

Effect of Heteroscedasticity on Ordinary Least Squares : Lecture XVIII

Charles B. Moss¹

¹University of Florida

February 27, 2012

1 Readings

2 Heteroscedasticity and Autocorrelation

- GLS and Feasible Generalized Least Squares
- Example - Production Economics
- Effect on Ordinary Least Squares
- Theorem 9.1 - Finite Sample Properties
- Theorem 9.2 - Consistency of OLS

3 Generalized Least Squares

- Transformed Regressions
- Theorem 9.5 - Properties of the GLS Estimator
- Testing Generalized Least Squares

4 Readings

Readings

- * Greene, W.H. 2012. *Econometric Analysis* Seventh Edition (Chapter 9 Section 9.1 9.4 pp. 257 274)
- * Hansen, B.E. 2011. *Econometrics* Draft Textbook Website: <http://www.ssc.wisc.edu/bhansen>. (Chapter 9.1)
- * Morgan, M.S. 1992. *The History of Econometric Ideas*. Cambridge University Press. (Chapter 8: Haavelmos Probability Model pp. 229-258).
- * Theil, H. 1971. *Principles of Econometrics*. John Wiley & Sons. (Chapter 6 Sections 6.1 6.3 pp. 236-250).

Heteroscedasticity and Autocorrelation

- A basic proposition in the Gauss-Markov proof was homoscedasticity (or independently and identically distributed).
- If we generalize

$$\begin{aligned}y &= X\beta + \epsilon \\ E[\epsilon|X] &= 0 \\ E[\epsilon\epsilon'|X] &= \sigma^2\Omega = \Sigma\end{aligned}\tag{1}$$

where Ω is positive definite.

- Under homoscedasticity $\Omega = \sigma^2 I$.

Heteroscedasticity and Autocorrelation - Continued

- Taking heteroscedasticity first we assume that each individual observation has a different variance

$$\Omega = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} \quad (2)$$

- Empirical applications:
 - A traditional example of this form of nonspherical errors is the errors of a production function when the randomness is a function of the level of one of the inputs. In the Just and Pope production function an input such as nitrogen may make production more risky.
 - In systems of equation, it is unlikely that the each equation has the same variance.

Heteroscedasticity and Autocorrelation - Continued

- Autocorrelation occurs when the error in one period is correlated with the error in the previous period

$$\epsilon_t = \rho\epsilon_{t-1} + \nu_t \Rightarrow \sigma_t^2 = \rho^2\sigma_{t-1}^2 + \omega_t^2 \quad (3)$$

if $E[\omega_t^2] = E[\omega_s^2] = \sigma_\omega^2$ for all t, s

$$\sigma^2 \rightarrow \frac{\sigma_\omega^2}{1 - \rho^2} \quad (4)$$

if $\rho \ll 1$.

Heteroscedasticity and Autocorrelation - Continued

- The general form of the variance matrix in this case (first-order autocorrelation) becomes

$$\Omega = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^N \\ \rho & 1 & \rho & \dots & \rho^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^N & \rho^{N-1} & \rho^{N-2} & \dots & 1 \end{bmatrix} \quad (5)$$

- Any time series specification is typically suspect of at least first-order autocorrelation.

GLS and Feasible Generalized Least Squares

- The storyline over the next three lectures is that if heteroscedasticity or autocorrelation occurs then Ordinary Least Squares is still unbiased, but inefficient.
- Generalized Least Squares is unbiased and efficient.
 - In fact since Ordinary Least Squares is a subset of Generalized Least Squares, the question would be why do we ever worry about OLS?
 - The point is that if we know the Ω matrix we would always incorporate this information.
 - But since we don't know the information, we are always dealing with Feasible Generalized Least Squares (FGLS).
- The point is that we don't know the Ω matrix – it must be estimated.
- Since estimation implies error, **FGLS does not always dominate OLS.**

Example - Production Economics

- For example, take the standard quadratic production function for corn

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \frac{1}{2} (A_{11} x_1 x_1 + 2A_{12} x_1 x_2 + 2A_{13} x_1 x_3 + A_{22} x_2 x_2 + 2A_{23} x_2 x_3 + A_{33} x_3 x_3) \quad (6)$$

- First estimating the original model using ordinary least squares yields

Production Economics - Continued

Parameter	Estimate	Std. Dev.	t-Ratio	p-value
α_0	93.44560	14.01369	6.66817	0.00000
α_1	0.03240	0.13397	0.24181	0.80892
α_2	0.24683	0.27178	0.90820	0.36377
α_3	0.00252	0.22487	0.01122	0.99105
A_{11}	0.00066	0.00078	0.84936	0.39568
A_{12}	-0.00194	0.00191	-1.01553	0.30986
A_{13}	0.00120	0.00140	0.85769	0.39106
A_{22}	0.00147	0.00194	0.76073	0.44682
A_{23}	-0.00030	0.00100	-0.29673	0.76667
A_{33}	-0.00066	0.00180	-0.36610	0.71429

Production Economics - Continued

- Following the general outline of Just and Pope, we generate the estimated squared residuals and regress these estimates on the independent variables

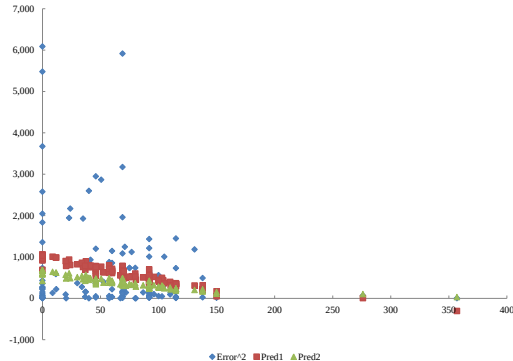
$$\hat{\epsilon}_i^2 = (y_i - X_i\hat{\alpha})^2 = \gamma_0 + \gamma_1 x_{1i} + \gamma_2 x_{2i} + \gamma_3 x_{3i} \quad (7)$$

- yielding the estimated values

Table: Estimated Effect on Residual Squared

Parameter	Estimate	Std. Dev.	t-Ratio	p-Value
γ_0	1061.300	260.420	4.075	0.000
γ_1	-0.159	1.548	-0.103	0.918
γ_2	-3.717	1.998	-1.860	0.065
γ_3	-1.705	1.648	-1.035	0.303

Production Economics - Continued



Production Economics - Continued

- An alternative regression

$$\widehat{(\epsilon_i^2)}^{1/2} = \left[(y_i - X_i \hat{\alpha})^2 \right]^{1/2} = \gamma_0 + \gamma_1 x_{1i} + \gamma_2 x_{2i} + \gamma_3 x_{3i} \quad (8)$$

Table: Estimated Effect on Residual Squared

Parameter	Estimate	Std. Dev.	t-Ratio	p-Value
γ_0	25.814	3.951	6.533	0.000
γ_1	0.001	0.023	0.046	0.964
γ_2	-0.058	0.030	-1.906	0.059
γ_3	-0.028	0.025	-1.100	0.273

Effect on Ordinary Least Squares

- Intuitively, the ordinary least squares estimator is still unbiased, but inefficient.
- Also, the estimate of the variance matrix is biased.
- To demonstrate the unbiasedness of the estimator, we return to the ordinary least squares result

$$\hat{\beta} = \beta + (X'X)^{-1} X'\epsilon \quad (9)$$

as long as the X is fixed in repeated sample $E[\hat{\beta} - \beta] = 0$.
Further, even if the sample is not fixed in repeated sample as long as $E[\epsilon|X] = 0$ we have

$$E[\hat{\beta}] = E_X[E[\hat{\beta}|X]] = \beta \quad (10)$$

Theorem 9.1 - Finite Sample Properties

If the regressors and disturbances are uncorrelated, then the unbiasedness of least squares is unaffected by violations of homoscedasticity. The least squares estimator is unbiased in the generalized regression model. With nonstochastic regressors, or conditional on X , the sampling variance of the least squares estimator is

$$\begin{aligned}
 V[\hat{\beta}|X] &= E\left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'|X\right] \\
 &= (X'X)^{-1} X' \epsilon \epsilon' X (X'X)^{-1} \\
 &= (X'X)^{-1} X' (\sigma^2 \Omega) X (X'X)^{-1} \\
 &= \frac{\sigma^2}{n} \left(\frac{1}{n} X'X\right)^{-1} \left(\frac{1}{n} X' \Omega X\right) \left(\frac{1}{n} X'X\right)^{-1}
 \end{aligned} \tag{11}$$

Theorem 9.1 - Finite Sample Properties – Continued

If the regressors are stochastic, then the unconditional variance is $E_X \left[\text{Var} \left[\hat{\beta} | X \right] \right]$. In Equation 9, $\hat{\beta}$ is a linear function of ϵ . Therefore, if ϵ is normally distributed, then

$$\hat{\beta} | X \sim N \left(\beta, \sigma^2 (X'X)^{-1} (X'\Omega X) (X'X)^{-1} \right) \quad (12)$$

Theorem 9.2 - Consistency of Ordinary Least Squares

If $Q = \text{plim}(X'X/n)$ and $\text{plim}(X'QX/n)$ are both finite positive definite matrices then $\hat{\beta}$ is consistent for β . Under the assumed conditions,

$$\text{plim}\hat{\beta} = \beta \quad (13)$$

1. The smallest eigenvalue of $X'X$ increases without bound as $n \rightarrow \infty$, which implies that $\text{plim}(X'X)^{-1} = 0$.
2. The largest eigenvalue of Ω is finite for all n .

Generalized Least Squares

- The basic generalized least squares estimator is then defined as

$$\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}y \quad (14)$$

where Ω is the **known** heteroscedasticity or autocorrelation.

- The variance for the estimated coefficients is then

$$V_{GLS}(\beta) = \sigma^2 (X'\Omega^{-1}X)^{-1} \quad (15)$$

- Note that in both equations the variance term in $\sigma^2\Omega$ cancels out of the expression.

Transformed Regression

- We can reformulate the generalized least squares problem by considering a decomposition of the Ω matrix

$$\Omega = C\Lambda C' \quad (16)$$

which could be any trianglization of the matrix (Cholesky decomposition, single valued decomposition, etc.).

- Using this decomposition, we could derive $P' = C\Lambda^{-1/2}$ which could be used to transform the regression to

$$\begin{aligned} Py &= PX\beta + P\epsilon \\ y_* &= X_*\beta + \epsilon_* \end{aligned} \quad (17)$$

Transformed Regression

- Thus,

$$\begin{aligned}\hat{\beta} &= (X'_*X_*)^{-1} X'_*y_* \\ &= (X'P'PX)^{-1} X'P'Py \\ &= (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}y\end{aligned}\tag{18}$$

Theorem 9.5 - Properties of the Generalized Least Squares Estimator

If $E[\epsilon_*|X_*] = 0$, then

$$\begin{aligned} E[\hat{\beta}|X_*] &= E\left[(X_*'X_*)^{-1} X_*'y_*|X_*\right] \\ &= \beta + E\left[(X_*'X_*)^{-1} X_*'\epsilon_*|X_*\right] = \beta \end{aligned} \quad (19)$$

Then GLS estimator $\hat{\beta}$ is unbiased. The result is equivalent to $E[P\epsilon|PX] = 0$, but because P is a matrix of known constants, we return to the familiar requirement $E[\epsilon|X] = 0$. The requirement that the regressors and disturbances be uncorrelated.

Theorem 9.5 - Continued

The GLS estimator is consistent if $\text{plim}(1/n)X'_*X_* = Q_*$, where Q_* is a finite positive definite matrix. Making the substitution, we see that this implies

$$\text{plim} [(1/n)X'\Omega^{-1}X]^{-1} = Q_*^{-1}. \quad (20)$$

We require the transformed data $X_* = PX$, not the original data X , to be well behaved. Under the assumption of heteroscedasticity the following hold:

The GLS estimator is asymptotically normally distributed, with mean β and sampling variance

$$V[\hat{\beta}|X_*] = \sigma^2 (X'_*X_*)^{-1} = \sigma^2 (X'\Omega^{-1}X). \quad (21)$$

Theorem 9.5 - Continued

The GLS estimator $\hat{\beta}$ is the minimum variance linear unbiased estimator in the generalized regression model. This statement follows by applying the Gauss-Markov theorem to the transformed model. The result in Theorem 9.5 is **Aitken's Theorem** (1935), and $\hat{\beta}$ is sometimes called the Aitken estimator. This broad result includes the Gauss-Markov theorem as a special case when $\Omega = I$.

Testing Generalized Least Squares

- The F -Test under Generalized Least Squares becomes

$$F[J, n - K] = \frac{(R\hat{\beta} - r)' [R\hat{\sigma}^2 (X_*' X_*)^{-1} R']^{-1} (R\hat{\beta} - r)}{\frac{(\hat{\epsilon}'_c \hat{\epsilon}_c - \hat{\epsilon}' \hat{\epsilon}) / J}{\hat{\sigma}^2}} \quad (22)$$

where J is the number of restrictions and the residual vector is defined as

$$\hat{\epsilon} = y_* - X_* \hat{\beta}. \quad (23)$$

Testing Generalized Least Squares - Continued

- The estimated variance is defined as

$$\hat{\sigma}^2 = \frac{\hat{\epsilon}'\hat{\epsilon}}{n-K} = \frac{(y - X\hat{\beta})' \Omega^{-1} (y - X\hat{\beta})}{n-K} \quad (24)$$

- Note that since $X'_*X_* = X'\omega^{-1}X$ the F -Test in Equation 22 can be rewritten as

$$F[J, n-K] = \frac{(R\hat{\beta} - r)' [R\hat{\sigma}^2 (X'\Omega^{-1}X)^{-1} R']^{-1} (R\hat{\beta} - r)}{J} \quad (25)$$

Readings - Next Time

- * Cameron, A. C. and P. K. Trivedi. 2010. *Microeconometrics Using Stata Revised Edition*. Stata Press. (Chapter 5).
- * Greene, W.H. 2012. *Econometric Analysis* Seventh Edition (Chapter 9 Section 9.1 9.4 pp. 257 274)
- * Hansen, B.E. 2011. *Econometrics* Draft Textbook Website: <http://www.ssc.wisc.edu/bhansen>. (Chapter 9.2)