Monitoria Econometria II

REC 2312 - 1º Semestre 2015

Monitor(a):Victória Mazás Martinez

3ª Lista de exercícios - Dados em painel

*Lembrando que a resolução é apenas uma síntese simplificada das respostas esperadas.

1. Descreva as hipóteses utilizadas nos modelos Pooled, Efeito Fixo (PD e within) e Efeito Aleatório para que os coeficientes sejam estimados consistentemente

   HIPÓTESES DO MODELO POOLED:

   Considere o modelo : $y_t = \mathbf{x}_t\boldsymbol{\beta} + u_{it}$, t= 1, 2, ...T

   Duas Hipóteses suficientes para o pooled OLS estimar consistentemente o $\boldsymbol{\beta}$:

   Hipótese: POLS.1 : $E(\mathbf{x'}_t\mathbf{u}_t) = 0$, t = 1,2,...T . Esta hipótese não nós diz nada sobre a relação entre x_s e u_t para $s \neq t$. Hipótese: POLS.2 : $rank[\sum_{t=1}^{T} E(\mathbf{x'}_t\mathbf{x}_t)] = K]$. Esta hipótese essencialmente exclui a perfeita dependência linear entre as variáveis explicativas.

   HIPÓTESES DO MODELO DE EFEITOS FIXOS

   Considere o modelo : $y_t = \mathbf{x}_t\boldsymbol{\beta} + c_i + u_{it}$, t= 1, 2, ...T

   No modelo de efeitos fixos, supomos $E(\mathbf{x}'_t c) \neq 0$

   *Primeira diferença

   $\Delta y = \Delta x\boldsymbol{\beta} + \Delta u$

   As condições chaves para o OLS estimar consistentemente o $\boldsymbol{\beta}$ do modelo acima são:

   Condição de ortogonalidade : $E(\Delta x'\Delta u) = 0$ e, Condição de rank : $E(\Delta\mathbf{x}'\Delta\mathbf{x}) = K$

*transformação within

$y_{it} - \bar{y}_i = (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)\boldsymbol{\beta} + (u_{it} - \bar{u}_i)$

reescrevendo

$\ddot{y}_{it} = \ddot{\mathbf{x}}_{it}\boldsymbol{\beta} + \ddot{u}_{it}, \quad t = 1, 2, ..., T$

HIPÓTESE $E(u_{it}|\mathbf{x}_i, c_i) = 0, \quad t = 1, 2, ...T$ (exogeneidade estrita de $\mathbf{x}_{it} : t : 1, ..., T$ condicionada ao efeito não observado.

HIPÓTESES DO MODELO DE EFEITOS ALEATÓRIOS

Considere o modelo : $y_t = \mathbf{x}_t\boldsymbol{\beta} + c_i + u_{it}$, t= 1, 2, ...T

Hipóteses:

$E(u_{it}|x_i, c_i) = 0, \quad t = 1, ..., T$ (exogeneidade estrita) $E(c_i|x_i) = E(c_i) = 0$

rank $E(\mathbf{X}_i'\boldsymbol{\Omega}^{-1}\mathbf{X}_i) = \mathbf{K}$

2. Suponha um modelo para estimar o efeito de uma formação profissional sobre os salários futuros:

$$log(salario_{it}) = \theta_t + \mathbf{z}_{it}\boldsymbol{\gamma} + \delta_1 prog_{it} + c_i + u_{it}$$

onde i indexa o indivíduo e t indexa o período de tempo. O parâmetro θ_t denota um intercepto invariante no tempo, z_{it} é um conjunto de características observáveis que afetam salário e podem também ser correlacionados com a participação no programa e c_i é um efeito individual não observável.

Um conjunto de dados é coletado em dois períodos de tempo. Em $t = 1$,ninguém estava participando do programa, sendo assim, $prog_{i1} = 0$ para todo i. Então um subgrupo passa a frequentar o programa e o subsequente salário é observado para os grupos de controle(não participantes) e tratamento(participantes) em $t = 1$.

(a) Discuta porque podemos ter viés de variável omitida neste caso.

Neste caso, podemos argumentar que escolher participar do programa de formação profissional pode estar relacionado à habilidade, ou seja, indvíduos mais habilidosos se auto-selecionariam para participar do programa.

(b) Discuta se podemos assumir exogeneidade estrita neste modelo.

Podemos assumir que u_{it} é não correlacionado com $prog_it$, mas a participação futura pode pode depender de u_{it} se as pessoas escolhem participar no futuro baseado nos choques sobre seus salários no passado, portanto $prog_{i,t+1}$ será correlacionado com u_{it}.

3. Um modelo dinâmico simples de determinação de salario com heterogeneidade não observada é:

$$log(salario_{it}) = \beta_1 log(salario_{i,t-1}) + c_i + u_{it}, \quad t = 1, 2, ...T$$

(a) O interesse reside em como os salários são persistentes (mensurados pelo tamanho de β_1) depois de controlar pela heterogeneidade não observada (produtividade individual), c_i. Mostre que a hipótese de exogeneidade estrita não se mantém em modelos de efeitos não observados com variáveis dependentes defasadas.

Escreva $y_{it} = log(salario_{it})$, a hipótese padrão seria:

$$E(u_{it}|y_{i,t-1}, ..., y_{i0}, c_i) = 0 \tag{1}$$

o que significa que todas as dinâmicas são capturados pelo primeiro lag. Deixe $x_{it} = yi, t-1$. Então, sob a hipótese (1), u_{it} é não correlacionado com $(x_{it}, x_{i,t-1}, ..., x_{i1})$, mas u_{it} não pode ser não correlacionado com $(x_{i,t+1}, ..., x_{iT})$, pois $x_{i,t+1} = y_{it}$. Portanto,

$$E(y_{it}u_{it}) = \beta_1 E(y_{i,t-1}u_{it} + E(c_i, u_{it}) + E(u_{it}^2) > 0 \tag{2}$$

uma vez que, $E(y_{i,t-1}u_{it}) = 0$ e $E(c_i, u_{it}) = 0$ sob a hipótese (1). Portanto, a hipótese de exogeneidade estrita nunca se mantém no modelo de efeitos não observados com variável depedente defasadas.

4. Explique o que ocorre quando nosso modelo de efeito fixo apresenta uma variável explicativa que é constante no tempo para todos os indivíduos. Explicite que hipótese do modelo é quebranda quando isto ocorre.

par Se x_{it} contém um elemento que não varia ao longo do tempo para qualquer i, então o correspondente elemento em $\ddot{x}_{it}$ é identicamente zero para todo t. Uma vez que, $\ddot{\mathbf{X}}_i$ contém uma coluna de zeros para todo i, a Hipótese FE.2 não pode ser verdadeira

5. Apresente o teste de Hausman no contesto de dados em painel (EF e EA)

A escolha entre os modelo de Efeito fixo e Efeito aleatório vai depender se c_i e x_i são correlacionados. O teste de Hausmam testa esta hipótese, baseado na diferença entre as estimativas de efeitos fixos e efeitos aleatórios. Uma vez que FE (Efeitos fixos) é consistente quando c_i e x_it são correlacionados, mas RE (Random effects) é inconsistente, uma significante diferença estatística é interpretada como evidência contra a hipotese de efeito aleatório.

Usualmente implementamos o teste de Hausman assumindo que há homocedasticidade sob a hipótese nula, o que implica que o estimador de efeitos aleatórios será mais eficiente que o estimador de efeito fixos. Assumindo que as hipóteses de Efeito Aleatório são mantidas, considerando o caso onde x_{it} contém somente elementos que variam no tempo. A estatística de Hausman pode ser computada como segue. Deixe $\hat{\boldsymbol{\delta}}_{RE}$ denotar o vetor de Efeitos Aleatórios estimados, e deixe $\hat{\boldsymbol{\delta}}_{FE}$ denotas as correspondentes estimativas de Efeitos Fixos; seja estes um vetor Mx1 . Então o teste será :

$$H_0 = \hat{\boldsymbol{\delta}}_{FE} - \hat{\boldsymbol{\delta}}_{RE} = 0$$

$$W = (\hat{\boldsymbol{\delta}}_{FE} - \hat{\boldsymbol{\delta}}_{RE})'[A\hat{var}(\hat{\boldsymbol{\delta}}_{FE}) - A\hat{var}(\hat{\boldsymbol{\delta}}_{RE})]^{-1}(\hat{\boldsymbol{\delta}}_{FE} - \hat{\boldsymbol{\delta}}_{RE})$$

é distribuído assintóticamente como uma X^2_M sob as hipóteses de Efeito Aleatório.

*Obs: Quando nosso interesse reside sobre apenas um parâmetro podemos computar o teste de Hausman utilizando uma estatística t : $W = (\hat{\delta}_{FE} - \hat{\delta}_{RE})/\{[dp(\hat{\delta}_{FE})]^2 - [dp(\hat{\delta}_{RE})]^2\}^{1/2}$

Na prática, uma falha em rejeitar significa que ou as estimativas EA e EF são suficientemente próximas que não importa qual será usada, ou a variação amostralé tão grande nas estimativas EF que não se pode concluir se diferenças praticamente significantes são estatísticamente significantes