Considere a estrutura básica do modelo de regressão:

$$y_{it} = x'_{it}\beta + z'_i\alpha + \varepsilon_{it}$$

Há K regressores em x, não incluindo a constante. A heterogeneidade ou efeito individual é $z_i'\alpha$ onde z_i contém um termo constante e um conjunto de variáveis grupo específicas ou individuais, que podem ser observadas, como raça e sexo, ou não observadas, como habilidades, preferências, características familiares específicas, etc., todas consideradas constantes no tempo. Este é um modelo de regressão clássica. Se z_i é observado para todos os indivíduos, pode-se aplicar MQO.

1. Regressão Pooled – Se z_i contém somente um termo constante, então MQO fornece estimativas consistentes e eficientes de um α comum e de β.

2. <u>Efeitos Fixos</u> – Se z_i não é observável, mas correlacionado com x_{it}, então o estimador de mínimos quadrados de β é tendencioso e inconsistente em consequência da omissão de variável. Entretanto, o modelo,

$$y_{it} = x'_{it}\beta + \alpha_i + \varepsilon_{it} \qquad \text{onde} \quad \alpha_i = z'_i\alpha$$

contém todos os efeitos observáveis e especifica uma média condicional estimável. Este procedimento de efeito fixo considera α_i como termo constante grupo-específico no modelo de regressão

3. Efeito Aleatório – Se a heterogeneidade individual não observável pode ser pressuposta não correlacionada com as variáveis incluídas, então o modelo pode ser formulado como:

$$y_{it} = x'_{it}\beta + \alpha + u_i + \varepsilon_{it}$$

Este procedimento de efeito aleatório especifica u_i como um elemento aleatório grupo-específico, semelhante ao ε_{it} só que para cada grupo.

Modelo de Efeito Fixo

Pressupõe-se nesse modelo que as diferenças entre unidades cross-section podem ser obtidas por diferenças no termo constante. Portanto, considera-se α_i um parâmetro desconhecido a ser estimado.

Se Y_i e X_i são T observações da i-ésima unidade, então

$$Y_i = \iota\alpha_i + X_i\beta + \varepsilon_i$$

Exemplos:

1) Na equação de rendimentos, Y_{it} poderá medir rendimento do chefe do domicílio, enquanto X_{it} poderá conter as variáveis educação, experiência, sexo, local de residência, etc. Observe que α_i não varia com o tempo e representa um efeito individual específico que não foi incluído na regressão.

2) Na função de produção, Y_{it} medirá a produção e X_{it} os insumos. Os efeitos específicos das firmas não observados são capturados por α_i , por exemplo, a habilidade gerencial dos executivos da firma.

$$Y_i = \iota\alpha_i + X_i\beta + \varepsilon_i$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \iota & 0 & \cdots & 0 \\ 0 & \iota & & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \iota \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} + \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}_{(nT\times K)} \beta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

ou

$$Y = \begin{bmatrix} d_1 & d_2 & \cdots & d_n & X \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \varepsilon$$

onde d_i são variáveis binárias indicando a i-ésima unidade.

Considere a matriz:

$$D = \begin{bmatrix} d_1 & d_2 & \cdots & d_n \end{bmatrix}_{(nT \times n)}$$

então,

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

que é conhecido também como modelo de variável binária.

$$Y = \begin{bmatrix} D & X \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \varepsilon$$

Partindo da equação normal obtém-se:

$$\begin{bmatrix} D' \\ X' \end{bmatrix} \begin{bmatrix} D & X \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} D' \\ X' \end{bmatrix} Y$$

$$\begin{bmatrix} D'D & D'X \\ X'D & X'X \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} D'Y \\ X'Y \end{bmatrix}$$

$$D'Da + D'Xb = D'Y \quad \rightarrow a = \left(D'D\right)^{-1} D'Y - \left(D'D\right)^{-1} D'Xb$$

$$X'Da + X'Xb = X'Y$$

$$X'D\left[(D'D)^{-1} D'Y - (D'D)^{-1} D'Xb\right] + X'Xb = X'Y$$

$$X'D(D'D)^{-1} D'Y - X'D(D'D)^{-1} D'Xb + X'Xb = X'Y$$

$$X'Xb - X'D(D'D)^{-1} D'Xb = X'Y - X'D(D'D)^{-1} D'Y$$

$$X'\underbrace{\left[I - D(D'D)^{-1} D'\right]}_{M_d}Xb = X'\underbrace{\left[I - D(D'D)^{-1} D'\right]}_{M_d}Y$$

$$b = \left[X'\left(I - D(D'D)^{-1} D'\right)X\right]^{-1}\left[X'\left(I - D(D'D)^{-1} D'\right)Y\right]$$

Como M_d é uma matriz simétrica e idempotente, podemos escrever.

$$b = \left[X'M_d'M_dX\right]^{-1}\left[X'M_d'M_dY\right]$$

$$b = \left(X_*'X_*\right)^{-1}\left(X_*'Y_*\right)$$

onde

$$X_* = M_dX \qquad \text{e} \qquad Y_* = M_dY$$

$$M_d = \begin{bmatrix} M_0 & 0 & \cdots & 0 \\ 0 & M_0 & & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & M_0 \end{bmatrix} \qquad M_0 = I_T - \frac{1}{T}\iota\iota'$$

Observe que se *n*=2 e *T*=3 a matriz D fica:

$$D = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\left(D'D\right)^{-1} = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix}$$

$$D\left(D'D\right)^{-1} = \begin{bmatrix} 1/3 & 0 \\ 1/3 & 0 \\ 1/3 & 0 \\ 0 & 1/3 \\ 0 & 1/3 \\ 0 & 1/3 \end{bmatrix}$$

$$M_d = I_{nT\times nT} - \underbrace{\begin{bmatrix} 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \end{bmatrix}}_{D(D'D)^{-1}D'}$$

$$M_d X = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_{11} \\ X_{12} \\ X_{13} \\ X_{21} \\ X_{22} \\ X_{23} \end{bmatrix} - \begin{bmatrix} 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} X_{11} \\ X_{12} \\ X_{13} \\ X_{21} \\ X_{22} \\ X_{23} \end{bmatrix}$$

$$M_d X = \begin{bmatrix} X_{11} - \frac{\Sigma X_{1t}}{3} \\ X_{12} - \frac{\Sigma X_{1t}}{3} \\ X_{13} - \frac{\Sigma X_{1t}}{3} \\ X_{21} - \frac{\Sigma X_{2t}}{3} \\ X_{22} - \frac{\Sigma X_{2t}}{3} \\ X_{23} - \frac{\Sigma X_{2t}}{3} \end{bmatrix}$$

$$\iota = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \iota\iota' = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \qquad \sum X_i = \iota' X = T\overline{X}$$

$$\overline{X} = \frac{1}{T} \Sigma X_i = \frac{1}{T} \iota' X \qquad \iota\overline{X} = \iota \frac{1}{T} \iota' X = \frac{1}{T} \iota\iota' X = \begin{bmatrix} \overline{X} \\ \overline{X} \\ \overline{X} \end{bmatrix}$$

$$\begin{bmatrix} X_1 - \overline{X} \\ X_2 - \overline{X} \\ X_3 - \overline{X} \end{bmatrix} = X - \iota\overline{X} = XI - \frac{1}{T}\iota\iota' X = \left[I - \frac{1}{T}\iota\iota' \right] X$$

$$M_d = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} = \begin{bmatrix} M_0 & 0 \\ 0 & M_0 \end{bmatrix}$$

$$M_0 = I_T - \frac{1}{T}\iota\iota'$$

Considere qualquer vetor Z_i $(T\times 1)$, então

$$M_0 Z_i = Z_i - \bar{Z}\imath$$

Portanto, a regressão de $M_d Y$ em função de $M_d X$ é equivalente a regressão de $(Y_{it} - \bar{Y}_{i.})$ em função de $(X_{it} - \bar{X}_{i.})$, onde

$$\bar{Y}_{i.} = \frac{\sum_{t=1}^{T} Y_{it}}{T} \quad \text{e} \quad \bar{X}_{i.} = \frac{\sum_{t=1}^{T} X_{it}}{T}$$

Os coeficientes das variáveis binárias podem ser recuperados da outra equação normal da regressão particionada

$$D'Da + D'Xb = D'Y$$

ou

$$a = (D'D)^{-1} D'(Y - Xb)$$

ou

$$a_i = \bar{Y}_{i.} - b'\bar{X}_{i.}$$

$$Var(b) = s^2 \left(X' M_d X \right)^{-1}$$

$$Var(a) = \frac{s^2}{T} + \bar{X}_{i\cdot}' Var(b) \bar{X}_{i\cdot}$$

$$s^2 = \frac{\sum_i \sum_t \left(Y_{it} - a_i - b' X_{it} \right)^2}{nT - n - K}$$

$$e_{it} = Y_{it} - a_i - b' X_{it}$$

Este procedimento pode ser estendido para efeitos específicos de tempo, isto e',

$$Y_{it} = \alpha_i + \gamma_t + \beta' X_{it} + \varepsilon_{it}$$

As estimativas dos parâmetros β são obtidas da regressão de

$$Y_{*it} = Y_{it} - \overline{Y}_{i\cdot} - \overline{Y}_{\cdot t} + \overline{\overline{Y}}$$

em função de

$$X_{*it} = X_{it} - \overline{X}_{i\cdot} - \overline{X}_{\cdot t} + \overline{\overline{X}}$$

onde

$$\overline{Y}_{\cdot t} = \frac{1}{n}\sum_{i=1}^{n} Y_{it}$$

$$\overline{\overline{Y}} = \frac{1}{nT}\sum_{i}\sum_{t} Y_{it}$$

$$a_i = \left(\overline{Y}_{i\cdot} - \overline{\overline{Y}}\right) - b'\left(\overline{X}_{i\cdot} - \overline{\overline{X}}\right)$$

$$c_t = \left(\overline{Y}_{\cdot t} - \overline{\overline{Y}}\right) - b'\left(\overline{X}_{\cdot t} - \overline{\overline{X}}\right)$$

Teste F

H_0: todos os termos constante são iguais

$$F\left(n-1, nT-n-K\right) = \frac{SQres.restrito - SQres.nãorestrito / n-1}{SQres.nãorestrito / nT-n-K}$$

$$Y_{it} = \alpha + \beta' X_{it} + \varepsilon_{it}$$ (modelo restrito)

$$Y_{it} = \alpha_i + \beta' X_{it} + \varepsilon_{it}$$ (modelo não restrito)

Estimadores entre e dentro de grupos:

1) A formulação original é:

$$Y_{it} = \alpha_i + \beta' X_{it} + \varepsilon_{it}$$

2) em termos dos desvios da média

$$Y_{it} - \overline{Y}_{i\cdot} = \beta'\left(X_{it} - \overline{X}_{i\cdot}\right) + \varepsilon_{it} - \overline{\varepsilon}_{i\cdot}$$

3) em termos da média

$$\overline{Y}_{i\cdot} = \alpha_i + \beta' \overline{X}_{i\cdot} + \overline{\varepsilon}_{i\cdot}$$

	i=1	i=2	i=3
1970	10	10	5
1980	8	7	6
1990	4	5	6
2000	6	5	4

$$\overline{Y}_{\cdot t} = \frac{10+10+5}{3}$$ para t = 1970

$$\overline{Y}_{i\cdot} = \frac{10+8+4+6}{4}$$ para i = 1

$$\overline{\overline{Y}} = \frac{10+8+\ldots+4}{12}$$

$$S_{XX}^{t} = \sum_i \sum_t \left(X_{it} - \overline{\overline{X}}\right)\left(X_{it} - \overline{\overline{X}}\right)'$$

$$S_{XY}^{t} = \sum_i \sum_t \left(X_{it} - \overline{\overline{X}}\right)\left(Y_{it} - \overline{\overline{Y}}\right)$$

Em 2) como os dados estão em desvios, as médias de $\left(Y_{it} - \overline{Y}_{i\cdot}\right)$ e $\left(X_{it} - \overline{X}_{i\cdot}\right)$ são 0.

$$S_{XX}^{w} = \sum_i \sum_t \left(X_{it} - \overline{X}_{i\cdot}\right)\left(X_{it} - \overline{X}_{i\cdot}\right)'$$

$$S_{XY}^{w} = \sum_i \sum_t \left(X_{it} - \overline{X}_{i\cdot}\right)\left(Y_{it} - \overline{Y}_{i\cdot}\right)$$

Em 3)

$$S_{XX}^{b} = \sum_{i} T\left(\overline{X}_{i\cdot} - \overline{\overline{X}}\right)\left(\overline{X}_{i\cdot} - \overline{\overline{X}}\right)'$$

$$S_{XY}^{b} = \sum_{i} T\left(\overline{X}_{i\cdot} - \overline{\overline{X}}\right)\left(\overline{Y}_{i\cdot} - \overline{\overline{Y}}\right)$$

Observe que

$$S_{XX}^{t} = S_{XX}^{b} + S_{XX}^{w} \quad \text{e} \quad S_{XY}^{t} = S_{XY}^{b} + S_{XY}^{w}$$

Portanto, há 3 estimadores possíveis de mínimos quadrados de β

$$b^{t} = \left(S_{XX}^{t}\right)^{-1} S_{XY}^{t} = \left(S_{XX}^{w} + S_{XX}^{b}\right)^{-1}\left(S_{XY}^{w} + S_{XY}^{b}\right)$$

$$b^w = \left(S_{XX}^w\right)^{-1} S_{XY}^w \longrightarrow$$ Estimador de variáveis binárias = within estimator

$$b^b = \left(S_{XX}^b\right)^{-1} S_{XY}^b \longrightarrow$$ Between estimator

como $S_{XY}^w = S_{XX}^w b^w$ e $S_{XY}^b = S_{XX}^b b^b$

substituindo em b^t

$$b^t = \left(S_{XX}^w + S_{XX}^b\right)^{-1}\left(S_{XX}^w b^w + S_{XX}^b b^b\right) = \left(S_{XX}^w + S_{XX}^b\right)^{-1} S_{XX}^w b^w + \left(S_{XX}^w + S_{XX}^b\right)^{-1} S_{XX}^b b^b$$

ou

$$b^t = F^w b^w + F^b b^b \qquad F^w = I - F^b$$

$$F^{w} = \left(S_{XX}^{w} + S_{XX}^{b}\right)^{-1} S_{XX}^{w} = I - \left(S_{XX}^{w} + S_{XX}^{b}\right)^{-1} \underbrace{S_{XX}^{b}}_{S_{XX}^{t} - S_{XX}^{w}}$$

$$= I - \left(\underbrace{S_{XX}^{w} + S_{XX}^{b}}_{S_{XX}^{t}}\right)^{-1} S_{XX}^{t} + \left(S_{XX}^{w} + S_{XX}^{b}\right)^{-1} S_{XX}^{w}$$

$$= I - I + F^{w} = F^{w}$$

Exemplo:

Função custo para produção aérea nos EU.

6 firmas observadas durante 15 anos

$$\ln \text{cost}_{it} = \beta_1 + \beta_2 \ln \text{output}_{it} + \beta_3 \ln \text{fuel price}_{it} + \beta_4 \text{load factor}_{it} + \varepsilon_{it}$$

Output = receita por passageiro por milha

Fuel price = preço do combustível

Load factor = taxa de utilização de capacidade

Modelo de Efeito Aleatório

Este modelo pressupõe que o erro aleatório associado com cada unidade *cross-section* não é correlacionado com outros regressores. Ademais, α_i's são aleatórios, isto é,

$$Y_{it} = \alpha_i + \beta' X_{it} + \varepsilon_{it}$$

onde

$$\alpha_i = \alpha + u_i$$

então,

$$Y_{it} = \alpha + \beta' X_{it} + u_i + \varepsilon_{it}$$

Pressupõe-se que:

$$E(\varepsilon_{it}) = E(u_i) = 0$$

$$E(\varepsilon_{it}^2) = \sigma_\varepsilon^2$$

$$E(u_i^2) = \sigma_u^2$$

$$E(\varepsilon_{it} u_j) = 0$$

$$E(\varepsilon_{it}\varepsilon_{js}) = 0 \qquad \text{se} \qquad t \neq s \quad \text{ou} \quad i \neq j$$

O modelo em blocos de T observações para Y_i X_i $u_i i$ e ε_i fica:

$$w_{it} = \varepsilon_{it} + u_i$$

$$w_i = \left[w_{i1}, w_{i2}, \ldots, w_{iT}\right]'$$

então

$$E\left(w_{it}^2\right) = \sigma_\varepsilon^2 + \sigma_u^2,$$

$$E\left(w_{it}, w_{is}\right) = E(\varepsilon_{it} + u_i, \varepsilon_{is} + u_i) = \sigma_u^2$$

Para as T observações da unidade i,

$$V = E(w_i, w_i') = \begin{bmatrix} \sigma_\varepsilon^2 + \sigma_u^2 & \sigma_u^2 & \cdots & \sigma_u^2 \\ \sigma_u^2 & \sigma_\varepsilon^2 + \sigma_u^2 & & \sigma_u^2 \\ \vdots & & & \\ \sigma_u^2 & \sigma_u^2 & \cdots & \sigma_\varepsilon^2 + \sigma_u^2 \end{bmatrix}$$

$$V = \sigma_\varepsilon^2 I + \sigma_u^2 ii'$$

A matriz de variância e covariância dos erros para as *nT* observações é:

$$\Omega = \begin{bmatrix} V & 0 & \cdots & 0 \\ 0 & V & & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & V \end{bmatrix} = I \otimes V$$

$$\hat{\beta} = \left(X'\Omega^{-1}X\right)^{-1}\left(X'\Omega^{-1}Y\right)$$

$$= \left(X'\left(I \otimes V\right)^{-1}X\right)^{-1}\left(X'\left(I \otimes V\right)^{-1}Y\right)$$

$$= \left(X'\Omega^{-1/2\prime}\Omega^{-1/2}X\right)^{-1}\left(X'\Omega^{-1/2\prime}\Omega^{-1/2}Y\right)$$

onde,

$$\Omega^{-1/2} = I \otimes V^{-1/2}$$

$$V^{-1/2} = I - \frac{\theta}{T} ii'$$

$$\theta = 1 - \frac{\sigma_\varepsilon}{\left(T\sigma_u^2 + \sigma_\varepsilon^2\right)^{1/2}} = 1 - \frac{\sigma_\varepsilon}{\sigma_1}$$

$$V^{-1} = \frac{D}{\sigma_\varepsilon^2} + \frac{ii'}{T\sigma_1^2} \qquad D = I - \frac{ii'}{T}$$

$$\sigma_1^2 = T\sigma_u^2 + \sigma_\varepsilon^2$$

$$VV^{-1} = \left(\sigma_\varepsilon^2 I + \sigma_u^2 ii'\right)\left(\frac{D}{\sigma_\varepsilon^2} + \frac{ii'}{T\sigma_1^2}\right)$$

$$= \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2} D + \frac{\sigma_\varepsilon^2}{T\sigma_1^2} ii' + \frac{\sigma_u^2}{\sigma_\varepsilon^2} Dii' + \frac{\sigma_u^2}{T\sigma_1^2} i\underbrace{i'i}_{T}i'$$

$$= D + \frac{\sigma_\varepsilon^2}{T\sigma_1^2} ii' + \frac{\sigma_u^2}{\sigma_\varepsilon^2} \underbrace{\left(I - \frac{ii'}{T}\right) ii'}_{0} + \frac{\sigma_u^2}{\sigma_1^2} ii'$$

$$= I - \frac{ii'}{T} + \frac{\sigma_\varepsilon^2}{T\sigma_1^2} ii' + \frac{\sigma_u^2}{\sigma_1^2} ii'$$

$$= I + \frac{1}{T\sigma_1^2}\left[-\sigma_1^2 ii' + \sigma_\varepsilon^2 ii' + T\sigma_u^2 ii'\right] = I$$

Lembrando que:

$$\sigma_1^2 = T\sigma_u^2 + \sigma_\varepsilon^2$$

$$V^{-1/2}V^{-1/2} = \sigma_\varepsilon^2 V^{-1} \qquad\qquad \theta = 1 - \frac{\sigma_\varepsilon}{\sigma_1}$$

$$\left(I - \frac{\theta}{T}ii'\right)\left(I - \frac{\theta}{T}ii'\right) = I - \frac{2\theta}{T}ii' + \frac{\theta^2}{T}ii'$$

$$= I - \frac{ii'}{T}\left[2\left(1 - \frac{\sigma_\varepsilon}{\sigma_1}\right) - \left(1 - 2\frac{\sigma_\varepsilon}{\sigma_1} + \frac{\sigma_\varepsilon^2}{\sigma_1^2}\right)\right]$$

$$= I - \frac{ii'}{T}\left[2 - 2\frac{\sigma_\varepsilon}{\sigma_1} - 1 + 2\frac{\sigma_\varepsilon}{\sigma_1} - \frac{\sigma_\varepsilon^2}{\sigma_1^2}\right]$$

$$= I - \frac{ii'}{T}\left(1 - \frac{\sigma_\varepsilon^2}{\sigma_1^2}\right) = \sigma_\varepsilon^2\left[\frac{ii'}{T\sigma_1^2} + \frac{1}{\sigma_\varepsilon^2}\left(I - \frac{ii'}{T}\right)\right]$$

$$V^{-1/2}Y_i = \begin{bmatrix} Y_{i1} - \theta\overline{Y}_{i\cdot} \\ Y_{i2} - \theta\overline{Y}_{i\cdot} \\ \vdots \\ Y_{iT} - \theta\overline{Y}_{i\cdot} \end{bmatrix}$$

MQG é obtido fazendo a regressão de $V^{-1/2}Y_i$ em função de $V^{-1/2}X_i$

Observe a similaridade deste modelo com o de variáveis binárias onde θ =1.

Pode-se demonstrar que o estimador de MQG também é uma matriz de médias ponderadas dos estimadores dentro (within) e entre (between) unidades.

$$\hat{\beta} = \hat{F}^{w} b^{w} + \left(I - \hat{F}^{w}\right) b^{b}$$

onde,

$$\hat{F}^{w} = \left[S_{XX}^{w} + \lambda S_{XX}^{b}\right]^{-1} S_{XX}^{w}$$

$$\lambda = \frac{\sigma_{\varepsilon}^{2}}{\sigma_{\varepsilon}^{2} + T\sigma_{u}^{2}} = \left(1 - \theta\right)^{2}$$

Observe que:

Se λ=1 então MQG=MQO $\Rightarrow \sigma^2_u$=0

Se λ=0 então σ^2_ε=0

Mínimos Quadrados Generalizados quando Ω é desconhecido

$$Y_{it} = \alpha + \beta' X_{it} + \varepsilon_{it} + u_i$$

$$Y_{i\cdot} = \alpha + \beta' X_{i\cdot} + \bar{\varepsilon}_{i\cdot} + u_i$$

$$Y_{it} - \bar{Y}_{i\cdot} = \beta' \left(X_{it} - \bar{X}_{i\cdot} \right) + \varepsilon_{it} - \bar{\varepsilon}_{i\cdot}$$

$$b = \left(X' M_d X \right)^{-1} \left(X' M_d Y \right)$$

$$\hat{\sigma}_{\varepsilon}^{2} = \frac{\sum_{i}\sum_{t}\left(e_{it} - \bar{e}_{i\cdot}\right)^{2}}{nT - n - K}$$

$$\hat{\sigma}_{u}^{2} = \hat{\sigma}_{**}^{2} - \frac{\hat{\sigma}_{\varepsilon}^{2}}{T}$$

$$\hat{\sigma}_{**}^{2} = \frac{e_{**}' e_{**}}{n - k}$$

$$\overline{Y}_{i\cdot} - \alpha - \beta'\overline{X}_{i\cdot} = \bar{\varepsilon}_{i\cdot} + u_{i} = \varepsilon_{**i}$$

$$\theta = 1 - \frac{\sigma_{\varepsilon}}{\left(T\sigma_u^2 + \sigma_{\varepsilon}^2\right)^{1/2}}$$

$$\frac{\hat{\sigma}_{\varepsilon}^2}{\hat{\sigma}_{\varepsilon}^2 + T\hat{\sigma}_u^2} = \left(1 - \hat{\theta}\right)^2$$

$$\hat{\theta} = 1 - \left(\frac{\hat{\sigma}_{\varepsilon}^2}{\hat{\sigma}_{\varepsilon}^2 + T\hat{\sigma}_4^2}\right)^{1/2}$$

Para obter u_i,

$$\hat{u}_i = \frac{\hat{\sigma}_u^2}{T\sigma_u^2 + \sigma_\varepsilon^2} \sum_{t=1}^{T} \left(Y_i - \hat{\hat{\beta}} X_i \right)$$

<u>Teste do multiplicador de Lagrange.</u>

$$H_0 = \sigma_u^2 = 0$$

$$H_1 = \sigma_u^2 \neq 0$$

$$LM = \frac{nT}{2(T-1)} \left[\frac{\sum_i \left(\sum_t e_{it} \right)^2}{\sum_i \sum_t e_{it}^2} - 1 \right]^2 \sim \chi^2(1)$$

Teste de Hausman:

H_0: não existe correlação entre u_i e X.

$$m = \left(b_F - b_A\right)'\left(Var\left(b_F\right) - Var\left(b_A\right)\right)^{-1}\left(b_F - b_A\right) \sim X^2\left(K'\right)$$

b_F = coeficiente efeito fixo

b_A = coeficiente efeito aleatório

K' = número de inclinações

$$\operatorname{cov}\left(\left(b_F - b_A\right), b_A\right) = \operatorname{cov}\left(b_F, b_A\right) - Var\left(b_A\right) = 0$$

$$\operatorname{cov}\left(b_F, b_A\right) = \operatorname{var}\left(b_A\right)$$

$$Var(b_F - b_A) = Var(b_F) + Var(b_A) - 2\operatorname{cov}(b_F, b_A)$$

$$= Var(b_F) - Var(b_A)$$