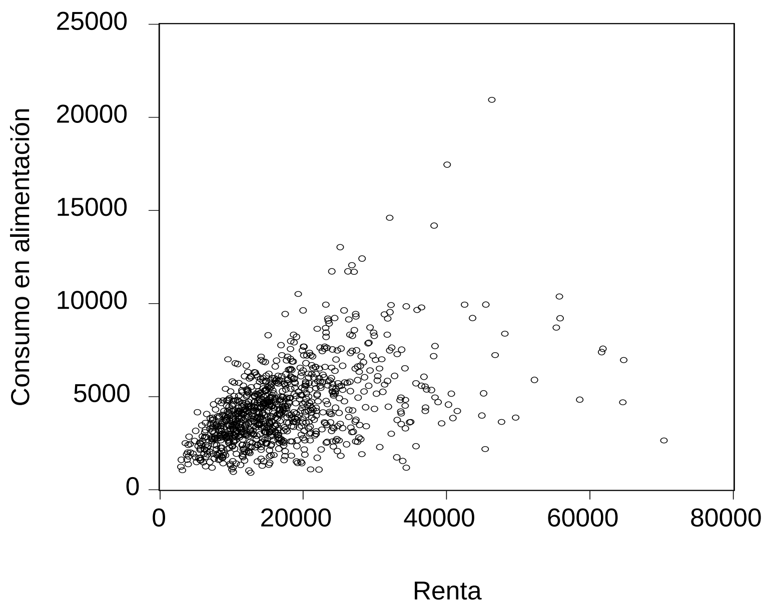
TEMA 3: HETEROCEDASTICIDAD

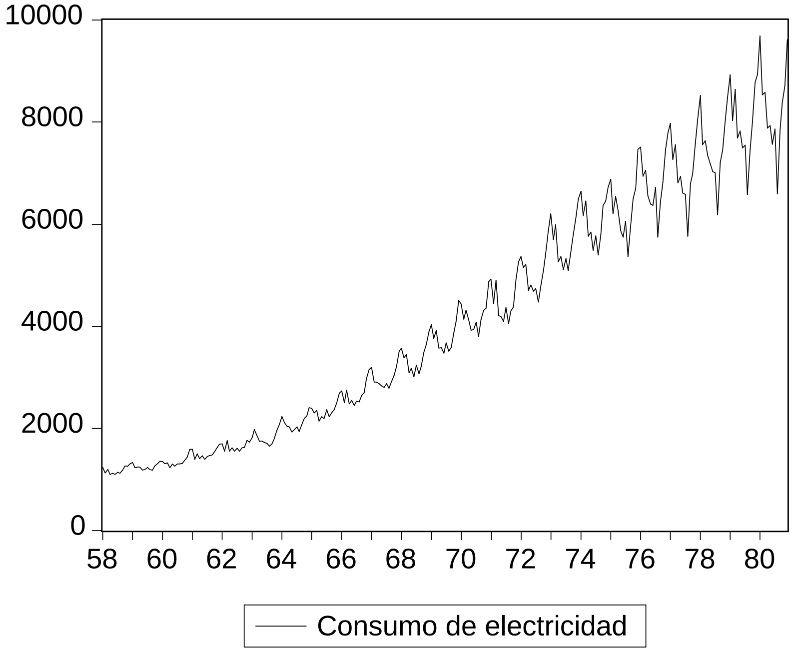
Wooldridge Capítulo 8

Gujarati: Capítulo 13

- **El análisis de muchos fenómenos económicos exige relajar el supuesto de HOMOCEDASTICIDAD que se hace en el modelo clásico de regresión.**

- **Tanto en el contexto de datos de sección cruzada como de datos de series temporales es habitual encontrarse con situaciones en las que existe HETEROCEDASTICIDAD.**





MODELO DE REGRESIÓN LINEAL CON HETEROCEDASTICIDAD

Para simplificar el análisis consideraremos el caso del modelo de regresión simple:

$$Y = \alpha + \beta X + \varepsilon$$

donde:

$$E(\varepsilon | X) = 0$$

$$\rightarrow E(Y|X) = \alpha + \beta X \rightarrow \alpha = E(Y) - \beta E(X) \quad y \quad \beta = \frac{C(XY)}{V(X)}$$

y:

$$V(\varepsilon|X) = V(Y|X) = g(X) \rightarrow$$ **HETEROCEDASTICIDAD**

1. ORIGEN DE LA HETEROCEDASTICIDAD

- **Por la naturaleza de los fenómenos analizados.**

  Ejemplo: En relación Consumo - Renta

$$C = \alpha + \beta R + \varepsilon \ \textbf{ con } \ E(\varepsilon | R) = 0$$

C = Consumo

R =Renta

Es sensato esperar que $V(C|R) = V(\varepsilon|R) = f(R) \neq cte$

Los individuos con mayores niveles de renta tienen más posibilidades de elección

- **Por la agregación de los datos**

  Ejemplo: Consideremos el modelo

$$Y = \alpha + \beta X + \varepsilon \quad con \ E(\varepsilon | X) = 0 \quad \rightarrow E(Y / X) = \alpha + \beta X$$

y además tenemos que $V(\varepsilon|X) = \sigma^2$

Sin embargo se dispone de datos medios por subgrupos, por ejemplo de edad, de la población

$$Y_j = \alpha + \beta X_j + \varepsilon_j \quad con \ Y_j = \frac{\sum_1^{n_j} Y_i}{n_j} \quad X_j = \frac{\sum_1^{n_j} X_i}{n_j} \quad \varepsilon_j = \frac{\sum_1^{n_j} \varepsilon_i}{n_j}$$

Dado $E(\varepsilon|X) = 0 \quad \rightarrow \quad E(\varepsilon_j|X_j) = 0$

Pero dado $V(\varepsilon|X) = \sigma^2 \quad \rightarrow \quad V(\varepsilon_j|X_j) = \sigma^2 / n_j \quad \rightarrow$ **HETEROCEDASTICIDAD**

- **La existencia de heterocedasticidad puede ser un síntoma de existencia de un error de especificación (forma funcional errónea, omisión de variables relevantes, existencia de parámetros cambiantes, etc.) → Es muy importante formular bien los modelos.**

Ejemplo: Forma funcional errónea

<u>Modelo correcto</u>

$$Y = \alpha + \beta X + \gamma Z + \varepsilon \quad con \quad E(\varepsilon | X, Z) = 0 \quad \rightarrow \quad E(Y / X, Z) = \alpha + \beta X + \gamma Z$$

***y* además tenemos que** $V(\varepsilon | X, Z) = \sigma^2$

<u>Modelo incorrecto</u>

$$Y = \alpha + \beta X + U \rightarrow \quad E(Y / X) = \alpha + \beta X + E(U / X) \quad con \quad E(U|X) \neq 0 \quad y \quad V(U|X) = f(Z)$$

2. PROPIEDADES DEL ESTIMADOR MCO

- **¿Qué propiedades tendrá el estimador de MCO en este contexto? Centrémonos en el análisis de la estimación de β.**

  Suponiendo que tenemos una muestra aleatoria:

- **Insesgadez.**

Sabemos que: $\hat{\beta} = \sum_i c_i Y_i \quad con \quad c_i = \dfrac{x_i}{\sum_i x_i^2} \qquad x_i = X_i - \bar{X}$

Dado que $\sum_i c_i = 0 \quad y \quad \sum_i c_i X_i = 1$ **, tenemos que:**

$$E(\hat{\beta} | X_i) = E(\sum_i c_i Y_i | X_i) = \sum_i c_i E(Y_i | X_i) = \sum_i c_i (\alpha + \beta X_i) = \beta$$

● Consistencia

Sabemos que $\hat{\beta} = \dfrac{\sum_i x_i y_i}{\sum_i x_i^2} = \dfrac{\sum_i x_i y_i / n}{\sum_i x_i^2 / n}$**, con** $x_i = X_i - \overline{X} \quad y_i = Y_i - \overline{Y}$**, luego**

$$p\lim \hat{\beta} = p\lim b = \frac{p\lim\left(\sum_i x_i y_i / n \right)}{p\lim\left(\sum_i x_i^2 / n \right)} = \frac{C(XY)}{V(X)} = \beta \Rightarrow Consistencia$$

● Expresión de la varianza

$$V(\hat{\beta}|X_i) = V(\sum_i c_i Y_i | X_i) = \sum_i c_i^2 V\left(Y_i | X_i\right) = \sum_i c_i^2 \sigma_i^2 = \frac{\sum_i x_i^2 \sigma_i^2}{\left(\sum_i x_i^2\right)^2}$$

Nótese que esta expresión no es la habitual → La construcción de intervalos de confianza y la contrastación de hipótesis que se realizan a partir de la expresión $V(\hat{\beta}|X_i) = \sigma_{\hat{\beta}}^2 = \dfrac{\sigma^2}{\sum_i x_i}$ **no son válidas. La importancia del "sesgo" dependerá de la magnitud de la Heterocedasticiad.**

● **Eficiencia**

MCO no es en este contexto el Estimador Lineal Insesgado de Mínima Varianza

→ Existe un estimador alternativo que es el Estimador Lineal e Insesgado de Mínima Varianza denominado el estimador de Mínimos Cuadrados Ponderados (MCP) o Mínimos Cuadrados Generalizados (MCG).

● **Muy importante:**

Todos estos resultados son ampliables al contexto de regresión lineal múltiple

3. Inferencia Robusta a la heterocedasticidad: Propuesta de White o de Eiker-White

● **En cualquier caso y a partir de resultados asintóticos, puede llevarse a cabo la construcción "aproximada" de intervalos de confianza y la contrastación de hipótesis, empleando la propuesta de estimación para** $\sigma^2_{\hat{\beta}} = \left(\dfrac{\sum_i x_i^2 \sigma_i^2}{\left(\sum_i x_i^2 \right)^2} \right)$ **de White o Eiker-White.**

● **La propuesta consiste en calcular**

$$\hat{\sigma}^2_{\hat{\beta}} = \frac{\sum_i x_i^2 \hat{\varepsilon}_i^2}{\left(\sum_i x_i^2\right)^2}$$

En donde $\hat{\varepsilon}_i$ **son los residuos de MCO del modelo. Es decir:**

$$\hat{\varepsilon}_i = e_i = Y_i - \hat{Y}_i = Y_i - \left(\hat{\alpha} + \hat{\beta} X_i\right)$$

● **Se puede demostrar que:**

$$\frac{\hat{\beta} - \beta}{\sqrt{\sum_i x_i^2 \hat{\varepsilon}_i^2 \Big/ \left(\sum_i x_i^2\right)^2}} \underset{a}{\sim} N(0,1)$$

es decir

$$\frac{\hat{\beta} - \beta}{s(\hat{\beta})} \underset{a}{\sim} N(0,1)$$

● **Se puede emplear este resultado para construir intervalos de confianza y realizar contrastes de hipótesis como se hace habitualmente.**

Ejemplo

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

***Y:* Gasto en alimentos**

***X:* Renta**

Dependent Variable: Y
Method: Least Squares
Sample: 1 965
Included observations: 965

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2472.729	125.0474	19.77433	0.0000
X	0.110691	0.006365	17.39069	0.0000

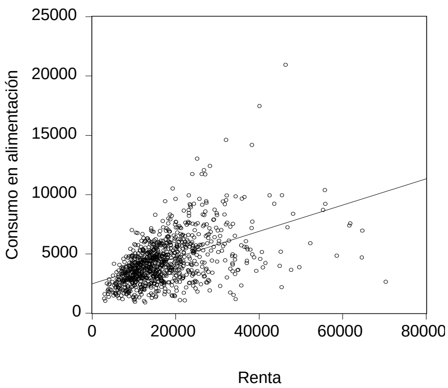
R-squared	0.238998	Mean dependent var	4396.906
Adjusted R-squared	0.238207	S.D. dependent var	2073.714
S.E. of regression	1809.953	Akaike info criterion	17.84206
Sum squared resid	3.15E+09	Schwarz criterion	17.85216
Log likelihood	-8606.794	F-statistic	302.4361
Durbin-Watson stat	1.960853	Prob(F-statistic)	0.000000

INTERVALO NO ROBUSTO A LA HETEROCEDASTICIDAD

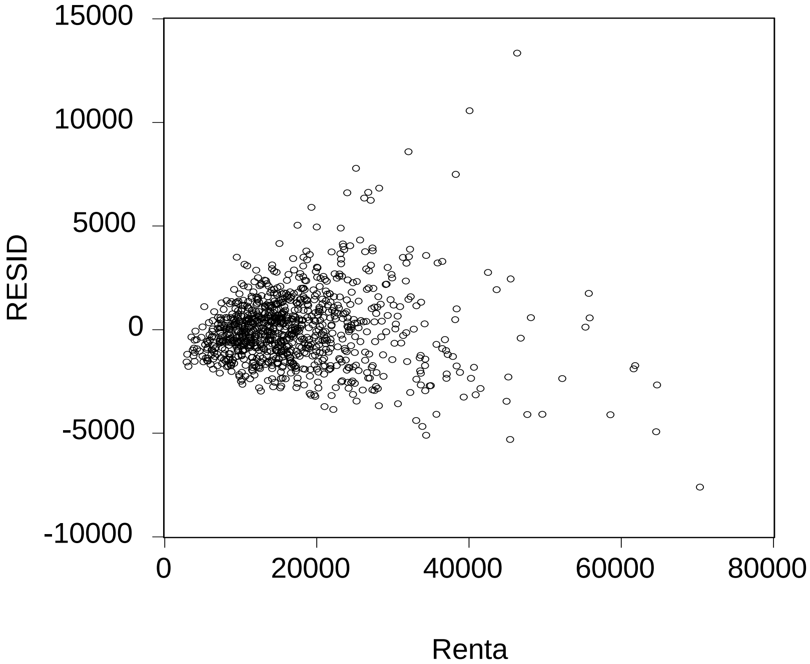
$$\hat{\beta} \pm 1{,}96 \times s(\hat{\beta})$$

$$0{,}11 \pm 1{,}96 \times 0{,}006 \Rightarrow \left[0{,}10\,;\,0{,}12\right]$$

MODELO AJUSTADO



RESIDUO VERSUS EL NIVEL DE RENTA



CORRECCIÓN POR HETEROCEDASTICIDAD

Dependent Variable: Y
Method: Least Squares
Sample: 1 965
Included observations: 965
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2472.729	167.5569	14.75755	0.0000
X	0.110691	0.011098	9.974151	0.0000

R-squared	0.238998	Mean dependent var	4396.906
Adjusted R-squared	0.238207	S.D. dependent var	2073.714
S.E. of regression	1809.953	Akaike info criterion	17.84206
Sum squared resid	3.15E+09	Schwarz criterion	17.85216
Log likelihood	-8606.794	F-statistic	302.4361
Durbin-Watson stat	1.960853	Prob(F-statistic)	0.000000

INTERVALO ROBUSTO

$$\hat{\beta} \pm 1{,}96 \times s(\hat{\beta})$$

$$0{,}11 \pm 1{,}96 \times 0{,}011 \Rightarrow [0{,}09;\ 0{,}13]$$

4. PROCEDIMIENTOS DE DETECCIÓN

- **Los métodos gráficos ya ilustrados**

- **Contrastes formales**

Hay múltiples propuestos en la literatura → Glejser, Harvey, Spearman, Breusch-Pagan, White

- **Vamos a considerar uno muy general y ampliamente utilizado → Contraste de White**

CONTRASTE DE WHITE

- **Sin pérdida de generalidad consideremos el modelo:**

$$Y = \alpha + \beta X_1 + \gamma X_2 + \varepsilon$$

en el que deseamos contrastar si $V(\varepsilon \mid X_1, X_2) = \sigma^2$

1º) Se estima el modelo por MCO y obtenemos $\hat{\varepsilon}^2$

2º) Se estima por MCO el modelo (niveles, cuadrados y productos cruzados)

$$\hat{\varepsilon}^2 = \delta_0 + \delta_1 X_1 + \delta_2 X_2 + \delta_3 X_1^2 + \delta_4 X_2^2 + \delta_5 X_1 X_2 + w$$

y se obtiene el coeficiente de determinación $R^2_{\hat{\varepsilon}^2}$

3º) Se contrasta

$H_0\!:\delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = 0\ (HOMOCEDASTICIDAD)$ **Si la hipótesis nula es cierta**

$$n \times R^2_{\hat{\varepsilon}^2} \underset{asy}{\sim} \chi^2_5 \qquad F = \frac{R^2_{\hat{\varepsilon}^2}/5}{\left(1 - R^2_{\hat{\varepsilon}^2}\right)/n-6} \underset{asy}{\sim} F_{5,n-6}$$

En general habrá k regresores y q parámetros δ

EJEMPLO

$Y_i = \alpha + \beta X_i + \varepsilon_i$

Y: **Gasto en alimentos**

X: **Renta**

White Heteroskedasticity Test:

F-statistic	69.96223	Probability	0.000000
Obs*R-squared	122.5375	Probability	0.000000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Sample: 1 965
Included observations: 965

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-995169.9	1009591.	-0.985716	0.3245
X	165.7302	89.31492	1.855571	0.0638
X^2	0.003584	0.001662	2.156254	0.0313

R-squared	0.126982	Mean dependent var	3269140.
Adjusted R-squared	0.125167	S.D. dependent var	9076724.
S.E. of regression	8489688.	Akaike info criterion	34.74971
Sum squared resid	6.93E+16	Schwarz criterion	34.76485
Log likelihood	-16763.73	F-statistic	69.96223
Durbin-Watson stat	2.028781	Prob(F-statistic)	0.000000

$$n \times R^2_{\hat{\varepsilon}^2} = 965 \times 0{,}126982 = 122{,}53 \rightarrow (\chi^2_2) \textbf{ Rechazo } \mathbf{H_0}$$

$$F = \frac{0{,}126982 / 2}{(1 - 0{,}126982) / 965 - 3} = 69{,}96 \rightarrow (F_{2,962}) \textbf{ Rechazo } \mathbf{H_0}$$

5. ESTIMADOR DE MÍNIMOS CUADRADOS PONDERADOS

- **Consideremos el modelo**

$$Y_i = \alpha + \beta X_i + \varepsilon_i \qquad i = 1 \ldots n$$

con

$$E(\varepsilon_i | X_i) = 0$$

$$V(Y_i | X_i) = V(\varepsilon_i | X_i) = g(X_i) = \sigma_i^2$$

- **Consideremos el modelo**

$$\frac{Y_i}{\sigma_i} = \alpha \frac{1}{\sigma_i} + \beta \frac{X_i}{\sigma_i} + \frac{\varepsilon_i}{\sigma_i} \quad \rightarrow \quad Y_i^* = \alpha X_{0i}^* + \beta X_{1i}^* + \varepsilon_i^*$$

Nótese:

1º) Los parámetros de este modelo (α y β) son los mismos que los del modelo original.

2º)

$$E(\varepsilon_i^*|X_i) = E\left(\frac{\varepsilon_i}{\sigma_i}|X_i\right) = \frac{1}{\sigma_i}E(\varepsilon_i|X_i) = 0$$

$$V(\varepsilon_i^*|X_i) = V\left(\frac{\varepsilon_i}{\sigma_i}|X_i\right) = \frac{1}{\sigma_i^2}V(\varepsilon_i|X_i) = 1 \Rightarrow Homocedasticidad$$

3º) El estimador MCO de este modelo transformado será el Estimador Lineal Insesgado de Mínima Varianza. Este estimador es conocido como estimador de Mínimos Cuadrados Ponderados (MCP) o Mínimos Cuadrados Generalizados.

Ejemplo

$Y_i = \alpha + \beta X_i + \varepsilon_i$

con

$V(\varepsilon_i|X_i) = V(Y_i|X_i) = \sigma_i^2 = \sigma^2 X_i^2$ **con *X_i>0***

Modelo Transformado

$$\frac{Y_i}{\sigma X_i} = \alpha\frac{1}{\sigma X_i} + \beta\frac{X_i}{\sigma X_i} + \frac{\varepsilon_i}{\sigma X_i} \Rightarrow \quad \frac{Y_i}{\sigma X_i} = \alpha\frac{1}{\sigma X_i} + \beta\frac{1}{\sigma} + \frac{\varepsilon_i}{\sigma X_i}$$

con: $$V\left(\frac{\varepsilon_i}{\sigma X_i}|X_i\right) = \frac{1}{\sigma^2 X_i^2}V(\varepsilon_i|X_i) = 1 \Rightarrow Homocedasticidad$$

o

$$\frac{Y_i}{X_i} = \alpha \frac{1}{X_i} + \beta + \frac{\varepsilon_i}{X_i}$$

con: $$V\left(\frac{\varepsilon_i}{X_i} \middle| X_i\right) = \frac{1}{X_i^2} V(\varepsilon_i | X_i) = \sigma^2 \Rightarrow Homocedasticidad$$

<u>Algunas consideraciones sobre el Estimador de Mínimos Cuadrados Ponderados</u>

● **Es fácil comprobar que MCP se obtiene también tras resolver el problema**

$$\underset{\alpha,\beta}{Min} \sum_i \hat{\varepsilon}_i^{*2} = \sum_i \left(\frac{Y_i - \left(\hat{\alpha} + \hat{\beta} X_i\right)}{\sigma_i} \right)^2$$

● **El estimador de MCP es el Estimador Lineal de Mínima Varianza.**

● **La Inferencia estadística se puede realizar sin problemas en el modelo transformado.**

Estimador de Mínimos Cuadrados Ponderados Factibles

- **La obtención del estimador de MCP requiere el conocimiento de σ_i^2 o al menos de ella multiplicada por un factor de proporcionalidad. Pero σ_i^2 no siempre es conocido.→ Debe estimarse, en cuyo caso se habla de Mínimos Cuadrados Ponderados Factibles.**

- **A partir de $\hat{\sigma}_i^2$ estimamos por MCO el modelo**

$$\frac{Y_i}{\hat{\sigma}_i} = \alpha \frac{1}{\hat{\sigma}_i} + \beta \frac{X_i}{\hat{\sigma}_i} + \frac{\varepsilon_i}{\hat{\sigma}_i}$$

Obtenemos así, el Estimador de Mínimos Cuadrados Ponderados Factibles.

Econometría. Tema 3. Pilar Abad

- **Si $\hat{\sigma}_i^2$ es un estimador consistente de σ_i^2 entonces MCP factibles tendrá asintóticamente las mismas propiedades que MCP.**

- **¿Cómo obtengo $\hat{\sigma}_i^2$?**

A partir de una estimación consistente obtenida a partir de los residuos de MCO.

Ejemplo 1

$Y_i = \alpha + \beta X_i + \varepsilon_i$ **con** $V(Y_i / X_i) = \sigma_i^2 = \gamma_0 + \gamma_1 X_i + \gamma_2 X_i^2$

Pasos:

1º) Estimo por MCO el modelo original, y obtengo $\hat{\varepsilon}_i = e_i = Y_i - \left(\hat{\alpha} + \hat{\beta} X_i\right)$

2º) Calculo $\hat{\varepsilon}_i^2 = e_i^2$

3º) Ajusto por MCO un modelo para $\hat{\varepsilon}_i^2 = e_i^2$

$$\hat{e}_i^2 = \hat{\gamma}_0 + \hat{\gamma}_1 X_i + \hat{\gamma}_2 X_i^2$$

4º) Estimo por MCO el modelo transformado con $\hat{e}_i^2 = \hat{\sigma}_i^2$**.**

$$\frac{Y_i}{\hat{\sigma}_i} = \alpha \frac{1}{\hat{\sigma}_i} + \beta \frac{X_i}{\hat{\sigma}_i} + \frac{\varepsilon_i}{\hat{\sigma}_i}$$

Ejemplo 2

$Y_i = \alpha + \beta X_i + \varepsilon_i$ **con** $V(\varepsilon_i / X_i) = V(Y_i / X_i) = \sigma_i^2 = EXP(\gamma_0 + \gamma_1 X_i)$

Pasos:

1º) Estimo por MCO el modelo original, y obtengo $\hat{\varepsilon}_i = Y_i - \left(\hat{\alpha} + \hat{\beta} X_i\right)$

2º) Calculo $\hat{\varepsilon}_i^2$

3º) Ajusto por MCO un modelo para $\log(\hat{\varepsilon}_i^2)$

$$\widehat{\log}(\hat{\varepsilon}_i^2) = \hat{\gamma}_0 + \hat{\gamma}_1 X_i$$

4º) Estimo por MCO el modelo transformado con $\exp(\widehat{\log}(\hat{\varepsilon}_i^2)) = \hat{\sigma}_i^2$

$$\frac{Y_i}{\hat{\sigma}_i} = \alpha \frac{1}{\hat{\sigma}_i} + \beta \frac{X_i}{\hat{\sigma}_i} + \frac{\varepsilon_i}{\hat{\sigma}_i}$$

MCP versus MCO

- **Cuando la varianza condicional está bien especificada, el estimador MCP es más eficiente que MCO.**

- **Cuando la especificación de la varianza condicional es incorrecta el estimador MCP no tiene por qué ser más eficiente que MCO y las inferencias no son válidas.**