# ECONOMETRIA II

## Modelos para dados em Painel ou Pooling cross-section time series

$$Y_{it} = \beta_{1i} + \sum_{k=2}^{K} \beta_{ki} X_{kit} + \varepsilon_{it}$$

Todos os coeficientes
são constantes

$$\beta_{ki} = \beta_k$$

heterocedástico
autocorrelacionado
corr. contemporânea

Inclinações constantes
Intercepto varia

$\beta_{1i}$ fixos
Modelo de efeito fixo
ou de variáveis binárias

$\beta_{1i}$ aleatório
modelo de efeito
aleatório ou
componente de erro

Inclinações e
intercepto
variam

$\beta_{ki}$ fixo
SUR

$\beta_{ki}$ aleatório
Modelo de
coeficiente
aleatório de
Swamy

## 1. Modelos para dados que usam *cross-section* e séries temporais

Considere o conjunto de equações:

$$\boldsymbol{Y}_1 = \boldsymbol{X}_1 \boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_1$$
$$\boldsymbol{Y}_2 = \boldsymbol{X}_2 \boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}_2$$
$$\dots$$
$$\boldsymbol{Y}_n = \boldsymbol{X}_n \boldsymbol{\beta}_n + \boldsymbol{\varepsilon}_n$$

Os modelos a serem examinados possuem $n$ unidades *cross-section*, denotados $i$ =1,...,$n$, observados a cada $T$ períodos de tempo, $t$ =1,...,$T$.

Exemplo:

$$I_{it} = \beta_1 + \beta_2 F_{it} + \beta_3 C_{it} + \varepsilon_{it}$$

Dados:

5 firmas ($i$=5):

General Motors (produz automóveis)

Chrysler (produz automóveis)

General Eletric (produz material elétrico)

Westinghouse (produz material elétrico)

US Steel (produz aço)

20 anos ($t$=20):

1935 a 1954

Variáveis:

$I_t$ = investimento bruto

$F_t$ = valor de mercado da firma no final do ano anterior à análise (lucros antecipados)

$C_t$ = valor do estoque de equipamentos no final do ano anterior à análise (investimento necessário para reposição de equipamentos)

Considere o modelo:

$$Y_{it} = \boldsymbol{\beta}'\boldsymbol{X}_{it} + \varepsilon_{it} \qquad i = 1,\dots,n \quad t = 1,\dots,T$$

$$\text{pressupõe-se que } \beta_1 = \beta_2 = \dots = \beta_n$$

Agrupando os *t*'s em vetores obtém-se :

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{\varepsilon}_i \qquad\qquad i = 1, \dots, n$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{(nT\times 1)} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}_{(nT\times K)} \beta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{(nT\times 1)}$$

cada sub-matriz ou sub-vetor tem *T* observações.

Considere a matriz de variância e covariância:

$$\Omega = E(\varepsilon\varepsilon') = \begin{bmatrix} \sigma_{11}V_{11} & \sigma_{12}V_{12} & \cdots & \sigma_{1n}V_{1n} \\ \sigma_{21}V_{21} & \sigma_{22}V_{22} & \cdots & \sigma_{2n}V_{2n} \\ \vdots & & & \\ \sigma_{n1}V_{n1} & \sigma_{n2}V_{n2} & \cdots & \sigma_{nn}V_{nn} \end{bmatrix}_{nT\times nT}$$

O modelo de regressão clássica especifica que:

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

$$Var(\varepsilon_{it}) = \sigma^2$$

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

Então

$$\Omega = \begin{bmatrix} \sigma^2 I & 0 & \cdots & 0 \\ 0 & \sigma^2 I & & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \sigma^2 I \end{bmatrix}$$

Modelo $Y = X\beta + \varepsilon$

Neste caso, aplica-se MQO aos dados agrupados (*nT* observações).

## 2. Heteroscedasticia

É provável que existam diferenças entre variâncias nas unidades cross-section.

Podemos então permitir que $\sigma^2$ varie com $i$, isto é,

$$\Omega = \begin{bmatrix} \sigma_1^2 I & 0 & \cdots & 0 \\ 0 & \sigma_2^2 I & & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \sigma_n^2 I \end{bmatrix}$$

o estimador será:

$$\hat{\beta} = \left(X'\Omega^{-1}X\right)^{-1}\left(X'\Omega^{-1}Y\right) = \left(\Sigma \frac{1}{\sigma_i^2} X_i' X_i\right)^{-1} \Sigma \frac{1}{\sigma_i^2} X_i' Y_i$$

$$X'\Omega^{-1} = \begin{bmatrix} X_1' & X_2' & \cdots & X_n' \end{bmatrix}_{(K\times nT)} \begin{bmatrix} \frac{1}{\sigma_1^2}I & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2^2}I & & 0 \\ \vdots & & \ddots & \\ 0 & & \cdots & \frac{1}{\sigma_n^2}I \end{bmatrix}_{(nT\times nT)} = \begin{bmatrix} \frac{X_1'}{\sigma_1^2} & \frac{X_2'}{\sigma_2^2} & \cdots & \frac{X_n'}{\sigma_n^2} \end{bmatrix}_{(K\times nT)}$$

$$X'\Omega^{-1}X = \begin{bmatrix} \frac{X_1'}{\sigma_1^2} & \frac{X_2'}{\sigma_2^2} & \cdots & \frac{X_n'}{\sigma_n^2} \end{bmatrix}_{(K\times nT)} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}_{(nT\times K)} = \begin{bmatrix} \frac{X_1'X_1}{\sigma_1^2} + \frac{X_2'X_2}{\sigma_2^2} + \cdots + \frac{X_n'X_n}{\sigma_n^2} \end{bmatrix}_{(K\times K)}$$

$$X'\Omega^{-1}Y = \begin{bmatrix} \dfrac{X_1'}{\sigma_1^2} & \dfrac{X_2'}{\sigma_2^2} & \cdots & \dfrac{X_n'}{\sigma_n^2} \end{bmatrix}_{(K\times nT)} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{(nT\times 1)} = \begin{bmatrix} \dfrac{X_1'Y_1}{\sigma_1^2} + \dfrac{X_2'Y_2}{\sigma_2^2} + \cdots + \dfrac{X_n'Y_n}{\sigma_n^2} \end{bmatrix}_{(K\times 1)}$$

$$\left(X'\Omega^{-1}X\right)^{-1}\left(X'\Omega^{-1}Y\right)_{(K\times 1)} = \left[\sum \frac{1}{\sigma_i^2} X_i'X_i\right]^{-1}\left[\sum \frac{X_i'Y_i}{\sigma_i^2}\right]$$

Para estimar $\sigma_i^2$ utiliza-se:

$$s_i^2 = \frac{e_i' e_i}{T}$$

para cada unidade *cross-section* $i$ com base em $b$ (MQO) obtido para as $nT$ observações, onde,

$$e_i = Y_i - bX_i$$

## 3. Teste para detectar heterocedasticidade

Teste do multiplicador de Lagrange:

Este teste é derivado da maximização da função de verossimilhança restrita.

Modelo:

$$Y_{it} = \boldsymbol{\beta}'\boldsymbol{X}_{it} + \varepsilon_{it}$$

Considere a função de densidade da distribuição normal:

$$f = \frac{1}{\sqrt{2\Pi\sigma_i^2}} e^{-\frac{1}{2}\frac{\varepsilon_i'\varepsilon_i}{\sigma_i^2}}$$

então, a função de verossimilhança é:

$$L = \prod_{i=1}^{nT} f = (2\pi)^{-nT/2} \left(\prod_i \sigma_i^2\right)^{-T/2} e^{-\frac{1}{2}\sum_i \frac{\varepsilon_i'\varepsilon_i}{\sigma_i^2}}$$

aplicando logaritmo,

$$\ln L = -\frac{nT}{2}\ln(2\pi) - \frac{T}{2}\sum_{i=1}^{n}\ln\sigma_i^2 - \frac{1}{2}\sum_{i=1}^{n}\frac{\varepsilon_i'\varepsilon_i}{\sigma_i^2}$$

$$\frac{\partial \ln L}{\partial \beta} = \sum_i \frac{1}{\sigma_i^2} X_i'(Y_i - X_i\beta) = \sum_i \frac{1}{\sigma_i^2} X_i'\varepsilon_i$$

Observe que:

$$\varepsilon_i'\varepsilon_i = (Y_i - X_i\beta)'(Y_i - X_i\beta) = (Y_i' - \beta' X_i')(Y_i - X_i\beta)$$

$$= Y_i'Y_i - Y'_{i(1\times T)} X_{i(T\times K)} \beta_{(K\times 1)} - \beta' X_i' Y_i + \beta' X_i' X_i \beta$$

$$= Y_i'Y_i - 2\beta' X_i' Y_i + \beta' X_i' X_i \beta$$

$$\frac{\partial \varepsilon_i'\varepsilon_i}{\partial \beta} = -2X_i'Y_i + 2X_i'X_i\beta = -2X_i'\left(Y_i - X_i\beta\right)$$

Continuando,

$$\frac{\partial \ln L}{\partial \sigma_i^2} = -\frac{T}{2\sigma_i^2} + \frac{1}{2\sigma_i^4}\varepsilon_i'\varepsilon_i$$

$$\frac{\partial^2 \ln L}{\partial \beta \partial \beta'} = -\sum_i \frac{1}{\sigma_i^2} X_i' X_i$$

$$\frac{\partial^2 \ln L}{\partial \left(\sigma_i^2\right)^2} = \frac{T}{2\sigma_i^4} - \frac{\varepsilon_i'\varepsilon_i}{\sigma_i^6}$$

$$\frac{\partial^2 \ln L}{\partial \beta \partial \sigma_i^2} = -\sum_i \frac{1}{\sigma_i^4} X_i' \varepsilon_i$$

A matriz de derivadas de primeira ordem é:

$$g = \begin{bmatrix} \partial \ln L / \partial \beta \\ \partial \ln L / \partial \sigma_i^2 \end{bmatrix} = \begin{bmatrix} \Sigma\left(\dfrac{1}{\sigma_i^2}\right) X_i' \varepsilon_i \\ -\dfrac{T}{2\sigma_i^2} + \dfrac{\varepsilon_i' \varepsilon_i}{2\sigma_i^4} \quad \text{para } i = 1,\ldots,n \end{bmatrix}$$

A matriz de derivadas de segunda ordem é:

$$S = \begin{bmatrix} -\sum_i \frac{1}{\sigma_i^{2}} X_i' X_i & -\sum_i \frac{1}{\sigma_i^{4}} X_i' \varepsilon_i \\ -\sum_i \frac{1}{\sigma_i^{4}} X_i' \varepsilon_i & \frac{T}{2\sigma_i^{4}} - \frac{\varepsilon_i ' \varepsilon_i}{\sigma_i^{6}} \end{bmatrix}$$

A matriz de informação (H= –E(S)) é:

$$H\left(\beta, \sigma^2\right) = \begin{bmatrix} \sum \frac{1}{\sigma_i^{2}} X_i' X_i & 0 \\ 0 & -\frac{T}{2\sigma_i^{4}} + \frac{T\sigma_i^{2}}{\sigma_i^{6}} \end{bmatrix} = \begin{bmatrix} \sum_i \frac{1}{\sigma_i^{2}} X_i' X_i & 0 \\ 0 & \left(\frac{T}{2\sigma_i^{4}}\right) I \end{bmatrix}$$

e a inversa da matriz de informação ($H^{-1}$) é a matriz de variâncias.

Sob a hipótese de que as variâncias são iguais tem-se:

$$\frac{\partial \ln L}{\partial \sigma_i^2} = \frac{T}{2\sigma^2}\left[\frac{s_i^2}{\sigma^2} - 1\right]$$

onde

$$s_i^2 = \frac{e_i ' e_i}{T}$$

Se $\beta$ é o estimador de máxima verossimilhança, da condição de 1ª. ordem:

$$\frac{\partial \ln L}{\partial \beta} = 0 \quad \Rightarrow \quad g = \begin{bmatrix} 0 \\ -\frac{T}{2s^2} + \frac{Ts_i^2}{2s^4} \end{bmatrix}$$

$$s^2 = \frac{e'e}{nT} = \frac{\Sigma e_i' e_i}{nT} = \frac{1}{n}\sum_i s_i{}^2$$

O multiplicador de Lagrange $g'H^{-1}g$ é:

$$LM = \sum_i \left[ \frac{T}{2s^2}\left(\frac{s_i^2}{s^2} - 1\right)\right]^2 \left(\frac{2s^4}{T}\right) = \frac{T}{2}\sum_i \left(\frac{s_i^2}{s^2} - 1\right)^2$$

Esta estatística testa a hipótese $H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_n^2$ e tem distribuição de qui-quadrado com *n-1* graus de liberdade.

## Correlação contemporânea

Há correlação entre as unidades cross-section, i.e.,

$$\operatorname{cov}\left(\varepsilon_{it}\varepsilon_{js}\right)=\sigma_{ij} \qquad \text{se } t=s$$

$$=0 \qquad \text{se } t\neq s$$

$$\operatorname{cov}\left(\varepsilon_{i}\varepsilon'_{j}\right)=\sigma_{ij} I$$

A matriz de variância e covariância neste caso é:

$$\Omega=\begin{bmatrix} \sigma_{11}I & \sigma_{12}I & \cdots & \sigma_{1n}I \\ \sigma_{21}I & \sigma_{22}I & & \sigma_{2n}I \\ \vdots & & & \\ \sigma_{n1}I & \sigma_{n2}I & \cdots & \sigma_{nn}I \end{bmatrix}$$

Se $n$ = 3, a matriz de covariância será:

$$\Omega = E(\varepsilon\varepsilon') = E\left[\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\end{pmatrix}\begin{pmatrix}\varepsilon_1' & \varepsilon_2' & \varepsilon_3'\end{pmatrix}\right] = \begin{bmatrix} E(\varepsilon_1\varepsilon_1') & E(\varepsilon_1\varepsilon_2') & E(\varepsilon_1\varepsilon_3') \\ E(\varepsilon_2\varepsilon_1') & E(\varepsilon_2\varepsilon_2') & E(\varepsilon_2\varepsilon_3') \\ E(\varepsilon_3\varepsilon_1') & E(\varepsilon_3\varepsilon_2') & E(\varepsilon_3\varepsilon_3') \end{bmatrix} = \begin{bmatrix} \sigma_{11}I & \sigma_{12}I & \sigma_{13}I \\ \sigma_{12}I & \sigma_{22}I & \sigma_{23}I \\ \sigma_{13}I & \sigma_{23}I & \sigma_{33}I \end{bmatrix}$$

$$E(\varepsilon_1\varepsilon_2') = E\left[\begin{pmatrix}\varepsilon_{11}\\ \varepsilon_{12}\\ \vdots \\ \varepsilon_{1T}\end{pmatrix}\begin{pmatrix}\varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2T}\end{pmatrix}\right] = E\begin{bmatrix} \varepsilon_{11}\varepsilon_{21} & \varepsilon_{11}\varepsilon_{22} & \cdots & \varepsilon_{11}\varepsilon_{2T} \\ \varepsilon_{12}\varepsilon_{21} & \varepsilon_{12}\varepsilon_{22} & \cdots & \varepsilon_{12}\varepsilon_{2T} \\ \vdots & & & \\ \varepsilon_{1T}\varepsilon_{21} & \varepsilon_{1T}\varepsilon_{22} & \cdots & \varepsilon_{1T}\varepsilon_{2T} \end{bmatrix}$$

$$\operatorname{cov}(\varepsilon_{1t}, \varepsilon_{2t}) = E(\varepsilon_{1t}\varepsilon_{2t}) = \sigma_{12} \qquad t = 1,2,\ldots,T$$

$$\operatorname{cov}(\varepsilon_{1t}, \varepsilon_{2s}) = E(\varepsilon_{1t}\varepsilon_{2s}) = 0 \qquad t \neq s$$

$$E(\varepsilon_1\varepsilon_2') = \begin{bmatrix} \sigma_{12} & 0 & \cdots & 0 \\ 0 & \sigma_{12} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \sigma_{12} \end{bmatrix} = \sigma_{12} I_T$$

Para estimar o modelo, utiliza-se MQO usando as $nT$ observações para obter $b$ e os resíduos ($e_i$), onde

$$e_i = Y_i - bX_i$$

e

$$s_{ij} = \frac{e_i' e_j}{T}$$

é um estimador consistente de $\sigma_{ij}$

Estimando-se $\Omega$ utiliza-se o estimador:

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

para as $nT$ observações.

Observe que:

$$\Omega = \Sigma_{n\times n} \otimes I$$

e

$$\Omega^{-1} = \Sigma^{-1} \otimes I$$

Produto de Kronecker

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

$$A \otimes B = \begin{bmatrix} 1\begin{bmatrix} 2 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix} & 3\begin{bmatrix} 2 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix} \\ 2\begin{bmatrix} 2 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix} & 0\begin{bmatrix} 2 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 2 & 2 & 0 & 6 & 6 & 0 \\ 1 & 0 & 3 & 3 & 0 & 9 \\ 4 & 4 & 0 & 0 & 0 & 0 \\ 2 & 0 & 6 & 0 & 0 & 0 \end{bmatrix}$$

Seja $\sigma^{ij}$ o *ij* ésimo elemento de $\Sigma^{-1}$, então,

$$\Omega^{-1} = \begin{bmatrix} \sigma^{11}I & \sigma^{12}I & \cdots & \sigma^{1n}I \\ \sigma^{21}I & \sigma^{22}I & & \sigma^{2n}I \\ \vdots & & & \\ \sigma^{n1}I & \sigma^{n2}I & \cdots & \sigma^{nn}I \end{bmatrix}$$

A utilidade disto é que

$$X'\Omega^{-1}X = \sum_i \sum_j \sigma^{ij} X_i' X_j$$

e

$$X'\Omega^{-1}Y = \sum_i \sum_j \sigma^{ij} X_i' Y_j$$

e o estimador fica:

$$\hat{\beta} = \left[ \sum_i \sum_j \sigma^{ij} X_i' X_j \right]^{-1} \left[ \sum_i \sum_j \sigma^{ij} X_i' Y_j \right]$$

Exemplo:

$$\begin{bmatrix} X_1' & X_2' \end{bmatrix} \begin{bmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{21} & \sigma^{22} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} X_1'\sigma^{11} + X_2'\sigma^{21} & X_1'\sigma^{12} + X_2'\sigma^{22} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

$$= X_1' X_1 \sigma^{11} + X_2' X_1 \sigma^{21} + X_1' X_2 \sigma^{12} + X_2' X_2 \sigma^{22}$$

## Testando correlação contemporânea:

Teste do multiplicador de Lagrange desenvolvido por Breusch e Pagan.

$$\lambda_{LM} = T\sum_{i=2}^{n}\sum_{j=1}^{i-1} r_{ij}^{\,2}$$

onde

$$r_{ij}^2 = \frac{s_{ij}^{\,2}}{s_{ii}s_{jj}} \qquad s_{ij} = \frac{e_i'e_j}{T}$$

Esta estatística tem distribuição de qui-quadrado com $\frac{n(n-1)}{2}$

graus de liberdade e $H_0$ = não há corr. contemporânea,

i.e., fora da diagonal principal os valores são zero.

**Autocorrelação**

Vamos pressupor que não há correlação nos erros entre as unidades cross-section.

$$\operatorname{cov}\left(\varepsilon_{it}, \varepsilon_{js}\right) = 0 \qquad \text{se} \qquad i \neq j$$

e vamos permitir autocorrelação dentro das unidades cross-section.

$$E\left(\varepsilon_i \varepsilon_j\right) = 0$$

$$E\left(\varepsilon_i \varepsilon_i\right) = \sigma_i^2 V_i$$

Então, a matriz de variância e covariância fica:

$$\Omega = \begin{bmatrix} \sigma_1^2 V_1 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 V_2 & & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \sigma_n^2 V_n \end{bmatrix}$$

Considere um AR(1)

$$\varepsilon_{it} = \rho_i \varepsilon_{it-1} + u_{it} \qquad 0 < |\rho_i| < 1$$

Revisão de Autocorrelação do Mestrado.

$$Y_t = X_t\beta + \varepsilon_t$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

$$0 \le |\rho| < 1$$

$$\varepsilon_{t-1} = \rho\varepsilon_{t-2} + u_{t-1}$$

$u_t \sim N(0, \sigma_u^2)$ são independentes entre si e com os ε.

$$\varepsilon_t = \rho(\rho\varepsilon_{t-2} + u_{t-1}) + u_t \longrightarrow \varepsilon_t = \rho^2\varepsilon_{t-2} + \rho u_{t-1} + u_t$$

$$\varepsilon_{t-2} = \rho\varepsilon_{t-3} + u_{t-2}$$

$$\varepsilon_t = \rho^2(\rho\varepsilon_{t-3} + u_{t-2}) + \rho u_{t-1} + u_t$$

$$\varepsilon_t = \rho^3\varepsilon_{t-3} + \rho^2 u_{t-2} + \rho u_{t-1} + u_t$$

Generalizando,

$$\varepsilon_t = u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \ldots = \sum_{i=0}^{\infty} \rho^i u_{t-i}$$

$$Var(\varepsilon) = E\left[\left(u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \ldots\right)\left(u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \ldots\right)\right]$$

$$\sigma_\varepsilon^2 = \sigma_u^2 + \rho^2 \sigma_u^2 + \rho^4 \sigma_u^2 + \rho^6 \sigma_u^2 + \ldots$$

$$= \sigma_u^2 \left(1 + \rho^2 + \rho^4 + \rho^6 + \ldots\right) = \frac{\sigma_u^2}{1 - \rho^2}$$

$$\mathrm{cov}\left(\varepsilon_t, \varepsilon_{t-s}\right) = E\left(\varepsilon_t, \varepsilon_{t-s}\right) = E\left[\left(u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \ldots\right)\left(u_{t-s} + \rho u_{t-s-1} + \rho^2 u_{t-s-2} + \ldots\right)\right]$$

Para $s$ = 1

$$\operatorname{cov}(\varepsilon_t, \varepsilon_{t-1}) = E\left[\left(u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \dots\right)\left(u_{t-1} + \rho u_{t-2} + \rho^2 u_{t-3} + \dots\right)\right]$$

$$= \rho\sigma_u^2 + \rho^3\sigma_u^2 + \dots$$

Para $s$ = 2

$$\operatorname{cov}(\varepsilon_t, \varepsilon_{t-2}) = E\left[\left(u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \rho^3 u_{t-3} + \dots\right)\left(u_{t-2} + \rho u_{t-3} + \rho^2 u_{t-4} + \dots\right)\right]$$

$$= \rho^2\sigma_u^2 + \rho^4\sigma_u^2 + \dots$$

Generalizando,

$$\operatorname{cov}(\varepsilon_t, \varepsilon_{t-s}) = \rho^s \sigma_u^2 + \rho^{s+2} \sigma_u^2 + \ldots$$

$$\rho^s \sigma_u^2 (1 + \rho^2 + \rho^4 + \ldots) = \frac{\rho^s \sigma_u^2}{1 - \rho^2}$$

$$\Omega = \frac{\sigma_u^2}{1-\rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \rho^2 & & \rho^{T-2} \\ \rho^2 & \rho & 1 & \rho & & \rho^{T-3} \\ \vdots & & & & & \\ \rho^{T-1} & \rho^{T-2} & & & \cdots & 1 \end{bmatrix}$$

$$Y_t = \alpha + \beta X_t + \varepsilon_t \qquad \varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

$$Y_{t-1} = \alpha + \beta X_{t-1} + \varepsilon_{t-1}$$

$$\underbrace{Y_t - \rho Y_{t-1}}_{Y^*} = \underbrace{\alpha(1-\rho)}_{\alpha^*} + \beta\underbrace{\left(X_t - \rho X_{t-1}\right)}_{X^*} + \underbrace{\varepsilon_t - \rho\varepsilon_{t-1}}_{u_t}$$

Para $t = 2,\ldots, n$

$$Var\left(\varepsilon_t - \rho\varepsilon_{t-1}\right) = \mathrm{var}(\varepsilon_t) + \rho^2\,\mathrm{var}(\varepsilon_{t-1}) - 2\rho\,\mathrm{cov}(\varepsilon_t\varepsilon_{t-1})$$

$$= \sigma_\varepsilon^2 + \rho^2\sigma_\varepsilon^2 - 2\rho^2\sigma_\varepsilon^2 = (1-\rho^2)\sigma_\varepsilon^2 = \sigma_u^2$$

uma vez que $\sigma_\varepsilon^2 = \dfrac{\sigma_u^2}{1-\rho^2}$

Para $t = 1$

$$\sqrt{1-\rho^2}Y_t = \sqrt{1-\rho^2}\alpha + \sqrt{1-\rho^2}X_t + \sqrt{1-\rho^2}\varepsilon_t$$

$$Var\left[\left(\sqrt{1-\rho^2}\right)\varepsilon_t\right] = \left(1-\rho^2\right)\sigma_\varepsilon^2 = \sigma_u^2$$

Então, voltando aos dados pooling,

$$\varepsilon_{it} = \rho_i \varepsilon_{it-1} + u_{it} \qquad 0 < \left|\rho_i\right| < 1$$

$$Var\left(\varepsilon_{it}\right) = \sigma_i^2 = \frac{\sigma_{ui}^2}{1-\rho_i^2} \qquad \operatorname{cov}(\varepsilon_{it}, \varepsilon_{i\,t-s}) = \frac{\rho_i^{\ s}\sigma_{ui}^2}{1-\rho_i^{\ 2}}$$

$$\sigma_i^2 V_i = \underbrace{\frac{\sigma_{ui}^2}{1-\rho_i^2}}_{\sigma_i^2}\begin{bmatrix} 1 & \rho_i & \rho_i^2 & \cdots & \rho_i^{T-1} \\ \rho_i & 1 & \rho_i & & \rho_i^{T-2} \\ \vdots & & & & \vdots \\ \rho_i^{T-1} & \rho_i^{T-2} & \rho_i^{T-3} & \cdots & 1 \end{bmatrix}$$

Suponha que $r_i$ é um estimador consistente de $\rho_i$, então podemos transformar os dados usando a transformação de Prais-Winsten.

$$y_{*i} = \begin{bmatrix} \sqrt{1-r_i^2}\,Y_{i1} \\ Y_{i2} - r_i Y_{i1} \\ \vdots \\ Y_{iT} - r_i Y_{iT-1} \end{bmatrix} \qquad X_{*i} = \begin{bmatrix} \sqrt{1-r_i^2}\,X_{i1} \\ X_{i2} - r_i X_{i1} \\ \vdots \\ X_{iT} - r_i X_{iT-1} \end{bmatrix}$$

Como estimar $\rho_i$:

Aplica-se MQO as $nT$ observações e obtém-se $b$, então estima-se

$e_i = Y_i - bX_i$ para cada $i$

$$r_i = \frac{\sum_{t=2}^{T} e_{it} \cdot e_{i,t-1}}{\sum_{t=2}^{T} e_{i,t-1}^2} \qquad t = 2.,\ldots,T$$

que é o coeficiente da regressão:

$$e_{it} = \rho_i e_{i,t-1} + \xi_{it}$$

Se $|r_i| > 1$ então usar $r_i = \dfrac{\Sigma e_{it} e_{it-1}}{\sqrt{\Sigma e_{it}^2} \sqrt{\Sigma e_{it-1}^2}}$ que estará sempre entre 1 e –1

Com os dados transformados, o modelo é agora somente heterocedástico, isto é, a matriz de variância e covariância é:

$$\Omega = \begin{bmatrix} \sigma_1^2 I & 0 & \cdots & 0 \\ 0 & \sigma_2^2 I & & \vdots \\ & \vdots & & \\ 0 & 0 & \cdots & \sigma_n^2 I \end{bmatrix}$$

e o estimador é:

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

Para obter $\hat{\Omega}$ , aplica-se MQO aos *nT* dados transformados e obtém-se $b^*$, então utiliza-o para encontrar:

$$e_i^* = Y_i^* - b^* X_i^*$$

Com base nestes resíduos estima-se $\sigma_i^2$

$$\hat{\sigma}_{ui}^2 = \frac{e_i^{*}{}' e_i^*}{T} \qquad \text{e} \qquad \hat{\sigma}_i^2 = \frac{\hat{\sigma}_{ui}^2}{1 - r_i^2}$$

Após estimar a matriz Ω pode-se estimar os parâmetros utilizando:

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

Ou, com $\hat{\sigma}_i^2$ aplica-se MQO ao modelo:

$$Y_i^{**} = \beta X_i^{**} + \varepsilon_i^{**}$$

onde

$$Y_i^{**} = \frac{Y_i^*}{\hat{\sigma}_i} \quad \text{e} \quad X_i^{**} = \frac{X_i^*}{\hat{\sigma}_i} \quad \text{e} \quad \varepsilon_i^{**} = \frac{\varepsilon_i^*}{\hat{\sigma}_i}$$

Teste para autocorrelação:

$$\mathrm{H}_0 : \rho_i = 0 \qquad\qquad \frac{(T-1)r_i^2}{1-r_i^2} \sim \chi^2(1)$$

Permitindo ainda a existência de correlação entre as unidades cross-section.

$$\operatorname{cov}(u_{it}, u_{jt}) = \sigma_{u_{ij}}$$

então a matriz $\Omega$ que era

$$\Omega = \begin{bmatrix} \sigma_{11}I & \sigma_{12}I & \cdots & \sigma_{1n}I \\ \sigma_{21}I & \sigma_{22}I & & \sigma_{2n}I \\ \vdots & & & \\ \sigma_{n1}I & \sigma_{n2}I & & \sigma_{nn}I \end{bmatrix}$$

ficará ao invés de $\sigma_{ij}I$, $\sigma_{ij}V_{ij}$ fora da diagonal principal e na diagonal $\sigma_i^2 V_i$

$$\sigma_{ij}V_{ij} = \underbrace{\frac{\sigma_{u_{ij}}}{1-\rho_i\rho_j}}_{\sigma_{ij}} \begin{bmatrix} 1 & \rho_j & \rho_j^2 & \cdots & \rho_j^{T-1} \\ \rho_i & 1 & \rho_j & & \rho_j^{T-2} \\ \rho_i^2 & \rho_i & 1 & & \rho_j^{T-3} \\ \vdots & & & & \\ \rho_i^{T-1} & \rho_i^{T-2} & \rho_i^{T-3} & \cdots & 1 \end{bmatrix}$$

onde $\sigma_{uij} = \operatorname{cov}(u_{it}, u_{jt})$

Para $s$ = 1

$$\operatorname{cov}\left(\varepsilon_{it}, \varepsilon_{jt-1}\right) = E\left[\left(u_{it} + \rho_i u_{it-1} + \rho_i^{\,2} u_{it-2} + \ldots\right)\left(u_{jt-1} + \rho_j u_{jt-2} + \rho_j^{\,2} u_{jt-3} + \ldots\right)\right]$$

$$= \rho_i \sigma_{uij} + \rho_i^{\,2} \rho_j \sigma_{uij} + \rho_i^3 \rho_j^2 \sigma_{uij} \ldots$$

Para $s$ = 2

$$\operatorname{cov}\left(\varepsilon_{it}, \varepsilon_{jt-2}\right) = E\left[\left(u_{it} + \rho_i u_{it-1} + \rho_i^{\,2} u_{it-2} + \rho_i^{\,3} u_{it-3} + \ldots\right)\left(u_{jt-2} + \rho_j u_{jt-3} + \rho_j^{\,2} u_{jt-4} + \ldots\right)\right]$$

$$= \rho_i^{\,2} \sigma_{uij} + \rho_i^3 \rho_j \sigma_{uij} + \rho_i^4 \rho_j^2 \sigma_{uij} + \ldots$$

Generalizando:

$$\operatorname{cov}\left(\varepsilon_{it}, \varepsilon_{jt-s}\right) = \rho_i^s \rho_j^0 \sigma_{uij} + \rho_i^{\,s+1} \rho_j^1 \sigma_{uij} + \rho_i^{\,s+2} \rho_j^2 \sigma_{uij} + \ldots$$

$$= \rho_i^{\,s} \sigma_{uij} \left(1 + \rho_i \rho_j + \rho_i^2 \rho_j^2 + \ldots\right) = \frac{\rho_i^s \sigma_{uij}}{1 - \rho_i \rho_j}$$

Estimação:

Com as $nT$ observações estima-se $b$ e $e_i$ .

Utiliza-se $e_i$ para estimar $\rho_i$ e transformar os dados (transformação de Prais-Winsten).

Aplica-se MQO aos dados transformados para obter $e_{*i}$ e $\hat{\sigma}_{ij}$

e utiliza o estimador:

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

onde

$$\hat{\sigma}_{uij} = \frac{e_{*i}'e_{*j}}{T} \qquad \text{e} \qquad \hat{\sigma}_{ij} = \frac{\hat{\sigma}_{uij}}{1 - r_i r_j}$$