Heteroscedasticity

Econ 671

Purdue University

Heteroscedasticity

- In this lecture we discuss methods for dealing with *heteroscedasticity* - that is, the case where the variance of each error term is different, and potentially depends on x.

- We begin by discussing consequences for the OLS estimator, and then move on to discuss other estimation approaches for the general problem where the covariance matrix is unknown.

Heteroscedasticity

Dealing with the heteroscedasticity problem involves substituting

-

with

-

That is, each variance of the error vector ϵ is potentially different. We continue to assume, however, that the errors are uncorrelated so that the off-diagonals of Ω are zero.

Heteroscedasticity

The first question to address is, perhaps, reasonably obvious:

What happens when we apply our OLS estimator $\hat{\beta} = (X'X)^{-1}X'y$ when heteroscedasticity is present?

Is the OLS estimator still unbiased? Note:

-

Heteroscedasticity

One can also show that the OLS estimator is consistent in general. So what's the problem? OLS is unbiased and consistent, so why does heteroscedasticity get any press at all?
Note

-

so that

-

When $\Omega = \sigma^2 I_n$, this reduces to the "traditional" formula: $\sigma^2 E_x \left[(X'X)^{-1} \right]$. Thus,

-

Heteroscedasticity

Any ideas how to fix this problem?

It would seem as if replacing Ω with an estimate, say

$$\hat{\Omega} \equiv diag\{\hat{\epsilon}_i^2\}$$

is pointless since Ω contains as many unknown parameters as observations!

Any thoughts on why this might work despite this problem?

- White (1980, *Econometrica*) notes all that is required to consistently estimate the variance of the asymptotic distribution is a consistent estimate of the $k \times k$ matrix $plim \left[\frac{1}{n} X'\Omega X \right]$, for which (under a few additional assumptions)

  -

Heteroscedasticity

Concerning the asymptotics, note:

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

which implies

$$\sqrt{n}(\hat{\beta} - \beta) = \left(\frac{X'X}{n}\right)^{-1} \frac{1}{\sqrt{n}} X'\epsilon.$$

For the second term above we can show, similar to the OLS derivation:

-

where, for the case of heteroscedasticity,

-

Heteroscedasticity

Putting all these pieces together, then, we obtain:

-

The limiting covariance matrix is not known since: (1) it is defined in terms of plims and (2) σ_i^2 are not known. We can estimate this matrix by replacing the plim terms with their sample counterparts, as well as replacing σ_i^2 with $\hat{\sigma}_i^2$. This gives:

-

[Note that it is incorrect to write that the above is consistent; rather, we can consistently estimate the covariance matrix appearing in the right-hand side of $\sqrt{n}(\hat{\beta} - \beta)$ above.]

Heteroscedasticity

So, let's review: In the presence of heteroscedasticity, *of unknown form*, the preceding derivations suggest the following prescription:

1. Run OLS, since it remains unbiased and consistent.

2. Get the fitted residual vector $\hat{\epsilon} = y - X\hat{\beta}$ and, from it, form

$$\hat{\Omega} = diag\{\hat{\epsilon}_i^2\}_{i=1}^n.$$

3. Calculate the heteroscedasticity-robust standard errors to correct any inferences you make using the OLS estimator in the presence of heteroscedasticity.

Heteroscedasticity

Moving beyond the pure heteroscedasticity case, we might ask the following, related question:

Suppose that $E(\epsilon\epsilon'|X) = \Omega$ is known. Even though OLS retains some desirable properties, and we can "fix" our inferences in the presence of heteroscedasticity,
Is it possible to do *better* than OLS?

That is, can we come up with some other estimator (that is both unbiased and consistent) yet *more efficient* than OLS?

The short answer is "yes" and a theorem which speaks to this result is known as *Aitken's Theorem*. This theorem states that an estimator, known as *generalized least squares* is, in fact, more efficient than OLS.

Theorem

In the linear regression model

$$y = X\beta + \epsilon, \quad E(\epsilon\epsilon'|X) = \Omega$$

with Ω known, the generalized least squares estimator:

-

is the best linear unbiased estimator of β (i.e., it is efficient among this class of estimators).

Proof.

Since Ω is positive definite and symmetric, it follows from earlier lectures that we can write

-

where $CC' = C'C = I_n$ and Λ is diagonal with positive elements.
Define Λ^* as the diagonal matrix whose elements are the reciprocal square roots of the elements of Λ. Finally, define

-

noting that H is invertible. Then, note

-

Proof.

The last step follows since, for example

-

Likewise, we can show $\Omega(C\Lambda^{-1}C') = I_n$.

Now, let us go back to our regression model:

$$y = X\beta + \epsilon, \quad E(\epsilon\epsilon'|X) = \Omega$$

and premultiply everything by H:

-

$\square$

Proof.

We represent this transformed regression model as

-

where $\tilde{z}$ is just notation to reflect that the variable z has been premultiplied by H. Note that

-

So, the transformed model meets all the requirements of the Gauss-Markov theorem, and thus we can apply it. □

Proof.

Specifically, we can state that

-

is the minimum variance linear unbiased estimator of β.
We continue by noting that

-

Since H is invertible, the argument reverses and we can show that the class of linear estimators in the (X, y) space is equivalent to those in the $(\tilde{X}, \tilde{y})$ space. It follows that the GLS estimator is, in fact, the most efficient linear unbiased estimator of β in the case of general (but known) covariance matrix. $\qquad\square$