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 dos 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 da variáveis do modelo. Os dados se no tópico a seguir (Tabela 2).

Primeiramente fazer a análise não considerando os dados com transformação.

Tabela 2 – Análise descritiva das variáveis da pesquisa

Fonte: dados 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 .

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).

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.