AFRE 835: Introductory Econometrics

Chapter 8: Heteroskedasticity

Spring 2017

Introduction

- Up until this point, we have relied on the assumption of homoskedasticity (MLR.5); i.e., $Var(u|\boldsymbol{x}) = \sigma^2$.
- Since u captures all other factors affecting our dependent variable, this is a strong assumption that is unlikely to hold in a number of settings; e.g.,
  - Wages modeled as a function of age, experience, and education;
  - Housing prices as a function on housing attributes and local amenities;
  - GRE test scores as a function of undergraduate GPA/coursework.
- We saw in the case of the linear probability model, the homoskedasticity assumption is necessarily violated.
- This chapter examines approaches to addressing violations of the homoskedasticity assumption.
- In particular, we now allow $Var(u_i|\boldsymbol{x}_i) = \sigma_i^2$, where the i subscript is used to emphasize that the variance depends on the particular value of $\boldsymbol{x}_i$.

Outline

1. Consequences of Heteroskedasticity for OLS
2. Heteroskedasticity-Robust Inference after OLS Estimation
3. Testing for Heteroskedasticity
4. Weighted Least Squares Estimation
5. Feasible GLS
6. Heteroskedasticity in the Linear Probability Model

Consequences of Heteroskedasticity for OLS

- A key first question is: How are our earlier results regarding the OLS estimator impacted by violations of MLR.5?
- Importantly, OLS remains unbiased and consistent. Theorem 3.1 used only assumptions MLR.1 through MLR.4.
- Likewise, our goodness of fit measures (R^2 and $\bar{R}^2$) remain unchanged.
- However, homoskedasticity was used to :
  - construct $Var(\hat{\beta}_j)$.
  - derive the distributions of the various test statistics (t-stat, F-stat, and LM statistic) used in hypothesis testing.

    ... These will no longer be valid if homoskedasticity does not hold.

What to Do?

- There are two basic approaches for dealing with heteroskedasticity.
  1. Adjust the standard errors of the OLS estimator to account for heteroskedasticity;
  2. Develop more efficient estimators.

Heteroskedasticity-Robust Procedures

- A major innovation in econometrics in the past three decades or so, has been the development of procedures to adjust the OLS standard errors so that they are valid in the presence of *heteroskedasticity of unknown form*, at least asymptotically.
- Note: The OLS estimator itself does not change, just the variance that is associated with the OLS estimator.
- Parallel adjustments are available for the t-, F- and LM-statistics used in hypothesis testing.
- The argument is easier to see in the context of a single independent variable, say

$$y_i = \beta_0 + \beta_1 x_i + u_i \tag{1}$$

with $Var(u_i|x_i) = \sigma_i^2$.

The Robust Variance

- The OLS estimator in this case is given by

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \beta_1 + \frac{\sum_{i=1}^{n}(x_i - \bar{x})u_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2}$$

- The conditional variance of the OLS estimator then becomes:

$$Var(\hat{\beta}_1|\boldsymbol{x}) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2\sigma_i^2}{[\sum_{i=1}^{n}(x_i - \bar{x})^2]^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2\sigma_i^2}{SST_x^2} \tag{3}$$

- It can be shown that a consistent estimator of this variance is given by:

$$\widehat{Var}(\hat{\beta}_1|\boldsymbol{x}) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2\hat{u}_i^2}{SST_x^2} \tag{4}$$

- Note: This estimator is valid even if homoskedasticity holds.

The Robust Variance (cont'd)

- Similar robust standard errors can be constructed in the more general multiple regression setting, with

$$\widehat{Var}(\hat{\beta}_j|\boldsymbol{x}) = \frac{\sum_{i=1}^{n}\hat{r}_{ij}^2\hat{u}_i^2}{SSR_j^2} = \frac{\sum_{i=1}^{n}\hat{r}_{ij}^2\hat{u}_i^2}{[SST_j(1-R_j^2)]} \tag{5}$$

...where $\hat{r}_{ij}$ denotes the i^{th} residual from regressing the x_j's on all the other regressors and SSR_j^2 and R_j^2 denote the SSR and R^2, respectively, from that regression.

The Robust Variance in Matrix Form

- Recall that in matrix notation, our model takes the form

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \tag{6}$$

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{u} \tag{7}$$

  with $E(\boldsymbol{u}|\boldsymbol{X}) = 0$.
- In the simplest form of heteroskedasticity, with no correlation in the errors across observations, we have

$$Var(\boldsymbol{u}|\boldsymbol{X}) = E(\boldsymbol{uu}'|\boldsymbol{X}) = \boldsymbol{\Omega} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} \tag{8}$$

The Robust Variance in Matrix Form

- Our OLS Estimator takes the form:

$$\begin{aligned} \hat{\beta} &= (\boldsymbol{X'X})^{-1}\boldsymbol{X'y} \\ &= (\boldsymbol{X'X})^{-1}\boldsymbol{X'}(\boldsymbol{X}\beta + \boldsymbol{u}) \\ &= (\boldsymbol{X'X})^{-1}\boldsymbol{X'X}\beta + (\boldsymbol{X'X})^{-1}\boldsymbol{X'u} \\ &= \beta + (\boldsymbol{X'X})^{-1}\boldsymbol{X'u} \end{aligned} \tag{9}$$

- From this we can see that

$$\begin{aligned} Var(\hat{\beta}) &= (\boldsymbol{X'X})^{-1}E[\boldsymbol{X'uu'X}](\boldsymbol{X'X})^{-1} \\ &= \boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1} \end{aligned} \tag{10}$$

  where $\boldsymbol{A} = \boldsymbol{X'X}$ and $\boldsymbol{B} = E[\boldsymbol{X'uu'X}] = \boldsymbol{X'}\Omega\boldsymbol{X}$

The Robust Variance in Matrix Form

- In constructing a estimator for this variance, we replace $\boldsymbol{B}$ with

$$\hat{\boldsymbol{B}} = (n-k-1)^{-1} \sum_{i=1}^{n} \hat{u}_i^2 \boldsymbol{x}_i' \boldsymbol{x}_i \tag{11}$$

  where $\boldsymbol{x}_i$ is a row vector of regressors for observation i.
- This gives us our estimator of the asymptotic variance of our OLS estimator:

$$\widehat{Var(\hat{\beta})} = \frac{n}{n-k-1} \boldsymbol{A}^{-1} \hat{\boldsymbol{B}} \boldsymbol{A}^{-1} \tag{12}$$

- Note that if homoskedasticity holds, then $\boldsymbol{B} = \sigma^2 \boldsymbol{A}$ and the usual variance estimator applies.

The Robust Variance (cont'd)

- The heteroskedasticity robust-standard errors are referred to as *White's sandwich* standard errors - based on the form of the estimator.
- Robust *t*-stat's are formed by just replacing the usual standard errors with their robust counterpart.
- Hypothesis testing proceeds as usual.
- Note: The interpretation of the estimated model proceeds just as it would in OLS with homoskedasticity.

Example 8.1 in Wooldridge

$$
\begin{aligned}
\widehat{\log(wage)} = \underset{\substack{(.100)\\[.109]}}{.321} + \underset{\substack{(.055)\\[.057]}}{.213}\, marrmale - \underset{\substack{(.058)\\[.058]}}{.198}\, marrfem - \underset{\substack{(.056)\\[.057]}}{.110}\, singfem \\
+ \underset{\substack{(.0067)\\[.0074]}}{.0789}\, educ + \underset{\substack{(.0055)\\[.0051]}}{.0268}\, exper - \underset{\substack{(.00011)\\[.00011]}}{.00054}\, exper^2 \\
+ \underset{\substack{(.0068)\\[.0069]}}{.0291}\, tenure - \underset{\substack{(.00023)\\[.00024]}}{.00053}\, tenure^2
\end{aligned}
$$

$n = 526$, $R^2 = .461$.

Example #2: Housing Prices

Housing Price

	non-robust	robust
bdrms	13.853	13.853
	(9.010)	(8.479)
sqrft	0.123	0.123
	(0.013)**	(0.018)**
lotsz100	0.207	0.207
	(0.064)**	(0.125)
_cons	-21.770	-21.770
	(29.475)	(37.138)
R^2	0.67	0.67
N	88	88

* $p < 0.05$; ** $p < 0.01$

Example #2: log(Housing Prices)

	Log Housing Price	
	non-robust	robust
bdrms	0.037	0.037
	(0.028)	(0.031)
lsqrft	0.700	0.700
	(0.093)**	(0.104)**
llotsize	0.168	0.168
	(0.038)**	(0.041)**
_cons	-1.297	-1.297
	(0.651)*	(0.781)
R^2	0.64	0.64
N	88	88

* $p < 0.05$; ** $p < 0.01$

Heteroskedasticity Robust F-Stats

- There heteroskedasticity-robust counterpart of the usual F-statistic (or at least a transformation of it) known as the heteroskedasticity-robust Wald statistic.
- Most computer package will compute the test statistics for your.
- To understand the test statistic, it helps to first consider the *Wald* test in the context of homoskedasticity.
- Suppose that our null hypotheses take the form $H_0 : \boldsymbol{R\beta} = \boldsymbol{r}$, with the alternative hypothesis $H_1 : H_0$ is not true.
- This form for the null hypothesis allows for all manner of linear restrictions on the parameter space, including exclusion restrictions.
- With large samples, the OLS estimator $\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ will be approximately distributed $\mathcal{N}(0, \sigma^2 \boldsymbol{A}^{-1})$, where $\boldsymbol{A} = E(\boldsymbol{x}_i' \boldsymbol{x}_i)$.
- So, under the null hypothesis $\sqrt{n}(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r}) = \sqrt{n}(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{R\beta})$ is approximately distributed $\mathcal{N}(0, \sigma^2 \boldsymbol{RA}^{-1}\boldsymbol{R}')$

Heteroskedasticity Robust F-Stats (cont'd)

- The Wald test statistic then becomes

$$\left[\sqrt{n}(\boldsymbol{R}\hat{\beta}-\boldsymbol{r})\right]'(\sigma^2\boldsymbol{R}\boldsymbol{A}^{-1}\boldsymbol{R}')^{-1}\left[\sqrt{n}(\boldsymbol{R}\hat{\beta}-\boldsymbol{r})\right]\overset{a}{\sim}\chi^2_q \qquad (13)$$

where q denotes the number of restrictions.
- Note: this is essentially standardizing a series of q normal random variables, squaring them and summing them, which is precisely how we form a χ^2_q random variable.
- In obtaining the heteroskedasticity-robust version of the Wald test, we simply replace the usual OLS standard errors with their robust counterpart (see (12) above).
- The usual Chow test we saw earlier cannot be conducted through separate regressions.

Heteroskedasticity-Robust LM Tests

- Suppose we are interested in testing q exclusion restriction associated with the multiple regression model

$$y=\beta_0+\beta_1x_1+\cdots+\beta_kx_k+u \qquad (14)$$

with $H_0:\beta_1=\beta_2=\cdots=\beta_q=0$.
- The homoskedastic version of the LM test involves
  - running the restricted version of the model to obtain residuals $\tilde{u}$.
  - regressing $\tilde{u}$ on all the independent variables to obtain the resulting $R^2_{\tilde{u}}$.
  - Testing the null hypothesis, using the fact that under the null hypothesis, $LM=nR^2_{\tilde{u}}\sim\chi^2_q$.
- The robust version of this test statistic (detailed on pp. 275-275 of Wooldridge) involves additional steps, but can be implemented using a series of simple OLS regressions.

Testing for Heteroskedasticity

- A reason to test for heteroskedasticity is present, OLS is no longer efficient and one can potential do better by modeling the heteroskedasticity.
- Consider a linear model satisfying MLR.1 through MLR.4, with

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + u \tag{15}$$

- Our null hypothesis will be $H_0 : Var(u|\boldsymbol{x}) = \sigma^2$
- Since we assume $E(u|\boldsymbol{x}) = 0$, this is equivalent to $H_0 : E(u^2|\boldsymbol{x}) = \sigma^2$.
- If we observed u, we might consider allowing u^2 to be a linear function of the independent variables; i.e.,

$$u^2 = \delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k + v \tag{16}$$

...We could then test the hypothesis $\tilde{H}_0 : \delta_j = 0 \ \forall j = 1, \ldots, k$.

Testing for Heteroskedasticity (cont'd)

- If we let $R^2_{\hat{u}^2}$ denote the $R-squared$ from this second regression, then we can form either an F-test or use an LM test, with
  - $F = \frac{R^2_{\hat{u}^2}/k}{(1-R^2_{\hat{u}^2})/(n-k-1)} \overset{a}{\sim} F_{k,n-k-1}$
  - $LM = nR^2_{\hat{u}^2} \overset{a}{\sim} \chi^2_k$, which is known as the **Breausch-Pagan test for heteroskedasticity**.
- In examining the housing market model in examples #2 and #3, Wooldridge (Example 8.4) conducts both the F and LM tests.
  - For the level-level version of the model, the F-test has a p-value of 0.002, while the LM test has a p-value of 0.0028, clearly rejecting homoskedasticity.
  - For the log-log version of the model, the F-test has a p-value of 0.245, while the LM test has a p-value of 0.239 clearly failing to reject homoskedasticity.

Variants on the Heteroskedasticity Test

- There are a number of variants on the homoskedasticity tests above:
  - One can add terms to the second stage regression equation (16), including quadratic terms and interaction terms; e.g., with $k = 2$

$$u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_2 x_1^2 + \delta_4 x_2^2 + \delta_5 x_1 x_2 + v \quad (17)$$

  - If the corresponding LM test is used, this is referred to as the **White test for heteroskedasticity**
  - Alternatively, one can replace the regressors on the right-hand side of (16) with linear and quadratic fitted values for y; i.e.,

$$u^2 = \delta_0 + \delta_1 \hat{y} + \delta_2 \hat{y}^2 + \tilde{v} \quad (18)$$

  testing $H_0 : \delta_1 = 0, \delta_2 = 0$ using either an LM or F statistic.
  ... This approach avoids the proliferation of parameters in (17).
  - One caution: all of these tests assume that we have correctly specified $E(y|\boldsymbol{x})$. A rejection of the null could be reflecting errors in this specification.

Weighted Least Squares

- As suggested above, one reason to test for heteroskedasticity is that, by explicitly (and correctly) modeling the heteroskedasticity, one can obtain more efficient estimator of the parameters.
- Suppose we know that $Var(u|\boldsymbol{x}) = \sigma^2 h(\boldsymbol{x})$.
- The a simple transformation of our original model gets us back to a model that satisfies MLR.1 through MLR.5.
- To see this, consider our standard multiple regression model

$$y_i = \beta_0 x_{i0} + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + u_i \quad (19)$$

where $x_{i0} = 1 \ \forall i$ is use to reflect the constant.

Weighted Least Squares (cont'd)

- If we divide (19) through by $\sqrt{h_i}$ where $h_i = h(\boldsymbol{x}_i)$, then we get the transformed model:

$$\frac{y_i}{\sqrt{h_i}} = \beta_0 \left(\frac{x_{i0}}{\sqrt{h_i}}\right) + \beta_1 \left(\frac{x_{i1}}{\sqrt{h_i}}\right) + \cdots + \beta_k \left(\frac{x_{ik}}{\sqrt{h_i}}\right) + \left(\frac{u_i}{\sqrt{h_i}}\right)$$

or

$$y_i^* = \beta_0 \cdot x_{i0}^* + \beta_1 x_{i1}^* + \cdots + \beta_k x_{ik}^* + u_i^* \tag{20}$$

  where the * simply denotes the original variable divided by $\sqrt{h_i}$.
- This transformed model satisfies the GAUSS-Markov assumptions (MLR.1 to MLR.5).
- In particular: $E(u_i^*|\boldsymbol{x}_i^*) = E(\frac{u_i}{\sqrt{h_i}}|\boldsymbol{x}_i) = \frac{1}{\sqrt{h_i}}E(u_i|\boldsymbol{x}_i) = 0$

  ...and $Var(u_i^*|\boldsymbol{x}_i^*) = Var(\frac{u_i}{\sqrt{h_i}}|\boldsymbol{x}_i) = \frac{1}{h_i}Var(u_i|\boldsymbol{x}_i) = \frac{\sigma^2 h_i}{h_i} = \sigma^2$.

  $\Rightarrow$ The OLS estimator applied to (20) is *BLUE assuming* the variance specification is correct.

Weighted Least Squares (cont'd)

- Note that the model in (20) has no intercept.
- The resulting estimator is called **Weighted Least Squares**, since it weights each observation by the inverse of its standard deviation.
- The estimator falls into a broader class of estimators known as **Generalized Least Squares (GLS) Estimators**.
- Note: The resulting parameter estimates can then be used in original model, since, even with heteroskedasticity, we still have the population regression function

$$E(y_i|\boldsymbol{x}_i) = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} \tag{21}$$

The Housing Price Example

	Housing Price		
	non-robust	robust	WLS robust
bdrms	13.853	13.853	18.005
	(9.010)	(8.479)	(11.860)
sqrft	0.123	0.123	0.157
	(0.013)**	(0.018)**	(0.026)**
lotsz100	0.207	0.207	0.097
	(0.064)**	(0.125)	(0.031)**
_cons	-21.770	-21.770	-87.252
	(29.475)	(37.138)	(55.194)
R^2	0.67	0.67	0.76
N	88	88	88

* $p < 0.05$; ** $p < 0.01$

Log Housing Price Example

	Log Housing Price		
	non-robust	robust	robust
bdrms	0.037	0.037	0.040
	(0.028)	(0.031)	(0.030)
lsqrft	0.700	0.700	0.703
	(0.093)**	(0.104)**	(0.101)**
llotsize	0.168	0.168	0.167
	(0.038)**	(0.041)**	(0.042)**
_cons	-1.297	-1.297	-1.318
	(0.651)*	(0.781)	(0.764)
R^2	0.64	0.64	0.66
N	88	88	88

* $p < 0.05$; ** $p < 0.01$

Applicability

- It is rarely the case that one knows the exact form of the heteroskedasticity.
- However, there are situations in which it is reasonable to hypothesize an approximate form.
- Specifically, when the dependent variable consists of averaged values, it may be reasonable to assume that $Var(y_i) = \sigma^2/m_i$, where m_i denotes the sample size used in constructing y_i.

Estimating the Heteroskedasticity Function

- In the absence of a known heteroskedasticity structure, an alternative approach is to try to hypothesize a structure for $h(\boldsymbol{x})$ and to then estimate it.
- A popular structure is the linear exponential form, with

$$Var(u|\boldsymbol{x}) = \sigma^2 exp(\delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k) \tag{22}$$

- The advantage is that it insures that the variance is positive.
- Aside: The constant in the linear exponential term is redundant, since we can rewrite (22) as

$$\begin{aligned} Var(u|\boldsymbol{x}) &= \sigma^2 exp(\delta_0) exp(\delta_1 x_1 + \cdots + \delta_k x_k) \\ &= exp(\tilde{\delta}_0 + \delta_1 x_1 + \cdots + \delta_k x_k) \\ &\tilde{\sigma}^2 exp(\delta_1 x_1 + \cdots + \delta_k x_k) \end{aligned} \tag{23}$$

where $\tilde{\delta}_0 \equiv ln[\sigma^2 exp(\delta_0)]$ and $\tilde{\sigma}^2 = \sigma^2 exp(\delta_0)$. Either form works.

The Feasible GLS (FGLS) Estimator

- The **Feasible GLS** proceeds by using an estimated variance model and using it in weighted least squares. Specifically,
  1. Obtain residuals $\hat{u}$ from OLS regression of y on $x_1, \ldots, x_k$, including a constant.
  2. Create $ln(\hat{u}^2)$ and regress it on x_1 through x_k (with or without a constant) and obtain fitted values, $\hat{g}$.
  3. Construct $\hat{h} = exp(\hat{g})$.
  4. Estimate original model by WLS, using $\frac{1}{\hat{h}}$ as weights.

     ...Note: WLS with $\frac{1}{\hat{h}}$ as weights is equivalent to dividing y, x_0, x_1 through x_k by $\sqrt{\hat{h}}$ prior to OLS estimation.
- It still will typically make sense to use robust standard errors for this last stage, in case our heteroskedasticity specification is incorrect.

Example: Housing Prices

Housing Price

	non-robust	robust	WLS robust	FGLS robust
bdrms	13.853	13.853	18.005	7.013
	(9.010)	(8.479)	(11.860)	(20.703)
sqrft	0.123	0.123	0.157	0.167
	(0.013)**	(0.018)**	(0.026)**	(0.032)**
lotsz100	0.207	0.207	0.097	0.130
	(0.064)**	(0.125)	(0.031)**	(0.046)**
_cons	-21.770	-21.770	-87.252	-82.471
	(29.475)	(37.138)	(55.194)	(66.252)
R^2	0.67	0.67	0.76	0.76
N	88	88	88	88

* $p < 0.05$; ** $p < 0.01$

Housing Price - Variance Est.

	ln(Var). Est.
bdrms	0.3846
	(0.2292)
sqrft	0.0005
	(0.0004)
lotsz100	0.0034
	(0.0008)**
_cons	3.9007
	(0.9157)**
R^2	0.09
N	88

* $p < 0.05$; ** $p < 0.01$

Note: With only 88 observations, *FGLS* is asking a lot of the data.

Example #2: WLS uses enroll100 (Problem C12)

Math10

	non-robust	robust	WLS robust	FGLS robust
expend100	0.139	0.139	0.281	0.207
	(0.082)	(0.118)	(0.139)*	(0.145)
totcomp100	0.006	0.006	-0.018	-0.016
	(0.012)	(0.015)	(0.019)	(0.019)
enroll100	-0.019	-0.019	-0.005	-0.002
	(0.020)	(0.021)	(0.018)	(0.030)
lnchprg	-0.300	-0.300	-0.417	-0.318
	(0.037)**	(0.035)**	(0.038)**	(0.046)**
_cons	23.679	23.679	29.056	29.197
	(3.965)**	(4.101)**	(4.671)**	(5.018)**
R^2	0.18	0.18	0.36	0.16
N	408	408	408	408

Example #2: Variance Estimation

	Math10
	Variance Est
expend100	0.038
	(0.020)
totcomp100	0.002
	(0.003)
enroll100	-0.014
	(0.005)**
lnchprg	-0.007
	(0.008)
_cons	1.030
	(0.812)
R^2	0.04
N	408

$p < 0.05$; $p < 0.01$

Final Notes of WLS vs OLS

- If the assumed structure for the heteroskedasticity is incorrect,
  - the relative efficiency of WLS versus OLS is uncertain.
    ... however, WLS will often perform better if there is moderate to strong heteroskedasticity.
  - WLS is still consistent, as long as MLR.1 through MLR.4 still hold.
  - The usual WLS standard errors are incorrect, as will be the accompany test statistics that we might normally use.
    ... However, we can simply use the robust standard errors in this case.
- Since both OLS and WLS are consistent estimators under MLR.1 through MLR.4, if the two yield *substantially* different estimates, this suggests a potential problem with assumptions MLR.1 through MLR.4.

Heteroskedasticity in the Linear Probability Model (LPM)

- Recall that in the context of the LPM, the conditional variance of y is necessarily heteroskedastic, with

$$Var(y|\boldsymbol{x}) = Var(u|\boldsymbol{x}) = p(\boldsymbol{x})[1 - p(\boldsymbol{x})] \tag{24}$$

where

$$p(\boldsymbol{x}) = E(y|\boldsymbol{x}) = \beta_0 + \beta_1 + \cdots + \beta_k x_k. \tag{25}$$

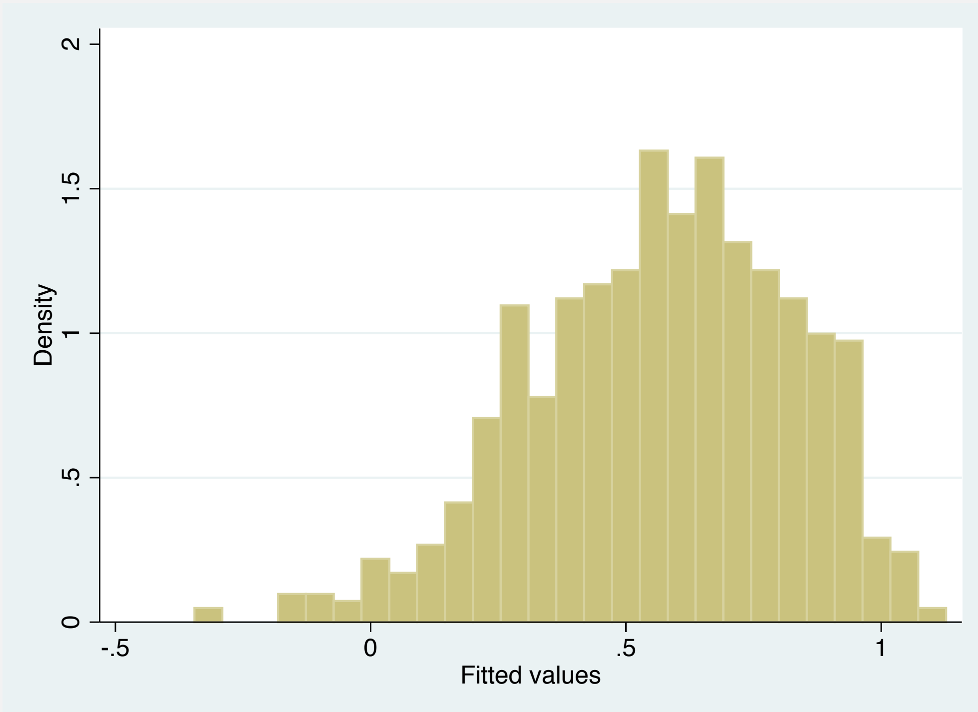
- FGLS in this case will involve
  - estimating the LPM via OLS to obtain fitted values $\hat{y}$
  - construct estimates $\hat{h}_i$ of $Var(y_i|\boldsymbol{x}_i)$.
  - Estimate the LPM using WLS with fitted weights $1/\hat{h}$ and robust standard errors.
- The obvious problem here is that $\hat{h}$ need not lie in the unit interval.
- Wooldridge suggests modifications to *hath* to force lie strictly in the unit interval (say through censoring) and proceeding.

Wooldridge Example 8.8

	inlf		
	non-robust	robust	FGLS robust
nwifeinc	-0.003	-0.003	-0.005
	(0.001)*	(0.002)*	(0.002)**
educ	0.038	0.038	0.045
	(0.007)**	(0.007)**	(0.009)**
exper	0.039	0.039	0.042
	(0.006)**	(0.006)**	(0.006)**
expersq	-0.001	-0.001	-0.001
	(0.000)**	(0.000)**	(0.000)**
age	-0.016	-0.016	-0.017
	(0.002)**	(0.002)**	(0.003)**
kidslt6	-0.262	-0.262	-0.284
	(0.034)**	(0.032)**	(0.042)**
kidsge6	0.013	0.013	0.016
	(0.013)	(0.014)	(0.017)
_cons	0.586	0.586	0.542
	(0.154)**	(0.152)**	(0.182)**
R^2	0.26	0.26	0.17
N	753	753	753

* $p < 0.05$, ** $p < 0.01$

Wooldridge Example 8.8 - $\hat{h}$



Example #2 Loan Application

	approve		
	non-robust	robust	FGLS robust
hispan	-0.142	-0.142	-0.162
	(0.031)**	(0.040)**	(0.042)**
black	-0.213	-0.213	-0.218
	(0.024)**	(0.034)**	(0.034)**
hrat	0.002	0.002	0.001
	(0.001)	(0.002)	(0.002)
obrat	-0.008	-0.008	-0.011
	(0.001)**	(0.001)**	(0.002)**
appinc1000	-0.241	-0.241	-0.271
	(0.085)**	(0.114)*	(0.135)*
married	0.042	0.042	0.056
	(0.015)**	(0.016)**	(0.022)*
_cons	1.093	1.093	1.243
	(0.035)**	(0.041)**	(0.067)**
R^2	0.08	0.08	0.11
N	1,986	1,986	1,986

* $p < 0.05$; ** $p < 0.01$

Example #2 Loan Application