Cairo University

Faculty of Graduate Studies for Statistical Research

Heteroskedasticity in Multiple Regression Analysis: What it is, How to Detect it and How to Solve it

Name:

Ghofran Talaat sayed Muhamed

Under supervision of:

Dr. Heba Fathy

DEC-2023

Table of contents

Table of Contents

Heteroskedasticity in Multiple Regression Analysis: What it is, How to Detect it and How to Solve it

Ghofran Talaat

Faculty of Graduate Studies for Statistical Research, Cairo University, 12613, Giza, Egypt

Abstract

Within the realm of psychology and the social sciences, Ordinary Least Squares (OLS) regression stands out as one of the most widely used techniques for conducting data analysis. To ensure the appropriateness of the inferences drawn from the application of this method, it is imperative that a number of assumptions are met including the one of constant error variance (i.e. homoskedasticity). By drawing inspiration from the field of econometrics, the purpose of this paper is to provide a comprehensive explanation of the meaning of heteroskedasticity and use one of the real world situations where heteroskedasticity exists, and then the comprehensive explanation continues to include how to detect heteroskedasticity through statistical tests specifically designed for this purpose and the consequences of ignoring it. Additionally, this paper delves into the methods of addressing heteroskedasticity through the application of heteroskedastic-consistent standard errors.

Keywords: Linear Regression Model, OLS, OLS Assumptions, heteroscedasticity, HC standard errors, Efficiency, Gauss Markov theorem.

Introduction

Ordinary least squares (OLS) regression is the most widely used method for fitting linear statistical models and it takes the following form: $Y_i = B_0 + B_1 X_{i1} + B_2 X_{i2} + \dots + B_K X_{iP} + \varepsilon_i$ Where Y_i is the i^{th} value of the outcome variable, B_0 is the regression constant, x_{ij} is case i score on the j^{th} of p explanatory variables, B_j is predictor j's partial regression weight, and ε_i is the error for case i. We can represent the OLS regression model as follows: $Y = XB + \varepsilon$

This form of OLS is called the matrix form of the OLS model, here Y is a vector of outcome observations and since we can denote the sample size by n then Y will be of order $(n \times 1)$, X is a matrix of the independent variables' values and since it includes a column of ones in order to represent the regression constant X's order will be $n \times (P+1)$, notice that we have P explanatory variables, and ε is vector of order $(n \times 1)$ and it represents the errors on the model. The regression coefficients $(B_i \ \ \forall i = 1, 2, \ldots, P)$ in the model provide information about the relationship between each predictor variable and the outcome variable.

Researchers are often interested in testing the significance of a specific regression coefficient/ coefficients or making a confidence interval for it using an estimate derived from the sample combined with an estimate of the sampling variance of the estimate (Hayes & Cai, 2007).

The validity of inferential statistics represented by hypothesis tests and confidence intervals depends on the extent to which the model's assumptions are met. The OLS model has a set of assumptions and according to the Gauss Markov theorem if these assumptions are met; the ordinary least squares (OLS) estimator of the coefficients of a linear regression model will be the best linear unbiased estimator (BLUE).

The assumptions of the OLS regression model

1. Y_i's are generated according to the OLS model basic form; meaning that the linear regression model must be **"linear in parameters".**
2. X values are fixed (rather than random); meaning that they should be **exogenous** $(cov(x, \epsilon) = 0)$.
3. Errors are uncorrelated random variables.
4. The expected value for the errors should be zero in order to ensure un-biasness.
5. Constant error's variance which is known as homoscedasticity assumption (White, 1980).

Assumption (2) can be relaxed without causing significant consequences when the predictors are random variables rather than fixed, through interpreting the partial regression coefficients as "conditional" on values of X, given that the conditional expectation of ε given the values of X is zero. Unfortunately, this conditional expectation may be nonzero, and that may be due to various reasons like for example the existence of measurement error for the predictors which may lead to a bias in the OLS regression estimates (Hayes & Cai, 2007) however this is beyond the scope of the present paper.

The focus of this paper is **“Heteroskedasticity in Multiple Regression Analysis”** since, homoskedasticity assumption, is an assumption that cannot be treated lightly. We briefly overview the consequences of heteroskedasticity on hypothesis testing and confidence intervals based on OLS regression coefficient estimates, some of the methods for detecting violations of this assumption, and a suggested remedy **(heteroskedasticity-consistent (HC) standard error)** that helps us to deal with the problem without changing the estimation method.

A significant number of scholars devote their attention to the concept of heteroskedasticity and construct various models to address specific issues related to heteroscedasticity thus we are going to present some of the studies made before in this subject.

Literature review

In an integrative review paper carried out by, (Hayes & Cai, 2007) highlights the significant consequence of Heteroskedasticity, and he showed that this violation, in turn, leads to biased and inconsistent estimators of the covariance matrix of the parameter estimates, thereby affecting inferential statistics. To mitigate the impact of heteroskedasticity on OLS regression, the paper recommends the adoption of heteroskedasticity-consistent (HC) standard error estimators, such as HC3 and HC4. These estimators have the potential to enhance the validity of inferences and the power of hypothesis tests in the presence of heteroscedasticity.

A study carried out by (Zietz, 2001) demonstrated that in the case of cross-sectional data and a linear regression model, the rejection of the null hypothesis of homoskedasticity may be due to neglected parameter heterogeneity. The paper highlighted that the neglect of parameter heterogeneity poses a significant problem for regression analysis as it leads to a significant bias in the estimated regression coefficients, rendering them economically meaningless. The findings of this paper indicate that the presence of heteroskedasticity should not be dismissed as unimportant and/or routinely treated with the application White's (1980) heteroskedasticity-consistent variance covariance matrix estimator. Instead, it is crucial to investigate the underlying causes of heteroskedastic residuals, with neglected parameter heterogeneity being one potential factor that must be considered.

(Linton & Xiao, 2019) conducted a study that explores the effective estimation of nonparametric regression in the presence of heteroskedasticity. Their analysis specifically focuses on local polynomial estimation of nonparametric regressions with conditional heteroskedasticity in a time series context. The researchers introduced a weighted local polynomial regression smoother that takes into consideration the dynamic heteroskedasticity. Their findings indicate that, in many commonly used nonparametric regression models, their proposed method exhibits a lower asymptotic variance compared to the conventional unweighted procedures. However, it is worth mentioning that there is a traditional recommendation against weighting for heteroskedasticity in nonparametric regressions.

What Is Heteroskedasticity?

The assumption of homoskedasticity in regression analysis states that the variance of the regression errors is constant across different values of X (please refer to Figure 1). In other words, the variance of the errors is unrelated to any predictor or any linear combination of the predictor variables.

This assumption is important because it allows us to make valid statistical inferences. When this assumption is violated, we say that the errors are heteroskedastic, which implies that the variance of the errors is not constant across different values of X (please refer to Figure 2 or Figure 3 which is more general case where Auto-correlation exists as well). ((Astivia & Zumbo, 2019); (Hayes & Cai, 2007); (White, 1980))

$$Var(\epsilon) = \mathbb{E}(\epsilon\epsilon') = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 & 0 \\ 0 & \sigma^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \sigma^2 & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \sigma^2 \end{bmatrix} = \sigma_\epsilon^2 \boldsymbol{I}$$

Figure (1): Homoscedasticity (Astivia & Zumbo, 2019)

$$Var(\epsilon) = \mathbb{E}(\epsilon\epsilon') = \begin{bmatrix} {\sigma_1}^2 & 0 & \cdots & 0 & 0 \\ 0 & {\sigma_2}^2 & \cdots & 0 & 0 \\ \vdots & \vdots & {\sigma_3}^2 & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & 0 & {\sigma_i}^2 \end{bmatrix}$$

Figure(2): Heteroskedasticity (Astivia & Zumbo, 2019)

$$Var(\epsilon) = \mathbb{E}(\epsilon\epsilon') = \begin{bmatrix} \sigma_{1,1}^2 & \sigma_{1,2} & \cdots & \sigma_{1,i-1} & \sigma_{1,i} \\ \sigma_{2,1} & \sigma_{2,2}^2 & \cdots & \sigma_{2,i-1} & \sigma_{2,i} \\ \vdots & \vdots & \sigma_{3,3}^2 & \vdots & \vdots \\ \sigma_{i-1,1} & \sigma_{i-1,2} & \sigma_{i-1,3} & \ddots & \sigma_{i-1,i} \\ \sigma_{i,1} & \sigma_{i,2} & \cdots & \sigma_{i,i-1} & \sigma_{i,i}^2 \end{bmatrix}$$

Figure (3): Heteroscedasticity and Auto-correlation (Astivia & Zumbo, 2019)

Real-World Example

One common example of heteroskedasticity can be observed in the relationship between food expenses and income. Individuals with lower incomes often face limitations in their food expenses due to their restricted budget. As income levels rise, individuals tend to allocate more resources towards food consumption as they have a wider range of options and fewer budgetary constraints. In this scenario, the variance of the residuals (taking into consideration that the variance of residuals equivalent to the variance of expenditure) is not equal across the independent variable of income. If one were to conduct a regression analysis using this dataset, the presence of heteroskedasticity would be observed.

Consequences of Heteroskedasticity

The unbiased and strongly consistent OLS regression estimator of the regression coefficients remains robust even in cases where the assumption of homoskedasticity is violated (Cribari-Neto, 2004). However, it exhibits lower efficiency as evidenced by a larger sampling variance compared to certain alternative estimators. Nevertheless, statistical inference can be compromised when this assumption is violated, as it introduces bias to the standard errors and test-statistics ((Hayes & Cai, 2007);(Godfrey, 2006)).

The standard errors of the regression coefficients are a function of the variance-covariance matrix of the error terms $\left(S.E\left(\hat{B}\right)=\sigma^2\left(X'X\right)^{-1}\right)$. If we neglected the fact of the existing heteroskedasticity, we will underestimate or overestimate the variance of the error terms which will lead to smaller/higher standard errors of the regression coefficients. (Godfrey, 2006)

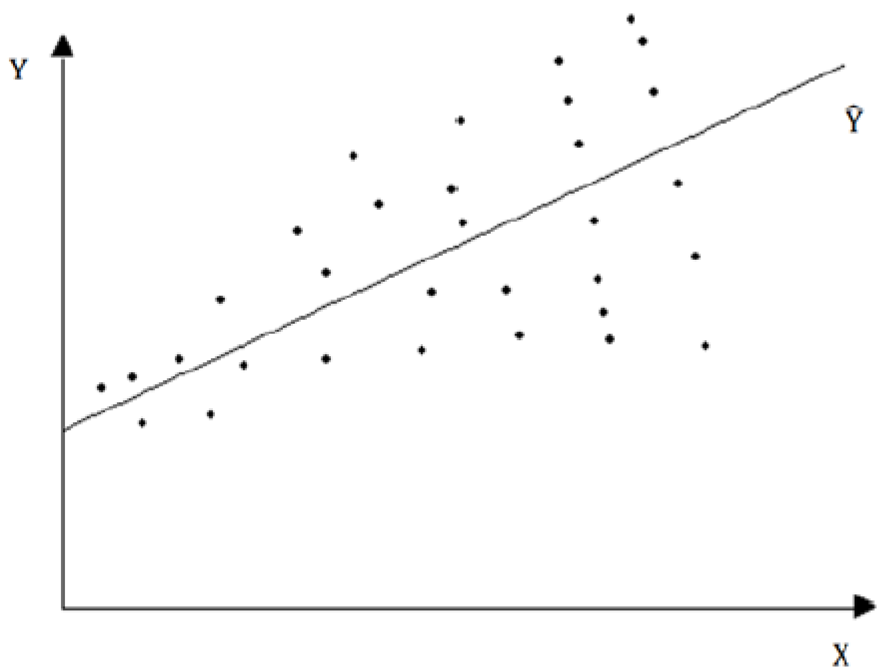

Figure (4): Diagram of distribution of points and its effect on estimation (Đalić & Terzić, 2021)

From Figure(4) it is clear that if we neglected the fact of the existing heteroskedasticity, and assumed to have a constant variance for the residuals across different value of X, we will underestimate (in case of large values of X) or overestimate (in case of small values of X) the variance of the error terms as mentioned before, as mentioned before the variance of the error terms is the same as the variance of the dependent variable Y, for more mathematical illustration: $\mathrm{var}(XB+\varepsilon)=\mathrm{var}(\varepsilon) \because X \text{ is fixed due to OLS assumptions}$.

The construction of confidence intervals and test-statistics for assessing the significance of regression coefficients relies on the accurate specification of standard errors. Consequently, if there is an incorrect specification of the standard errors, both the confidence intervals and test-statistics will become invalid. This, in turn, can result in incorrect decisions being made regarding the null hypothesis ((Godfrey, 2006);(Fox, 1997)).

The extent of the problem heteroskedasticity produces depends on both the form and the severity of heteroscedasticity. Consequently, this leads to an increase in the occurrence of Type I errors or a decrease in the statistical power for hypotheses testing that encompass the regression coefficients, as well as inaccuracies in estimating the upper and lower bounds of confidence intervals ((Hayes, 1996);(Long & Ervin, 2000)).

Consideration must be given to the fact that the presence of heteroscedasticity leads to bias, which is not eradicated by increasing the sample size (Long & Ervin, 2000). In reality, the issues arising from heteroskedasticity can be exacerbated when employing OLS regression.

Detecting heteroscedasticity

A frequently suggested technique found in introductory textbooks for identifying heteroskedasticity in ordinary least squares (OLS) regression models is to create a graph displaying the residuals derived from the sample data in relation to the fitted values, and subsequently analyze for any observable trends or patterns. If the plot appears to be a random cloud of noise with no discernible pattern (please refer to Figure 5), the assumption of homoskedasticity is likely to hold. (Cohen et al., 2013)

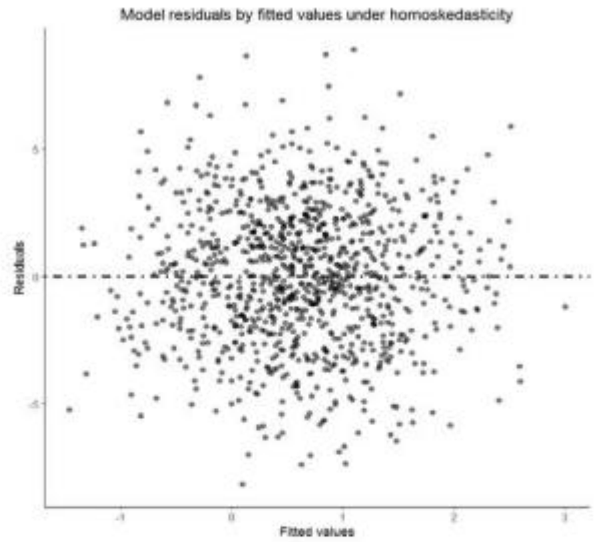

Figure (5): Homoskedasticity (Astivia & Zumbo, 2019)

However, If any kind of clustering or trend is detected (please refer to Figure 6), then the assumption is suspected and further assessment is needed. (Cohen et al., 2013)

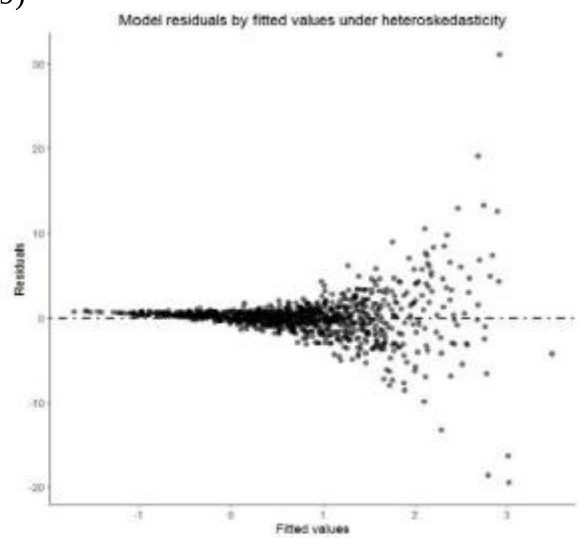

Figure (6): Heteroskedasticity (Astivia & Zumbo, 2019)

There is another opinion (please refer to Figure 7) that is replacing residual values on the scattering plot diagram by their squared values and draw them against X (this type is made for each x_i however we can consider the different Xs all at once by using fitted values instead as in Figure 5 and 6):

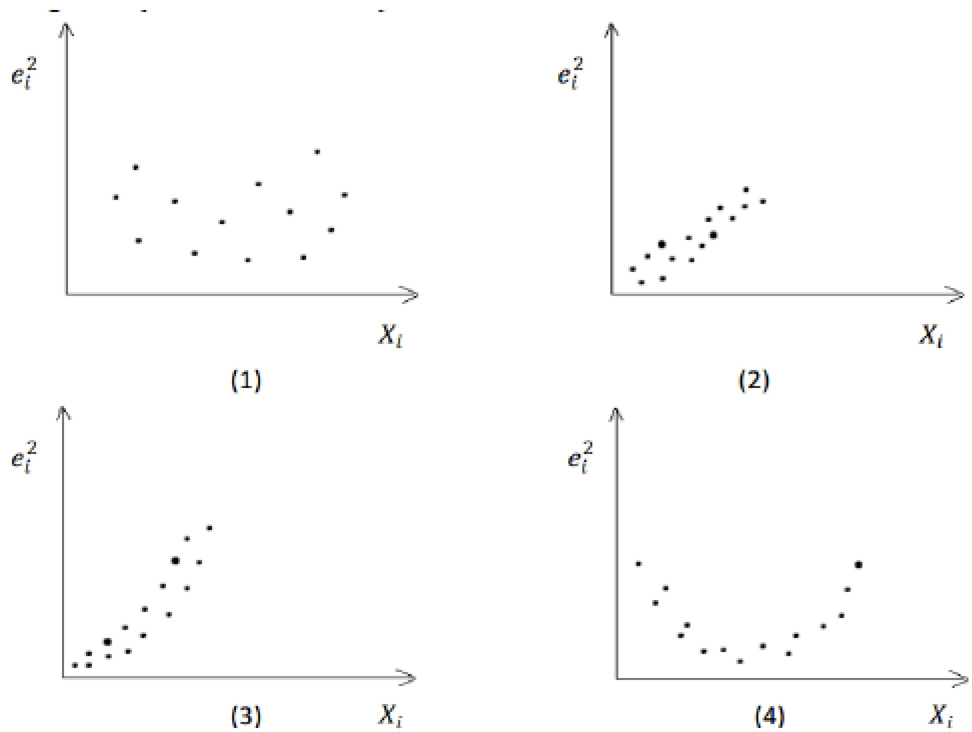

Figure (7): Scatter Plot Diagrams (Đalić & Terzić, 2021)

Although residual plots have the potential to be valuable in identifying heteroskedasticity in ordinary least squares (OLS) regression models, they may not provide a comprehensive assessment. In order to further investigate potential violations of assumptions, we will introduce two widely employed tests in the field of econometrics for identifying heteroscedasticity: **Breusch-Pagan test** and **White Test**.

Breusch–Pagan test which can be considered as one of the most classical heteroskedasticity test explicitly assesses whether the model errors (squared error) are associated with any of the model explanatory variables in a linear relationship through an auxiliary regression of the following form $\varepsilon_i^2 = \propto_0 + \propto_1 X_{1i} + \propto_2 X_{2i} + \ldots + \propto_P X_{Pi} + u_i$ for the original regression models of the form $Y_i = B_0 + B_1 X_{1i} + B_2 X_{2i} + \ldots + B_P X_{Pi} + \varepsilon_i$ (Astivia & Zumbo, 2019).

And then the null hypothesis must be stated and it has the following form $H_0 = \propto_1 = \ldots = \propto_p = 0$ (Astivia & Zumbo, 2019). Taking into consideration that this null is equivalent to saying that homoscedasticity exists in the original regression model, that is because

$\varepsilon_i^{\ 2}$ is used to represent the error variance so if $\propto_i = 0 \quad \forall i = 1, 2, \ldots, P$ then it is equivalent to saying that X's don't affect the variance of the error then the variance is constant across the different x's values (homoscedasticity).

Then R^2 is obtained from the auxiliary regression, in order to get the test statistic which is equivalent to $\left(nR^2\right)$ where n is the sample size, this test follows $chi - squared$ distribution with p-1 degrees of freedom (Astivia & Zumbo, 2019).

Breusch-Pagan test can only identify linear associations between the model residuals and the model predictors. However, the **White Test** can detect higher-order, non-linear functional forms of the predictors, such as quadratic and cross-product interactions among them. In this case, the regression model for the (squared) error terms (Auxiliary Regression) will take the following form: $\varepsilon_i^{\ 2} = \propto_0 + \propto_1 X_{1i} + \propto_2 X_{2i} + \ldots + \propto_K X_{Pi} + \gamma_1 X_{1i}^2 + \ldots + \gamma_p X_{pi}^2 + \ldots + \delta_1 (X_{1i}, X_{2i}) + \ldots + v_i$ (Astivia & Zumbo, 2019).

White test has the same null hypothesis and test statistic as the Breusch-Pagan test. However, it is more general in detecting other functional forms of heteroskedasticity. Nevertheless, there are important limitations to consider. Firstly, if many predictors are present, the regression of the linear, quadratic, and interaction terms in the same equation can quickly use up all the degrees of freedom present in the sample. Secondly, the White test does not exclusively test for heteroskedasticity. Model misspecifications could be detected through it, so a statistically significant p-value cannot be used as absolute evidence that heteroskedasticity is present (Astivia & Zumbo, 2019).

Fixing heteroscedasticity

Researchers have traditionally utilized data transformation as a means to address heteroscedasticity (Osborne, 2005). Among the various options available, the logarithmic transformation and other "variance stabilizing" techniques such as the square root have gained popularity. Nonetheless, the selection of a specific transformation may be arbitrary, potentially leading to a fundamental alteration in meaning of the variables and the regression model itself. Consequently, the interpretation of parameter estimates becomes reliant upon the newly induced scaling resulting from the transformation. Furthermore, it is not uncommon for transformations to exhibit limited or no discernible impact (Mueller, 1949).

A method that is highly favored as alternative to reduce the impact of heteroskedasticity on inference is the use of a heteroscedasticity consistent standard error (HCSE) estimator for OLS parameter estimates, as suggested by Long & Ervin, 2000. This technique is also referred to as heteroskedasticity-robust standard errors or robust standard errors. It finds its application in statistics and econometrics, particularly in linear regression and time series analysis.

In this method, the regression model is estimated using the standard OLS, but a different method is used to estimate the standard errors, which does not make the assumption of homoscedasticity. The attractiveness of this method is that, unlike other methods such as WLS, it requires neither knowledge about nor a model of the functional form of the heteroscedasticity as well as it does not require intensive computer simulation such as that required when performing bootstrapping (Astivia & Zumbo, 2019) i.e. another solution for heteroscedasticity used for small sample sizes or when the violations of parametric assumptions are present (Mooney et al., 1993).

Before we move forward, it's crucial to understand that heteroskedasticity might signal a mis-specified model. Yet, as (MacCallum, 2003) points out, all models are intrinsically imperfect to some extent. The most we can aspire to with a mathematical model of a dependent variable is to mimic the process that produces our observations. If we insist on perfect specification before making inferences from models, we may as well stop testing hypotheses using models, as achieving perfect specification is unattainable.

When performing inferential tests using OLS regression, we operate under the assumption of homoskedasticity, we presume that the errors adhere to a specific model (i.e., the elements in e have constant variance). On the other hand, when WLS or GLS are employed, the model of heteroskedasticity is assumed to be accurate (Astivia & Zumbo, 2019).

By advocating for the use of an HCSE estimator in OLS regression, we are not denying the value of other methods; in fact, they are beneficial. We are simply acknowledging an alternative approach to address the issue of inference in OLS regression, one that does not enforce a constraint (either assumed or modeled) on the error structure and does not require us to change the estimation method.

Heteroskedastic consistent standard errors fundamentally acknowledge the existence of non-unique variance and provide a different method for calculating the variance of the sample regression coefficients. Assuming a general form of heteroskedasticity, the variance-covariance matrix of the regression model errors takes a specific structure as follows:

$$\mathrm{var}(\varepsilon) = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \cdots & \sigma_{1i} \\ \sigma_{21} & \sigma_{22}^2 & \cdots & \sigma_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{i1} & \sigma_{i2} & \cdots & \sigma_{ii}^2 \end{bmatrix} = \Omega$$

For the usual OLS regression model expressed in the matrix form $Y = XB + \varepsilon$ the variance of the estimated regression coefficients $(\hat{B}s)$ when heteroskedasticity is present is as follows:

$\hat{\Sigma}_{\hat{B}} = \left(X'X\right)^{-1} X'\Omega X\left(X'X\right)^{-1}$. However, the usual form of it under the condition of homoscedasticity ($\Omega = \sigma^2 I$) is simply $\hat{\Sigma}_{\hat{B}} = \sigma^2 \left(X'X\right)^{-1}$ where σ^2 is the mean squared residuals (Astivia & Zumbo, 2019). It now apparent that the matrix **Ω** plays a pivotal role in the calculation of having proper standard errors that account for non-constant variance.

The rationale behind the creation of the new standard errors

Just as with any given sample, all we have is an estimate $\hat{\Omega}$ that will be needed to obtain these new uncertainties and since $\hat{\Omega}$ stems from the residuals of the regression model i.e. as estimates of the population regression errors, the conceptual idea behind heteroskedastic-consistent standard errors is to use the variance of each sample residual (r_i) to estimate the variance of the population errors (ϵ_{ii}). (i.e. the diagonal elements of **Ω**). Now, because there is only one residual r_i per person, per sample, this is a one-sample estimate, so $var\left(\epsilon_{ii}\right) = \frac{\left(r_i - 0\right)^2}{1} = r_i^2$ (Astivia & Zumbo, 2019).

Notice that, by assumptions stated back then, the mean of the residuals is zero, and regarding the division by 1 in the expression it is basically a reflection of the general formula for variance i.e. Variance is typically calculated as the average of the squared differences from the mean.

Therefore, if we let $\hat{\Omega} = diag\left(r_i^2\right)$ and back-substituting it in $\left(X'X\right)^{-1} X'\Omega X\left(X'X\right)^{-1}$ this will implies $\hat{\Sigma}_{\hat{B}} = \left(X'X\right)^{-1} (X' diag\left(r_i^2\right)X)\left(X'X\right)^{-1}$ this is the oldest and most widely used form of heteroskedastic-consistent standard errors and it has been shown to be a consistent estimator of $var\left(\hat{B}\right)$ even if the form of heteroskedasticity is not known (Astivia & Zumbo, 2019).

There are other versions of this standard error that offer alternative adjustments which perform better for small sample sizes, but they all follow a similar pattern to the modified standard errors of the estimated parameters; $\hat{\Sigma}_{\hat{B}} = \left(X'X\right)^{-1} (X' diag\left(r_i^2\right)X)\left(X'X\right)^{-1}$ nevertheless, we can say that the key issue is to obtain a better estimate of $\hat{\Omega}$ so that the new standard errors yield the correct Type I error rate and thus avoid the consequences of the assumptions' violation (Astivia & Zumbo, 2019).

Results

The paper presents a comprehensive examination of heteroskedasticity in regression analysis, a condition that arises when the variance of the errors is not constant across different values of X, the violation of homoscedasticity assumption, exemplified in the paper through a real-world scenario of the relationship between food expenses and income. Heteroskedasticity can, introduce bias to the standard errors and thus the test-statistics, and lead to incorrect decisions regarding the null hypothesis due to inaccurate specification of standard errors. The paper suggests techniques for identifying heteroskedasticity, such as creating a residual plot and conducting tests like the Breusch-Pagan test and White Test. To address heteroskedasticity, the paper recommends one method that helps us to continue using the OLS estimation method after addressing the violation.

Conclusion

In conclusion, this paper has underscored the significance of the assumption of homoskedasticity in Ordinary Least Squares (OLS) regression, a widely used technique in psychology and social sciences. It has provided a detailed understanding of heteroskedasticity, its real-world implications, and the consequences of overlooking it. The paper has also highlighted the importance of detecting heteroskedasticity through specially designed statistical tests.

Furthermore, it has explored the application of heteroskedastic-consistent standard errors as a method to address heteroskedasticity. The assumption of homoskedasticity is not to be taken lightly, and this paper emphasizes the need for careful consideration of this assumption in the application of OLS regression. The insights drawn from this paper contribute to a more robust and reliable application of regression analysis in the field of psychology and social sciences.

References

Astivia, O. L. O., & Zumbo, B. D. (2019). Heteroskedasticity in Multiple Regression Analysis: What it is, How to Detect it and How to Solve it with Applications in R and SPSS. *Practical Assessment, Research, and Evaluation*, *24*(1), 1.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*: Routledge.

Cribari-Neto, F. (2004). Asymptotic inference under heteroskedasticity of unknown form. *Computational Statistics & Data Analysis, 45*(2), 215-233.

Đalić, I., & Terzić, S. (2021). Violation of the assumption of homoscedasticity and detection of heteroscedasticity. *Decision Making: Applications in Management and Engineering, 4*(1), 1-18.

Fox, J. (1997). *Applied regression analysis, linear models, and related methods*: sage publications, Inc.

Godfrey, L. (2006). Tests for regression models with heteroskedasticity of unknown form. *Computational Statistics & Data Analysis, 50*(10), 2715-2733.

Hayes, A. F. (1996). Permutation test is not distribution-free: Testing H_0: $\rho= 0$. *Psychological Methods, 1*(2), 184.

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

Linton, O., & Xiao, Z. (2019). Efficient estimation of nonparametric regression in the presence of dynamic heteroskedasticity. *Journal of Econometrics, 213*(2), 608-631.

Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician, 54*(3), 217-224.

MacCallum, R. C. (2003). 2001 presidential address: Working with imperfect models. *Multivariate behavioral research, 38*(1), 113-139.

Mooney, C. Z., Duval, R. D., & Duvall, R. (1993). *Bootstrapping: A nonparametric approach to statistical inference*: sage.

Mueller, C. G. (1949). Numerical transformations in the analysis of experimental data. *Psychological Bulletin, 46*(3), 198.

Osborne, J. (2005). Notes on the use of data transformations. *Practical assessment, research and evaluation, 9*(1), 42-50.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: journal of the Econometric Society*, 817-838.

Zietz, J. (2001). Heteroskedasticity and neglected parameter heterogeneity. *Oxford Bulletin of Economics and Statistics, 63*(2), 263-273.