Heteroscedasticity in Linear Regression Analysis

Here I use the Multreg.dat data from my StatData page with SAS.

```
data grades; infile 'F:\StatData\multreg.dat';
input GPA GRE_Q GRE_V MAT AR;
proc reg; model GPA = GRE_Q -- AR / SPEC HCC HCCMETHOD=3 stb scorr2 vif; run; quit;
```

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Heteroscedasticity Consistent Standard Error	Heteroscedasticity Consistent t Value	Heteroscedasticity Consistent Pr > |t|	Standardized Estimate	Variance Inflation
Intercept	1	-1.73811	0.95074	-1.83	0.0795	0.85811	-2.03	0.0536	0	0
GRE_Q	1	0.00400	0.00183	2.18	0.0385	0.00151	2.65	0.0136	0.32406	1.53084
GRE_V	1	0.00152	0.00105	1.45	0.1593	0.00136	1.12	0.2737	0.21091	1.46947
MAT	1	0.02090	0.00955	2.19	0.0382	0.01181	1.77	0.0891	0.32214	1.50690
AR	1	0.14423	0.11300	1.28	0.2135	0.11074	1.30	0.2046	0.20160	1.73478

HCC Approximation Method: HC3

In yellow here are the traditional tests of the regression parameters. In blue are the tests using a variance/covariance matrix which has been corrected to adjust for any heteroscedasticity that may be present. The results of the two methods are very similar, but the partial effect of MAT moves from significant to not significant with the correction.

Test of First and Second Moment Specification

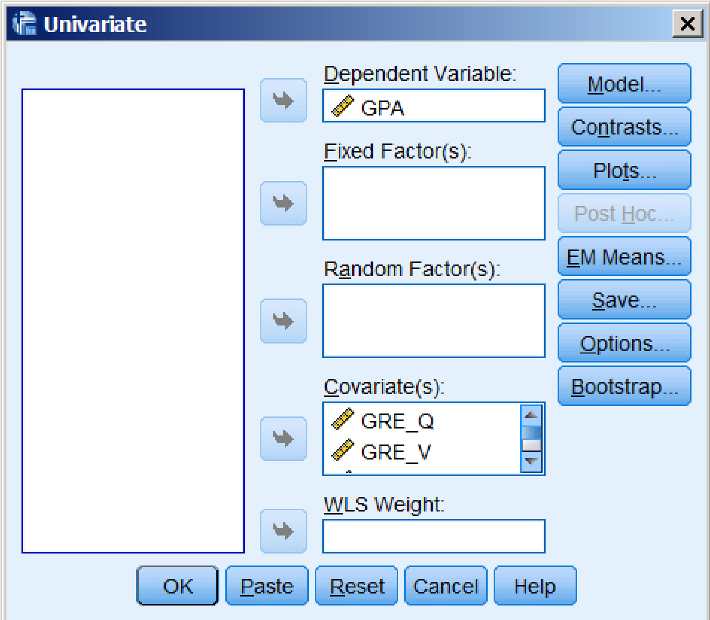
DF	Chi-Square	Pr > ChiSq
14	20.84	0.1057

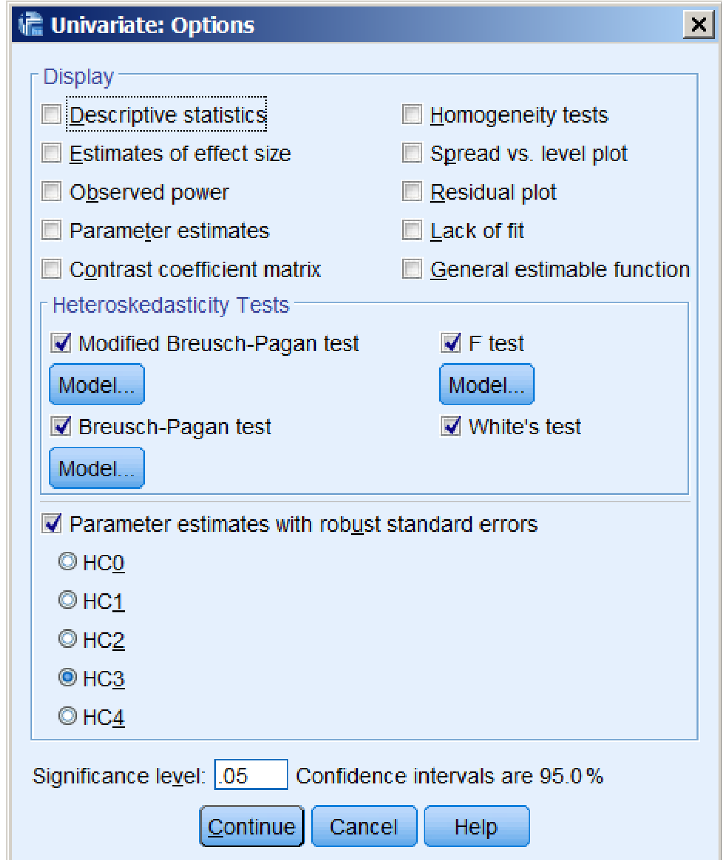
I usually use residuals plots to determine whether or not there is a problem with heteroscedasticity, but this statistic (obtained with the SPEC option) tests the null hypothesis that in the population there is homoscedasticity and the error terms are not correlated (that is, the residuals are not correlated with the value of the predictors). The non-significant result obtained here indicates that there are no such problems.

Yes, SPSS can be used in a similar fashion, starting with Version 25. Bring the data into SPSS and click Analyze, General Linear Model, Univariate.

Scoot GPA into the Dependent Variable box and the predictors in the Covariates box.

Click Options. Check everything in the Heteroskedasticity Tests box and then Continue, OK.






```
GET
  FILE='F:\SPSS\MultReg.sav'.

UNIANOVA GPA WITH GRE_Q GRE_V MAT AR
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /PRINT MBP WHITE F BP
  /CRITERIA=ALPHA(.05)
  /ROBUST=HC3
  /DESIGN=GRE_Q GRE_V MAT AR.
```

Univariate Analysis of Variance

Tests for Heteroskedasticity

White Test for Heteroskedasticity[a,b,c]

Chi-Square	df	Sig.
30.000	14	.008

a. Dependent variable: GPA

b. Tests the null hypothesis that the variance of the errors does not depend on the values of the independent variables.

c. Design: Intercept + GRE_Q + GRE_V + MAT + AR + GRE_Q * GRE_Q + GRE_Q * GRE_V + GRE_Q * MAT + GRE_Q * AR + GRE_V * GRE_V + GRE_V * MAT + GRE_V * AR + MAT * MAT + MAT * AR + AR * AR

Modified Breusch-Pagan Test for Heteroskedasticity[a,b,c]

Chi-Square	df	Sig.
6.437	1	.011

a. Dependent variable: GPA
b. Tests the null hypothesis that the variance of the errors does not depend on the values of the independent variables.
c. Predicted values from design: Intercept + GRE_Q + GRE_V + MAT + AR

Breusch-Pagan Test for Heteroskedasticity[a,b,c]

Chi-Square	df	Sig.
5.045	1	.025

a. Dependent variable: GPA
b. Tests the null hypothesis that the variance of the errors does not depend on the values of the independent variables.
c. Predicted values from design: Intercept + GRE_Q + GRE_V + MAT + AR

F Test for Heteroskedasticity[a,b,c]

F	df1	df2	Sig.
7.649	1	28	.010

a. Dependent variable: GPA
b. Tests the null hypothesis that the variance of the errors does not depend on the values of the independent variables.
c. Predicted values from design: Intercept + GRE_Q + GRE_V + MAT + AR

Notice that all of the SPSS tests indicate that there is a problem with heteroscedasticity.

Here are the results with the traditional tests.

Tests of Between-Subjects Effects

Dependent Variable: GPA

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	6.683[a]	4	1.671	11.134	.000
Intercept	.502	1	.502	3.342	.079
GRE_Q	.716	1	.716	4.770	.039
GRE_V	.316	1	.316	2.105	.159
MAT	.719	1	.719	4.789	.038
AR	.244	1	.244	1.629	.214
Error	3.752	25	.150		
Total	339.780	30			
Corrected Total	10.435	29			

a. R Squared = .640 (Adjusted R Squared = .583)

Here are the results with the correction for heteroscedasticity.

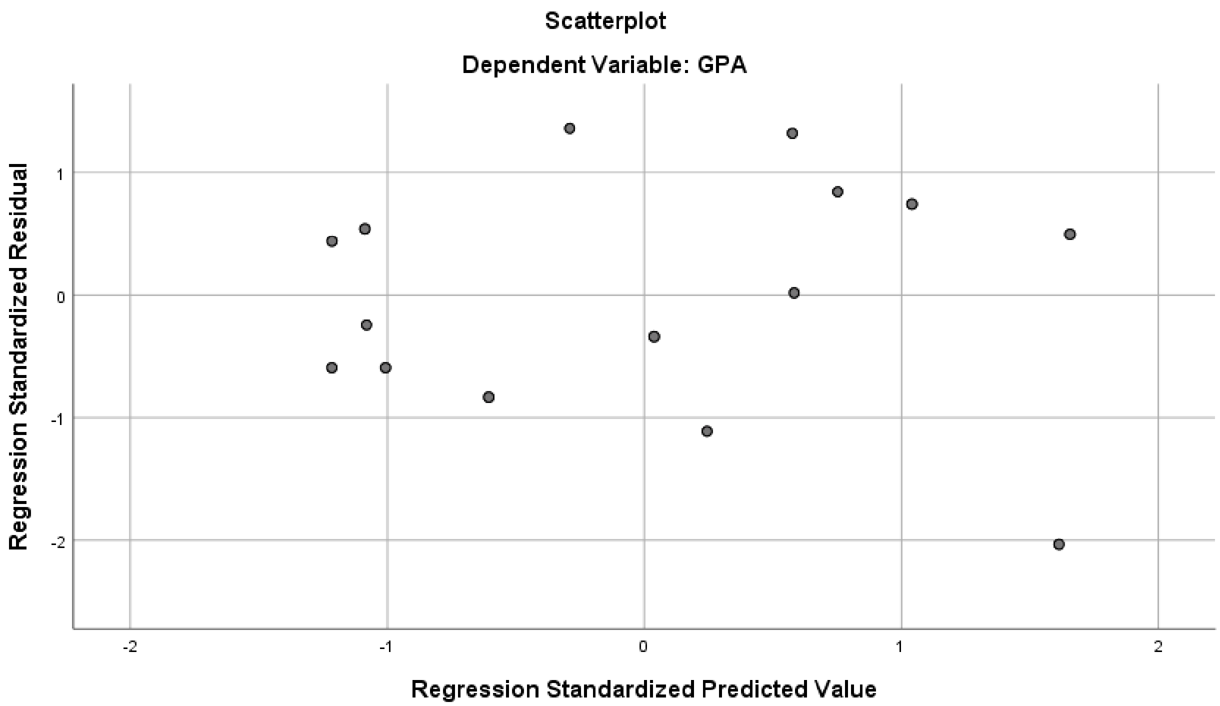
Parameter Estimates with Robust Standard Errors

Dependent Variable: GPA

Parameter	B	Robust Std. Error[a]	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	-1.738	.858	-2.026	.054	-3.505	.029
GRE_Q	.004	.002	2.654	.014	.001	.007
GRE_V	.002	.001	1.119	.274	-.001	.004
MAT	.021	.012	1.769	.089	-.003	.045
AR	.144	.111	1.302	.205	-.084	.372

a. HC3 method

A picture is worth a thousand *p* values?



More on Heteroscedasticity

References

Hayes, A. F. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods*, *39*,, 709-722.

Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, *54*, 217-224.

SAS (2019). Prog reg: Testing for heteroscedasticity.

SPSS. Unianova.

Williams, R. (2015). Heteroskedasticity. – Stata code, but a good read even for those who do not use Stata.