REPORT OF THE

HOUSEHOLD RATING WORKING PARTY 1993

Members

Andrew Button (Chairman)
Martyn Bennett
Craig Pettengell
Nylesh Shah
Budhi Singh
Philip Webb

Table of Contents

1 Summary

The work of the Household Rating Working Party (HRWP) for 1993 has concentrated on three main areas. Two of these are continuations of the work of the previous HRWP presented at GISG 1992, namely a market overview and analysis of the British Geological Survey Subsidence Database. The third area of work has been an attempt to determine if the frequency of some types of claims are affected by the current state of the economy.

It has now been some time since insurance companies started to rate by postcode. However this change has lead to fairly significant differences in the published rates of the insurance companies for specific postcode areas. The differences may be attributable to differences in loadings for:

expenses
commission
re-insurance costs
past losses recoveries
profit loadings
minor differences in cover

Differences may also be caused by each company's policy on smoothing between neighbouring postcode regions and the use of additional factors such as year of construction.

The published rates will be influenced by marketing considerations. This may be partly responsible for the differences in the published rates.

Not withstanding the above the working party analysed the tabulated rates for buildings insurance contracts by postcode of thirteen companies to see whether there were any patterns across the country. Companies which did not have a published rate for any particular postcode were ignored in the calculation. First of all the difference between the highest and lowest rate was calculated. Chart 1 shows the number of postcodes falling in each range. It is fairly surprising that the market rates are so widely spread.

On closer examination it was evident that some of the large spreads were due to one or two companies being out of line with the majority of the companies. This may be to discourage business being written in the area or to encourage growth in the area possibly by writing at rates which are subsidised by other areas. To get a better idea of the differences the coefficient of

variation (the sample standard deviation divided by the mean) for each of the postcodes was calculated. Chart 2 shows the number of postcodes falling into each range. Again it was surprising to see the number of postcodes where the coefficients were large.

The map of England and Wales shows the geographical distribution of the coefficients. The differences in Scotland were not marked. It appears that the majority of the large differences are concentrated in the south and east. It is possible that some companies have higher rates in this area due to the recovery of losses resulting from the 1987 and 1990 storms in this area. Coefficients greater than 0.25 are not concentrated in an particular area of the country.

Further analysis needs to be carried out to determine whether any information can be extracted from published rates in relation to the relative rates charged by each company. It may be that although there are great differences in the actual rates the relative rates of companies are similar.

In conclusion it appears that as far as the published rates are concerned the buildings insurance market is fairly competitive in the sense that a customer can find differing prices. However most of the difference is probably attributable to insurance companies wishing to collect an adequate overall premium and not to differing assessments of the underlying risk.

Other developments

Use of detailed postcodes

With the introduction of rating by outer postcodes (first four characters) some companies have started to take account of the next character in specific areas. Given the big differences in rates illustrated above new entrants into the market may be the first to select risks based on the first five characters.

Year of construction

Companies have started to use this as an additional factor to differentiate risks.

Age Discounts

These are becoming more popular, however they are usually allowed for by the means of special products rather than by means of a discount.

No claims discounts

Some companies have started to offer these. It will be interesting to see how the market as a whole reacts and develops.

Loyalty Discounts

Given the high costs of acquiring new business it is surprising that only a few companies offer these.

Subsidence Excess

Most companies have now increased the subsidence excess to £1,000.

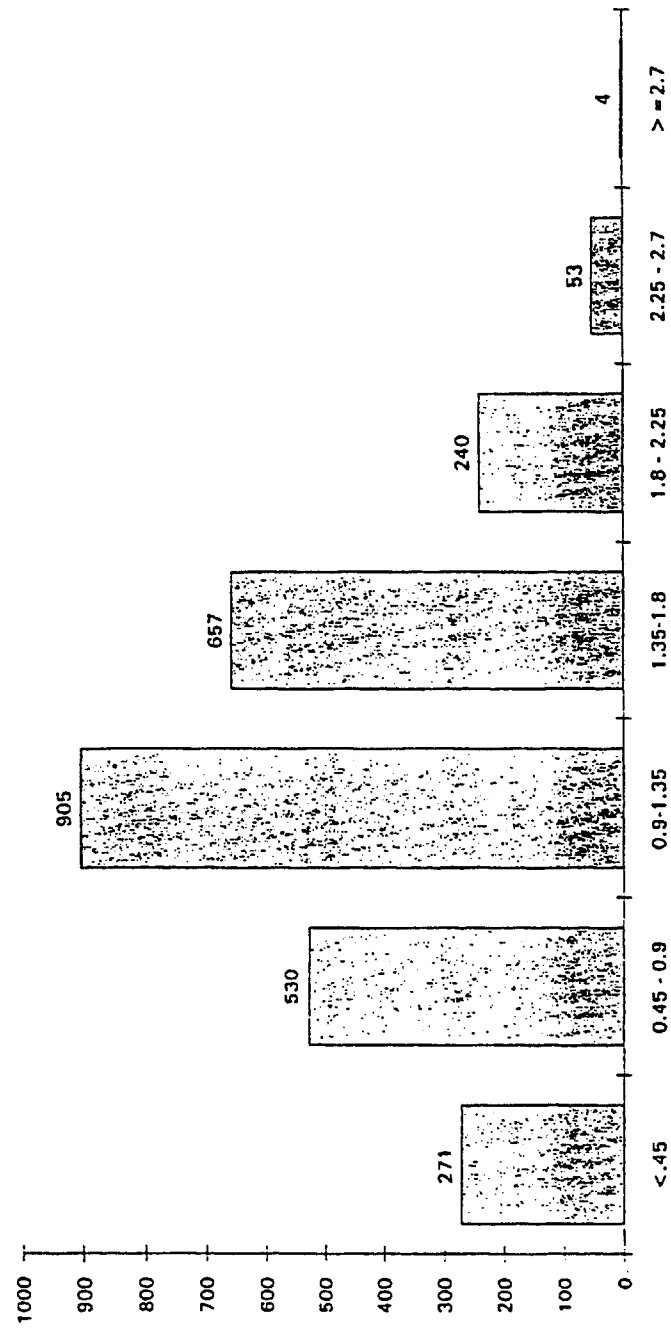Chart 1

Number Of Postcodes Falling Into Each Spread Range

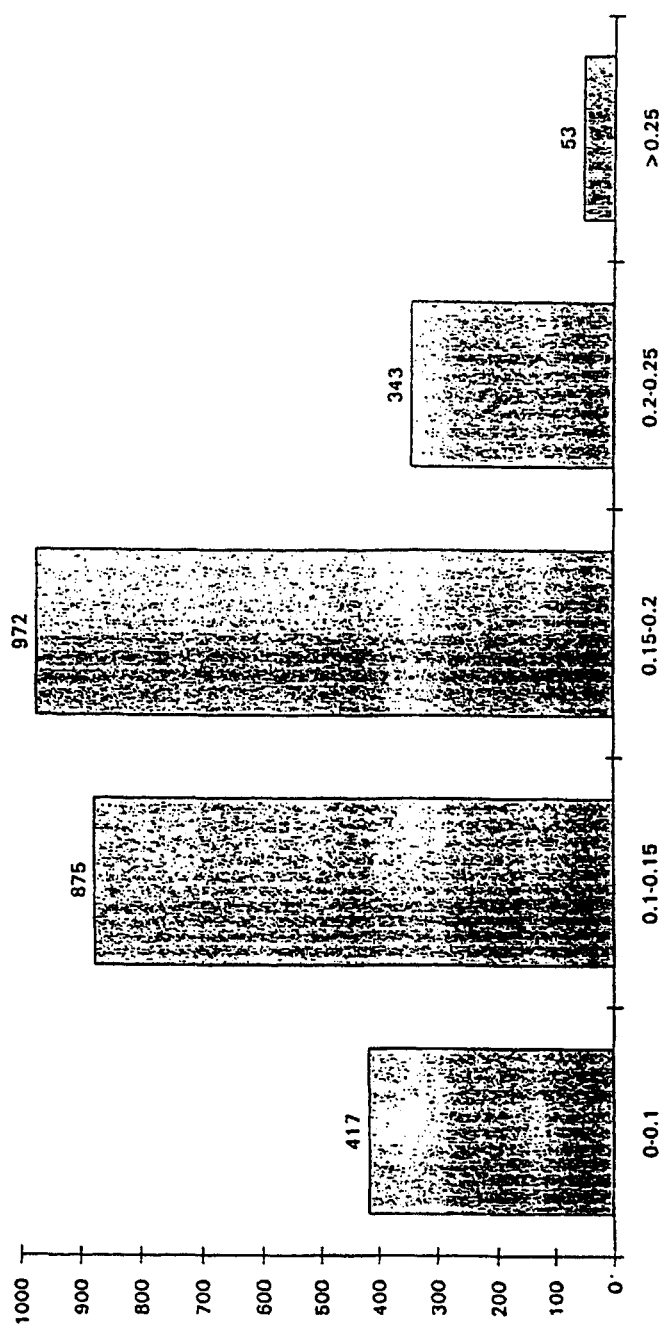**Chart 2**

Number Of Postcodes Falling Into Each Coefficient Of Variation Range

Coefficient of Variation of Building Rates

COV
0 - 0.1	0.1 - 0.15
0.15 - 0.2	0.2 - 0.25
>0.25	

Purpose and Scope

The purpose of the investigations were to understand how economic conditions affect claims experience. This would enable companies to reflect anticipated economic changes in their pricing levels and underwriting criteria, together with maintaining an ongoing claims operation consistent with the expected incidence of fraudulent claims.

The claim types that were expected to be most affected by economic conditions were

- theft of contents
- contents accidental damage (fraudulent elements)

Summary

Although some insurance data was provided, it was not sufficient for us to perform meaningful analyses. However, we were able to review an investigation on the effects of the economy on general property crime. In addition, we considered the type of insurance data that would be required for a meaningful analysis. The rest of this section discusses these together with our preliminary conclusions on how analysis in this area could be of benefit to household insurers.

Lessons Obtained From a General Property Crime Investigation

The Home Office Research and Statistics Department have an interest in the effect of the economy upon general crime rates due to their need to provide adequate judicial and custodial facilities. We now summarise the study of Crime and Consumption [1] performed by Dr Simon Field.

This study investigated relationships between changes in reported property crime and economic indicators.

The crime types examined included:

- residential burglary
- non residential burglary
- theft from person
- theft from vehicle
- theft of vehicle
- shop theft
- criminal damage
- fraud
- forgery

The economic indicators examined included:

- personal consumption per capital
- unemployment
- personal disposable income
- gross domestic product

The time period of data examined ranged from 40 to 100 years depending upon the crime type and economic indicator pair examined.

The conclusion was that all property crime types excluding fraud and forgery were correlated to economic conditions. The key economic driver appeared to be consumption. The graph illustrates that property crime appears to increase as consumption decreases and vice versa.

Some of the suggested drivers behind these correlations include the following:

Motivation to Steal. As employment opportunities decrease preventing to some extent the legal acquisition of goods, there is an increased motivation to steal amongst the unemployed.

Availability of Items to Steal. With economic growth comes a greater availability of items to steal.

PROPERTY CRIME AND CONSUMPTION

Annual growth in recorded property crime and annual decline in personal consumption per capita (+3). Logarithmic growth rates.

— Consumption   —•— Property Crime
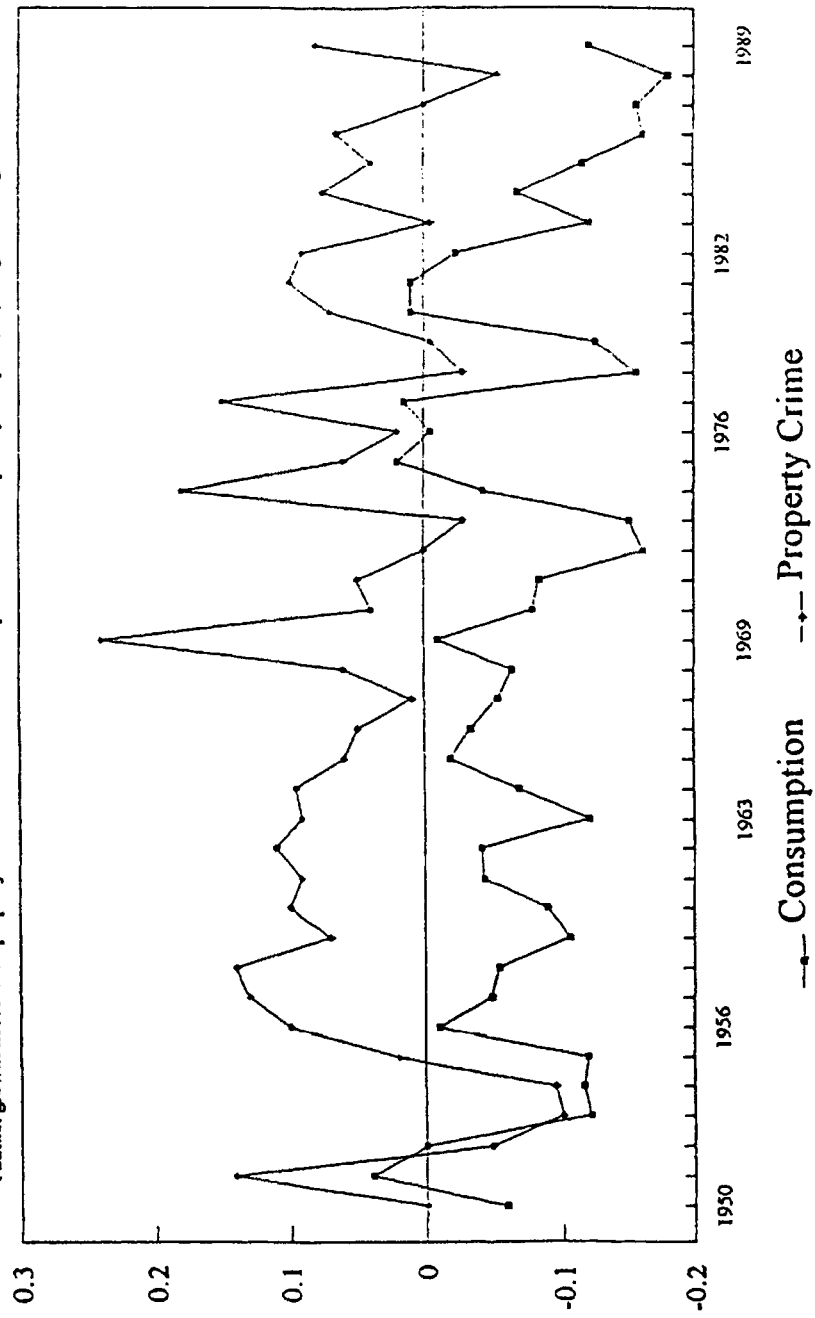
Theft Opportunities. Properties are unoccupied more often in times of economic growth creating greater theft opportunities.

The first driver would cause increases in recorded property crime to be associated with decreases in consumption, whilst the other two act in the opposite way. Therefore, the first of the drivers appears to be dominant as the graph shows that increases in recorded property crime are associated with decreases in consumption.

Model Insurance Analysis

We were provided with data from a limited number of sources. However, the data were not in a format that would enable us to perform meaningful analyses. However, we set out below our considerations regarding the type of data required, the relationships to be examined and some of the wider issues associated with such an analysis.

Data. The data should split claims experience by numbers and amounts and should include exposure information, both in terms of insured years and sum insured years. The economic data is most readily available in Economic Trends [2].

Both the claims experience and the economic data need to be collected for a period of time relating to a number of economic cycles. If this is not the case any apparent correlation could be due to short term movements rather than a real underlying effect. This could prevent an analysis of accidental damage claims at present due to their limited history.

The data we were provided with either had sufficient historical extent but excluded exposure information, or included exposure information but had an insufficient historical extent. The exposure information was required to ensure that the analysis would not be invalidated due to changes in the insured population and the inflation of insured values.

Relationships to Examine. It is our experience that there are many factors that are likely to affect a long term relationship between claims experience and the state of the economy. These might include changes in insureds propensity to claim, erosion of moral values, socio-economic changes, general level of premium rates, amongst others. Given that economic cycles last more than a few years, these long term effects are likely to inhibit a model that can relate to absolute level of recession affected claims to particular economic indicators. However, these secular effects are unlikely to be of significance over a period of say one year. Therefore, a more realistic model is likely to be obtained from investigation of the relationship between increases in claims between one year and the next to changes in economic indicators over the same period.

Economic Indicators. Each of the economic indicators will either lag, lead or be coincident with the economy in general. For example, decreases in unemployment tend to lag behind increases in gross domestic product. Therefore, consideration will need to be given to the lag between each economic indicator and the claims experience when determining which gives the best correlation. Of course, the leading indicators will be of most interest in a business context.

Insurance. The claims experience will be affected by changes in claims handling procedures, changes in insurance coverage and redlining amongst others. A knowledge of how these issues have affected past claims experience may be necessary.

Socio-Economic Effects. An interesting extension to the analysis would be to consider socio-economic effects explicitly. Experience has shown that young unemployed males tend to be a major cause of recession-related claims. Information is available that could be used in conjunction with geographic information systems for this purpose.

Conclusions

It is clear from the work performed by Dr Simon Field that there is a link between property crime and the economy. We are not aware of any studies that indicate that this is also the case for insurance claims experience.

The knowledge of the extent of such a relationship could be significant for household contents insurance. For example, supposing the theft rate increases by 20% in a year then if theft related claims account for 50% of the overall claims costs, the economic effects may account for up to 10% of the premium increase required.

Clearly, in order to use a knowledge of links between the economy and insurance claims experience, one will need an opinion of economic prospects. Experience has shown that precise forecasts are in fact rarely accurate! However, a general understanding of the current state of the economy and the direction in which it may be moving could help a company to make appropriate business decisions regarding increases in premium rates, tightening of underwriting controls and increased attention to possible fraudulent claims in the claims handling process. Indeed, perhaps the more interesting scenario to consider is one where the economy appears likely to improve in the short term and whether knowledge of such a relationship would enable a company to decrease premium rates ahead of its competitors.

The Working Party would very much like to examine appropriate insurance claims data in the coming year. We heartily request those that have access to such data to provide it.

Acknowledgements

The Working Party are very grateful to Dr Simon Field for the insight he has provided.

References

1   Crime and Consumption, Dr Simon Field. Home Office Research and Statistics Department Research Bulletin Number 29, 1990.

2   Economic Trends. Central Statistical Office.

This database was first described in the report of the 1992 HRWP, and the results of an initial high level analysis were presented at GISG. During the past year BGS have continued to develop this database, which should be completed by the end of August 1993.

The database is called the Geo-Hazard Susceptibility Package (GHASP), and BGS define its purposes as

i)   to aid the insurance industry in establishing more accurate and detailed building insurance ratings

ii)   to give underwriters a better understanding of their detailed exposure to the risks of subsidence

iii)   to help insurers in diagnosing what type of subsidence is present in each claim, since this is crucial in arriving at the correct remedial measures for each property

GHASP contains information for each postcode sector of the country (except the Scottish Highlands), on susceptibility to ground movements which insurers would define as subsidence. It does this by looking at six factors within each postcode sector

- shrink / swell clays

- landslips

- cambering and gullying

- shallow mining

- natural dissolution

- compressible deposits

Each geological hazard is assessed across each postcode sector as a whole and an average risk factor produced for each sector.

The risk factors have been assessed by field and engineering geologists with long experience of BGS surveying in their own particular regions. The result is a quantitative assessment of the average risk in each sector, calculated by applying weighting factors at both a local and national level.

Where susceptibility to a geological hazard is indicated, this does not necessarily mean that it has occurred in the past or will occur in the immediate future, as many hazards are triggered by changes in the local environment for example rainfall levels, changes in the watertable, excavations, construction etc. The local weightings within each postcode sector have been used to try to accommodate such changes so that the susceptibilities reflect not only the past history of the area but also have predictive value for the future.

Future updates to the GHASP database which are currently being considered include the following:

i)   Inclusion of factors to cover running sands, land fill and made ground.

ii)   An amalgamation of GHASP and the National Rivers Authority flooding database, covering both inland and coastal flooding.

iii)   Site specific data. This is aimed at niche markets and will provide much more localised and detailed information based on 1:10000 maps, which will allow the rating of individual buildings, for example commercial or industrial buildings. Site specific data is obviously much more expensive to produce than postcode sector data and it is envisaged that it would be produced on demand for relatively small areas.

iv)   A derivation of the system to provide surveyors and loss adjusters with guidance on the cause of subsidence damage to buildings, and may be available on a site specific basis. It is necessary to correctly identify the cause of subsidence damage, for example landslip, shallow mining etc, so that remedial work can be carried out properly. It is felt that some surveyors and loss adjusters may take a blanket approach, in that once a property in a particular area has been identified as suffering from a certain type of subsidence damage, then all subsequent occurrences of subsidence in that area will be diagnosed similarly, without proper investigation. It is envisaged that this enhancement to the system will assist in reducing this perceived problem.

5 Subsidence Claims Analysis

This analysis is a continuation of the work carried out by the HRWP and presented at GISG in 1992.

Due to problems in obtaining the factors from BGS, the work has not been completed at the time of going to print, but will be presented at GISG in October 1993.

As part of the analysis it is intended to use GLIM (Generalised Linear Interactive Modelling) which is a computer program for fitting linear models and more general models relating to them. A brief description is given in the Appendix to this paper.

6   Future Work

For the past two years, the members of the HRWP have tried to carry out analyses relating to topical issues in the household market and to provide updates on sources of information such as the BGS GHASP database and suitable analytical tools, for example GLIM.

We have had considerable problems in obtaining suitable data in sufficient volumes to carry out meaningful analyses, and unless insurers are more willing in the future to contribute data, the HRWP must consider its future.

Therefore, we seek views on whether the HRWP is carrying out a useful function in its present format. If so, will the insurers be willing to contribute data more freely than at present? If not, whether a more useful role could be undertaken by, for example, adopting a more theoretical approach or concentrating on longer term issues.

Examples might include investigations into customer rating, such as credit scoring and behavioural scoring, and reporting on other databases such as that being produced by the National Rivers Authority.

<u>Overview</u>

GLIM (Generalised Linear Interactive Modelling) is a computer program for fitting linear models and more general models relating to them, ie Generalised Linear Models (GLMs).

GLMs are a natural generalisation of linear models, which have been extensively used in actuarial work, for example in describing claim frequency rates and an average claims costs in motor insurance.

Although their use in actuarial work is relatively new, they have a wide area of applications beyond modelling household insurance statistics.

Other applications within actuarial work to which they have been applied are:

i)   Fitting loss distributions in general insurance work.

ii)   Modelling excess mortality in life assurance underwriting.

iii)   Modelling lapse rates by age at entry and duration in life assurance.

The theory behind GLMs is very general and widely used outside the actuarial world, for example in Credit Scoring by applying points to levels of a number of factors with acceptance / rejection dependent on whether a prescribed pass mark is reached.

<u>The Model</u>

For a series of observations, a GLM aims to investigate their relationship with any factors which might be thought to affect them. Each observation is known as a response and the factors affecting it are contained in a linear predictor.

A GLIM is characterised by:

i)   A modelling distribution.
ii)   A linear predictor.
iiI) A link function.

A suitable modelling distribution has to be chosen as the responses are independent random variables. The GLIM computer package offers a choice restricted to the exponential family of distributions which include the normal, binomial, poisson and gamma distributions.

The linear predictor is a linear function of the unknown parameters of the modelling distribution, the predictor parameters being estimated by maximum likelihood. It is linked into the model via a function of the means of the independent response variables, called the link function.

It is possible to use more than one link function with a modelling distribution. In the special case of a normal modelling distribution with the identity link, this is identical to estimation by least squares using the classical model.

Testing the Fit

The number of parameters in the model can be varied from one (the null model) to the number of observations (the saturated model). In the case of the null model all the variation in the data is attributed to the error structure of the model. At the other extreme, the saturated model, the fitted values are the data themselves. The aim is to find a model between these two extremes which involves as few parameters as possible whilst leaving a small pattern-free set of residuals. This is resolved by interpreting the difference between model deviances, using the saturated model as a benchmark.

The statistical test used for this purpose is the chi-square. In the special case of the normal modelling distribution and he identity link, this leads to F-statistics associated with analysis of variance tests.

The GLIM computer package displays the deviance and degrees of freedom for any fitted model. Its suitability compared with any other model is indicated by the significance of the reduction in the deviance using the appropriate statistical test. The test is therefore not of the goodness of fit of a particular model taken in isolation, but compares two models, one of which is a simplified version of the other. However, starting with the saturated model means that all factors which genuinely affect the response are included initially so the fit will be good.

Plots of residual errors play an important role in checks on the adequacy or otherwise of the fit of any model. They help to highlight the following departures from the model:

i)   Systematic - where the model is a poor fit of the data.

ii)   Isolated points - the presence of outliers, which need investigation.

If the model is good, there should be an evenly spread scatter of values around the mean value of zero in any plot of the standardised residuals.