

## 9

## Checking the models II: the other three assumptions

### 9.1 Homogeneity of variance

Plotting the residuals against the fitted values is one of the standard techniques introduced in the main text to check the assumption of homogeneity of variance. It is illustrated here with the *diets* dataset. The notation for different kinds of residuals is inconsistent among statistical packages. The standardised residuals of the main text are called studentised residuals by SPSS. We use them rather than raw residuals for these model checking plots, as they are easier to interpret (see main text).

Here is how to get SPSS to produce its automatic residual plot. Of the nine squares produced, the relevant plot is the one in the middle of the bottom row. It has no scales on its axes, but it does show the general shape. The other plots are best ignored.

#### SPSS COMMANDS FOR BOX 9.1 (AUTOMATIC PLOT) **Weight change with diet**

Syntax

```
glm WGHTCH by DIET
  /print parameters
  /plot residuals
  /design DIET.
```

Menu route

Analyze > General Linear Model > Univariate

WGHTCH → Dependent Variable

DIET → Fixed Factor(s)

Options

- Residual plot
- Parameter estimates

SPSS OUTPUT FOR BOX 9.1 AND FIGURE 9.1 (a) (AUTOMATIC PLOT)  
**Weight change with diet, with a graph of standardised residuals against fitted values**

**General linear model**

**Between-Subjects Factors**

		N
DIET	1	5
	2	5
	3	5

**Tests of Between-Subjects Effects**

Dependent Variable: WGHTCH

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	173.563 <sup>a</sup>	2	86.782	4.973	.027
Intercept	55.527	1	55.527	3.182	.100
DIET	173.563	2	86.782	4.973	.027
Error	209.424	12	17.452		
Total	438.514	15			
Corrected Total	382.987	14			

a. R Squared = .453 (Adjusted R Squared = .362)

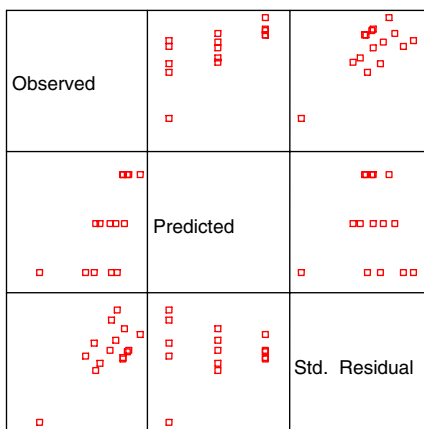
**Parameter Estimates**

Dependent Variable: WGHTCH

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	-2.258	1.868	-1.209	.250	-6.329	1.813
[DIET=1]	4.214	2.642	1.595	.137	-1.543	9.971
[DIET=2]	8.332	2.642	3.154	.008	2.575	14.089
[DIET=3]	0 <sup>a</sup>	.	.	.	.	.

a. This parameter is set to zero because it is redundant

Dependent Variable: WGHTCH



Model: Intercept + DIET

We can obtain a plot with scales on the axes by saving the relevant ‘temporary variables’ `pred` (for predicted values) and `sresid` (for standardised residuals), and plotting them, as follows:

In the syntax route, we make up new names `WTPR` and `WTRES` for the variables to hold the predicted values and residuals. If we omit these new names and their brackets, SPSS calls them `pre_1` and `sre_1`. We then ask SPSS to draw a scatter plot of these two variables. This analysis is in two steps — after the first step, the ANOVA and parameter estimates shown in SPSS Box 9.1 will be produced. The graph will be produced after the second step.

In the menu route we introduce the notation of a ‘subpane’. The pane produced by the ‘Save’ button has a number of subpanes. We refer to a subpane by giving its heading followed by a colon, then underneath, the action required. For example,

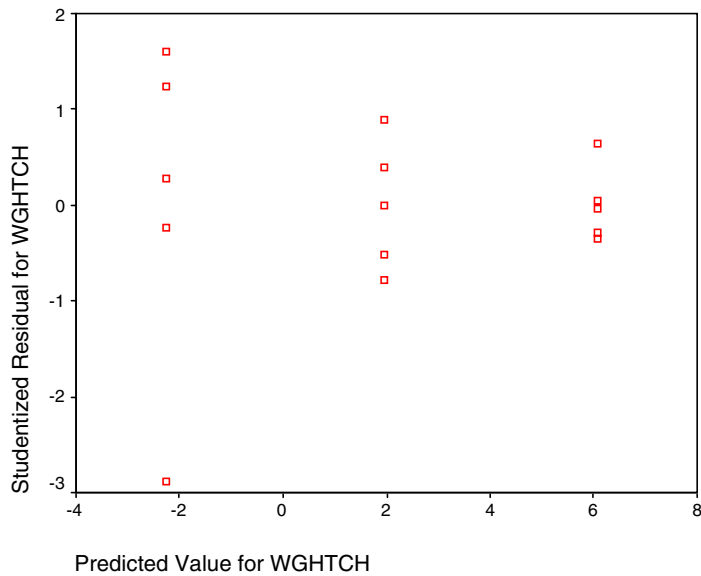
Predicted Values:

Unstandardized

should indicate that in the Predicted Values subpane, you check the Unstandardised box.

SPSS COMMANDS FOR BOX 9.1 (NON-AUTOMATIC PLOT) <b>Weight change with diet</b>	
Syntax	<pre>glm WGHTCH by DIET   /print parameters   /save pred(WTPR) sresid(WTRES)   /design DIET.  graph   /scat WTPR with WTRES.</pre>
Menu route	<p>Analyze &gt; General Linear Model &gt; Univariate</p> <p>WGHTCH → Dependent Variable</p> <p>DIET → Fixed Factor(s)</p> <p><input type="button" value="Save"/></p> <p>Predicted Values:  <input checked="" type="checkbox"/> Unstandardized</p> <p>Residuals:  <input checked="" type="checkbox"/> Studentized</p> <p><input type="button" value="Options"/></p> <p><input checked="" type="checkbox"/> Parameter estimates</p> <p>Graphs &gt; Scatter (then select ‘Simple’ and click ‘Define’)</p> <p>Studentized Residuals for WGHTCH (<code>sre_1</code>) → Y Axis</p> <p>Predicted Values for WGHTCH (<code>pre_1</code>) → X Axis</p>

SPSS OUTPUT FOR BOX 9.1 (b) (NON-AUTOMATIC PLOT)  
**A graph of standardised residuals against fitted values**

**Graph**


---

## 9.4 Model criticism and solutions

The Normality of error may be examined by plotting a histogram of residuals or a Normal order plot. Normal Order plots (referred to as type q-q by SPSS) allow a more quantitative assessment of the Normality of the distribution of residuals. We are interested in the Normal Q-Q plot produced (not the detrended plot that is also produced). This method will first of all be described with some Normally distributed residuals. Using the *plantlets* data set, the height of two species are compared.

SPSS COMMANDS FOR BOX 9.2 AND FIGURE 9.6 **Analysis of height in the plantlets dataset**

```
Syntax      glm HGHT by SPECIES
            /save sresid(HRES)
            /design SPECIES.

            graph
            /histogram(normal) HRES.

            pplot
            /variables HRES
            /type q-q.
```

Menu route     Analyze > General Linear Model > Univariate  
                  HGHT → Dependent Variable  
                  SPECIES → Fixed Factor(s)  
                    
                  Residuals:  
                   Studentized

                  Graphs > Histogram  
                  Studentized Residual for HGHT → Variables  
                   Display Normal curve

                  Graphs > Q-Q  
                  Studentized Residual for HGHT → Variables

SPSS OUTPUT FOR BOX 9.2 **Analysis of height in the plantlets dataset**

---

**General linear model**

**Between-Subjects Factors**

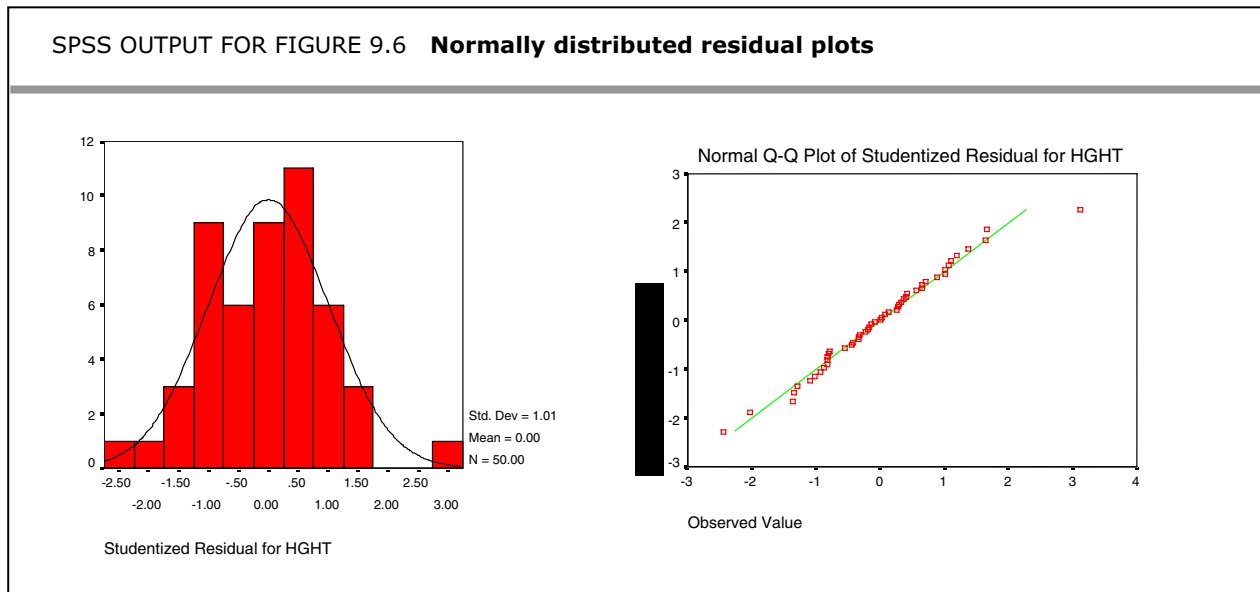
		N
SPECIES	1	25
	2	25

**Tests of Between-Subjects Effects**

Dependent Variable: HGHT

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	260.833 <sup>a</sup>	1	260.833	46.277	.000
Intercept	6849.181	1	6849.181	1215.173	.000
SPECIES	260.833	1	260.833	46.277	.000
Error	270.546	48	5.636		
Total	7380.560	50			
Corrected Total	531.379	49			

a. R Squared = .491 (Adjusted R Squared = .480)



The residuals in Figure 9.6 appear to be Normally distributed and originate from a data set size 50. Imagine these standardised residuals being ranked. If these 50 values have been drawn at random from a Standard Normal Distribution (see revision section), then we can calculate the expected value of the lowest residual — which in this case we would expect to be  $-2.24$  standard deviations below the mean of zero. Corresponding expected residuals can be calculated for all 50 points. These can be calculated in SPSS by the use of the RANK command. For example, the following set of commands will produce the 50 ‘expected’ standardised residuals corresponding to the observed residuals:

SPSS COMMANDS FOR ALTERNATIVE FIGURE 9.6 **Normal order plot**

```
Syntax      rank HRES
            /normal into NQ.

            graph
            /scatter NQ with HRES.
```

Menu route Transform > Rank Cases

Studentized residuals for HGHT → Variable(s)

Rank Types

More >>

Rank

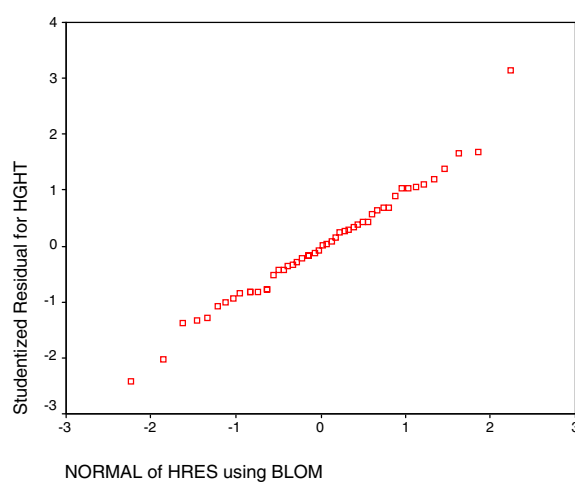
Normal scores

Graph > Scatter (then select 'Simple' and click 'Define')

Normal of sre\_2 → Y Axis

Studentised residual for HGHT (sre\_2) → X Axis

which would produce the following graph:

SPSS ALTERNATIVE FIGURE 9.6 **Normal order plot****Graph**

We now turn to the remaining variables in the dataset, and use the automatic plots.

SPSS COMMANDS FOR BOX 9.3 AND FIGURE 9.6 **Analysing longevity in the plantlets dataset**

```
Syntax      glm LONGEV by SPECIES
            /save sresid(LRES)
            /design SPECIES.

            graph
            /histogram(normal) LRES.

            pplot
            /variables LRES
            /type q-q.
```

Menu route      Analyze > General Linear Model > Univariate  
                    LONGEV → Dependent Variable  
                    SPECIES → Fixed Factor(s)  
                      
                    Residuals:  
                     Studentized

                    Graphs > Histogram  
                            Studentized Residual for LONGEV → Variables  
                             Display Normal curve

                    Graphs > Q-Q  
                            Studentized Residual for LONGEV → Variables

SPSS OUTPUT FOR BOX 9.3 **Analysis of longevity in the plantlets dataset**

**General linear model**

**Between-Subjects Factors**

		N
SPECIES	1	25
	2	25

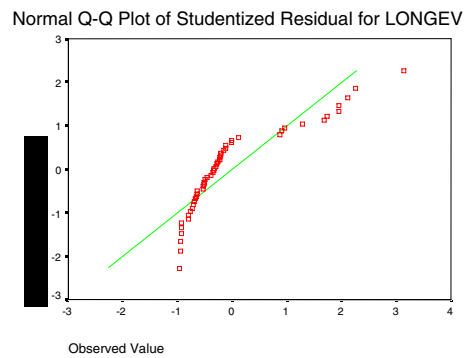
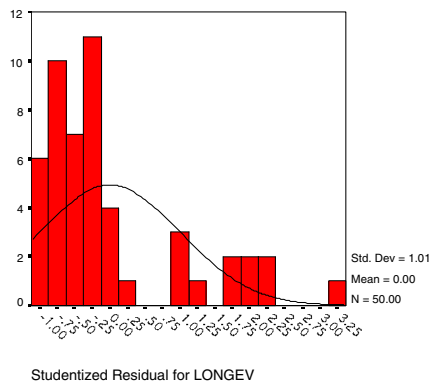
**Tests of Between-Subjects Effects**

Dependent Variable: LONGEV

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	3299.969 <sup>a</sup>	1	3299.969	.302	.585
Intercept	402914.691	1	402914.691	36.831	.000
SPECIES	3299.969	1	3299.969	.302	.585
Error	525102.620	48	10939.638		
Total	931317.280	50			
Corrected Total	528402.589	49			

a. R Squared = .006 (Adjusted R Squared = -.014)

SPSS OUTPUT FOR FIGURE 9.7 **Residual plots for right hand skewed residuals**



SPSS COMMANDS FOR BOX 9.4 AND FIGURE 9.8 **Analysing leaf size in the plantlets dataset**

```

Syntax      glm LEAFSIZE by SPECIES
            /save sresid(LFRES)
            /design SPECIES.

            graph
            /histogram(normal) LFRES.

            pplot
            /variables LFRES
            /type q-q.

```

Menu route      Analyze > General Linear Model > Univariate

                    LEAFSIZE → Dependent Variable

                    SPECIES → Fixed Factor(s)

                            Residuals:

Studentized

                    Graphs > Histogram

                            Studentized Residual for LEAFSIZE → Variables

Display Normal curve

                    Graphs > Q-Q

                            Studentized Residual for LEAFSIZE → Variables

SPSS OUTPUT FOR BOX 9.4 **Analysis of longevity in the plantlets dataset**

**General linear model**

**Between-Subjects Factors**

		N
SPECIES	1	25
	2	25

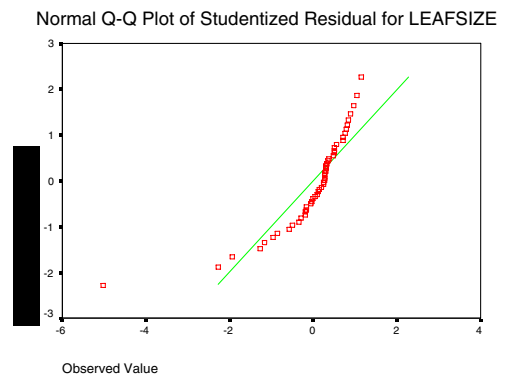
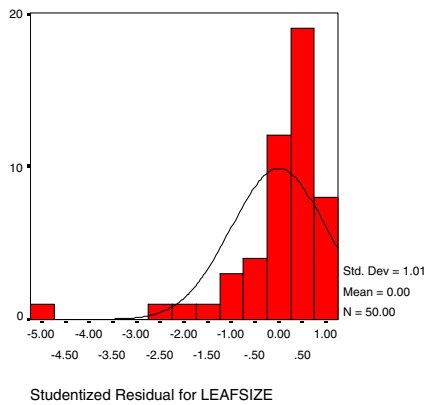
**Tests of Between-Subjects Effects**

Dependent Variable: LEAFSIZE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.115 <sup>a</sup>	1	.115	.252	.618
Intercept	703.875	1	703.875	1539.254	.000
SPECIES	.115	1	.115	.252	.618
Error	21.950	48	.457		
Total	725.940	50			
Corrected Total	22.065	49			

a. R Squared = .005 (Adjusted R Squared = -.016)

SPSS OUTPUT FOR FIGURE 9.8 **Residual plots for left hand skewed residuals**



### Plotting residuals against fitted values

This technique was introduced earlier in this chapter as a means of checking for homogeneity of variance. It can also illustrate any problems with linearity, as illustrated with the *3Ys* dataset.

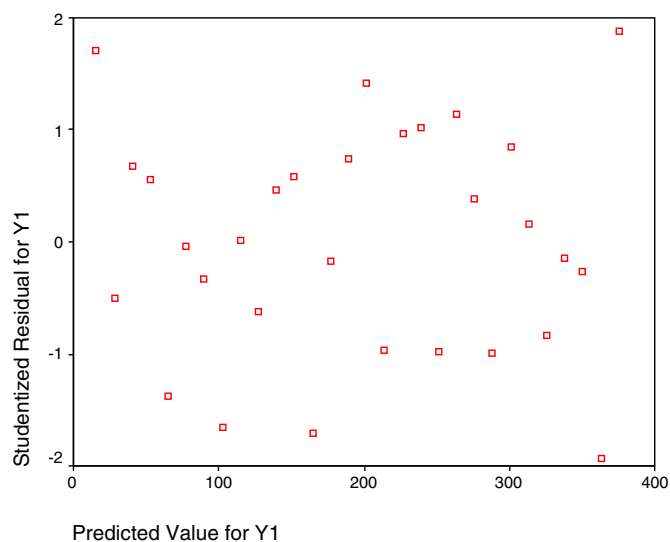
SPSS COMMANDS FOR BOX 9.5 AND FIGURE 9.9 <b>Dataset 3Ys, Y1</b>	
Syntax	<pre>glm Y1 with X       /save sresid(Y1RES) pred(Y1PRED)       /design X.  graph       /scatter Y1PRED with Y1RES.</pre>
Menu route	<p>Analyze &gt; General Linear Model &gt; Univariate</p> <p>Y1 → Dependent Variable</p> <p>X → Covariate</p> <div style="border: 1px solid black; padding: 2px; display: inline-block; margin: 5px 0;">Save</div> <p>Residuals:  <input checked="" type="checkbox"/> Studentized</p> <p>Predicted Values:  <input checked="" type="checkbox"/> Unstandardized</p> <p>Graphs &gt; Scatter (then select 'Simple' and click 'Define')</p> <p>Studentized Residual for Y1 → Y Axis</p> <p>Predicted Value for Y1 → X Axis</p>

SPSS OUTPUT FOR BOX 9.5 **Dataset 3Ys, Y1****General linear model****Tests of Between-Subjects Effects**

Dependent Variable: Y1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	345771.548 <sup>a</sup>	1	345771.548	108377.5	.000
Intercept	75.115	1	75.115	23.544	.000
X	345771.548	1	345771.548	108377.5	.000
Error	89.332	28	3.190		
Total	1492468.380	30			
Corrected Total	345860.880	29			

a. R Squared = 1.000 (Adjusted R Squared = 1.000)

SPSS OUTPUT FOR FIGURE 9.9 **Residual with homogeneous variance against fitted values****Graph**

The same commands may be repeated for the other two variables in this dataset, in which the residuals show increasing and decreasing variance with the fitted values (see Figure 9.10 and 9.11 in the main text).

### Plotting the residuals against each explanatory variable

We will illustrate this with the second response variable in the  $3Y_5$  dataset. In this instance, plotting against the fitted values, and plotting against the explanatory variable (X) will result in the same graph, as there is only one explanatory variable. However, in models with several explanatory variables, the same technique may allow you to pinpoint which explanatory variable may be responsible for the non-linearity.

#### SPSS COMMANDS FOR AN EXAMPLE OF FIGURE 9.13

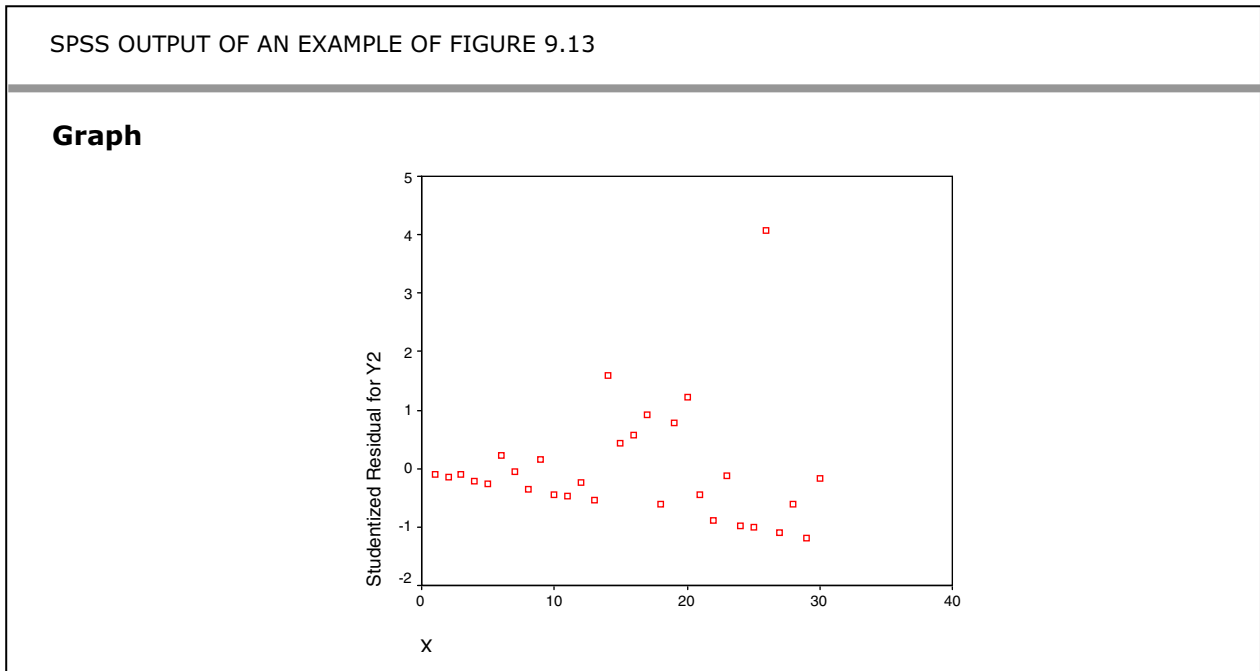
```
Syntax          glm Y2 with X
                /save sresid(Y2RES)
                /design X.

                graph
                /scatter X with Y2RES.
```

```
Menu route      Analyze > General Linear Model > Univariate
                Y2 → Dependent Variable
                X → Covariate
                Save
                Residuals:
                 Studentized

                Graphs > Scatter (then select 'Simple' and click 'Define')
                Studentized Residual for Y2 → Y Axis
                X → X Axis
```

In addition to the ANOVA table, this will produce the following graph:



These techniques may then be applied to the *bacteria* dataset, to resolve whether interactions or transformations are the better way of solving for non-linearity and heterogeneity of variance, as illustrated in the subsequent section of the main text.

---

## 9.5 Predicting the volume of merchantable wood: an example of model criticism

This example brings together both model criticism and basic data manipulation techniques, with an analysis based on the *merchantable timber* dataset. This case study starts off with an analysis of the untransformed data.

SPSS COMMANDS FOR BOX 9.11 AND FIGURE 9.19 **Untransformed analysis of tree volume**

Syntax

```
glm VOLUME with DIAM
  /print parameters
  /save sresid(DRES) pred(DPRED)
  /design DIAM.

graph
  /scatter DPRED with DRES.
```

Menu route

Analyze > General Linear Model > Univariate

VOLUME → Dependent Variable

DIAM → Covariate(s)

Residuals:  
 Studentized

Predicted Values:  
 Unstandardized

Parameter estimates

Graphs > Scatter (then select 'Simple' and click 'Define')

Studentized Residual for VOLUME → Y Axis

Predicted Value for VOLUME → X Axis

SPSS OUTPUT FOR BOX 9.11 **Untransformed analysis of tree volume**

**General linear model**

**Tests of Between-Subjects Effects**

Dependent Variable: VOLUME

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	7581.910 <sup>a</sup>	1	7581.910	419.471	.000
Intercept	2179.086	1	2179.086	120.558	.000
DIAM	7581.910	1	7581.910	419.471	.000
Error	524.173	29	18.075		
Total	36324.990	31			
Corrected Total	8106.084	30			

a. R Squared = .935 (Adjusted R Squared = .933)

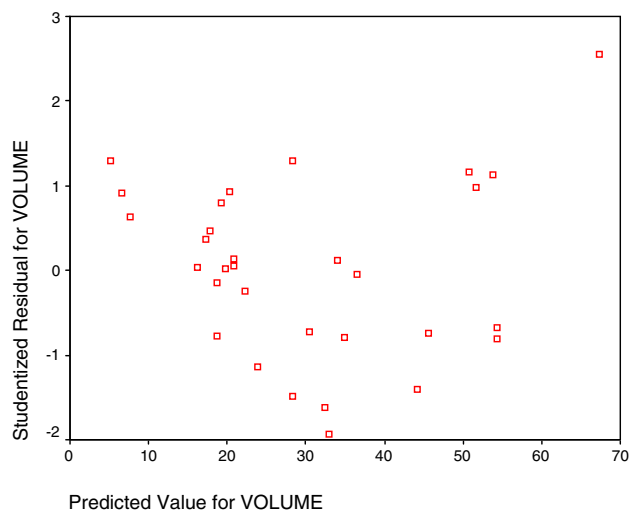
**Parameter Estimates**

Dependent Variable: VOLUME

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	-36.945	3.365	-10.980	.000	-43.827	-30.063
DIAM	60.791	2.968	20.481	.000	54.721	66.862

SPSS OUTPUT FOR FIGURE 9.19 **Residual plot for analysis of untransformed data**

**Graph**



The second analysis transforms the data and includes two explanatory variables. To take the natural logs of VOLUME, HEIGHT and DIAM, follow the syntax or menus in the next box. If using the syntax route, select all four lines and then press run (otherwise SPSS will only execute the last line). We recommend simply typing the formula in the menu route. If you prefer, and this is particularly useful if you aren't sure what functions are available, or what they are called, you can scroll down the list of functions and double-click on the one you want. The second analysis is then performed on the transformed data.

SPSS COMMANDS TO TRANSFORM DATA	
Syntax	<pre>compute LVOL=ln(VOLUME) . compute LHGHT=ln(HEIGHT) . compute LDIAM=ln(DIAM) . execute.</pre>
Menu route	<p>Window &gt; Data Editor</p> <p>Transform &gt; Compute</p> <p>LVOL → Target Variable</p> <p>LN(VOLUME) → Numeric Expression</p> <p>Transform &gt; Compute</p> <p>LHGHT → Target Variable</p> <p>LN(HEIGHT) → Numeric Expression</p> <p>Transform &gt; Compute</p> <p>LDIAM → Target Variable</p> <p>LN(DIAM) → Numeric Expression</p>

SPSS COMMANDS FOR BOX 9.12 AND FIGURE 9.20 **Comparing trees to cylinders**

```
Syntax      glm LVOL with LHGHT LDIAM
            /print parameters
            /save sresid(LDRES) pred (LDPRED)
            /design LHGHT LDIAM.

            graph
            /scatter LDPRED with LDRES.
```

Menu route      Analyze > General Linear Model > Univariate

                  LVOL → Dependent Variable

                  LHGHT LDIAM → Covariate(s)

                  Residuals:

Studentized

                  Predicted Values:

Unstandardized

Parameter estimates

                  Graphs > Scatter (then select 'Simple' and click 'Define')

                  Studentized Residual for LVOL → Y Axis

                  Predicted Values for LVOL → X Axis

SPSS OUTPUT FOR BOX 9.12 **Comparing trees to cylinders**

**General linear model**

**Tests of Between-Subjects Effects**

Dependent Variable: LVOL

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	8.123 <sup>a</sup>	2	4.062	613.289	.000
Intercept	2.475E-02	1	2.475E-02	3.737	.063
LHGHT	.198	1	.198	29.862	.000
LDIAM	4.628	1	4.628	698.739	.000
Error	.185	28	6.623E-03		
Total	340.343	31			
Corrected Total	8.309	30			

a. R Squared = .978 (Adjusted R Squared = .976)

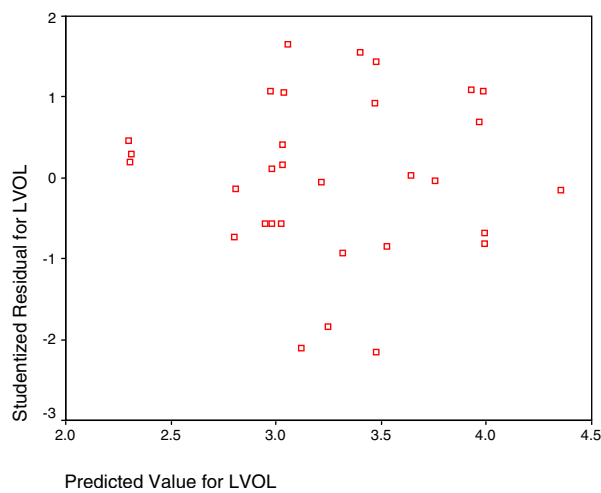
**Parameter Estimates**

Dependent Variable: LVOL

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	-1.705	.882	-1.933	.063	-3.511	.102
LHGHT	1.117	.204	5.465	.000	.698	1.536
LDIAM	1.983	.075	26.434	.000	1.829	2.136

SPSS OUTPUT FOR FIGURE 9.20 **Residual plot for tree volume treated as a cylinder**

**Graph**



The third step in this case study is to estimate the mean deviation of the intercept from the expected cylinder intercept. This is done by creating a new variable LDEVINT, and requesting its mean and standard error. At the same time, we estimate the proportion of the cylinder that represents useable timber in the variable LUSEFRAC.

**SPSS COMMANDS FOR BOXES 9.13 AND 9.14**  
**Estimating the mean deviation of the intercept from the expected cylinder intercept**

---

Syntax

```
compute LDEVINT=LVOL - LHGHT -2*LDIAM.
compute LUSEFRAC=LVOL - (-0.242 + LHGHT +2*LDIAM) .

descriptives
  variables = LDEVINT LUSEFRAC
  /STATISTICS = MEAN STDDEV MIN MAX SEMEAN.
```

Menu route

Window > Data Editor  
 Transform > Compute  
 LDEVINT → Target Variable  
 LVOL - LHGHT -2\*LDIAM → Numeric Expression

Transform > Compute  
 LUSEFRAC → Target Variable  
 LVOL - (-0.242 + LHGHT +2\*LDIAM) → Numeric Expression

Analyze > Descriptive Statistics > Descriptives  
 LDEVINT LUSEFRAC → Variable(s)  
  
 S.E. Mean

**SPSS OUTPUT FOR BOXES 9.13 AND 9.14**  
**Estimating the mean deviation of the intercept from the expected cylinder intercept**

---

**Descriptives**

**Descriptive Statistics**

	N	Minimum	Maximum	Mean		Std.
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic
LDEVINT	31	-1.37	-1.06	-1.1994	.0142	.07909
LUSEFRAC	31	-1.13	-.82	-.9574	.0142	.07909
Valid N (listwise)	31					

## 9.8 Exercises

### Stabilising the variance

The SPSS commands used in this exercise fall into three parts. When following the syntax route, select all commands, including execute, and click on run.

(1) Transforming the data

SPSS COMMANDS TO TRANSFORM DATA	
Syntax	<pre>compute SRDAM = sqrt(DAMAGE) . compute LOGDAM = ln(DAMAGE) . compute INVDAM = -1/ DAMAGE . execute.</pre>
Menu route	<p>Window &gt; SPSS Data Editor</p> <p>Transform &gt; Compute</p> <p>SRDAM → Target Variable</p> <p>sqrt(DAMAGE) → Numeric Expression</p> <p>Transform &gt; Compute</p> <p>LOGDAM → Target Variable</p> <p>ln(DAMAGE) → Numeric Expression</p> <p>Transform &gt; Compute</p> <p>INVDAM → Target Variable</p> <p>-1/DAMAGE → Numeric Expression</p>

## (2) Plotting the graphs

## SPSS COMMANDS FOR FIGURE 9.21

```
Syntax      graph
            /scatter WDKLR with DAMAGE.

            graph
            /scatter WDKLR with SRDAM.

            graph
            /scatter WDKLR with LOGDAM.

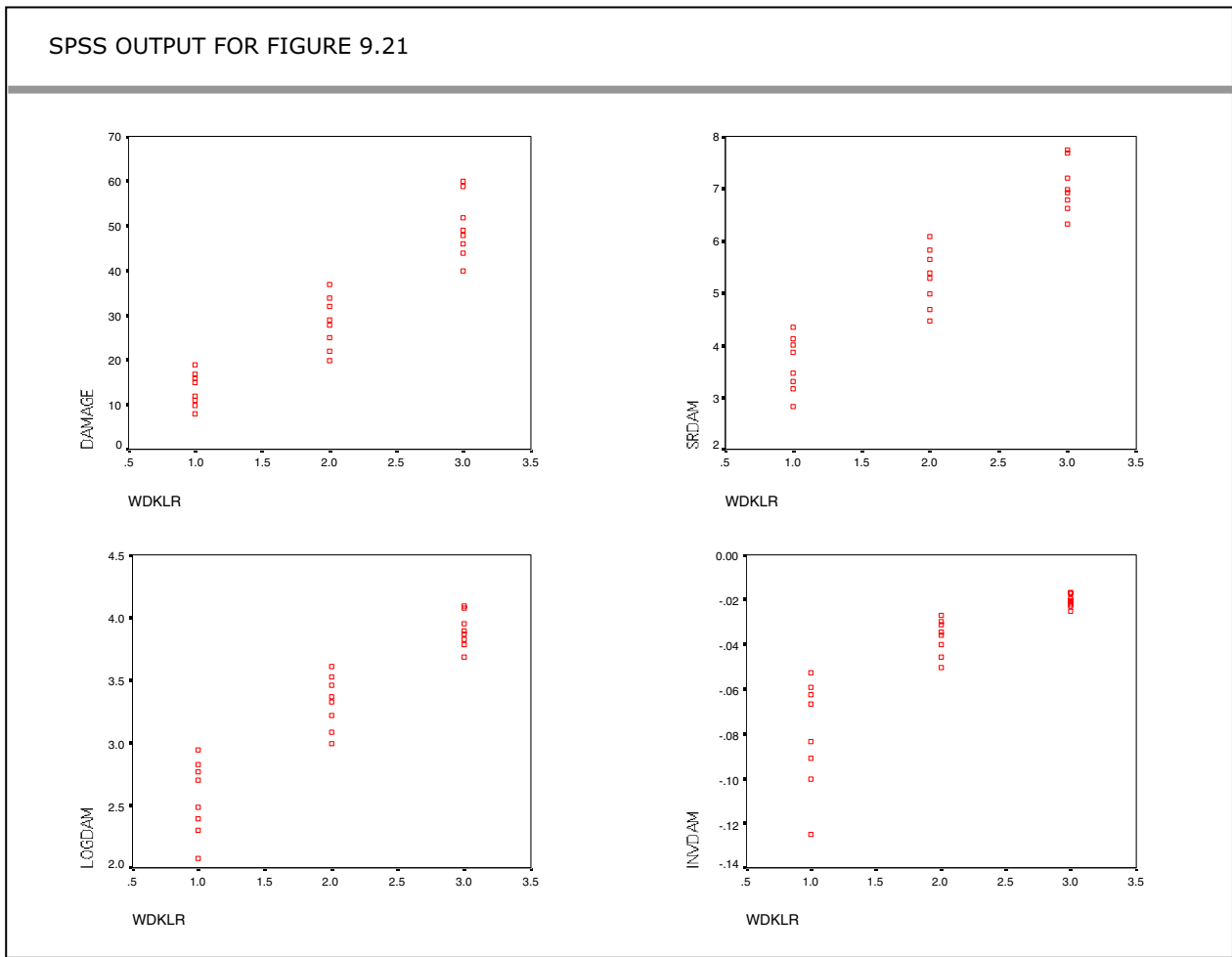
            graph
            /scatter WDKLR with INV DAM.
```

```
Menu route  Graph > Scatter (then select 'Simple' and click 'Define')
            DAMAGE → Y Axis
            WDKLR  → X Axis

            Graph > Scatter (then select 'Simple' and click 'Define')
            SRDAM  → Y Axis
            WDKLR  → X Axis

            Graph > Scatter (then select 'Simple' and click 'Define')
            LOGDAM → Y Axis
            WDKLR  → X Axis

            Graph > Scatter (then select 'Simple' and click 'Define')
            INV DAM → Y Axis
            WDKLR  → X Axis
```



(3) Obtaining descriptive statistics

SPSS COMMANDS FOR BOX 9.15	
Syntax	<code>means DAMAGE SRDAM LOGDAM INVDAM by WDKLR.</code>
Menu route	Analyze > Compare Means > Means... DAMAGE SRDAM LOGDAM INVDAM → Dependent List WDKLR → Independent List

## SPSS OUTPUT FOR BOX 9.15

**Means****Case Processing Summary**

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
DAMAGE * WDKLR	24	30.0%	56	70.0%	80	100.0%
SRDAM * WDKLR	24	30.0%	56	70.0%	80	100.0%
LOGDAM * WDKLR	24	30.0%	56	70.0%	80	100.0%
INVDAM * WDKLR	24	30.0%	56	70.0%	80	100.0%

**Report**

WDKLR		DAMAGE	SRDAM	LOGDAM	INVDAM
1	Mean	13.50	3.6408	2.5654	-.0800
	N	8	8	8	8
	Std. Deviation	3.817	.52867	.29723	.02460
2	Mean	28.38	5.3012	3.3260	-.0367
	N	8	8	8	8
	Std. Deviation	5.878	.55758	.21326	.00799
3	Mean	49.75	7.0383	3.8985	-.0204
	N	8	8	8	8
	Std. Deviation	6.985	.4923	.13932	.00282
Total	Mean	30.54	5.3268	3.2633	-.0457
	N	24	24	24	24
	Std. Deviation	16.149	1.50376	.59813	.02941

### Stabilising the variance in a blocked experiment

The SPSS output for this exercise may be found in the answers to exercises.

### Lizard skulls

The first analysis is with the raw data.

SPSS COMMANDS FOR BOX 9.16 AND FIGURE 9.22	
Syntax	<pre>glm BVOL with JAWL   /print parameters   /save sresid(BRES) pred(BPRED)   /design JAWL.  graph   /scatter BPRED with BRES.</pre>
Menu route	Analyze > General Linear Model > Univariate BVOL → Dependent Variable JAWL → Covariate(s) <input type="button" value="Save"/> Residuals: <input checked="" type="checkbox"/> Studentized Predicted Values: <input checked="" type="checkbox"/> Unstandardized <input type="button" value="Options"/> <input checked="" type="checkbox"/> Parameter estimates  Graphs > Scatter (then select 'Simple' and click 'Define') Studentized Residual for BVOL → Y Axis Predicted Values for BVOL → X Axis

SPSS OUTPUT FOR BOX 9.16 **Analysis of lizard skulls****General linear model****Tests of Between-Subjects Effects**

Dependent Variable: BVOL

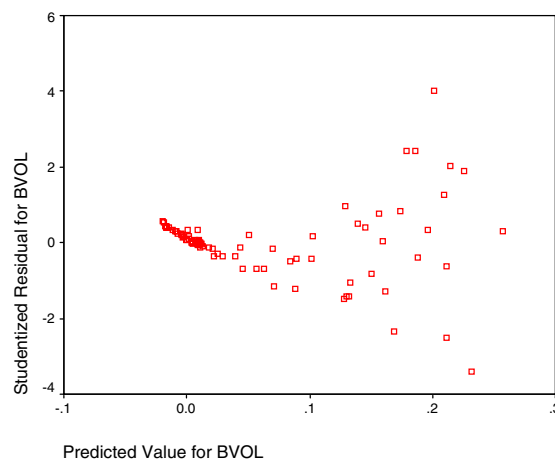
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.523 <sup>a</sup>	1	.523	284.655	.000
Intercept	4.888E-02	1	4.888E-02	26.617	.000
JAWL	.523	1	.523	284.655	.000
Error	.143	78	1.837E-03		
Total	1.033	80			
Corrected Total	.666	79			

a. R Squared = .785 (Adjusted R Squared = .782)

**Parameter Estimates**

Dependent Variable: BVOL

Parameter	B	Std.	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	-4.16E-02	.008	-5.159	.000	-5.763E-02	-2.554E-02
JAWL	.525	.031	16.872	.000	.463	.586

SPSS OUTPUT FOR FIGURE 9.22 **Residual plot of lizard skulls****Graph**

The second analysis is performed on the transformed allometric measurements, and includes SITE as an extra explanatory variable.

SPSS COMMANDS FOR BOX 9.17 AND FIGURE 9.23 **Alternative analysis for lizard skulls**

```
Syntax      compute LBVOL = ln(BVOL) .
            compute LJAWL = ln(JAWL) .

            glm LBVOL by SITE with LJAWL
              /print parameters
              /save sresid(LBRES) pred(LBPRED)
              /design SITE LJAWL.

            graph
              /scatter LBPRED with LBRES.
```

Menu route Window > SPSS Data Editor  
 Transform > Compute  
 LBVOL → Target Variable  
 ln(BVOL) → Numeric Expression

Transform > Compute  
 LJAWL → Target Variable  
 ln(JAWL) → Numeric Expression

Analyze > General Linear Model > Univariate  
 LBVOL → Dependent Variable  
 SITE → Fixed Factor(s)  
 LJAWL → Covariate(s)

Save  
 Residuals:  
 Studentized  
 Predicted Values:  
 Unstandardized

Options  
 Parameter estimates

Graphs > Scatter (then select 'Simple' and click 'Define')  
 Studentized Residual for LBVOL → Y Axis  
 Predicted Values for LBVOL → X Axis

SPSS OUTPUT FOR BOX 9.17 **Transformed lizard skulls****General linear model****Between-Subjects Factors**

	N
SITE 1	9
2	9
3	8
4	10
5	9
6	4
7	8
8	16
9	7

**Tests of Between-Subjects Effects**

Dependent Variable: LBVOL

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	203.850 <sup>a</sup>	9	22.650	114.333	.000
Intercept	1.087E-02	1	1.087E-02	.055	.815
SITE	.301	8	3.766E-02	.190	.992
LJAWL	10.405	1	10.405	52.520	.000
Error	13.867	70	.198		
Total	1385.696	80			
Corrected Total	217.717	79			

a. R Squared = .936 (Adjusted R Squared = .928)

**Parameter Estimates**

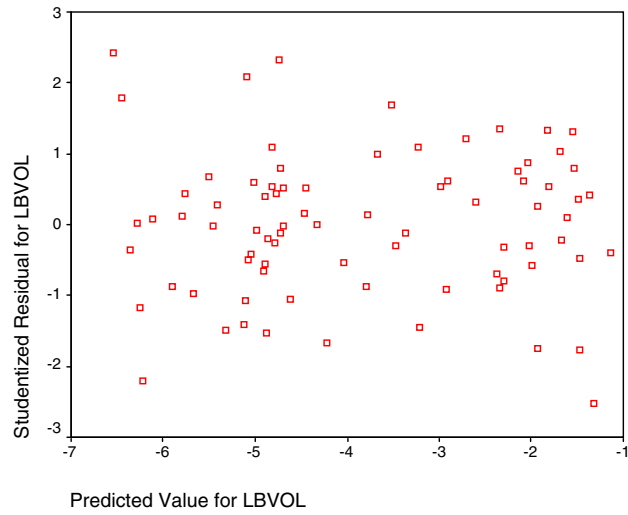
Dependent Variable: LBVOL

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	-.195	.837	-.233	.817	-1.863	1.474
[SITE=1]	9.106E-02	.593	.154	.878	-1.091	1.273
[SITE=2]	.129	.271	.475	.636	-.412	.669
[SITE=3]	.148	.334	.443	.659	-.518	.814
[SITE=4]	-2.27E-02	.299	-.076	.940	-.620	.574
[SITE=5]	4.800E-02	.479	.100	.920	-.907	1.003
[SITE=6]	7.423E-02	.329	.226	.822	-.582	.730
[SITE=7]	-5.60E-02	.240	-.233	.816	-.534	.422
[SITE=8]	.168	.621	.271	.787	-1.070	1.405
[SITE=9]	0 <sup>a</sup>	.	.	.	.	.
LJAWL	1.990	.275	7.247	.000	1.443	2.538

a. This parameter is set to zero because it is redundant

SPSS OUTPUT FOR FIGURE 9.23 **Residual plot for alternative analysis of lizard skulls**

**Graph**



**Checking the perfect model**

The SPSS output for this exercise may be found in the answers to exercises.