Descrição das variáveis de Estudo

Quadro com as principais variáveis utilizada

A pesquisa utilizou a regressão de dados em painel, com uma amostra de 40 empresas, envolvendo 320 observações (painel não-balanceado). A unidade básica de estudo é representada por empresas, observadas em diferentes instantes do tempo (de 2010 a 2018). O objetivo da utilização da regressão em painel (ou dados longitudinais) é verificar a relação ente legibilidade da norma contábil e a legibilidade da nota explicativa das empresas, mas levando-se em consideração o tempo e as características individuais das mesmas.

Acredito que deveria ser feita um modelo para cada uma das categorias de nota explicativa. Nesse caso a estudo conta com 23 categorias de notas explicativas.

Assim, os modelos de dados em painel diferem dos modelos com dados temporais e *cross section* dado o caráter duplo que atribui a cada variável. De acordo Hill, Judge e Griffiths (2010) o modelo geral para os dados em painel é representado por:

(5)

Com: i =1,..., N os indivíduos (N indivíduos, países, regiões, empresas, setores); t=1,...,T os períodos de tempo que está sendo analisado (T períodos); β0 = parâmetro de intercepto; βk = coeficiente angular correspondente à k-ésima variável explicativa do modelo.

Se para cada indivíduo *i* dispõe-se do mesmo número de dados temporais, o painel chama-se *balanceado* (ou equilibrado). Se o número de dados temporais não é o mesmo para todos os indivíduos, o painel denomina-se de *não-balanceado*.

O teste de Hausman (1978) foi utilizado para decidir qual dos modelos é o mais apropriado: o modelo de efeitos aleatórios (Ho) ou o modelo de efeitos fixos (HA). O teste apresenta-se da seguinte forma:

H0: Cov (ai, Xit) = 0 (efeitos aleatórios)

HA: Cov (ai, Xit) ≠ 0 (efeitos fixos)

Sob a hipótese nula, os estimadores do modelo com efeitos aleatórios são consistentes e eficientes. Sob a hipótese alternativa, os estimadores MQG com efeitos aleatórios (e MQO) são não consistentes, mas os estimadores com efeitos fixos são. Esta é uma das vantagens dos modelos com efeitos fixos, uma vez que permite a endogeneidade dos regressores.

Desta forma, existem alguns modelos diferentes que podem ser utilizados para dados em painel. Assim, para a escolha dos modelos em painel, nesse artigo, por efeito fixo, aleatório ou *Pooled* foi aplicado os testes de Breusch-Pagan, Chow e Hausman. Para o primeiro teste rejeita-se a menos de 1% a hipótese nula. Portanto, o modelo estimado por efeitos aleatórios mostra-se mais adequado do que o modelo *pooled* (*pooled cross-section*).

Posteriormente, foi aplicado o teste de Chow. Rejeita-se a menos de 1% a hipótese nula. Portanto, o modelo estimado por efeitos fixos mostra-se mais adequado do que o modelo *pooled*. Após Teste de Breusch-Pagan e Chow, descarta-se o modelo *pooled*. Por último, demonstra que o efeito fixo foi a melhor opção, comparado o aleatório. Foi utilizado o software Stata 13 para a realização das análises empíricas. Todas as estimações foram realizadas, utilizando-se o comando *robust* para correção de qualquer tipo de heterocedasticidade.

Análise Descritiva dos variáveis utilizadas no Modelo

Nesse sentido, o próximo passo consiste em analisar os resultados das principais medidas de análise em estatística descritiva, que se encontram no tópico a seguir (Tabela 2).

Tabela 2 – Análise descritiva dos indicadores

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | | | | |
|  | N | Mínimo | Máximo | Média | Erro Desvio |
| LegNE | 4614 | 0 | 85 | 28,67 | 12,999 |
| LegCPC | 7360 | 29 | 80 | 46,87 | 10,485 |
| TAM | 7360 | 804536,00 | 900135000,00 | 52917180,1906 | 121971307,00073 |
| COMPLEX | 7360 | 1 | 7 | 3,44 | 1,576 |
| CAPIT | 7360 | 0 | 31 | 18,62 | 9,166 |
| EXT | 7360 | 20 | 231 | 86,33 | 35,491 |
| ADR | 7360 | 0 | 1 | ,38 | ,484 |
| TAMLOG | 7360 | 5,91 | 8,95 | 7,3179 | ,53960 |
| N válido (de lista) | 4614 |  |  |  |  |

Fonte: Fonte da Pesquisa

Análise Descritiva das variáveis por países e setores

1. Teste de Normalidade SW







Fazer os teste para inserir nessa seção.

Tratamento de outliers



Após o tratamento







1. Teste de Normalidade Assimetria e Curtose

Com a finalidade de avaliar se a amostra coletada é oriunda de uma população a qual apresenta normalidade, Hair Jr. *et al.* (2009) e Kline (2011) recomendam que sejam analisadas as medidas de assimetria e curtose da amostra. Fávero *et al.* (2009, p. 58), explica que a assimetria corresponde ao “grau de desvio ou achatamento, da simetria de uma distribuição, ou seja, se uma distribuição for simétrica, os valores da média, da mediana e da moda serão iguais”. Em relação à curtose, ou achatamento, os autores definem que trata-se da “altura do ponto máximo da curva de distribuição” (FÁVERO *et al*, 2009, p. 60).

De acordo com Hair Jr. *et al.* (2009) e Kline (2011), a amostra segue uma distribuição normal quando os valores de assimetria são inferiores a 3 e os de curtose são menores do que 8. Os resultados das análises dos cálculos dos construtos usados nesta pesquisa indicam que seus valores de assimetria e curtose estão dentro dos limites apontados pelos autores, conforme se pode verificar Tabela 1. Foi necessário utilizar a transformação logarítmica na base 10 para a variável TAM .

No caso dos dados que são Omissos teria que ver se realmente iremos trabalhar com esses no modelo ...

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Teste de curtose e Assimetria** | | | | | | | | | |
|  | | LegNE | LegCPC | COMPLEX | CAPIT | EXT | ADR | TAMLOG |
| N | Válido | 4614 | 7360 | 7360 | 7360 | 7360 | 7360 | 7360 |
| Omisso | 2746 | 0 | 0 | 0 | 0 | 0 | 0 |
| Assimetria | | 0,218 | 1,340 | -0,064 | -0,199 | 0,933 | 0,517 | 0,411 |
| Erro de assimetria padrão | | 0,036 | 0,029 | 0,029 | 0,029 | 0,029 | 0,029 | 0,029 |
| Curtose | | 0,302 | 2,001 | -0,885 | -1,475 | 2,095 | -1,734 | 0,650 |
| Erro de Curtose padrão | | 0,072 | 0,057 | 0,057 | 0,057 | 0,057 | 0,057 | 0,057 |

1. Teste de Verificação de Multicolinearidade

Refere-se a existência de correlação alta entre duas ou mais variáveis independentes.  
Resultado do teste de Multicolinearidade com as variáveis explicativas do modelo.

|  |  |  |  |
| --- | --- | --- | --- |
| Modelo | |  | |
| VIF | 1/VIF |
|  | LegCPC | 1,006 | 0,9939627 |
| TAMLOG | 2,829 | 0,3534215 |
| COMPLEX | 1,309 | 0,7641322 |
| ADR | 1,938 | 0,5160561 |
| RevCPC | 1,005 | 0,9950947 |
| GC | 1,512 | 0,6615079 |
| EXT | 1,424 | 0,7021447 |
| CAPIT | 2,055 | 0,4866306 |
| AUDIT | 1,218 | 0,8208785 |
|  | Media VIF | 1,588 |  |

Fonte: dados da pesquisa

Critério VIF < 10 e Vif médio < 10

Como observado não a problemas relacionados a multicolinearidade entre as variáveis, sendo assim, não houve indicativo para remoção de nenhuma das variáveis do modelo.

1. Teste de autocorrelação

Será utilizado o Teste de wooldridge para testar se existe autocorreção ou correlação serial quando os erros ou perturbações da regressão são correlacionados ao longo do tempo violando a hipótese de que os erros são aleatórios ou não correlacionados.

1. Teste heterocedasticidade
2. Analise da correlação múltipla
   1. Qual modelo devera ser escolhido. São 3.
3. Resultados e considerações sobre o modelo

**Referencias**

FÁVERO, Luiz Paulo et al. **Análise de Dados:**Modelagem multivariada para tomada de decisões. Rio de Janeiro: Elsevier, 2009.

HAIR JR.; Joseph F.; BLACK, William C.; BABIN, Barry J.; ANDERSON, Rolph E.; TATHAM, Ronald L. **Análise Multivariada de Dados**. 6. ed. Porto Alegre: Bookman, 2009.

KLINE, Rex B. **Principles and Practice of Structural Equation Modeling**. 3. ed. New York: The Guilford Press, 2011.