=rnormal(0,5)
g x4=4*runiform()
g s2=x4^2+e
g u=rnormal(0,s2)
g x2=rnormal(10,40)
g x3=rnormal(-5,25)
g y=1+2*x2+3*x3+4*x4+u
```

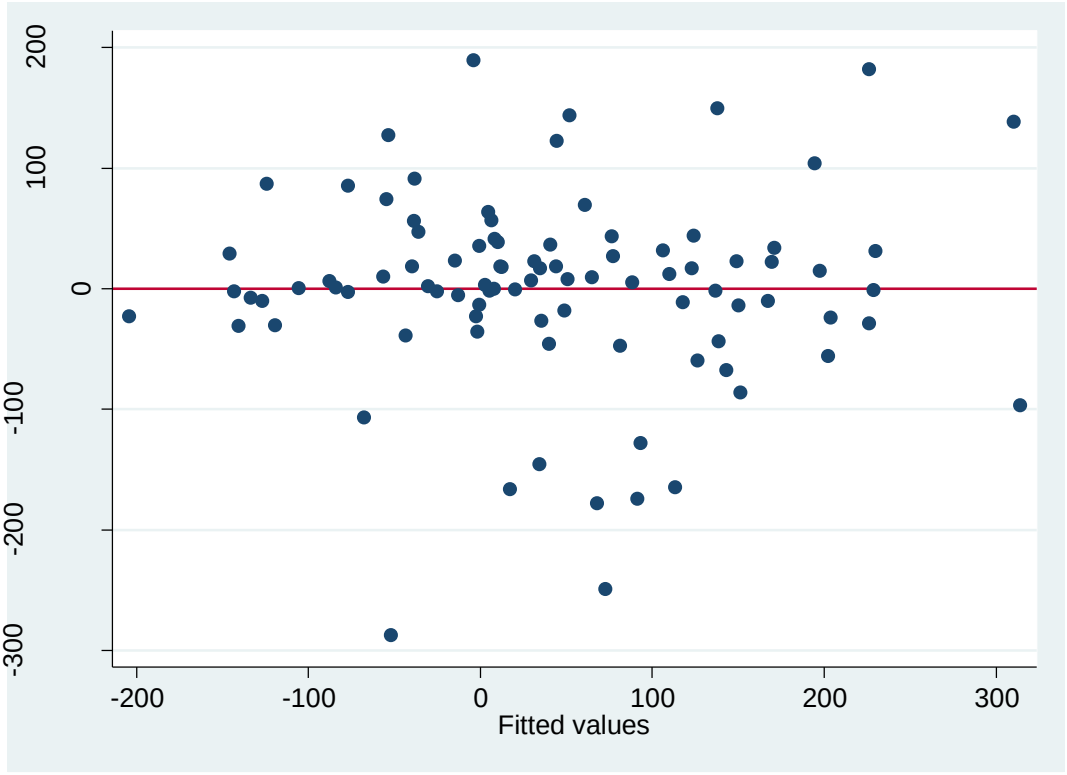
Detecting Heteroscedasticity Problem

Model: $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + u_i$

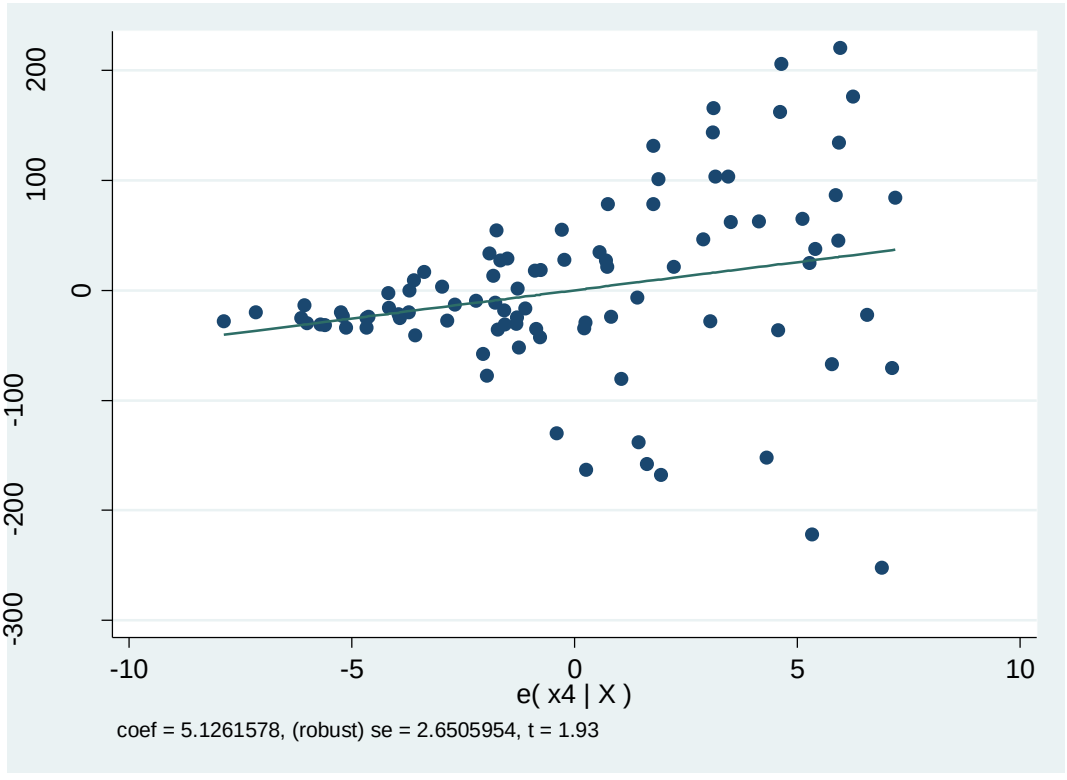
```
. reg y x2 x3 x4

      Source |       SS           df       MS      Number of obs   =        92
-------------+----------------------------------   F(3, 88)        =     51.38
       Model |  1037462.95         3  345820.985   Prob > F        =    0.0000
    Residual |  592246.119        88  6730.06954   R-squared       =    0.6366
-------------+----------------------------------   Adj R-squared   =    0.6242
       Total |  1629709.07        91  17908.8909   Root MSE        =    82.037

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x2 |    1.90836   .1920766     9.94   0.000     1.526648    2.290072
          x3 |   2.822761   .4036704     6.99   0.000     2.020551    3.624972
          x4 |   5.126158   2.227739     2.30   0.024     .6989941    9.553322
       _cons |  -5.892517   19.20263    -0.31   0.760     -44.0537    32.26867
------------------------------------------------------------------------------
```

Informal Test:

```
. rvfplot, yline(0)
```




```
. avplot x5, recast(scatter)
```



Formal Test – White's Test:

```
. whitetst

White's general test statistic :  40.56354  Chi-sq( 9)  P-value =  6.0e-06
```

Solving Heteroscedasticity Problem

From OLS Estimated Results:

```
. reg y x2 x3 x4

      Source |       SS           df       MS      Number of obs   =        92
-------------+----------------------------------   F(3, 88)        =     51.38
       Model |  1037462.95         3  345820.985   Prob > F        =    0.0000
    Residual |  592246.119        88  6730.06954   R-squared       =    0.6366
-------------+----------------------------------   Adj R-squared   =    0.6242
       Total |  1629709.07        91  17908.8909   Root MSE        =    82.037

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x2 |    1.90836   .1920766     9.94   0.000     1.526648    2.290072
          x3 |   2.822761   .4036704     6.99   0.000     2.020551    3.624972
          x4 |   5.126158   2.227739     2.30   0.024     .6989941    9.553322
       _cons |  -5.892517   19.20263    -0.31   0.760     -44.0537    32.26867
------------------------------------------------------------------------------
```

Determine Variance equation

Functional form: $\hat{u}_i^2 \approx \hat{\sigma}_i^2 = \alpha x_{4i}^2 + e_i$

```
. predict uhat, r
(8 missing values generated)

. g x42=x4^2

. g uhat2=uhat^2
(8 missing values generated)

. reg uhat2 x42, nocon

      Source |       SS           df       MS      Number of obs   =        92
-------------+----------------------------------   F(1, 91)        =     59.85
       Model |  7.8311e+09         1  7.8311e+09   Prob > F        =    0.0000
    Residual |  1.1907e+10        91   130846720   R-squared       =    0.3967
-------------+----------------------------------   Adj R-squared   =    0.3901
       Total |  1.9738e+10        92   214544911   Root MSE        =     11439

------------------------------------------------------------------------------
       uhat2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         x42 |   96.22585   12.43834     7.74   0.000     71.51862    120.9331
------------------------------------------------------------------------------
```

Weighted Least Squares:

Remedy by using Weighted Least Squares

Functional form: $w \cdot RD = \beta_1 w + \beta_2 w \cdot SALE + w \cdot u$

where $w = \frac{1}{\hat{\sigma}_i}$

```
. predict shat2, xb

. g shat=sqrt(shat2)

. g wy=w*y

. g wx2=w*x2

. g wx3=w*x3

. g wx4=w*x4

. reg wy wx2 wx3 wx4 w, nocon
```

Source	SS	df	MS	Number of obs	=	92
Model	19191.8747	4	4797.96866	F(4, 88)	=	6990.14
Residual	60.4024158	88	.686391089	Prob > F	=	0.0000
				R-squared	=	0.9969
				Adj R-squared	=	0.9967
Total	19252.2771	92	209.263881	Root MSE	=	.82849

wy	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
wx2	1.965304	.0449396	43.73	0.000	1.875996	2.054612
wx3	2.968128	.1579418	18.79	0.000	2.654252	3.282004
wx4	3.77476	1.023696	3.69	0.000	1.740379	5.80914
w	2.924305	2.810356	1.04	0.301	-2.660687	8.509298

Using command Weighted Least Squares

```
. vwls y x2 x3 x4, sd(shat)
```

Variance-weighted least-squares regression			Number of obs	=	92
Goodness-of-fit chi2(88)	=	60.40	Model chi2(3)	=	1807.52
Prob > chi2	=	0.9892	Prob > chi2	=	0.0000

y	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
x2	1.965304	.0542429	36.23	0.000	1.85899	2.071619
x3	2.968128	.1906388	15.57	0.000	2.594483	3.341773
x4	3.774759	1.235621	3.05	0.002	1.352987	6.196531
_cons	2.924306	3.392154	0.86	0.389	-3.724193	9.572805

Relieve the Problem: White's Robust Standard Error

From OLS Estimated Results:

```
. reg y x2 x3 x4

      Source |       SS           df       MS      Number of obs   =        92
-------------+----------------------------------   F(3, 88)        =     51.38
       Model |  1037462.95         3  345820.985   Prob > F        =    0.0000
    Residual |  592246.119        88  6730.06954   R-squared       =    0.6366
-------------+----------------------------------   Adj R-squared   =    0.6242
       Total |  1629709.07        91  17908.8909   Root MSE        =    82.037

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x2 |    1.90836   .1920766     9.94   0.000     1.526648    2.290072
          x3 |   2.822761   .4036704     6.99   0.000     2.020551    3.624972
          x4 |   5.126158   2.227739     2.30   0.024     .6989941    9.553322
       _cons |  -5.892517   19.20263    -0.31   0.760     -44.0537    32.26867
------------------------------------------------------------------------------
```

OLS with White's Robust Standard Error Estimated Results:

```
. reg y x2 x3 x4, vce(r)

Linear regression                               Number of obs     =         92
                                                F(3, 88)          =      56.51
                                                Prob > F          =     0.0000
                                                R-squared         =     0.6366
                                                Root MSE          =     82.037

------------------------------------------------------------------------------
             |               Robust
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x2 |    1.90836   .2115225     9.02   0.000     1.488004    2.328717
          x3 |   2.822761   .3495399     8.08   0.000     2.128124    3.517399
          x4 |   5.126158   2.650595     1.93   0.056    -.1413435    10.39366
       _cons |  -5.892517   12.69386    -0.46   0.644     -31.1189    19.33387
------------------------------------------------------------------------------
```