Descrição das variáveis de Estudo

Tratamento para variáveis ...

Proxy das variáveis

Categoria de Notas Explicativas

**Descrição da Variável Dependente**

Por meio das hipóteses de pesquisa e do referencial teórico estudado, foi definida a variável dependente bem como as variáveis de independentes com a finalidade de viabilizar a questão de pesquisa.

Somatório das **Notas Explicativas** / **Qtt** de notas explicativas encontradas naquele respectivo ano.

**Descrição da Variáveis Independente**

Somatório do valor da **LegCPC** / **Qtt** de explicativas encontradas naquele respectivo ano

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.

Como iremos inserir no modelo o setores de referencias..

Objetiva-se também verificar há possibilidade de existirem diferenças entre os setores, como por exemplo o Y, que possuem regulação além da própria CVM, realizou-se as análises dos modelos anteriores com a inserção de *dummies* para o controle por setor, a fim de eliminar possíveis impactos decorrentes dos mesmos.

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 das variáveis utilizadas no Modelo**

Nesse sentido, o próximo passo consiste em analisar os resultados das principais medidas de análise em estatística descritiva a partir da analise de tendência central, dispersão e amplitude que são dados relevantes para compreender o comportamento das variáveis do modelo. Na tabela **XI** retrata as estatísticas descritivas para as variáveis dependente, independente e de controle. De modo geral, nota-se variabilidade entre as métricas estudadas, além disso, ressalta-se que nessa primeira análise dos dados não foram feitos tratamento para os outliers. A variável

**Tabela XI:** Estatística descritiva das variáveis utilizadas na pesquisa

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variável** | **Média** | **Desvio Padrão** | **Mínimo** | **Máximo** |
| LegNEMedio | 28,754 | 3,892 | 15,750 | 38,53333 |
| LegCPCMedio | 45,901 | 1,188 | 43,308 | 49,26667 |
| RevCPC | 0,375 | 0,485 | 0,000 | 1 |
| TAM\_log | 52900000,000 | 1,220 | 804536,000 | 900000000 |
| COMPLEX | 3,444 | 1,579 | 1,000 | 7 |
| CAPIT | 18,625 | 9,180 | 0,000 | 31 |
| GC | 0,54375 | 0,4988623 | 0 | 1 |
| AUDIT | 0,938 | 0,242 | 0,000 | 1 |
| EXT | 86,328 | 35,544 | 20,000 | 231 |
| ADR | 0,375 | 0,485 | 0,000 | 1 |

Fonte: Elaborado pelo autor

Primeiramente fazer a análise não considerando os dados com transformação.

Tabela 2 – Análise descritiva das variáveis da pesquisa

Fonte: dados da pesquisa

|  |
| --- |
|  |

**Teste de verificação de adequação para aplicação das técnicas**

Com a finalidade de minimizar os problemas que por ventura existirem na aplicação do modelo em painel, foram realizados testes para especificação para as variáveis e modelos de regressão em Painel.

**Teste de Normalidade**

Primeiramente foi verificado se os dados seguiam distribuição normal. Portanto, através do teste Shapiro-Wilk as variáveis dependente e independente foram submetidas à estatística W.

**Tabela XII:** Teste de Normalidade (Shapiro-Wilk)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variável** | **W** | **V** | **Z** | **Prob>z** |
| LegNEMedio | 0,990 | 2,240 | 1,899 | 0,029 |
| LegCPCMedio | 0,990 | 2,148 | 1,800 | 0,036 |
| TAM | 0,378 | 140,331 | 11,641 | 0,000 |
| COMPLEX | 0,982 | 4,101 | 3,323 | 0,000 |
| CAPIT | 0,921 | 17,894 | 6,792 | 0,000 |
| EXT | 0,949 | 11,421 | 5,735 | 0,000 |

Fonte: Elaborado pelo autor

Segundo os resultados apresentados na tabela XII, rejeitamos a hipótese nula de dados normalmente distribuídos ao nível de significância de 5%. Os valores de grande de V´ indicam a não normalidade dos dados.

**Teste de Assimetria e Curtose**

Para testar formalmente se os resíduos seguem distribuição normal foi utilizado o teste de assimetria e curtose dos resíduos. Assim, com base no “valor p”, análise conjunta, foi possível rejeitar a hipótese nulo de que os dados possuem distribuição normal. No teste, foi utilizado o parâmetro “noadjust” e também

**Tabela XIII:** Teste de Assimetria e Curtose

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variáveis** | **Pr(Skewness)** | **Pr(Kurtosis)** | **chi2(2)** | **Prob>chi2** |
| LegNEMedio | 0,148 | 0,120 | 4,520 | 0,105 |
| LegCPCMedio | 0,906 | 0,297 | 1,100 | 0,576 |
| TAM | 0,339 | 0,000 | 13,930 | 0,001 |
| COMPLEX | 0,633 | 0,000 | 35,240 | 0,000 |
| EXT | 0,103 | 0,000 | 22,630 | 0,000 |

Fonte: Elaborado pelo autor

Com a finalidade de corrigir *outlier* e foram

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **W** | **V** | **Z** | **Prob>z** |
| LegNEMedio | 0,990 | 2,240 | 1,899 | 0,029 |
| LegCPCMedio | 0,995 | 1,047 | 0,109 | 0,457 |
| \*wlTAM | 0,987 | 2,918 | 2,521 | 0,006 |
| COMPLEX | 0,982 | 4,101 | 3,323 | 0,000 |
| CAPIT | 0,921 | 17,894 | 6,792 | 0,000 |
| \*WsqEXT | 0,989 | 2,377 | 2,039 | 0,021 |

Fonte: Elaborado pelo autor

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 .

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.

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. Sendo assim, as hipóteses deste teste são:

H0: Ausência de autocorrelação no modelo e;

H1: Presença de autocorrelação no modelo.

Quando foi detectado o problema de autocorrelação, no primeiro momento, utilizou-se a estratégia de correção serial AR(1), processo auto regressivo de primeira ordem, com o objetivo de tratar esse problema (WOOLDRIDGE, 2002).

**Tabela XV:** Teste Wooldridge para autocorrelação

|  |  |  |
| --- | --- | --- |
| **Equação** | **F** | **Prob > F** |
| **01** | F( 1, 39) = 11.112 | 0.0019 |

Fonte: Elaborado pelo autor

**Tabela XVI:** Teste Wald para heterocedasticidade

|  |  |  |
| --- | --- | --- |
| **Equação** | **chi2** | **Prob > chi2** |
| **01** | chi2(40) = 2747.02 | 0.0000 |

Fonte: Elaborado pelo autor

**Tabela XVII:** Matriz de Correlação das variáveis de pesquisa

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | LegNEMedio | cLegCPCMedio | RevCPC | wlTAM | COMPLEX | CAPIT | GC | AUDIT | WsqEXT | ADR |
| LegNEMedio | 1.00 |  |  |  |  |  |  |  |  |  |
| LegCPCMedio | -0.1904\* | 1.00 |  |  |  |  |  |  |  |  |
| RevCPC | -0.0178 | 0.0219 | 1.00 |  |  |  |  |  |  |  |
| wlTAM | -0.0445 | 0.0392 | 0.0495 | 1.00 |  |  |  |  |  |  |
| COMPLEX | -0.0799 | 0.1985\* | 0.0072 | 0.3868\* | 1.00 |  |  |  |  |  |
| CAPIT | 0.0466 | -0.1671\* | 0.0423 | 0.6452\* | 0.2023\* | 1.00 |  |  |  |  |
| GC | 0.0554 | 0.0223 | 0.0097 | -0.4765\* | -0.1441\* | -0.4715\* | 1.00 |  |  |  |
| AUDIT | -0.2250\* | 0.2523\* | -0.0667 | -0.1245\* | 0.2611\* | 0.0458 | 0.0486 | 1.00 |  |  |
| WsqEXT | -0.1962\* | 0.3196\* | -0.0127 | 0.4142\* | 0.3763\* | 0.1109\* | -0.2729\* | 0.1911\* | 1.00 |  |
| ADR | -0.1520\* | -0.0609 | 0.0000 | 0.6552\* | 0.3430\* | 0.5078\* | -0.2236\* | 0.0400 | 0.3964\* | 1.00 |

Fonte: Elaborado pelo autor

Estimações Econométricas

Escolha do melhor modelo

Adotadas as verificações e correções descritas anteriormente, estimou-se o modelo de regressão por mínimos quadrados ordinários (MQO), dados empilhados, para análise da hipótese principal de que a legibilidade da norma contábil afeta a legibilidade das notas explicativas emitidas pelas companhias abertas brasileiras.

* 1. Qual modelo devera ser escolhido. São 3.

1. 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.