Luís Filipe Martins | Dpt. de Métodos Quantitativos | ISCTE - EG
luis.martins@iscte.pt | http://home.iscte.pt/~lfsm | Lisboa, 29/09/2006

MODELOS COM DADOS DE PAINEL

Wooldridge §13.3-13.5,14

0.1 Introdução e Motivação

Neste capitulo discutimos a modelização de variáveis medidas por dados de painel que normalmente combinam dados seccionais e temporais. Um painel de uma dada variável económica x é uma amostra (conjunto de observações) na qual os n individuos (ou firmas, ...) são observados ao longo de T periodos de tempo. Por exemplo, $Sal_{it}, i = 1, ..., n; t = 1, ..., T$ representa o montante de salário do trabalhador i no periodo t; enquanto que Q_{it} é o nível do output da firma i para o periodo t.

Devido à sua natureza, em dados de painel é natural que exista heterogeneidade para os diferentes individuos e também dependência nas observações porque a variável evolui cronológicamente. Neste capitulo, não estamos a estudar variáveis nas quais os individuos/firmas/... não são os mesmos ao longo do tempo. Um exemplo desta situação - dados *Pooled* seccionais - é os census em que se verifica uma independência amostral para diferentes periodos de tempo (assume-se independencia entre individuos). Nestes casos, a modelização é semelhante à de dados seccionais. A análise pode ser feita ano a ano (e depois comparar estatisticamente) ou então agregam-se os dados para os diferentes periodos (a amostra combinada seccional passa a ter dimensão nT) e usam-se variáveis Dummy para captar a informação respeitante a cada periodo. Note-se que ao combinarem-se T amostras seccionais, a dimensão disponível para a amostra é (muito) superior quando comparada com uma única amostra seccional.

Ao longo do capitulo vamos assumir que o painel não é *unbalaced*, isto é, que não há falta de observações para um conjunto de individuos num particular periodo de tempo; que a forma funcional do modelo é a "mais simples" - a linear; e que n é relativamente mais elevado que T. Em relação a esta última hipótese, considera-se que o número de individuos/firmas, ... n é grande e que o periodo em análise T é pequeno, e não o contrário ou serem ambos grande. Por outras palavras, qualquer resultado assimptótico é derivado quando $n \to \infty$ e $T > 1$ é fixo. Eis alguns exemplos de modelos lineares com dados de painel.

Example 1 *$Sal_{it} = \beta_1 + \beta_2 Exp_{it} + u_{it}, i = 1, ..., n; t = 1, ..., T$, onde Sal_{it} representa o montante de salário do trabalhador i no periodo t e Exp_{it} é o número de anos de experiência do trabalhador i no periodo t.*

Example 2 *$\ln(Q_{it}) = \beta_1 + \beta_2 \ln(L_{it}) + (\alpha_i + v_{it}), i = 1, ..., n; t = 1, ..., T$, onde Q_{it} é o nível do output da firma i para o periodo t; L_{it} é o respectivo input trabalho (numero de trabalhadores); α_i é uma medida da eficiência da firma i (assumido que permanece constante ao longo do tempo) e v_{it} é o choque tecnológico na firma i para o periodo t.*

Sem mencionar as propriedades dos erros u_{it}, α_i e v_{it} nos dois examplos anteriores, identificamos o primeiro como um modelo de efeitos aleatórios e o segundo de efeitos fixos. Como se observa, no segundo exemplo assume-se que os erros u_{it} podem ser decompostos em duas componentes, α_i e v_{it}. A componente não observável α_i é o efeito fixo ou efeito individual (heterogeneidade individual) porque, sendo constante para todo o t, apenas está indexado ao particular individuo i. O estudo destes dois modelos é feito em seguida.

0.2 O Modelo de Efeitos Aleatórios como uma SUR

O modelo para os salários, em que não se assume a existência de heterogeneidade individual, é de efeitos aleatórios. Em termos gerais, e para uma amostra $\{(x_{it1}, ..., x_{itk}, y_{it}) : i = 1, ..., n; t = 1, ..., T\}$, considere-se a seguinte especificação,

$$y_{it} = \beta_1 x_{it1} + ... + \beta_k x_{itk} + u_i = x_{it}^{'}\beta + u_{it}, i = 1, ..., n; t = 1, ..., T \tag{1}$$

que se reduz a

$$y_{it} = \beta_1 + \beta_2 x_{it2} + ... + \beta_k x_{itk} + u_{it} = x_{it}^{'}\beta + u_{it}, i = 1, ..., n; t = 1, ..., T \tag{2}$$

quando se inclui um termo independente $x_{it1} = 1, \forall i, t$.

Este modelo pode ser visto como um caso especial de um sistema de equações simultâneas

$$y_{im} = x_{im}^{'}\beta_m + u_{im}, i = 1, ..., n; m = 1, ..., M \tag{3}$$

em que existem M equações lineares para n observações e β_m é um vector k_m-dimensional em que k_m é o números de regressores da equação m (não tem portanto de ser igual para todo m). Este modelo não é alvo de estudo mas podem-se enumerar os métodos GMM, 3SLS, FIML e FIVE (entre outros) para estimar β_m. O modelo SUR (seemingly unrelated regressions) é o (3) mas impondo a hipótese de que os erros u_{im} são exógenos para **todo** o m (os regressores são pré-determinados em cada equação e **também** nas outras equações), $E(x_{im}u_{ih}) = 0, m, h = 1, ..., M$. Nestas condições pode-se estimar β_m equação a equação por OLS (ou o GLS=SUR que é mais eficiente quando existe heterocedasticidade e/ou autocorrelação dos erros, $E\left(u_1 u_1^{'}\right) \neq \sigma^2 I_M$.)[1]

[1]Pode-se demonstrar que o OLS e o GLS coincidem quando os regressores são os mesmos para todas as equações, como é o caso dos modelos com dados de painel. Veja-se, para isso, a expressão do estimador no modelo com efeitos aleatórios a derivar em seguida.

Em relação ao modelo (1), assume-se para o SUR que as variáveis dependente y_{im} e os regressores x_{im} são os mesmos para todas as M equações - MRLM. Consequentemente, $k_m = k$ para todo m e, por outro lado, é imposta a restrição de coeficientes comuns para todas as equações, $\beta_m = \beta$.

Tal como para o MRLM, podemos especificar o modelo de efeitos aleatórios (1) em termos de matrizes. Como estamos a considerar o caso de n grande e T fixo, (1) é equivalente a

$$Y_i = X_i\beta + U_i = \beta_1\iota_T + \beta_2\underline{x}_{i2} + ... + \beta_k\underline{x}_{ik} + U_i, i = 1, ..., n, \tag{4}$$

onde

$$\begin{aligned}
U_{i_{T\times 1}} &= (u_{i1}, ..., u_{iT})^{'}; \beta_{k\times 1} = (\beta_1, \beta_2, ..., \beta_k)^{'}; \iota_{T_{T\times 1}} = (1, ..., 1)^{'}; \qquad (5)\\
Y_{i_{T\times 1}} &= \begin{pmatrix} y_{i1} \\ ... \\ y_{iT} \end{pmatrix}; X_{i_{T\times k}} = \begin{pmatrix} x_{i1}{}^{'} \\ ... \\ x_{iT}{}^{'} \end{pmatrix} = \begin{pmatrix} 1 & x_{i12} & ... & x_{i1k} \\ ... & ... & & ... \\ 1 & x_{iT2} & ... & x_{iTk} \end{pmatrix} = (\iota, \underline{x}_{i2}, ..., \underline{x}_{ik});\\
x_{it_{k\times 1}} &= (1, x_{it2}, ..., x_{itk})^{'}, \underline{x}_{ij_{T\times 1}} = (x_{i1j}, x_{i2j}, ..., x_{iTj})^{'}, i = 1, ..., n; j = 2, ..., k.
\end{aligned}$$

Note que (4) pode ser interpretado como um MRLM de dados seccionais $i = 1, ..., n$. A diferença é que devido à componente temporal, $t = 1, ..., T$, a variável dependente Y_i (e os erros U_i) é um vector e X_i é uma matriz para todo i. Desta forma, as hipóteses básicas do modelo (4) e os estimadores são uma extensão natural das do MRLM.

P1 Para cada $i = 1, ..., n$ fixo, o processo estocástico $(x_{i1}, ..., x_{iT}, y_{i1}, ..., y_{iT})$ é conjuntamente estacionário que possui memória que se dissipa (conjuntamente fracamente dependente (*weakly dependent*)). Por outro lado, a amostra é *i.i.d.* ao longo dos individuos/firmas/... i : $\{(x_{it1}, ..., x_{itk}, y_{it}) : i = 1, ..., n; t = 1, ..., T \text{ fixo}\}$ ou $\{(x_{1t}, ..., x_{nt}, y_{1t}, ..., y_{nt}) : t = 1, ..., T$ fixo$\}$. Note que se considerarmos heterocedasticidade, $\{(x_{it1}, ..., x_{itk}, y_{it}) : i = 1, ..., n; t = 1, ..., T \text{ fixo}\}$ é apenas independente pois $\{y_{it} : i = 1, ..., n; t = 1, ..., T \text{ fixo}\}$ deixa de ser identicamente distribuido.

P2 (ortogonalidade) Para cada $t = 1, ..., T$, os k regressores são ortogonais (não correlacionados) aos erros:

$$E(x_{it}u_{it}) = 0_{k\times 1}, t = 1, ..., T \Leftrightarrow E(X_i^{'}U_i) = \sum_{t=1}^{T} E(x_{it}u_{it}) = 0_{k\times 1}. \tag{6}$$

Existem, portanto, Tk condições de ortogonalidade. Sob a condição de *i.d.*, temos que $E(X_1^{'}U_1) = 0_{k\times 1}$.

P3 (identificação de β) Para cada $i = 1, ..., n$ (ou fixando $i = 1$ para dados *i.d.*), a caracteristica da matriz $\begin{pmatrix} x_{i1}x_{i1}^{'} \\ ... \\ x_{iT}x_{iT}^{'} \end{pmatrix}_{Tk\times k}$ é k. Para x estocástico, a matriz é $\begin{pmatrix} E\left(x_{i1}x_{i1}^{'}\right) \\ ... \\ E\left(x_{iT}x_{iT}^{'}\right) \end{pmatrix}$.

P4 (homocedasticidade condicional e existência de autocorrelação nos erros): Para cada individuo $i = 1, ..., n$,

$$E(u_{it}u_{is}|x_{it}, x_{is}) = E(u_{it}u_{is}|x_{it2}, ..., x_{itk}, x_{is2}, ..., x_{isk}) = \sigma_{ts}, t, s = 1, ..., T, \tag{7}$$

que **não** depende de i nem dos regressores x. Consequentemente, pode-se escrever esta condição como $E(u_{1t}u_{1s}|x_{1t}, x_{1s}) = \sigma_{ts}, t, s = 1, ..., T$. Por outro lado, quando $t = s$ (homocedasticidade para cada equação), $E(u_{1t}^2|x_{1t}) = \sigma_t^2, t = 1, ..., T$. A forma matricial de (7), $E(U_1U_1'|X_1)$, é

$$\Omega_{T\times T} = \begin{pmatrix} E(u_{11}^2|x_{11}) & ... & E(u_{11}u_{1T}|x_{11}, x_{1T}) \\ ... & & ... \\ E(u_{1T}u_{11}|x_{1T}, x_{11}) & ... & E(u_{1T}^2|x_{1T}) \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & ... & \sigma_{1T} \\ ... & & ... \\ \sigma_{T1} & ... & \sigma_T^2 \end{pmatrix}, \tag{8}$$

uma matriz que é comum para todo o individuo $i = 1, ..., n$ e daí implicar homocedasticidade. Note que esta condição é ainda mais dificil de se verificar do que a correspondente condição de homocedasticidade no MRLM com dados seccionais. A justificação é que em (7) estamos a condicionar em todo o passado e futuro dos regressores! Como o modelo de interesse é de dados de painel, é natural considerar dependência entre erros: $Cov(u_{it}, u_{is}|x_{it}, x_{is}) = \sigma_{ts} \neq 0, t \neq s$.

Finda a discussão das hipóteses do modelo[2], apresentamos o estimador FGLS (estimador SUR) para β que também é denominado por estimador de efeitos aleatórios:

$$\widehat{\beta}_{EA} = \left(\frac{1}{n}\sum_{i=1}^{n} X_i'\widehat{\Omega}^{-1}X_i\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} X_i'\widehat{\Omega}^{-1}Y_i = \left(\sum_{i=1}^{n} X_i'\widehat{\Omega}^{-1}X_i\right)^{-1} \sum_{i=1}^{n} X_i'\widehat{\Omega}^{-1}Y_i, \tag{9}$$

em que X_i é uma matriz e Y_i um vector e onde

$$\widehat{\Omega} = \frac{1}{n}\sum_{i=1}^{n} \widehat{U}_i\widehat{U}_i', \tag{10}$$

em que $\widehat{U}_i = Y_i - X_i\widehat{\beta}_{OLS}$ são os (vector) residuos OLS no qual

$$\widehat{\beta}_{OLS} = \left(\sum_{i=1}^{n} X_i'X_i\right)^{-1} \sum_{i=1}^{n} X_i'Y_i. \tag{11}$$

É interessante verificar que o OLS $\widehat{\beta}_{OLS}$ é muito simplesmente o estimador do modelo com dados *Pooled* seccionais pois, neste caso, $\Omega_{T\times T} = E(U_1U_1'|X_1) = \sigma^2 I_T$. No modelo de efeitos aleatórios com dados de painel, apesar de serem ambos consistentes (sob $P1 - P3$), o estimador FGLS $\widehat{\beta}_{EA}$ é assimptóticamente mais eficiente do que o OLS $\widehat{\beta}_{OLS}$:

$$AV\left(\widehat{\beta}_{EA}\right) = \left[E\left(X_1'\Omega^{-1}X_1\right)\right]^{-1} \tag{12}$$

$$AV\left(\widehat{\beta}_{OLS}\right) = \left[E\left(X_1'X_1\right)\right]^{-1} E\left(X_1'U_1U_1'X_1\right)\left[E\left(X_1'X_1\right)\right]^{-1} \tag{13}$$

$$= \left[E\left(X_1'X_1\right)\right]^{-1} E\left(X_1'\Omega X_1\right)\left[E\left(X_1'X_1\right)\right]^{-1} \tag{14}$$

[2]Não consideramos a normalidade dos erros pois assumimos que n é elevado para um fixo T.

No entanto, é importante salientar o facto de que a temática da eficiencia com dados de painel não é tão decisiva como a de dados seccionais pois em dados de painel o número de observações é T vezes maior para o mesmo número de individuos n. Finalmente, pode ser demonstrado que, sob $P1 - P4$,

$$\sqrt{n}\left(\widehat{\beta}_{EA} - \beta\right) \xrightarrow{d} N\left(0, AV\left(\widehat{\beta}_{EA}\right)\right), \tag{15}$$

onde $AV\left(\widehat{\beta}_{EA}\right)$ pode ser consistentemente estimado por

$$\widehat{AV\left(\widehat{\beta}_{EA}\right)} = \left(\frac{1}{n}\sum_{i=1}^{n} X_i'\widehat{\Omega}^{-1}X_i\right)^{-1}. \tag{16}$$

Apesar da sua simplicidade em termos de dados de painel, o modelo de efeitos aleatórios tem algumas limitações. Em primeiro lugar, o número de parâmetros a estimar na matriz Ω pode ser elevado para T grande. Sob a hipótese $P4$, esta matriz tem $\frac{T^2-T}{2} + T = \frac{T(T+1)}{2}$ coeficientes diferentes! Ainda em relação à condição $P4$ em conjunto com $P2$, estas são muito diffceis de se verificar na prática. Estamos a assumir que os regressores, que são variáveis observáveis, não dependem contemporâneamente dos não-observáveis erros e nem do passado e futuro! Por outro lado, o modelo (1) não considera a existência de um efeito individual (heterogeneidade individual) não-observável para todos os periodos α_i. No modelo $\ln(Q_{it})$, mas de efeitos aleatórios, $\ln(Q_{it}) = \beta_1 + \beta_2 \ln(L_{it}) + u_{it}$, se não se considerar a eficiência da firma i, α_i, este é um modelo cuja estimação estará esviesada ao se omitir a informação relevante α_i.

Uma especificação alternativa a (1) é dada pelo modelo de efeitos fixos que será alvo de estudo na próxima secção. No modelo de efeitos fixos, assume-se que os erros $u_{it} = \alpha_i + v_{it}$ são explicados por duas componentes não observáveis: α_i, um efeito individual que é constante ao longo do tempo; e v_{it} uma variável ruído para cada individuo i e periodo t. Formalmente, e para uma amostra $\{(x_{it1}, ..., x_{itk}, y_{it}) : i = 1, ..., n; t = 1, ..., T\}$, y_{it} é gerado por

$$y_{it} = \beta_1 x_{it1} + ... + \beta_k x_{itk} + \alpha_i + v_{it} = x_{it}'\beta + \alpha_i + v_{it}, i = 1, ..., n; t = 1, ..., T \tag{17}$$

que se reduz a

$$y_{it} = \beta_1 + \beta_2 x_{it2} + ... + \beta_k x_{itk} + \alpha_i + v_{it} = x_{it}'\beta + \alpha_i + v_{it}, i = 1, ..., n; t = 1, ..., T \tag{18}$$

quando se inclui um termo independente. Em termos matriciais,

$$Y_i = X_i\beta + \iota_T \alpha_i + V_i, i = 1, ..., n, \tag{19}$$

onde $Y_i, X_i, \beta, \alpha_i$ e ι_T foram definidos anteriormente e $V_{i_{T\times 1}} = (v_{i1}, ..., v_{iT})'$. Em seguida, apresentam-se as hipóteses para α_i e v_{it} e, consequentemente, para $u_{it} = \alpha_i + v_{it}$, as quais determinam o método de estimação do modelo de efeitos fixos.

Condition 1 (EF) *O processo v_{it} é i.i.d. $\left(0,\sigma_v^2\right)$ para todo $i=1,...,n$ e $t=1,...,T$; o efeito individual α_i (constante ao longo do tempo, $\alpha_i=\alpha_{it}=\alpha_{is},\forall t,s$) é i.i.d. $\left(0,\sigma_\alpha^2\right)$ para todo $i=1,...,n$; e v_{it} e α_i não estão autocorrelacionados para todo i. Deste modo, $E(v_{it}^2)=E(v_{11}^2)=\sigma_v^2$ e $E(\alpha_i^2)=E(\alpha_1^2)=\sigma_\alpha^2$. Note-se que α_i, sendo constante em t, tem correlação igual a 1 para diferentes periodos.*

Com estas condições em relação a v_{it} e α_i conseguimos "resolver" a questão de existir um significativo número de coeficientes na matriz variancias-covariâncias para $U_1, E(U_1U_1'|X_1)$. Sob a condição 1, a estimação de Ω reduz-se à estimação de dois parâmetros, σ_v^2 e σ_α^2 :

P4F A homocedasticidade condicional (comum para todo individuo $i=1,...,n$) é dada por

$$\Omega_{T\times T}=\begin{pmatrix}\sigma_1^2 & ... & \sigma_{1T}\\ ... & & ...\\ \sigma_{T1} & ... & \sigma_T^2\end{pmatrix}=\begin{pmatrix}\sigma_\alpha^2+\sigma_v^2 & ... & \sigma_\alpha^2\\ ... & ... & ...\\ \sigma_\alpha^2 & ... & \sigma_\alpha^2+\sigma_v^2\end{pmatrix}=\sigma_\alpha^2\iota_T\iota_T'+\sigma_v^2I_T, \qquad (20)$$

isto é,

$$E(u_{it}u_{is}|x_{it},x_{is})=E(u_{1t}u_{1s}|x_{1t},x_{1s})=\left\{\begin{array}{c}\sigma_\alpha^2+\sigma_v^2, t=s\\ \sigma_\alpha^2, t\neq s\end{array}\right. \quad t,s=1,...,T. \qquad (21)$$

Como era de esperar, os erros u_{it} para um individuo i estão autocorrelacionados no tempo $\left(\sigma_\alpha^2\right)$ devido ao efeito individual α_i, que é constante no tempo. De salientar que esta hipótese ainda é mais restrictiva do que $P4$! O objectivo foi simplificar a dimensão dos parâmetros em $\Omega_{T\times T}$.

Note-se que na condição 1 nada foi dito em relação à exogeneidade/endogeneidade (ortogonalidade) entre erros u_{it} (via α_i) e regressores x_{it}. Num primeiro cenário, admitamos como válida a hipótese $P2$,

$$E(X_i'U_i)=E(X_1'U_1)=E(X_1'\iota_T\alpha_1)+E(X_1'V_1)=E(X_1'\iota_T\alpha_1)=0 \qquad (22)$$

em que erros (via α_i) e regressores são ortogonais contemporâneamente e não contemporâneamente. Então, o método de estimação e a distribuição assimptótica do estimador coincide com o FGLS=EA mas onde

$$\widehat{\Omega}=\widehat{\sigma}_\alpha^2\iota_T\iota_T'+\widehat{\sigma}_v^2I_T, \qquad (23)$$

em que, com os resíduos OLS $\widehat{u}_{it}$,

$$\widehat{\sigma}_\alpha^2 = \frac{2}{nT\left(T-1\right)}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{s=t+1}^{T}\widehat{u}_{it}\widehat{u}_{is} \qquad (24)$$

$$\widehat{\sigma}_v^2 = \left(\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{u}_{it}^2\right)-\widehat{\sigma}_\alpha^2. \qquad (25)$$

Porque as condições impostas são muito dificeis de se verificar na prática, este método de estimação é de evitar. Nomeadamente, estamos a assumir que o efeito individual α_i que é idêntico para todos os periodos não está correlacionado com os regressores nem contemporâneamente!

0.3 Modelo de Efeitos Fixos

Num cenário mais realista vamos admitir que o efeito individual α_i (e não v_{it}) está correlacionado com os regressores x_{it} (não existe exogeneidade estrita entre u_{it} e x_{it}). No exemplo em estudo, assume-se que a eficiência da firma i, α_i, está correlacionada com o número de trabalhadores, L_{it}, independentemente do periodo em causa[3]. Como se sabe, o OLS é enviesado e inconsistente sob a hipótese de endogeneidade,

$$E(X_i^{'}U_i) = E(X_1^{'}U_1) = E(X_1^{'}\iota_T\alpha_1) \neq 0. \tag{26}$$

Uma forma de estimar o modelo (17) em que α_i não é observado e existe endogeneidade é através de primeiras-diferenças, um procedimento que, obviamente, não pode ser aplicado para modelos em que os regressores não dependem do tempo, $x_{it} = x_i, t = 1, ..., T$. Transformando o modelo (17) em primeiras-diferenças, "elimina-se" a causa da endogeneidade (α_i) em que se assume exogeneidade entre v_{it} e $x_{it}, E(X_1^{'}V_1) = 0$, e obtem-se

$$y_{it} - y_{i,t-1} = (x_{it} - x_{i,t-1})^{'}\beta + (v_{it} - v_{i,t-1}), i = 1, ..., n; t = 2, ..., T. \tag{27}$$

No modelo às primeiras-diferenças, se o erro $v_{it} - v_{i,t-1}$ não está correlacionado com os regressores $x_{it} - x_{i,t-1}$ para todo i e t, pode-se utilizar o OLS (11) para β mas aplicado ao modelo transformado (27). Identifique-se esse estimador como $\widehat{\beta}_{PD}$. Neste caso, x_{it} é substituido por $x_{it} - x_{i,t-1}$ e y_{it} por $y_{it} - y_{i,t-1}$. Note-se que, por exemplo, $y_{it} - y_{i,t-1}, i = 1, ..., n; t = 2, ..., T$ é equivalente a

$$\begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} y_{i1} \\ ... \\ y_{i,T} \end{pmatrix} = A_{T-1\times T}Y_{i_{T\times 1}} = Y^*_{i_{T-1\times 1}}, i = 1, ..., n \tag{28}$$

Quando $v_{it} - v_{i,t-1}$ estão correlacionados com os regressores $x_{it} - x_{i,t-1}$ utiliza-se o método das variáveis instrumentais. Desde que $E(x_{it}v_{is}) = 0, 1 \leq t \leq s \leq T$, o passado dos regressores $x_{i1}, ..., x_{i,t-1}$ pode ser utilizado como instrumentos pois estão correlacionados com $x_{it} - x_{i,t-1}$ e são ortogonais a $v_{it} - v_{i,t-1} : E(x_{is}(v_{it} - v_{i,t-1})) = 0, 1 \leq s \leq t-1$.

Em muitas aplicações, o investigador tem dois anos de observações $(T = 2)$, normalmente consecutivos, para n individuos. Então, o modelo às primeiras-diferenças (27) é muito simplesmente um MRLM de tipo seccional e no qual se pode utilizar o OLS (dadas as hipóteses de v_{it}) :

$$\triangle y_i = \beta_1\triangle x_{i1} + ... + \beta_k\triangle x_{ik} + \triangle v_i = \triangle x_i^{'}\beta + \triangle v_i, i = 1, ..., n, \tag{29}$$

em que $\triangle y_i = y_{iT} - y_{i,T-1}$; $\triangle x_{ij} = x_{iTj} - x_{i,T-1,j}$ e $\triangle v_i = v_{iT} - v_{i,T-1}$.

[3] Mais evidente seria se o regressor fosse constante no tempo tal como o efeito individual.

Quando exite endogeneidade, podemos utilizar métodos alternativos ao de primeiras-diferenças. Entre outros inconvenientes, salienta-se que primeiras-diferenças implica a eliminação dos regressores que não dependem de t e, por outro lado, a perda de eficiência/informação ao reduzir-se a dimensão da amostra em n observações (um ano de observações). O método de variáveis instrumentais (GMM) de Hausman-Taylor é uma metodologia alternativa. Se os regressores que não dependem de t não estiverem correlacionados com α_i podemos usar um método semelhante ao de FGLS=EA. Nenhum destes dois métodos alternativos será alvo de estudo nestes apontamentos.

Podemos estar perante um modelo no qual os efeitos individuais α_i são observáveis. Por exemplo, no modelo $\ln(Q_{it})$, a eficiência da firma i, α_i, pode ser quantificada por uma medida de produtividade. Como é obvio, se α_i é mensurável, então estima-se por FGLS=EA o modelo $\ln(Q_{it}) - \alpha_i$ e onde os erros são v_{it}! O objecto de interesse nesta secção é tratar os $\alpha_i' s$ como coeficientes desconhecidos e que podem ou não ser alvo de estimação. De qualquer forma, o erro do modelo é apenas v_{it} que não tem necessariamnte de ter $\Omega, (8)$, como matriz de variâncias-covariâncias. Mais importante, é podermos assumir que α_i é correlacionado com x_{it} pois ambos são tidos como regressores. Nestas condições, e ao contrário do modelo anterior, não estamos a tratar de um modelo com endogeneidade.

Se os $\alpha_i' s$ não são objecto de estimação, então o método de primeiras-diferenças pode ser aplicado para estimar β. Existem, no entanto, formas alternativas de estimar o modelo de efeitos fixos. Em primeiro lugar, vamos considerar a transformação do modelo (17) em diferenças em relação à média temporal (modelo *demeaned*). O modelo (17) é equivalente a

$$y_{it} - \overline{y}_i = (x_{it} - \overline{x}_i)' \beta + (\alpha_i - \alpha_i) + (v_{it} - \overline{v}_i) = (x_{it} - \overline{x}_i)' \beta + (v_{it} - \overline{v}_i), i = 1, ..., n; t = 1, ..., T, \tag{30}$$

onde $\overline{y}_i = \frac{1}{T} \sum_{t=1}^{T} y_{it}$ e $\overline{x}_i, \overline{v}_i$ são definidos de uma forma análoga. Como se vê, também com o modelo *demeaned* se consegue retirar o efeito individual α_i da regressão. Sob a condição 1, v_{it} é $i.i.d.\left(0, \sigma_v^2\right)$, ou seja, $E\left(V_1 V_1' | X_1\right)_{T \times T} = \sigma_v^2 I_T$. Suponha-se (por absurdo, muito provavelmente) que não há endogeneidade no modelo (30) : $v_{it} - \overline{v}_i$ e $x_{it} - \overline{x}_i$ não estão autocorrelacionados. Então, o OLS (11) para β aplicado ao modelo transformado (30) é o mais eficiente[4] e identifique-se esse estimador como $\widehat{\beta}_{EF}$. Neste caso, x_{it} é substituido por $x_{it} - \overline{x}_i$ e y_{it} por $y_{it} - \overline{y}_i, i = 1, ..., n; t = 1, ..., T$ que pode ser escrito como

$$I_T \begin{pmatrix} y_{i1} \\ ... \\ y_{i,T} \end{pmatrix} - \frac{1}{T} \iota_T \iota_T' \begin{pmatrix} y_{i1} \\ ... \\ y_{i,T} \end{pmatrix} = \left(I_T - \frac{1}{T} \iota_T \iota_T' \right)_{T \times T} Y_{i_{T \times 1}} = Y_{i_{T \times 1}}^{**}, i = 1, ..., n. \tag{31}$$

O problema é que a hipótese de exogeneidade muito dificilmente se verifica na prática. A explicação é a seguinte: Mesmo que v_{it} e x_{it} não estejam autocorrelacionados contemporâneamente (pode não acontecer, no entanto), $v_{it} - \overline{v}_i$ e $x_{it} - \overline{x}_i$ podem-no estar pois $\overline{v}_i$ e $\overline{x}_i$ dependem das

[4] Ou o GLS caso se admita que $E\left(V_1 V_1' | X_1\right)_{T \times T} = \Omega \neq \sigma_v^2 I_T$.

observações i para todos os periodos - passado, presente e futuro. Por exemplo, x_{it} pode depender do passado dos erros $v_{i,t-1}, v_{i,t-2}, ..., v_{i,1}$! A solução em termo de estimação passa pelo uso de IV (GMM). A dificuldade é encontrar instrumentos que são ortogonais a v_{it} para todos os periodos t! Por tudo isto, talvez seja mais preferível na prática o recurso ao estimador de primeiras-diferenças $\widehat{\beta}_{PD}$ em vez de $\widehat{\beta}_{EF}$. Note-se que quando $T = 2$, $\widehat{\beta}_{PD} = \widehat{\beta}_{EF}$ pois (27) é equivalente a (30). Quando $T > 2$, estes dois métodos já não coincidem. O método de efeitos fixos pode ser aplicado a paineis *unbalaced.*

É interessante notar que o método de diferenças em relação à média temporal (modelo *demeaned* (30)) é idêntico a um método em que o modelo tem um termo independente para cada observação/individuo i que resulta da existência de efeitos individuais α_i. De (17), em que α_i pode estar correlacionado com x_{it},

$$y_{it} = x_{it}^{'}\beta + d_i^{'}\alpha_i + v_{it}, i = 1, ..., n; t = 1, ..., T, \tag{32}$$

onde $\alpha_{n\times 1} = (\alpha_1, ..., \alpha_n)^{'}$ e $d_{i_{n\times 1}}$ é uma coluna de zeros excepto o $i-$ésimo que é um. Em termos matriciais,

$$Y_i = X_i\beta + \iota_T d_i^{'}\alpha_i + V_i, i = 1, ..., n. \tag{33}$$

Dada a hipótese 1 onde v_{it} e α_i não estão autocorrelacionados para todo i e sob a condição de exogeneidade estrita onde

$$E(X_i^{'}V_i) = E(X_1^{'}V_1) = 0 \Leftrightarrow E(x_{1t}^{'}v_{1s}), t, s = 1, ..., T, \tag{34}$$

o OLS (11) para β[5] é consistente e assimptóticamente normal. Pode-se demostrar[6] que a expressão do estimador é

$$\widehat{\beta}_{EF} = \left(\sum_{i=1}^{n} X_i^{'}M_\iota X_i\right)^{-1} \sum_{i=1}^{n} X_i^{'}M_\iota Y_i, \tag{35}$$

onde

$$M_\iota = I_T - \iota_T\left(\iota_T^{'}\iota_T\right)^{-1}\iota_T^{'} = I_T - \frac{1}{T}\iota_T\iota_T^{'}, \tag{36}$$

e para o qual

$$AV\left(\widehat{\beta}_{EF}\right) = \left[E\left(X_1^{'}M_\iota X_1\right)\right]^{-1} E\left(X_1^{'}M_\iota V_1 V_1^{'}M_\iota X_1\right)\left[E\left(X_1^{'}M_\iota X_1\right)\right]^{-1} \tag{37}$$

que pode ser consistentemente estimado por

$$\widehat{AV\left(\widehat{\beta}_{EF}\right)} = \left(\frac{1}{n}\sum_{i=1}^{n} X_i^{'}M_\iota X_i\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^{'}M_\iota\widehat{\Omega}M_\iota X_i\right)\left(\frac{1}{n}\sum_{i=1}^{n} X_i^{'}M_\iota X_i\right)^{-1}, \tag{38}$$

[5] Ou o GLS caso se admita que $E\left(V_1V_1^{'}|X_1\right)_{T\times T} = \Omega \neq \sigma_v^2 I_T$.

[6] Estimação a dois passos (β e α) e com recurso a matrizes de projecção.

onde

$$\widehat{\Omega} = \frac{1}{n} \sum_{i=1}^{n} \left(M_\iota Y_i - M_\iota X_i \widehat{\beta}_{EF} \right) \left(M_\iota Y_i - M_\iota X_i \widehat{\beta}_{EF} \right)^{'} . \tag{39}$$

Se $\widehat{\Omega} = \widehat{\sigma}_v^2 I_T$ quando se assume que $E\left(V_1 V_1^{'} | X_1 \right) = \sigma_v^2 I_T$, então

$$\widehat{AV\left(\widehat{\beta}_{EF}\right)} = \widehat{\sigma}_v^2 \left(\frac{1}{n} \sum_{i=1}^{n} X_i^{'} M_\iota X_i \right)^{-1} . \tag{40}$$

Caso exita interesse na estimação consistente dos $\alpha_i^{\prime}s$ esta pode ser feita num segundo passo após a obtenção de $\widehat{\beta}_{EF}$ desde que dispomos de um suficiente número de graus de liberdade. Note-se que a dimensão de α aumenta com n e que apenas T observações são utilizadas para estimar cada α_i. Com o recurso ao modelo com médias pode-se estimar α_i após obter-se $\widehat{\beta}_{EF}$: $\alpha_i = \overline{y}_i - \overline{x}_i^{'} \widehat{\beta}_{EF}, i = 1, ..., n.$

Para finalizar esta secção discutimos um pormenor na modelização de efeitos fixos. Se os efeitos individuais (podem ser vistos como "regressores") α_i e os regressores x_{it} não estão autocorrelacionados, então pode-se estimar β consistentemente pela partição da matriz de regressores em que a componente $d_i^{'} \alpha_i$ fará parte do erro (compósito e autocorrelacionado, neste caso). Mas então, estamos perante o modelo de efeitos fixos não observáveis da secção anterior e para o qual foi proposto o estimador FGLS=EA onde $\widehat{\Omega} = \widehat{\sigma}_\alpha^2 \iota_T \iota_T^{'} + \widehat{\sigma}_v^2 I_T$. Como se sabe, o (F)GLS é o OLS de um modelo transformado no qual os (novos) erros não estão autocorrelacionados (além da homocedasticidade condicional). Neste modelo de efeitos fixos, o modelo transformado é quase-*demeaned*:

$$\begin{aligned} y_{it} - \lambda \overline{y}_i &= \left(x_{it} - \lambda \overline{x}_i \right)^{'} \beta + \left(v_{it} - \lambda \overline{v}_i \right), i = 1, ..., n; t = 2, ..., T. && (41) \\ \lambda &= 1 - \sqrt{\frac{\sigma_v^2}{\sigma_v^2 + T \sigma_\alpha^2}} && (42) \end{aligned}$$

Se $\lambda = 0$, temos o método OLS; se $\lambda = 1$, o de EF. Portanto, quanto maior é a variabilidade dos efeitos individuais, σ_α^2, mais próximo de EF é o estimador. Quanto menor é a variabilidade dos efeitos individuais, σ_α^2, mais próximo do OLS está o estimador.

0.4 Efeitos Aleatórios versus Efeitos Fixos

O teste de Hausman (apresentado no capitulo de endogeneidade no MRLM) pode ser aplicado, neste contexto de dados de painel, para distinguir efeitos aleatórios de efeitos fixos. O modelo geral é

$$y_{it} = \beta_1 x_{it1} + ... + \beta_k x_{itk} + \alpha_i + v_{it} = x_{it}^{'} \beta + \alpha_i + v_{it}, i = 1, ..., n; t = 1, ..., T \tag{43}$$

e a questão é a de saber se na prática os efeitos individuais α_i não estão ou estão autocorrelacionados com os regressores x_{it}. Se no primeiro caso (ver a discussão no final na última secção) os efeitos são aleatórios; no segundo caso estes são fixos.

A hipótese nula de interesse é de efeitos aleatórios, isto é, de exogeneidade entre α_i e x_{it}. O estimador $\widehat{\beta}_C$ que é consistente sob ambas as hipóteses H_0 e H_1 é o de efeitos fixos $\widehat{\beta}_{EF}$, (35) . O estimador $\widehat{\beta}_E$ que apenas é consistente sob a nula H_0 é o de efeitos aleatórios $\widehat{\beta}_{EA}$, (9) , (23) . Por outro lado, sob $H_0, \widehat{\beta}_{EA}$ é (assimptóticamente) mais eficiente do que $\widehat{\beta}_{EF}$. Portanto, a estatística de Hausman é dada por

$$H = n\left(\widehat{\beta}_{EF} - \widehat{\beta}_{EA}\right)' \left(\widehat{AV\left(\widehat{\beta}_{EF}\right)} - \widehat{AV\left(\widehat{\beta}_{EA}\right)}\right)^{-1} \left(\widehat{\beta}_{EF} - \widehat{\beta}_{EA}\right) \xrightarrow[H_0]{d} \chi^2_k, \tag{44}$$

onde $\widehat{AV\left(\widehat{\beta}_{EF}\right)}$ é dada por (38) , (39) e $\widehat{AV\left(\widehat{\beta}_{EA}\right)}$ por (16) , (23) . Se o valor observado da estatistica H, H_{obs}, é inferior ao valor critico (na distribuição χ^2_k e para um dado nivel de significância - 5%) então $\widehat{\beta}_{EA}$ é preferível a $\widehat{\beta}_{EF}$ e assume-se que nos erros (não observáveis) do modelo não existem processos correlacionados com os regressores. Se H_{obs} se situa na região critica, prefere-se o estimador $\widehat{\beta}_{EF}$ pois admite-se como válida a hipótese de endogeneidade.

Aplicações

(...)

Exercícios

1. Prove (21) .
2. Verifique que quando $T = 2$ (27) é equivalente a (30) , e por isso, $\widehat{\beta}_{PD} = \widehat{\beta}_{EF}$.
3. ...