Tema 8. Heteroscedasticidad

Gustavo A. García

ggarci24@eafit.edu.co

Econometría para la Toma de Decisiones

Maestría en Economía Aplicada

Escuela de Finanzas, Economía y Gobierno

Universidad EAFIT

Link slides en formato html

Link slides en formato PDF

En este tema

- Generalidades
- Tipos frecuentes de heteroscedasticidad
- Criterios para detectar heteroscedasticidad
- Estimación por Mínimos Cuadrados Ponderados (MCP)
- La función de heteroscedasticidad debe ser estimada: MCG factibles
- Inferencia robusta a la heteroscedasticidad
- Ejercicio aplicado en R

Lecturas

- Wooldridge, Jeffrey (2013). *Introducción a la econometría*. 5a edición, Cengage Learning. Cap. 8

- Gujarati, D. y Porter, D. (2010). *Econometría*. 5a edición, Mc Graw Hill. Cap. 11

Generalidades

- El modelo de regresión visto hasta ahora descansa sobre varios supuestos:
  - la forma funcional del modelo es adecuada
  - todas las variables explicativas relevantes son incluidas
  - las variables explicativas son exogenas
  - los supuestos sobre el termino de perturbación se cumplen

- Nos cuestionamos ahora sobre estos últimos supuestos, en particular

Qué pasa si el error de la varianza no es constante para cada observación?

Se puede detectar una varianza cambiante y si es así, es necesario ajustar el procedimiento de estimación si tal detección ocurre?

Generalidades

Covarianza del error no escalar

Se ha trabajado con el siguiente modelo:

$$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \ldots + \beta_k X_{ik} + u_i$$

o, en notación matricial

$$\mathbf{Y} = \mathbf{XB} + \mathbf{u}$$

Donde los términos de error u_i se han asumido a ser variables aleatorias incorrelacionadas, cada una con media cero e idéntica (constante) varianza σ_u^2. En notación matricial los supuestos son:

$$E(\mathbf{u}) = 0$$

$$Cov(\mathbf{u}) = E(\mathbf{uu}') = \sigma_u^2 \mathbf{I}$$

$$\mathbf{u} \sim N(0, \sigma_u^2 \mathbf{I})$$

La pregunta es entonces bajo qué circunstancias el supuesto de perturbaciones esféricas es violado: autocorrelación y heteroscedasticidad

Generalidades

Covarianza del error no escalar

Analicemos las circunstancias económicas que pueden llevar a la heteroscedasticidad

- Los datos de sección cruzada involucran datos sobre unidades económicas de diferentes tamaños: los hogares y las firmas tienen diferentes tamaños
- Entre más grande sea la firma o el hogar más difícil es explicar las variaciones en $\mathbf{Y}$ dadas las variaciones en $\mathbf{X}$
- Firmas u hogares grandes son más probables a ser más diversos y flexibles respecto a la forma en la cual los valores de $\mathbf{Y}$ son determinados
- Lo que implica es que para el modelo $\mathbf{Y} = \mathbf{XB} + \mathbf{u}$ entre más grande sea el tamaño de la unidad económica analizada, más grande será el error y así la proporción de variación en $\mathbf{Y}$ atribuida a $\mathbf{u}$ será más grande
- Esto lleva a que la varianza del término de error sea más grande, cuando el tamaño de la unidad económica sea más grande $\Longrightarrow$ heteroscedasticidad
- La heteroscedasticidad no es exclusiva de datos de sección cruzada, con datos de series de tiempo también aparece. Por ejemplo, cuando existen choques externos o cambios en las circunstancias que hacen que la parte "explicable" (o debido a las $\mathbf{X}$) de $\mathbf{Y}$ sea más pequeña

Generalidades

Propiedades de los MCO bajo heteroscedasticidad

Supongamos que la matriz de covarianzas del error es de la forma:

$$Cov(\mathbf{u}) = E(\mathbf{uu}') = \sigma_u^2 \mathbf{\Omega} = \mathbf{\Sigma}$$

En términos matriciales la heteroscedasticidad se expresa como:

$$Cov(\mathbf{u}) = \begin{pmatrix} \sigma_{u1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{u2}^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_{un}^2 \end{pmatrix}$$

Si los errores son homoscedasticos y no correlacionados, entonces se sabe que $\mathbf{\Sigma} = \sigma_u^2 \mathbf{I}$

La cuestión entonces es: Cuáles son las propiedades de los estimadores MCO $\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ cuando se viola el supuesto de perturbaciones esféricas?

- Insesgamiento: $E(\hat{\mathbf{B}}) = E((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}) = E(\mathbf{B} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}) = \mathbf{B}$
- Matriz de covarianzas de $\hat{\mathbf{B}}$

$$\begin{aligned} Cov(\hat{\mathbf{B}}) &= E[(\hat{\mathbf{B}} - E(\hat{\mathbf{B}}))(\hat{\mathbf{B}} - E(\hat{\mathbf{B}}))'] \\ &= E[(\hat{\mathbf{B}} - \mathbf{B})(\hat{\mathbf{B}} - \mathbf{B})'] \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{uu}')\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma_u^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\Omega\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \neq \sigma_u^2(\mathbf{X}'\mathbf{X})^{-1} \end{aligned}$$

Se invalida la inferencia estadística y no se esta usando el estimador más eficiente, el de mínima varianza

Generalidades

Propiedades de los MCO bajo heteroscedasticidad

En resumen cuando no se cumple el supuesto de perturbaciones esféricas:

- los estimadores MCO son aún insesgados, pero estos no son eficientes
- los errores estándar computados para los estimadores MCO no son los apropiados, y por tanto los intervalos de confianza y las pruebas de hipótesis que utilizan estos errores estándar pueden ser erróneos

Criterios para detectar heteroscedasticidad

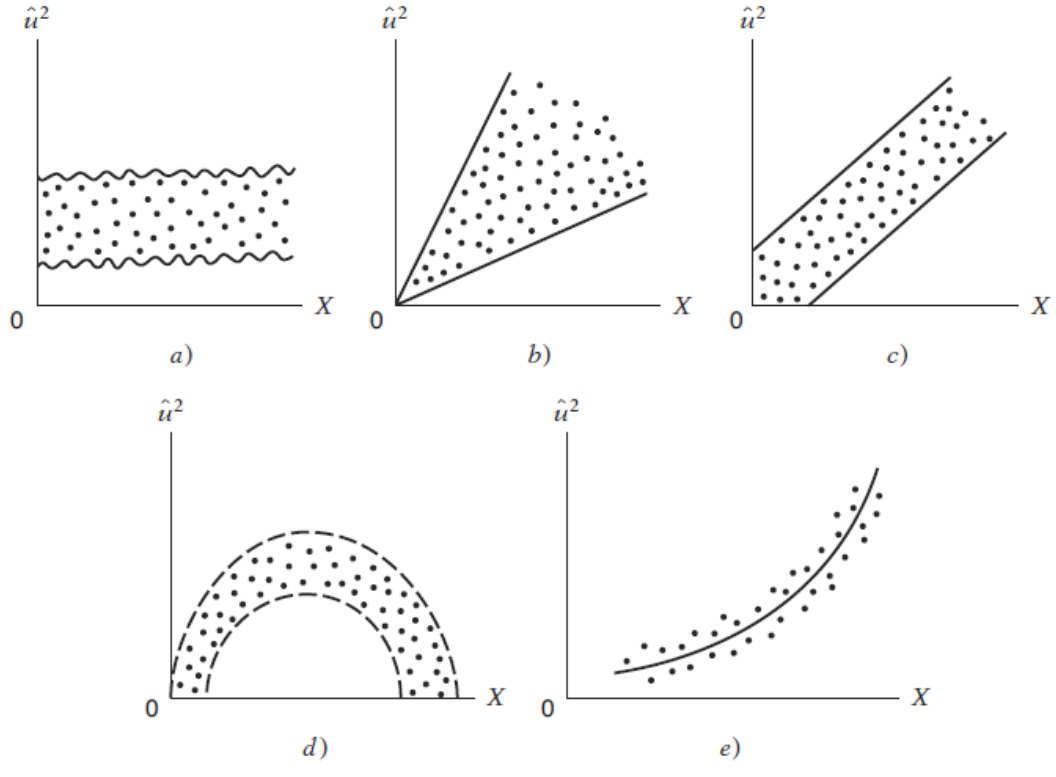
- ¿Para qué probar la existencia de heteroscedasticidad? Existen dos razones:
  - los t estadísticos tienen una distribución exacta a menos que se pruebe la existencia de heteroscedasticidad
  - si la heteroscedasticidad se presenta, los estimadores MCO no serán los mejores estimadores lineales insesgados

- Existen varios tests para corroborar la existencia de heteroscedasticidad. Aquí nos vamos a centrar en los tests más modernos, los cuales detectan el tipo de heteroscedasticidad que invalida la inferencia:
  - método gráfico (informal)
  - el test de Breusch-Pagan
  - el test de White

Criterios para detectar heteroscedasticidad

Método gráfico (informal)

Consiste en hacer un gráfico de dispersión entre los residuos $(\widehat{u}_i, \widehat{u}_i^2, |\widehat{u}_i|)$ y cada una de las variables explicativas (X_j) o Y o $\widehat{Y}$ y observar si hay algún tipo de relación o comportamiento sistemático



Criterios para detectar heteroscedasticidad

El test de Breusch-Pagan

1. Estime el modelo de Y sobre las Xs y obtenga los residuales al cuadrado, $\widehat{u}_i^2$
2. Corra los $\widehat{u}_i^2$ contra las Xs y salve el R^2 de esta regresión, $R^2_{\widehat{u}^2}$
3. Compute el estadístico F o LM de significancia conjunta de las Xs:

$$F = \frac{R^2_{\widehat{u}^2}/k}{(1 - R^2_{\widehat{u}^2})/(n-k-1)}$$

donde k es el número de regresores en la regresión de $\widehat{u}_i^2$ sobre las Xs

$$LM = nR^2_{\widehat{u}^2}$$

bajo H_0, LM se distribuye asintóticamente como una χ^2_k

Si se supone que la heteroscedasticidad sólo surge como consecuencia de algunas variables independientes, el test BP puede fácilmente ser modificado corriendo $\widehat{u}_i^2$ sobre los regresores que se supone generan el problema

Criterios para detectar heteroscedasticidad

El test de White

- White (1980) propone un test de heteroscedasticidad que adiciona los cuadrados y productos cruzados de todas las Xs en el punto dos del test BP. Para un modelo con $k = 3$ sería:

$$\widehat{u}_i^2 = \delta_1 + \delta_2 X_{i2} + \delta_3 X_{i3} + \delta_4 X_{i4} + \delta_5 X_{i2}^2 + \delta_6 X_{i3}^2 + \delta_7 X_{i2} X_{i3} + \delta_8 X_{i2} X_{i4} + \delta_9 X_{i3} X_{i4} + error$$

- El test de White para heteroscedasticidad es el estadístico F o LM para probar que todas la δ_j son cero, excepto el intercepto

- La abundancia de regresores es una debilidad del test de White: usa muchos grados de libertad para modelos con varias variables independientes

- un test más fácil de implementar y más conservador en los grados de libertad, es usar los valores MCO estimados en el test, estos es:

$$\widehat{Y}_i = \widehat{\beta}_1 + \widehat{\beta}_2 X_{i2} + \widehat{\beta}_3 X_{i3} + \ldots + \widehat{\beta}_k X_{ik}$$

  y estimar

$$\widehat{u}_i^2 = \delta_1 + \delta_2 \widehat{Y}_i + \delta_3 \widehat{Y}_i^2 + error$$

Luego se calcula el estadístico F o LM para corroborar la significancia conjunta de los δ_j

Estimación por Mínimos Cuadrados Ponderados (MCP)

Si la forma de la heteroscedasticidad es especificada (como una función de las variables explicatorias), entonces los MCP son más eficientes que los MCO

Sea $\mathbf{X}$ todas las variables explicatorias en el modelo regresión y se asume que

$$Var(u_i) = \sigma_u^2 h(\mathbf{X}) = \sigma_u^2 h_i$$

donde h_i es alguna función de las variables explicativas que determinan la heteroscedasticidad

La pregunta, entonces, es cómo podemos usar la información de la anterior ecuación para estimar los β_j? La idea es tomar el modelo original, el cual contiene heteroscedasticidad, y transformarlo en un modelo que tenga errores homocedásticos

Ya que h_i es una función de $\mathbf{X}_i$, $u_i/\sqrt{h_i}$ tiene un valor esperado de cero, además ya que $Var(u_i) = E(u_i^2) = \sigma_u^2 h_i$, la varianza de $u_i/\sqrt{h_i}$ es σ_u^2:

$$Var(u_i/\sqrt{h_i}) = E((u_i/\sqrt{h_i})^2) = E(u_i^2)/h_i = \sigma_u^2 h_i/h_i = \sigma_u^2$$

Con lo anterior, entonces, se tiene que se parte de un modelo con errores heteroscedásticos

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_k X_{ki} + u_i$$

el cual se divide por $\sqrt{h_i}$ para llegar a un modelo con errores homoscedásticos

$$Y_i/\sqrt{h_i} = \beta_1/\sqrt{h_i} + \beta_2(X_{2i}/\sqrt{h_i}) + \beta_3(X_{3i}/\sqrt{h_i}) + \ldots + \beta_k(X_{ki}/\sqrt{h_i}) + u_i/\sqrt{h_i}$$

$$Y_i^* = \beta_1 X_{1i}^* + \beta_2 X_{2i}^* + \beta_3 X_{3i}^* + \ldots + \beta_k X_{ki}^* + u_i^*$$

donde $X_{1i}^* = 1/\sqrt{h_i}$

Estimación por Mínimos Cuadrados Ponderados (MCP)

- La anterior ecuación parece algo peculiar, pero lo que es importante recordar es que se dedujo con el objeto de que se pudiera obtener estimadores de β_j que tuvieran mejores propiedades de eficiencia que los de MCO
- Cada parámetro de pendiente en β_j multiplica una nueva variable que rara vez tiene una interpretación útil
- Esto no deberá causar ningún problema si se recuerda que, para interpretar los parámetros del modelo, siempre se vuelve a la ecuación original
- Los parámetros del modelo modificado se estiman por MCO, dada las atractivas características que tiene. Estos estimadores, $\beta_1^*, \beta_2^*, \ldots, \beta_k^*$, serán diferentes de los estimadores MCO de la ecuación original
- Los β_j^* son ejemplos de estimadores de Mínimos Cuadrados Generalizados (MCG)
- En este caso los estimadores de MCG se utilizan para considerar la heteroscedasticidad de los errores

La función de heteroscedasticidad debe ser estimada: MCG factibles

- En la mayoría de los casos la forma exacta de la heteroscedasticidad no es obvia $\Longrightarrow$ es difícil encontrar la función h_i
- Sin embargo, en muchos casos puede modelarse la función h_i y utilizar los datos para estimar los parámetros desconocidos del modelo
- Esto da como resultado una estimación para cada h_i, que se denota como $\hat{h}_i$
- Usando $\hat{h}_i$ en lugar de h_i en la transformación de MCG, se obtiene un estimador llamado estimador de MCG factible
- Hay muchas maneras de modelar la heteroscedasticidad, pero aquí se estudiará un método particular, bastante flexible. Suponga que

$$Var(u_i) = \sigma_u^2 exp(\delta_1 + \delta_2 X_{2i} + \delta_3 X_{3i} + \ldots + \delta_k X_{ki})$$

Estos es que $h_i = exp(\delta_1 + \delta_2 X_{2i} + \delta_3 X_{3i} + \ldots + \delta_k X_{ki})$

- Pero por qué se utiliza una función exponencial?
  - Las alternativas lineales son adecuadas cuando se prueba heteroscedasticidad, pero pueden ser problemáticas cuando se trata de corregir la heteroscedasticidad empleando MCP
  - Los modelos lineales no aseguran que los valores predichos sean positivos, y para emplear el método de MCP las varianzas estimadas deben ser positivas

La función de heteroscedasticidad debe ser estimada: MCG factibles

Bajo el anterior estructura de la heteroscedasticidad, se puede escribir lo siguiente

$$u_i^2 = \sigma_u^2 exp(\delta_1 + \delta_2 X_{2i} + \delta_3 X_{3i} + \ldots + \delta_k X_{ki}) v$$

donde v tiene media igual a la unidad y también se asume que es independiente de las Xs. La anterior ecuación puede escribirse de la siguiente forma:

$$log(u_i^2) = a_1 + \delta_2 X_{2i} + \delta_3 X_{3i} + \ldots + \delta_k X_{ki} + e$$

donde e tiene media cero y es independiente de X

Como es usual, se debe reemplazar u por los residuales MCO. Por tanto, la idea es estimar la regresión:

$$log(\hat{u}_i^2) \text{ contra } X_2, X_3, \ldots, X_k$$

De hecho, lo que se necesita de esta regresión son los valores estimados, llamemos a estos $\hat{g}_i$, y luego la estimación de h_i es simplemente $\hat{h}_i = exp(\hat{g}_i)$. Luego se aplica MCP con ponderador a $1/\hat{h}_i$

La función de heteroscedasticidad debe ser estimada: MCG factibles

Procedimiento de los MCG factibles para corregir heteroscedasticidad

1. Estime la regresión de Y sobre $X_2, X_3, \ldots, X_k$ y obtenga los residuales, $\hat{u}_i$
2. Cree $log(\hat{u}_i^2)$
3. Estime el modelo de $log(\hat{u}_i^2)$ contra $X_2, X_3, \ldots, X_k$ y obtenga los valores estimados, $\hat{g}_i$
4. Calcule $\hat{h}_i = exp(\hat{g}_i)$
5. Estime el modelo original por MCP utilizando $1/\hat{h}_i$ como ponderador

Otra alternativa útil para estimar h_i es reemplazar las variables independientes Xs en el punto 3 por los valores estimados por MCO y su cuadrado. Entonces, la idea es obtener $\hat{g}_i$ como los valores estimados de la regresión

$$log(\hat{u}_i^2) \text{ contra } \hat{Y}_i, \hat{Y}_i^2$$

Inferencia robusta a la heteroscedasticidad

- Los estimadores MCO son aún útiles en presencia de heteroscedasticidad $\Longrightarrow$ los errores estándar pueden ser ajustados de tal forma que sean válidos ante la presencia de heteroscedasticidad de forma desconocida
- Esto es muy conveniente ya que implica que se pueden reportar nuevos estadísticos adecuados independientemente de heteroscedasticidad presente en la población
- Este método es conocido como estimadores robustos a la heteroscedasticidad $\Longrightarrow$ hay robustez a la heteroscedasticidad independientemente de la forma de ésta
- La aplicación de métodos robustos a la heteroscedasticidad son muy fáciles ya que muchos sofwares estadísticos incluyen esta corrección

Inferencia robusta a la heteroscedasticidad

En este punto, nos podemos hacer la siguiente pregunta:

> Si los errores estándar robustos a la heteroscedasticidad son validos más a menudo que los usuales errores estándar por MCO, por qué calculamos entonces estos últimos?

Los usuales errores estándar MCO son calculados ya que si el supuesto de homocedasticidad se mantiene y los errores son normalmente distribuidos, entonces los usuales t estadísticos tienen una exacta distribución t, independientemente del tamaño de la muestra

Los errores estándar robustos y los estadísticos t robustos son justificados sólo cuando el tamaño de la muestra crece. Con muestra pequeñas, los t estadísticos robustos puede tener una distribución que no es muy cercana a la t

En muestras grandes, se puede siempre reportar sólo los errores estándar robustos, lo cual es bastante convencional en los trabajos aplicados que utilizan datos de sección cruzada

En R, en el la función coeftest(modelo.ols, vcov = vcovHC(modelo.ols, "HC1")) del paquete sandwich

Ejercicio aplicado en R

En este ejercicio se usa información del Deparamento de Educación de California para el 200 sobre 400 escuelas de primaria. Se tiene información sobre medidas del rendimiento académico de la escuela, así como otros atributos de las escuelas primarias, como el tamaño de la clase, la inscripción, la pobreza, etc.

La idea es analizar los determinantes del desempeño de las escuelas y se estima la siguiente ecuación:

$$api00_i = \beta_1 + \beta_2 meals_i + \beta_3 ell_i + \beta_4 emer_i + u_i$$

donde $api00_i$ es un índice de desempeño académico de la escuela i, que oscila entre 200 y 1000, y entre más alto indica mejor desempeño; $meals_i$ es el porcentage de estudiantes recibiendo comida gratis; ell_i porcentage de estudiantes aprendiendo inglés; $emer_i$ porcentage de profesores con credenciales de profesores emergentes.

En el siguiente link se encuentra el código en R utilizado:

- Código en R