Example

Heteroscedasticity

R&D Expenditure, Sales, and Profits in 14 Industry Groupings in the United States, 2005 (all figures in millions of dollars)

Since the cross-sectional data presented in this table are quite heterogeneous, in a regression of R&D on sales, heteroscedasticity is likely

TABLE 11.5 Sales and Employment for Companies Performing Industrial R&D in the United States, by Industry, 2005 (values are in millions of dollars)

Source: National Science Foundation, Division of Science Resources Statistics, Survey of Industrial Research and Development: 2005 and the U.S. Census Bureau Annual Survey of Manufacturers, 2005.

	Industry	Sales	R&D	Profits
1	Food	374,342	2,716	234,662
2	Textiles, apparel, and leather	51,639	816	53,510
3	Basic chemicals	109,899	2,277	75,168
4	Resin, synthetic rubber, fibers, and filament	132,934	2,294	34,643
5	Pharmaceuticals and medicines	273,377	34,839	127,639
6	Plastics and rubber products	90,176	1,760	96,162
7	Fabricated metal products	174,165	1,375	155,80
8	Machinery	230,941	8,531	143,472
9	Computers and peripheral equipment	91,010	4,955	34,00
10	Semiconductor and other electronic components	176,054	18,724	81,317
11	Navigational, measuring, electromedical, and control instruments	118,648	15,204	73,250
12	Electrical equipment, appliances, and components	101,398	2,424	54,742
13	Aerospace products and parts	227,271	15,005	72,090
14	Medical equipment and supplies	56,661	4,374	52,44

Source	SS	df	MS	Number of obs =	14
Model	208733442	1	208733442	F(1, 12) =	2.48
Residual	1.0083e+09	12	84021567.1	Prob > F =	0.1410
				R-squared =	0.1715
Total	1.2170e+09	13	93614788.2	Adj R-squared =	0.1025
				Root MSE =	9166.3

rd	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
sales	.0437234	.0277404	1.58	0.141	-.0167178	.1041646
_cons	1337.874	5015.141	0.27	0.794	-9589.18	12264.93

$$\widehat{R\&D_i} = 1338 + 0.0437 Sales_i$$
$$se = (5015)\ (0.0277)$$
$$t = (0.27)\ (1.58)$$
$$r^2 = 0.172$$

There is a positive relationship between R&D and sales, although it is not statistically significant at the traditional levels

White Test

Source	SS	df	MS	Number of obs =	14
				F(2, 11) =	4.23
Model	9.2405e+16	2	4.6203e+16	Prob > F =	0.0435
Residual	1.2022e+17	11	1.0929e+16	R-squared =	0.4346
				Adj R-squared =	0.3318
Total	2.1263e+17	13	1.6356e+16	Root MSE =	1.0e+08

muhat2	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
sales	577.6563	1307.934	0.44	0.667	-2301.087	3456.4
sales2	.0008456	.0031711	0.27	0.795	-.006134	.0078253
_cons	-4.67e+07	1.12e+08	-0.42	0.685	-2.94e+08	2.00e+08

$$\widehat{u_i^2} = -46,746,325 + 578Sales_i + 0.000846Sales_i^2$$

$$se = (112,224,348) \quad (1308) \qquad (0.003171)$$

$$t = (-0.42) \qquad (0.44) \qquad (0.27)$$

$$R^2 = 0.435$$

Using the R^2 value and n=14, we obtain $nR^2 = 6.090$

Under the null hypothesis of no heteroscedasticity, this should follow a chi-square distribution with 2 df (because there are two regressors). The p-value of obtaining a chi-square value of as much as 6.090 or greater is about 0.0476. Since this is a low value, the White test also suggests that there is heteroscedsticity.

The true error variance is unknown, we cannot use the method of weighted least squares to obtain heteroscedasticity-corrected standard errors and t-values.

Therefore, we would have to make some educated guesses about the nature of the error variance.

White's heteroscedasticity-consistent standard errors

$$\widehat{R\&D_i} = 1337.87 + 0.0437Sales_i$$

$$se = (4892.447) \quad (0.0411)$$

$$t = (0.27) \qquad (1.06)$$

$$r^2 = 0.172$$

We see that the parameter estimates have not changed, the standard error of the intercept coefficient has decreased slightly, and the standard error of the slope coefficient has increased slightly. But remember that the White procedure is strictly a large-sample procedure, where as we have only 14 observations

STATA

Source	SS	df	MS	Number of obs =	14
				F(1, 12) =	2.48
Model	208733442	1	208733442	Prob > F =	0.1410
Residual	1.0083e+09	12	84021567.1	R-squared =	0.1715
				Adj R-squared =	0.1025
Total	1.2170e+09	13	93614788.2	Root MSE =	9166.3

RD	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Sales	.0437234	.0277404	1.58	0.141	-.0167178	.1041646
_cons	1337.874	5015.141	0.27	0.794	-9589.18	12264.93

```
. whitetst
White's general test statistic :    6.0842  Chi-sq( 2)  P-value =  .0477
```

H_0 : *Homoscedasticity*

H_1 : *Otherwise*

White's general test statistic is 6.0842.

Degree of freedom =2

Critical value of Chi-square at 5 percent significance level is 5.99147

6.0842 > 5.99147 Reject the null hypothesis

The White test also suggests that there is heteroscedsticity.