

A FEA e a USP respeitam os direitos autorais deste trabalho. Nós acreditamos que a melhor proteção contra o uso ilegítimo deste texto é a publicação online. Além de preservar o conteúdo motiva-nos oferecer à sociedade o conhecimento produzido no âmbito da universidade pública e dar publicidade ao esforço do pesquisador. Entretanto, caso não seja do interesse do autor manter o documento online, pedimos compreensão em relação à iniciativa e o contato pelo e-mail bibfea@usp.br para que possamos tomar as providências cabíveis (remoção da tese ou dissertação da BDTD).

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE CONTABILIDADE E ATUÁRIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS CONTÁBEIS

**AVALIAÇÃO DE EFICIÊNCIA DE EMPRESAS ATRAVÉS DE INDICADORES
CONTÁBEIS: UTILIZAÇÃO DA ANÁLISE POR ENVOLTÓRIA DE DADOS,
ANÁLISE DE AGRUPAMENTOS E ANÁLISE DE REGRESSÃO**

Luiz João Corrar

**T657.3
C823a**

SÃO PAULO

2009

Profa. Dra. Suely Vilela
Reitora da Universidade de São Paulo

Prof. Dr. Carlos Roberto Azzoni
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Fábio Frezatti
Chefe do Departamento de Contabilidade e Atuária

Prof. Dr. Edgard Bruno Cornachione Junior
Coordenador do Programa de Pós-Graduação em Ciências Contábeis

T657.3
C823a



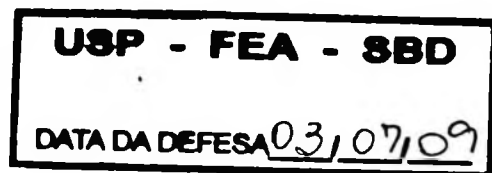
LUIZ JOÃO CORRAR

DEDALUS - ACERVO - FEA



**AVALIAÇÃO DE EFICIÊNCIA DE EMPRESAS ATRAVÉS DE INDICADORES
CONTÁBEIS: UTILIZAÇÃO DA ANÁLISE POR ENVOLTÓRIA DE DADOS,
ANÁLISE DE AGRUPAMENTOS E ANÁLISE DE REGRESSÃO**

Tese apresentada ao Departamento de Contabilidade e Atuária da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo como requisito para obtenção do título de Livre-Docente em Contabilometria.



SÃO PAULO

2009

98318

Tese defendida e aprovada no Departamento de Contabilidade e Atuária da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo - Programa de Pós-Graduação em Ciências Contábeis, pela seguinte banca examinadora:

FICHA CATALOGRÁFICA

Elaborada pela Seção de Processamento Técnico do SBD/FEA/USP

Corrar, Luiz João

Avaliação de eficiência de empresas através de indicadores contábeis: utilização da análise por envoltória de dados, análise de agrupamentos e análise de regressão / Luiz João Corrar. – São Paulo, 2009.

240 p.

Tese (Livre-Docência) – Universidade de São Paulo, 2009

Bibliografia.

1. Balanço contábil – Análise
2. Avaliação de desempenho organizacional
3. Contabilidade I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade II. Título.

Para Tania Tamer Corrar.

AGRADECIMENTOS

Em primeiro lugar, agradeço à minha esposa Tania Tamer Corrar, pelo carinho, amizade, dedicação, esforço, sacrifício, renúncia e incentivo.

Aos meus pais, Luiz Corrar e Carmelinda Loiero Corrar.

Aos meus amigos, Mauro B. Leiva, Milton C. S. Mello e Waldomiro Piedade Filho.

Aos professores que me deram oportunidade e incentivo na carreira profissional: Almir Ferreira de Souza, Klaus Dietmar Alvarez, Masayuki Nakagawa, Geraldo Barbieri, Marina Mitiyo Yamamoto, Ariovaldo dos Santos, Rubens Fama.

Às pessoas que ajudaram na elaboração deste trabalho: Maria Lucia Castiglia, Fernanda Meirelles Salotti, Nivaldo Lamac, Sofie T. Aversari Martins, Belinda Ludovici, Dulcinéia Jacomini.

Aos professores que contribuíram de alguma forma para a realização deste trabalho: Ariovaldo dos Santos, Marcelo Alvaro da Silva Macedo, Silvia Pereira de Castro Casanova, Gilberto de Andrade Martins, Marco Antonio Figueiredo Milani Filho, Patricia Varella, Romildo de Oliveira Moraes, Emanuel Rodrigues Junqueira de Matos.

E, enfim, a todos os meus colegas professores do Departamento de Contabilidade e Atuária da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo e em especial: Eliseu Martins, Iran Siqueira Lima, Reinaldo Guerreiro, Fabio Frezatti, José Roberto Kassai, Edson Luiz Riccio, Welington Rocha, L. Nelson Carvalho, Diogo Toledo do Nascimento, João Domiraci Paccez, Lazaro Plácido Lisboa, Nilton Cano Martin, Sergio de Iudícibus, Jerônimo Antunes, Bruno Meirelles Salotti, Carlos Alberto Pereira, Edgard Bruno Cornachione, Gerlando Lima, Alexsandro Broedel Lopes, Luis Eduardo Afonso, Joanilia Neide de Sales Cia, Roberta Carvalho de Alencar, pois todos, direta ou indiretamente, colaboraram para a execução deste trabalho.

Agradeço à Diretoria da Fundação Instituto de Pesquisas Contábeis, Atuariais e Financeiras – FIPECAFI pelo constante apoio e por ceder o banco de dados relativo ao anuário Melhores & Maiores.

RESUMO

A presente pesquisa busca abordar a avaliação de eficiência de empresas por meio do emprego de indicadores contábeis e do uso de diferentes instrumentos quantitativos como: Análise por Envoltória de Dados, Análise de Agrupamentos, Análise de Regressão e Testes de Hipóteses Paramétricos e Não Paramétricos. Para tanto, efetuam-se testes para avaliar se os indicadores contábeis obtidos com a técnica Delphi e com base nos dados do anuário Melhores e Maiores da Revista Exame são significativos para a avaliação de eficiência de empresas por usuários externos. O estudo também procura mostrar que é possível considerar a eficiência de empresas através de um pequeno número de indicadores contábeis selecionados e é apresentado um método para a separação de empresas em grupos com características homogêneas. Do mesmo modo, a pesquisa demonstra como usar técnicas estatísticas paramétricas em complementação ao modelo de análise por envoltória de dados para testar hipóteses sobre significância de indicadores contábeis. Além disso, demonstra que é possível criar modelos de previsão de níveis de eficiência através da utilização de indicadores contábeis conjugados com a técnica de análise por envoltória de dados e a análise de regressão. Este estudo tem caráter empírico, analítico e exploratório, e nele são adotados os resultados da pesquisa de Lyra (2008), que empregou a técnica Delphi na seleção de indicadores contábeis para avaliação de desempenho. Através da aplicação de análise por envoltória de dados, análise de agrupamentos e análise de regressão e com a utilização dos dados do anuário Melhores e Maiores, é possível constatar que a técnica Delphi, parte do estudo de Lyra (2008), mostra-se eficiente na seleção de indicadores contábeis significativos, podendo ser indicada para os futuros pesquisadores. De acordo com os resultados, não é necessário utilizar grande número de indicadores para avaliar eficiência de empresas desde que os indicadores utilizados sejam selecionados de forma eficiente. Ainda por meio dos dados obtidos, verifica-se que é plausível separar as empresas da economia em grupos homogêneos, através do método de análise de agrupamentos. Nota-se que o modelo de análise por envoltória de dados, além de eficiente para avaliar desempenho, pode ser complementado por outras técnicas como análise de regressão e técnicas estatísticas paramétricas e não paramétricas. Por fim, os resultados evidenciam que é possível criar modelos de previsão de níveis de eficiência através da utilização de indicadores contábeis conjugados com o modelo de análise por envoltória de dados e a análise de regressão.

ABSTRACT

This research looks at company efficiency assessment through the use of accounting indicators and different quantitative instruments, including: Data Envelopment Analysis, Cluster Analysis, Regression Analysis and Parametric and Nonparametric Tests of Hypotheses. Tests were performed to assess whether the accounting indicators obtained through the Delphi technique and based on the *Melhores e Maiores* database by the magazine *Exame* are significant for the sake of company efficiency assessment by external users. The study also attempts to show that company efficiency can be considered through a small number of selected accounting indicators. A method is presented to divide companies into groups with homogeneous characteristics. Likewise, the research demonstrates how parametric statistical techniques can be used to complement the data envelopment analysis model with a view to testing hypotheses on the significance of accounting indicators. Moreover, it is demonstrated that efficiency level forecasting models can be created through the use of accounting indicators in combination with the data envelopment analysis technique and regression analysis. This empirical, analytical and exploratory study adopted the research results reported in Lyra (2008), who used the Delphi technique to select accounting indicators for performance assessment. The application of data envelopment analysis, cluster analysis and regression analysis, using data from the *Melhores e Maiores* yearbook, shows that the Delphi technique, part of the study by Lyra (2008), is efficient to select significant accounting indicators and can be indicated to future researchers. According to the results, there is no need to use a large number of indicators to assess company efficiency. The obtained data also reveal that companies can be divided in homogeneous groups through the cluster analysis method. It is observed that the data envelopment analysis model is efficient for performance assessment and can be complemented by other techniques, such as regression analysis and parametric and nonparametric statistical techniques. Finally, the results evidence that efficiency level forecasting models can be created through the use of accounting indicators in combination with the data envelopment analysis model and regression analysis.

SUMÁRIO

LISTA DE ABREVIATURAS E SIGLAS	3
LISTA DE TABELAS	5
LISTA DE FIGURAS	9
1 INTRODUÇÃO	11
1.1 Contextualização	11
1.2 Questões de pesquisa	14
1.3 Objetivos	14
1.4 Metodologia	15
1.5 Organização do Trabalho	16
2 REVISÃO BIBLIOGRÁFICA	17
2.1 <i>Data Envelopment Analysis</i> (DEA)	19
2.2 A condição da homogeneidade das DMUs	27
2.3 <i>Cluster Analysis</i>	32
2.4 Indicadores contábeis	41
2.4.1 Seleção de indicadores contábeis através da técnica Delphi	42
2.4.2 Indicadores contábeis utilizados neste estudo	45
2.4.3 O banco de dados do anuário Melhores & Maiores	46
3 METODOLOGIA E ANÁLISE DOS RESULTADOS	49
3.1 Características da pesquisa	49
3.2 Revisão bibliográfica	50
3.3 Indicadores contábeis utilizados	50
3.4 O banco de dados do anuário Melhores & Maiores	52
3.5 O tratamento dos <i>outliers</i>	54
3.6 Utilização da <i>Cluster Analysis</i> para separar empresas em grupos homogêneos ..	55
3.7 <i>Cluster Analysis</i> e as opções do programa SPSS	57
3.8 Análise dos resultados da <i>Cluster Analysis</i>	58
3.9 Avaliação da eficiência das empresas	70
3.10 Avaliação dos resultados do <i>cluster</i> 1	72
3.11 Avaliação dos resultados do <i>cluster</i> 2	77
3.12 Avaliação dos resultados do <i>cluster</i> 3	81
3.13 Avaliação dos resultados do <i>cluster</i> 4	85
3.14 Avaliação dos resultados do <i>cluster</i> 5	89
3.15 Avaliação dos resultados do <i>cluster</i> 6	93
3.16 Avaliação dos resultados do <i>cluster</i> 7	97
3.17 Hipóteses a serem testadas	101
3.18 Regressão e análise dos resultados do <i>cluster</i> 1: variáveis estabelecidas pela Delphi	102
3.19 Regressão e análise dos resultados do <i>cluster</i> 2: variáveis estabelecidas pela Delphi	106
3.20 Regressão e análise dos resultados do <i>cluster</i> 3: variáveis estabelecidas pela Delphi	111
3.20.1 <i>Cluster</i> 3: Mais testes para a normalidade dos resíduos da regressão	116
3.21 Regressão e análise dos resultados do <i>cluster</i> 4: variáveis estabelecidas pela Delphi	120
3.21.1 <i>Cluster</i> 4: Mais testes para a normalidade dos resíduos da regressão	126
3.22 Regressão e análise dos resultados do <i>cluster</i> 5: variáveis estabelecidas pela Delphi	130
3.23 Regressão e análise dos resultados do <i>cluster</i> 6: variáveis estabelecidas pela Delphi	134

3.23.1	Cluster 6: Mais testes para a normalidade dos resíduos da regressão	138
3.24	Regressão e análise dos resultados do <i>cluster</i> 7: variáveis estabelecidas pela Delphi	142
3.25	Incluindo todas variáveis para regressão e análise dos resultados dos <i>clusters</i> ..	147
3.25.1	Regressão e análise dos resultados do <i>cluster</i> 1: incluindo todas as variáveis	148
3.25.1.1	Cluster 1: Análise mais detalhada de outras variáveis significativas .	153
3.25.1.1.1	Dummy Giro do ativo da indústria	153
3.25.1.1.2	Dummy Margem líquida da indústria	157
3.25.1.1.3	Dummy Liquidez corrente das empresas estatais	160
3.25.1.2	Cluster 1: Mais testes para a normalidade dos resíduos da regressão com variáveis adicionais	160
3.25.2	Regressão e análise dos resultados do <i>cluster</i> 2: incluindo todas as variáveis	163
3.25.3	Regressão e análise dos resultados do <i>cluster</i> 3: incluindo todas as variáveis	163
3.25.3.1	Cluster 3: Análise detalhada das outras variáveis significativas	169
3.25.3.1.1	Endividamento em longo prazo e endividamento geral	169
3.25.3.1.2	Dummy margem líquida do ramo indústria	170
3.25.3.1.3	Dummy giro do ativo do ramo serviços	172
3.25.3.1.4	Dummy de liquidez corrente de empresas com ações	173
3.25.4	Regressão e análise dos resultados do <i>cluster</i> 4: incluindo todas as variáveis	175
3.25.4.1	Cluster 4: Análise mais detalhada de outras variáveis significativas .	181
3.25.4.1.1	Endividamento em longo prazo	181
3.25.4.1.2	Capital Circulante Líquido	184
3.25.4.2	Cluster 4: Mais testes para a normalidade dos resíduos da regressão com variáveis adicionais	185
3.25.5	Regressão e análise dos resultados do <i>cluster</i> 5: incluindo todas as variáveis	189
3.25.5.1	Cluster 5: Análise detalhada das outras variáveis significativas	194
3.25.6	Regressão e análise dos resultados do <i>cluster</i> 6: incluindo todas as variáveis	196
3.25.6.1	Cluster 6: Análise detalhada das outras variáveis significativas	201
3.25.6.1.1	Variável Investimento no Imobilizado	201
3.25.6.1.2	Endividamento geral	203
3.25.6.2	Cluster 6: Mais testes para a normalidade dos resíduos da regressão com variáveis adicionais	204
3.25.7	Regressão e análise dos resultados do <i>cluster</i> 7: incluindo todas as variáveis	207
3.25.7.1	Cluster 7: Análise detalhada das outras variáveis significativas	211
3.25.7.1.1	Variável Dummy Composição do Endividamento de Empresas com Ações	212
3.25.7.1.2	Variável Dummy Giro do Ativo de Empresas com Ações ..	214
3.25.7.1.3	Variável Vendas	216
3.25.7.1.4	Capital circulante líquido	217
3.25.7.1.5	Ativo	218
4	CONCLUSÃO	221
	REFERÊNCIAS	227

LISTA DE ABREVIATURAS E SIGLAS

BCC: Banker, Charnes e Cooper
BIC: Bayesian Information Criterion
CCL7: Valor do Capital Circulante Líquido – CCL
CCR: Charnes, Cooper e Rhodes
COMPEN7: Composição do Endividamento
CREVEN7: Crescimento das Vendas
DACCUM: *Dummy* composição do endividamento e ações
DACCRES: *Dummy* crescimento das vendas e ações
DACGIR: *Dummy* giro do ativo e ações
DACLIQ: *Dummy* liquidez corrente e ações
DACMAR: *Dummy* margem líquida e ações
DACRENT: *Dummy* rentabilidade sobre o patrimônio líquido e ações
DB4COM: *Dummy* composição do endividamento e big four
DB4CRES: *Dummy* crescimento das vendas e big four
DB4GIR: *Dummy* giro do ativo e big four
DB4LIQ: *Dummy* liquidez corrente e big four
DB4MAR: *Dummy* margem líquida e big four
DB4RENT: *Dummy* rentabilidade sobre o patrimônio líquido e big four
DCOCOM: *Dummy* composição do endividamento e comércio
DCOCRES: *Dummy* crescimento das vendas e comércio
DCOGIR: *Dummy* giro do ativo e comércio
DCOLIQ: *Dummy* liquidez corrente e comércio
DCOMAR: *Dummy* margem líquida e comércio
DCORENT: *Dummy* rentabilidade sobre o patrimônio líquido e comércio
DEA: Data Envelopment Analysis
DESTCOM: *Dummy* composição do endividamento e empresa estatal
DESTCRES: *Dummy* crescimento das vendas e empresa estatal
DESTGIR: *Dummy* giro do ativo e empresa estatal
DESTLIQ: *Dummy* liquidez corrente e empresa estatal
DESTMAR: *Dummy* margem líquida e empresa estatal
DESTRENT: *Dummy* rentabilidade sobre o patrimônio líquido e empresa estatal
DINCOM: *Dummy* composição do endividamento e indústria
DINCRES: *Dummy* crescimento das vendas e indústria
DINGIR: *Dummy* giro do ativo e indústria
DINLIQ: *Dummy* liquidez corrente e indústria
DINMAR: *Dummy* margem líquida e indústria
DINRENT: *Dummy* rentabilidade sobre o patrimônio líquido e indústria
DMU: Decision Making Units
DNACCOM: *Dummy* composição do endividamento e empresa nacional
DNACCRES: *Dummy* crescimento das vendas e empresa nacional
DNACGIR: *Dummy* giro do ativo e empresa nacional
DNACLIQ: *Dummy* liquidez corrente e empresa nacional
DNACMAR: *Dummy* margem líquida e empresa nacional
DNACRENT: *Dummy* rentabilidade sobre o patrimônio líquido e empresa nacional
DSECOM: *Dummy* composição do endividamento e serviço
DSECRES: *Dummy* crescimento das vendas e serviço
DSEGIR: *Dummy* giro do ativo e serviço
DSELIQ: *Dummy* liquidez corrente e serviço

DSEMAR: *Dummy* margem líquida e serviço

DSERENT: *Dummy* rentabilidade sobre o patrimônio líquido e serviço

EBITIDA7: EBITDA

ENDGE7: Endividamento geral

ENDLP7: Endividamento de longo prazo

GIRO7: Giro do Ativo

INVIMOB7: Investimentos no imobilizado

LIQCOR7: Liquidez Corrente

LIQGE7: Liquidez geral

LLAJUS7: Lucro líquido ajustado

M&M: Melhores e Maiores

MARVEN7: Margem Líquida

PLAJUS7: Patrimônio líquido ajustado

RENTAJU7: Rentabilidade sobre o Patrimônio Líquido

VENDAS7: Valor das vendas

LISTA DE TABELAS

Tabela 1: Critério de Seleção do Número de <i>Clusters</i> (<i>Auto-Clustering</i>).....	58
Tabela 2: Distribuição dos <i>Clusters</i> (<i>Cluster Distribution</i>).....	60
Tabela 3: Perfil dos <i>Clusters</i> (<i>Centroids</i>).....	61
Tabela 4: escores das empresas componentes do <i>cluster 1</i>	72
Tabela 5: Escores das empresas componentes do <i>cluster 2</i>	78
Tabela 6: Escores das empresas componentes do <i>cluster 3</i>	82
Tabela 7: Escores das empresas componentes do <i>cluster 4</i>	86
Tabela 8: Escores das empresas componentes do <i>cluster 5</i>	90
Tabela 9: Escores das empresas componentes do <i>cluster 6</i>	94
Tabela 10: Escores das empresas componentes do <i>cluster 7</i>	98
Tabela 11: Resultados da regressão para o <i>cluster 1</i>	103
Tabela 12: Teste de correlação serial dos resíduos da regressão para o <i>cluster 1</i>	105
Tabela 13: Teste de heterocedasticidade de White para os resíduos do <i>cluster 1</i>	105
Tabela 14: Processo de White para heterocedasticidade referente ao <i>cluster 1</i>	106
Tabela 15: Resultados da regressão para o <i>cluster 2</i>	107
Tabela 16: Teste de correlação serial dos resíduos da regressão para o <i>cluster 2</i>	109
Tabela 17: Teste de heterocedasticidade de White para os resíduos do <i>cluster 2</i>	110
Tabela 18: Processo de White para heterocedasticidade referente ao <i>cluster 2</i>	111
Tabela 19: Resultados da regressão para o <i>cluster 3</i>	112
Tabela 20: Teste da normalidade dos resíduos da regressão para o <i>cluster 3</i>	114
Tabela 21: Teste de correlação serial dos resíduos da regressão para o <i>cluster 3</i>	114
Tabela 22: Teste de heterocedasticidade de White para os resíduos do <i>cluster 3</i>	115
Tabela 23: Processo de White para heterocedasticidade referente ao <i>cluster 3</i>	116
Tabela 24: <i>Cluster 3</i> : regressão com todas as variáveis da técnica DELPHI.....	117
Tabela 25: Teste de correlação serial.....	118
Tabela 26: Teste de heterocedasticidade dos resíduos.....	119
Tabela 27: Teste de heterocedasticidade de White para os resíduos.....	120
Tabela 28: Resultados da regressão para o <i>cluster 4</i>	121
Tabela 29: Teste de correlação serial dos resíduos da regressão para o <i>cluster 4</i>	123
Tabela 30: Teste de heterocedasticidade de White para os resíduos do <i>cluster 4</i>	125
Tabela 31: Processo de Newey-West para correlação serial e heterocedasticidade referente ao <i>cluster 4</i>	125
Tabela 32: <i>Cluster 4</i> : regressão com todas as variáveis da técnica Delphi.....	126
Tabela 33: Teste de correlação serial.....	127
Tabela 34: Teste de heterocedasticidade dos resíduos.....	128
Tabela 35: Teste de heterocedasticidade de White para os resíduos.....	129
Tabela 36: Teste de normalidade dos resíduos.....	130
Tabela 37 Resultados da regressão para o <i>cluster 5</i>	131
Tabela 38: Teste de correlação serial dos resíduos da regressão para o <i>cluster 5</i>	132
Tabela 39: Teste de heterocedasticidade de White para os resíduos do <i>cluster 5</i>	133
Tabela 40: Resultados da regressão para o <i>cluster 6</i>	134
Tabela 41: Teste de correlação serial dos resíduos da regressão para o <i>cluster 6</i>	136
Tabela 42: Teste de heterocedasticidade de White para os resíduos do <i>cluster 6</i>	137
Tabela 43: Processo de White para heterocedasticidade referente ao <i>cluster 6</i>	138
Tabela 44: <i>Cluster 6</i> : regressão com todas as variáveis da técnica Delphi.....	139
Tabela 45: Teste de correlação serial.....	140
Tabela 46: Teste de heterocedasticidade dos resíduos.....	141
Tabela 47: Teste de heterocedasticidade de White para os resíduos.....	142

Tabela 48: Resultados da regressão para o <i>cluster</i> 7.....	143
Tabela 49: Teste de correlação serial dos resíduos da regressão para o <i>cluster</i> 7.....	145
Tabela 50: Teste de heterocedasticidade de White para os resíduos do <i>cluster</i> 7.....	146
Tabela 51: Processo de White para heterocedasticidade referente ao <i>cluster</i> 7.....	147
Tabela 52: Resultados da regressão para o <i>cluster</i> 1 com todas as variáveis.....	149
Tabela 53: Teste de correlação serial dos resíduos da regressão para o <i>cluster</i> 1.....	150
Tabela 54: Teste de heterocedasticidade de White para os resíduos do <i>cluster</i> 1 com todas as variáveis	151
Tabela 55: Processo de White para heterocedasticidade referente ao <i>cluster</i> 1.....	152
Tabela 56: Teste de normalidade.....	153
Tabela 57: Teste de Homogeneidade das variâncias.....	154
Tabela 58: Teste ANOVA.....	154
Tabela 59: Testes de Welch e de Brown-Forsythe.....	154
Tabela 60: Teste de Kruskal-Wallis.....	155
Tabela 61: Teste de Tamhane.....	156
Tabela 62: Teste de normalidade.....	157
Tabela 63: Teste de Homogeneidade das variâncias.....	157
Tabela 64: Teste ANOVA.....	158
Tabela 65: Testes de Welch e de Brown-Forsythe.....	158
Tabela 66: Teste de Kruskal-Wallis.....	159
Tabela 67: Teste de Tamhane.....	159
Tabela 68: Liquidez corrente <i>cluster</i> 1.....	160
Tabela 69: Teste de normalidade dos resíduos da regressão.....	161
Tabela 70: Resultados da regressão para o <i>cluster</i> 3 com todas as variáveis.....	165
Tabela 71: Teste de correlação serial dos resíduos da regressão para o <i>cluster</i> 3.....	166
Tabela 72: Teste de heterocedasticidade de White para os resíduos do <i>cluster</i> 3 com todas as variáveis	167
Tabela 73: Processo de White para heterocedasticidade referente ao <i>cluster</i> 3.....	168
Tabela 74: Correlação entre Endividamento a longo prazo e endividamento geral.....	170
Tabela 75: Teste de normalidade.....	171
Tabela 76: Teste de Homogeneidade das variâncias.....	171
Tabela 77: Teste ANOVA.....	171
Tabela 78: Teste de Normalidade	172
Tabela 79: Teste de Homogeneidade das Variâncias.....	173
Tabela 80: Teste de Kruskal-Wallis.....	173
Tabela 81: Teste de normalidade.....	174
Tabela 82: Teste t para diferença de médias para amostras independentes.....	174
Tabela 83: Índice de liquidez corrente de empresas com ações (1) e sem ações (0).....	174
Tabela 84: Resultados da regressão para o <i>cluster</i> 4 com todas as variáveis.....	176
Tabela 85: Teste de correlação serial dos resíduos da regressão para o <i>cluster</i> 4 com todas as variáveis.....	178
Tabela 86: Teste de heterocedasticidade de White para os resíduos do <i>cluster</i> 4 com todas as variáveis.....	179
Tabela 87: Processo de Newey-West para heterocedasticidade referente ao <i>cluster</i> 4.....	180
Tabela 88: Regressão: escores em função do endividamento a longo prazo.....	181
Tabela 89: Regressão com as variáveis transformadas.....	182
Tabela 90: Teste e correlação serial dos resíduos da regressão.....	183
Tabela 91: Teste de heterocedasticidade dos resíduos da regressão.....	184
Tabela 92: Regressão: escores em função do capital circulante líquido.....	185

Tabela 93: <i>Cluster 4</i> : Regressão com log de todas as variáveis da técnica Delphi e das variáveis adicionais.....	186
Tabela 94: Teste de correlação serial.....	187
Tabela 95: Teste de heterocedasticidade dos resíduos.....	188
Tabela 96: Resultados da regressão para o <i>cluster 5</i> com todas as variáveis.....	191
Tabela 97: Teste de correlação serial dos resíduos da regressão para o <i>cluster 5</i> com todas as variáveis.....	192
Tabela 98: Teste de heterocedasticidade de White para os resíduos do <i>cluster 5</i> para todas as variáveis.....	193
Tabela 99: Teste de normalidade	194
Tabela 100: Teste de diferença de médias para amostras independentes.....	195
Tabela 101: Estatística descritiva.....	195
Tabela 102: Resultados da regressão para o <i>cluster 6</i> com todas as variáveis.....	197
Tabela 103: Teste de correlação serial dos resíduos da regressão para o <i>cluster 6</i> com todas as variáveis.....	198
Tabela 104: Teste de heterocedasticidade de White para os resíduos do <i>cluster 6</i> com todas as variáveis.....	199
Tabela 105: Processo de White para heterocedasticidade referente ao <i>cluster 6</i>	200
Tabela 106: Regressão: escores em função da variável investimento no imobilizado.....	201
Tabela 107: Teste e correlação serial dos resíduos da regressão.....	202
Tabela 108: Teste de heterocedasticidade dos resíduos da regressão.....	203
Tabela 109: Regressão: escores em função da variável endividamento geral.....	204
Tabela 110: Resultados da regressão para o <i>cluster 7</i> com todas as variáveis.....	208
Tabela 111: Teste de correlação serial dos resíduos da regressão para o <i>cluster 7</i> com todas as variáveis.....	209
Tabela 112: Teste de heterocedasticidade de White para os resíduos do <i>cluster 7</i> com todas as variáveis.....	210
Tabela 113: Teste de normalidade.....	212
Tabela 114: Teste de diferença de médias para amostras independentes.....	213
Tabela 115: Teste não paramétrico de Mann-Witney.....	214
Tabela 116: Teste de normalidade.....	215
Tabela 117: Teste de Levene para a igualdade das variâncias.....	215
Tabela 118: Teste não paramétrico de Mann-Witney.....	216
Tabela 119: Regressão: escores em função das vendas.....	217
Tabela 120: Regressão: escores em função do capital circulante líquido.....	218
Tabela 121: Regressão: escores em função do ativo.....	219

LISTA DE FIGURAS

Figura 1: Representação dos tamanhos dos <i>clusters</i> (<i>Cluster Size</i>).....	59
Figura 2: Variável Crescimento das Vendas para os 7 <i>clusters</i>	62
Figura 3: Variável Rentabilidade sobre o Patrimônio para os 7 <i>clusters</i>	63
Figura 4: Variável Margem Líquida para os 7 <i>clusters</i>	63
Figura 5: Variável Composição do Endividamento para os 7 <i>clusters</i>	64
Figura 6: Variável Giro do Ativo para os 7 <i>clusters</i>	65
Figura 7: Variável Liquidez Corrente para os 7 <i>clusters</i>	65
Figura 8: Variáveis de importância na formação do <i>cluster</i> 1.....	66
Figura 9: Variáveis de importância na formação do <i>cluster</i> 2.....	67
Figura 10: Variáveis de importância na formação do <i>cluster</i> 3.....	67
Figura 11: Variáveis de importância na formação do <i>cluster</i> 4.....	68
Figura 12: Variáveis de importância na formação do <i>cluster</i> 5.....	69
Figura 13: Variáveis de importância na formação do <i>cluster</i> 6.....	69
Figura 14: Variáveis de importância na formação do <i>cluster</i> 7.....	70
Figura 15: Plotagem das empresas do <i>cluster</i> 1.....	74
Figura 16: Distribuição dos escores do <i>cluster</i> 1.....	75
Figura 17: Empresas de referência do <i>cluster</i> 1.....	76
Figura 18: Melhorias potenciais para o <i>cluster</i> 1.....	77
Figura 19: Plotagem das empresas do <i>cluster</i> 2.....	79
Figura 20: Distribuição dos escores do <i>cluster</i> 2.....	80
Figura 21: Empresas de referência do <i>cluster</i> 2.....	80
Figura 22: Melhorias potenciais para o <i>cluster</i> 2.....	81
Figura 23: Plotagem das empresas do <i>cluster</i> 3.....	83
Figura 24: Distribuição dos escores do <i>cluster</i> 3.....	83
Figura 25: Empresas de referência do <i>cluster</i> 3.....	84
Figura 26: Melhorias potenciais para o <i>cluster</i> 3.....	85
Figura 27: Plotagem das empresas do <i>cluster</i> 4.....	87
Figura 28: Distribuição dos escores do <i>cluster</i> 4.....	88
Figura 29: Empresas de referência do <i>cluster</i> 4.....	88
Figura 30: Melhorias potenciais para o <i>cluster</i> 4.....	89
Figura 31: Plotagem das empresas do <i>cluster</i> 5.....	91
Figura 32: Distribuição dos escores do <i>cluster</i> 5.....	92
Figura 33: Empresas de referência do <i>cluster</i> 5.....	92
Figura 34: Melhorias potenciais para o <i>cluster</i> 5.....	93
Figura 35: Plotagem das empresas do <i>cluster</i> 6.....	95
Figura 36: Distribuição dos escores do <i>cluster</i> 6.....	96
Figura 37: Empresas de referência do <i>cluster</i> 6.....	96
Figura 38: Melhorias potenciais para o <i>cluster</i> 6.....	97
Figura 39: Plotagem das empresas do <i>cluster</i> 6.....	99
Figura 40: Distribuição dos escores do <i>cluster</i> 7.....	100
Figura 41: Empresas de referência do <i>cluster</i> 7.....	100
Figura 42: Melhorias potenciais para o <i>cluster</i> 7.....	101
Figura 43: Teste da normalidade dos resíduos da regressão para o <i>cluster</i> 1.....	104
Figura 44: Teste da normalidade dos resíduos da regressão para o <i>cluster</i> 2.....	108
Figura 45: Teste da normalidade dos resíduos da regressão para o <i>cluster</i> 3.....	113
Figura 46: Normalidade dos resíduos da regressão.....	117
Figura 47: Teste da normalidade dos resíduos da regressão para o <i>cluster</i> 4.....	122
Figura 48: Normalidade dos resíduos da regressão.....	127

Figura 49: Teste da normalidade dos resíduos da regressão para o <i>cluster</i> 5.....	132
Figura 50: Teste da normalidade dos resíduos da regressão para o <i>cluster</i> 6.....	135
Figura 51: Normalidade dos resíduos da regressão.....	139
Figura 52: Teste da normalidade dos resíduos da regressão para o <i>cluster</i> 7.....	144
Figura 53: Teste da normalidade dos resíduos da regressão para o <i>cluster</i> 1 com todas as Variáveis.....	150
Figura 54: Teste da normalidade dos resíduos da regressão para o <i>cluster</i> 3 com todas as variáveis	165
Figura 55: Teste da normalidade dos resíduos da regressão para o <i>cluster</i> 4 com todas as Variáveis.....	177
Figura 56: Teste de normalidade dos resíduos da regressão.....	182
Figura 57: Teste de normalidade dos resíduos da regressão.....	183
Figura 58: Normalidade dos resíduos da regressão.....	187
Figura 59: Teste da normalidade dos resíduos da regressão para o <i>cluster</i> 5 com todas as Variáveis.....	191
Figura 60: Teste da normalidade dos resíduos da regressão para o <i>cluster</i> 6 com todas as Variáveis.....	198
Figura 61: Teste de normalidade dos resíduos da regressão.....	202
Figura 62: Teste da normalidade dos resíduos.....	206
Figura 63: Teste da normalidade dos resíduos da regressão para o <i>cluster</i> 3 com todas as Variáveis.....	209

1 Introdução

Este estudo trata da avaliação de eficiência de empresas através da utilização de indicadores contábeis e do uso de diversos instrumentos quantitativos como: Análise por Envoltória de Dados (DEA), Análise de Agrupamentos (*Cluster Analysis*), Análise de Regressão e Testes de Hipóteses Paramétricos e Não Paramétricos.

1.1 Contextualização

O primeiro ponto discutido é se os indicadores contábeis podem ser usados na previsão de falências, como, por exemplo, em determinações que envolvam decisão de conceder ou não crédito a um determinado cliente. Kanitz (1976) responde essa questão através de pesquisa em que comparou indicadores contábeis de empresas solventes e insolventes. Através da utilização de testes estatísticos concluiu que diversos indicadores contábeis são significativos na discriminação entre esses dois tipos de empresas, independentemente dos balanços contábeis serem considerados como fidedignos ou não e/ou de terem seus valores monetários corrigidos pela variação do poder aquisitivo da moeda ou não. A partir disso, percebe-se que é possível tomar decisões a partir da análise de indicadores contábeis.

Outro problema analisado é se a avaliação das demonstrações contábeis tem um caráter objetivo ou se depende da experiência e habilidade do analista ou usuário da informação contábil. A solução para essa questão é fornecida por Iudicibus (2007, p. 5), ao salientar que analisar os balanços é a arte de saber extrair relações úteis dos relatórios contábeis, tendo em vista o objetivo pretendido. Isso significa que o analista tem papel fundamental na escolha dos indicadores, nas respectivas análises e nas conclusões. Dessa forma, os resultados da análise de balanços e as conclusões daí resultantes apresentam um caráter subjetivo e dependem muito da capacidade e experiência do analista ou do usuário da informação contábil.

Eliseu Martins (2005c, p. 3) também responde a questão da objetividade na análise das demonstrações contábeis ao considerar que a contabilidade não é uma ciência exata e, assim, os usuários da informação contábil, quando analisam os números contidos nas demonstrações contábeis devem ter não só conhecimento, mas também arte para avaliá-los e interpretá-los.

É necessário também reconhecer como reduzir a subjetividade na análise das demonstrações contábeis. Segundo Lyra (2008, p. 33), “É razoável admitir que a “arte” de analisar as

demonstrações contábeis possa ser mais bem consubstanciada com o auxílio de ferramentas quantitativas. O uso de métodos quantitativos concomitante aos trabalhos de análise de demonstrações contábeis permite uma melhor seleção e agrupamento de indicadores a serem utilizados em cada caso”.

De fato, a aplicação de algum instrumental quantitativo poderia ser útil na redução da subjetividade da análise das demonstrações contábeis. Porém, o instrumento mais adequado depende da finalidade da análise: se for para discriminar empresas solventes de insolventes, poderia ser aplicada a análise discriminante, ou regressão logística ou redes neurais etc. Se for para fazer previsão, poderia ser utilizada a análise de regressão, ou séries temporais etc. Se for para avaliar riscos, poderia ser empregada simulação probabilística, ou cálculo de probabilidades etc.

Por sua vez, quanto à avaliação da eficiência, Matarazzo (1998, p. 193) afirma que “A análise de Balanços através de índices só adquire consistência e objetividade quando os índices são comparados com padrões, pois, do contrário, as conclusões se sujeitam à opinião e, não raro, ao humor do analista”. Os padrões podem ser obtidos através de simples médias ou medianas obtidas do setor de atividades ao qual a empresa pertence.

Uma forma de avaliar a eficiência é comparar a produtividade da empresa com a produtividade que deveria ser alcançada por ela em condições de máxima eficiência. Os primeiros pesquisadores a desenvolverem esse tipo de modelo de avaliação foram Charnes, Cooper e Rhodes (1978), baseados no trabalho pioneiro de Farrell (1957). Charnes *et al* (1978) criaram o instrumento usado para a avaliação da eficiência, o modelo *Data Envelopment Analysis* (DEA).

O modelo DEA é uma técnica não paramétrica utilizada para estabelecer a fronteira eficiente de produção. Essa eficiência é determinada de maneira relativa. Considerando um conjunto de recursos à disposição das empresas (*inputs*), compara-se o que foi efetivamente produzido por elas (*outputs*) com aquilo que poderia ser feito, utilizando-se os mesmos recursos. Assim, são feitas comparações entre as empresas para determinar quais são as mais eficientes. Podem ser usados como *inputs* e *outputs* alguns indicadores contábeis.

Saen e Memariani (2005, p. 313) comentam que o DEA pode ser aplicado em qualquer organização ou indústria onde um conjunto de empresas ou organizações ou subunidades razoavelmente homogêneas usem o mesmo conjunto de *inputs*, possivelmente em diferentes combinações, para produzir um conjunto de *outputs* identificáveis, e provavelmente em diferentes combinações. O pressuposto do modelo DEA clássico é a existência de homogeneidade entre as unidades em análise.

A partir disso, a dúvida é sobre como conseguir homogeneidade quando se sabe que muitas vezes empresas de um mesmo setor de atividade apresentam comportamento muito diferente com relação à forma de conduzir os negócios. Pohlmann (2004 e 2007) esclarece tal dúvida ao elaborar pesquisa com a utilização de dados da edição Melhores & Maiores, da Revista Exame. Seu objetivo foi propor uma separação das empresas em agrupamentos através da utilização da análise de conglomerados ou *cluster analysis*. Assim, o valor primário da análise *cluster* repousa na classificação de dados, como sugerido pelo agrupamento natural de dados por si mesmo. (POHLMANN, 2004 e 2007)

Clusters analysis é uma técnica em que nenhuma definição prévia é executada com relação ao número de grupos ou à sua estrutura. A separação em grupos é feita com base em similaridades ou distâncias (dissimilaridades). Os *inputs* exigidos são medidas de similaridade ou dados, a partir dos quais as similaridades possam ser computadas (POHLMANN apud CORRAR, PAULO e DIAS FILHO, 2007, p. 326).

Devem-se definir também os indicadores contábeis a serem utilizados. De acordo com Matarazzo (1998, p. 154), “o importante não é o cálculo de grande número de índices, mas de um conjunto de índices que permita conhecer a situação da empresa, segundo o grau de profundidade desejada da análise”. Dessa maneira, Matarazzo sugere que indicadores devam ser selecionados em função do objetivo desejado.

Além de determinar os indicadores a serem utilizados, é também preciso selecionar indicadores significativos. Sobre esse assunto, Lyra (2008) utilizou a técnica Delphi e a participação de especialistas em sua tese de doutorado apresentada ao Departamento de Contabilidade e Atuária da Universidade de São Paulo. O procedimento utilizado pelo Professor Lyra (2008, p. 85- 94) consiste em: (a) elaboração de questionário e uma relação de indicadores contábeis definidos com base na literatura existente; (b) pré-teste junto a alguns

professores especializados (c) realização da pesquisa junto a um grupo de professores especialistas brasileiros e estrangeiros.

Os resultados do estudo de Lyra (2008) são utilizados como ponto de partida para este trabalho, que pretende encontrar respostas para as questões a seguir relacionadas.

1.2 Questões de pesquisa

Esta pesquisa se propõe a responder as seguintes questões:

- (a) Indicadores contábeis, obtidos através da técnica Delphi e com a participação de especialistas, são significativos para avaliação de eficiência de empresas?
- (b) É possível avaliar a eficiência das empresas através de um pequeno número de indicadores contábeis?
- (c) Como proceder para separar diversas empresas em grupos com características homogêneas?
- (d) Como podem ser usadas técnicas estatísticas paramétricas em complementação à técnica Análise por Envoltória de Dados (DEA), para testar hipóteses sobre significância de indicadores contábeis?
- (e) Como podem ser criados modelos de previsão de níveis de eficiência através da utilização de indicadores contábeis conjugados com a técnica DEA e a análise de regressão?

1.3 Objetivos

Este estudo tem como **objetivo geral**:

Avaliar a eficiência de empresas através de indicadores contábeis, com a utilização da análise por envoltória de dados, da análise de conglomerados e da análise de regressão.

A pesquisa apresenta os seguintes **objetivos específicos**:

1. Avaliar se os indicadores contábeis obtidos com a técnica Delphi e com base nos dados da Edição Melhores & Maiores da Revista Exame são significativos para a avaliação de eficiência de empresas por usuários externos, através da realização de testes;
2. Mostrar a possibilidade de avaliação da eficiência de empresas por meio de um pequeno número de indicadores contábeis selecionados;
3. Apresentar um método para a separação de diversas empresas em grupos com características homogêneas;
4. Demonstrar como usar técnicas estatísticas paramétricas em complementação ao modelo de Análise por Envoltória de Dados (DEA) para testar hipóteses sobre significância de indicadores contábeis;
5. Comprovar a possibilidade de modelos de previsão de níveis de eficiência através da utilização de indicadores contábeis conjugados com a técnica DEA e a análise de regressão.

Em síntese, o foco da pesquisa é a avaliação da eficiência de empresas tomando por base os indicadores contábeis obtidos a partir da técnica Delphi e também os dados publicados pela edição Melhores & Maiores da Revista Exame. Para tanto, são utilizados os instrumentos: Análise por Envoltória de Dados, Análise de Conglomerados, Análise de Regressão e Testes Estatísticos Paramétricos e Não Paramétricos.

1.4 Metodologia

Os detalhes da metodologia empregada neste trabalho encontram-se no capítulo 3: Metodologia e Análise dos Resultados. A pesquisa classifica-se como empírico-analítica, pois utiliza técnicas de coleta, tratamento e análise de dados. Há ainda uma revisão bibliográfica sobre o tema e sobre o instrumental quantitativo a ser utilizado. A amostra selecionada compreende dados de caráter secundário obtidos da edição Melhores & Maiores, referentes à edição de 2008.

1.5 Organização do Trabalho

Este estudo se organiza da seguinte forma:

Capítulo 1 – Introdução: tem por finalidade contextualizar o problema, apresentar as questões de pesquisa, os objetivos gerais e específicos e a organização do estudo.

Capítulo 2 – Revisão bibliográfica: efetua a revisão da literatura sobre o tema e busca referências conceituais que dão o suporte necessário para o desenvolvimento do trabalho.

Capítulo 3 – Metodologia e análise dos resultados: expõe a metodologia utilizada para a realização da pesquisa, apresenta e analisa os resultados da aplicação do instrumental quantitativo.

Capítulo 4 – Conclusão: traz as considerações finais sobre a pesquisa bem como as recomendações.

2 Revisão Bibliográfica

As características de eficácia e de eficiência são fundamentais na organização de uma empresa. Dessa maneira, Atkinson *et al* (1995, p. 46) afirmam que:

A mais importante função dos empregados de uma organização é administrar a cadeia de atividades de forma eficaz e eficiente. Dizemos que uma organização é eficaz se ela alcança seus objetivos. Dizemos que uma organização é eficiente se usa a menor quantidade de recursos possíveis para alcançar seus objetivos.

Para os mesmos autores (ATKINSON *et al*, 1995, p. 514), produtividade é a razão entre a quantidade produzida (*output*) e a quantidade de recursos (*input*) utilizada na produção.

Analisando esses três conceitos (eficácia, eficiência e produtividade), é possível estabelecer que a eficácia se resume em atingir o objetivo proposto; a produtividade é um quociente que mede como o objetivo é alcançado em função dos insumos utilizados; e a eficiência é uma medida relativa que envolve comparação entre produtividades. Dessa forma, uma organização é mais eficiente que outra se utiliza menor quantidade de recursos para obtenção da mesma quantidade de produtos. Ou ainda, em outras palavras, a eficiência envolve uma comparação entre o que foi produzido com o que deveria ser produzido envolvendo a mesma quantidade de recursos.

Segundo Cooper *et al* (1995, p. 1), uma empresa pode utilizar uma medida de produtividade parcial quando, por exemplo, mede a quantidade produzida por empregado. Ao considerar a razão entre todos os *outputs* e todos os *inputs*, trata-se de uma medida de produtividade total, sendo que tal medida envolve problemas relacionados com a escolha das variáveis a serem utilizadas e da forma de ponderá-las.

Singh *et al* (2000, p. 234) comentam que, dentro de um contexto geral, há três diferentes técnicas para medir a produtividade: (1) medição através de índices, (2) programação linear e (3) modelos econométricos.

(1) Medição através de índices

Quanto a essa técnica, existem diversos modelos construídos ao longo do tempo para medir a produtividade. Tais modelos utilizam principalmente indicadores obtidos através da contabilidade.

(2) Programação linear

Conforme Singh *et al* (2000, p. 236), a programação linear apresenta as seguintes vantagens: não exigir que as variáveis sejam avaliadas monetariamente e também não necessitar de especificação de uma forma funcional.

Entre os instrumentos de programação linear, o mais comumente utilizado é a técnica *Data Envelopment Analysis* (DEA). A tradução utilizada por alguns autores nacionais é análise por envoltória de dados (CASANOVA, 2002) ou análise de envoltória de dados (Mello *et al*, 2005). Outros usam simplesmente análise envoltória de dados (Macedo, 2004).

A DEA é um procedimento de estimação não-paramétrico flexível que identifica as contribuições de um conjunto de *inputs* para a obtenção do máximo de quantidades de um determinado conjunto de *outputs* (SINGH, 2000, p. 236).

(3) Modelos econométricos

Sudit (1995, in SINGH, 2000, p. 236-7) apresenta uma boa discussão sobre a utilização de modelos econométricos, conforme a síntese a seguir:.

O uso dos modelos econométricos requer o conhecimento *a priori* sobre o processo tecnológico de produção. Ou, em outras palavras, há necessidade de se conhecer a forma funcional do modelo. Formas funcionais muito flexíveis e com muitos parâmetros exigem uma amostra bastante grande para sua estimação. Quando não se conhece o processo tecnológico geralmente se emprega a função de produção translog popularizada por Christensen (1973, apud SINGH, 2000, p. 236-7). A função translog é flexível, mas, requer amostras grandes em função do número de parâmetros a serem estimados.

Quando são usados dados amostrais para estimação de funções de produção, não se pode assumir que todas as observações passadas sejam pontos de observações tecnicamente eficientes. Muitas observações podem estar abaixo da fronteira eficiente. Assim, quando são usados dados amostrais históricos nos modelos econométricos, não se alcança a fronteira eficiente, mas uma fronteira de produção de desempenho médio. Ao se usarem resultados

assim obtidos, como se estivessem na fronteira de produção, as possibilidades estão sendo subestimadas de forma significativa.

Singh (2000, p. 237) salienta que as três formas de avaliação da produtividade abordadas poderão ser úteis dependendo da questão que se pretende resolver. A seguir são apresentadas as três questões básicas em avaliação de produtividade formuladas por esse autor:

Qual foi o desempenho comparativo de uma determinada empresa nos últimos três anos?

Quais são os *inputs* utilizados de forma ineficiente no processo de produção?

Em que grau o capital pode ser substituído pelo trabalho, dada uma determinada especificação tecnológica?

Para solucionar tais questões, alguns recursos são possíveis. Quanto à primeira pergunta, sobre o desempenho comparativo de uma empresa específica nos últimos três anos, devem-se utilizar alguns índices para calcular as variações de produtividade passadas. Em relação à segunda, a respeito dos *inputs* usados de forma ineficiente no processo de produção, a solução pode ser obtida através do DEA. E para responder a terceira questão, relativa a como o capital pode ser substituído pelo trabalho, diante de uma determinada especificação tecnológica, a resposta pode ser obtida através dos modelos econométricos.

2. 1 Data Envelopment Analysis (DEA)

A denominação Análise por Envoltória de Dados é uma tradução feita por autores nacionais para *Data Envelopment Analysis*, cuja sigla internacionalmente conhecida é DEA. Essa técnica utiliza modelos de programação linear para mensurar a eficiência de entidades dos mais diversos tipos, denominados DMUs, *Decision Making Units*. Trata-se de uma técnica não paramétrica usada para especificar a fronteira eficiente de produção, uma vez que essa eficiência seja determinada de maneira relativa. Analisando um conjunto de recursos à disposição das DMUs (*inputs*), compara-se o que foi produzido por elas (*outputs*) com aquilo que poderia ser produzido com a utilização dos mesmos recursos. Assim, são feitas comparações entre as diversas DMUs para se definir quais são as mais eficientes. Nesta pesquisa, são considerados como *inputs* e *outputs* alguns indicadores obtidos a partir dos dados publicados pela Edição de Melhores e Maiores. Os primeiros pesquisadores a desenvolver esse modelo deles foram Charnes, Cooper e Rhodes (1978), baseados no trabalho pioneiro de Farrell (1957).

Segundo Cooper *et al* (2007, p. 2), os novos enfoques do DEA não exigem que sejam atribuídos pesos para cada *input* ou *output*, como ocorre com o enfoque dos índices, e também não demandam a especificação da forma funcional dos modelos econométricos.

O DEA utiliza técnicas como a programação matemática, que permite manipular grande número de variáveis e restrições e possibilita relaxar as exigências encontradas em outras técnicas limitadoras da escolha do número de *inputs* e *outputs*. (COOPER *et al*, 2007, p. 2)

Inspirado pelos conceitos de análise de atividades de Debreu (1951) e de Koopmans (1951), Farrel (1957) elaborou um modelo de medição da eficiência produtiva e introduziu o conceito de eficiência relativa. Esse modelo considera todos os *inputs* de um processo de avaliação e, segundo Farrel (1957), procura corrigir os problemas inerentes aos modelos de índices de produtividade geralmente utilizados até então e mostra ainda como esse modelo poderia ser usado na prática através de uma aplicação no setor da agricultura norte-americana.

Entretanto, o modelo não teve a devida atenção até 1978, quando surgiu o trabalho de Charnes, Cooper e Rhodes (1978) que, considerando as idéias básicas de Farrel (1957), desenvolveram o denominado *Data Envelopment Analysis* (DEA).

Charnes *et al* (1978) não só ampliaram e incrementaram o escopo do trabalho de Farrel (1957), como também contribuíram com a melhoria do poder e a facilidade computacional. Distinguiram os conceitos de desempenho fracamente eficiente (ou eficiência, de Farrel) e fortemente eficiente e introduziram a formulação de uma programação fracionária em que os índices de eficiência se situavam entre zero e a unidade. Essa formulação possibilitou ampliar as medidas de eficiência baseadas na razão entre *outputs* e *inputs* utilizada pela engenharia sem a necessidade de usar o recurso de atribuição de pesos, como costumeiramente eram utilizado, no trabalho com múltiplos *inputs* e *outputs*. (COOPER *et al*, 2005, p. 4)

Charnes *et al* (1978) utilizam pela primeira vez o termo *Decision Making Units* (DMU) para designar unidades tomadoras de decisão. As DMUs podem ser entidades públicas ou empresas privadas ou também departamentos e filiais de uma determinada organização.

Resumidamente, Charnes *et al* (1978, p. 429) criaram um modelo de programação linear não convexo e introduziram uma nova definição de eficiência para a avaliação de atividades de

entidades sem fins lucrativos participantes de programas públicos. Estabeleceram uma medida escalar de eficiência de cada DMU a partir de dados observados de múltiplos *outputs* e múltiplos *inputs* relativos a essas unidades e, para efeitos computacionais, instituíram as equivalências entre o referido modelo e o modelo de programação linear ordinária. Através do modelo dual proporcionado pelos modelos de programação linear, puderam constituir uma série de relações a partir dos dados observados nas DMUs. Eles ainda dispuseram novas formas de interpretar os resultados obtidos da aplicação do modelo aos programas de caráter público.

O modelo de Charnes *et al* (1978, p. 431), conhecido como modelo CCR, é uma extensão da formulação de programação não linear para um problema de programação fracional ordinária. Na programação fracional, os coeficientes (as razões) assumem valores entre zero e a unidade e também os problemas de programação fracional podem ser reduzidos à forma de programação linear. Essa transformação foi sugerida por esses autores para facilitar as dificuldades computacionais devido à utilização de uma grande quantidade de *inputs* e de *outputs*. Além disso, a programação linear possibilita a utilização das relações de dualidade, as quais têm a vantagem de proporcionar informações úteis para a tomada de decisões, como a avaliação de custos de oportunidade e taxas de incremento por unidade na eficiência.

Segundo o conceito de eficiência de Pareto-Koopmans, adotado por Charnes *et al* (1978, p. 433), a redução em qualquer *input* ou o acréscimo em qualquer *output* possui algum valor. Ou ainda, de acordo com a definição apresentada em Cooper *et al* (2005, p. 5), a eficiência de Pareto-Koopmans é obtida por uma DMU se e somente se não for possível melhorar qualquer de seus *inputs* ou *outputs* sem piorar alguns de seus demais *inputs* ou *outputs*. Essa definição evidencia que: (1) não é necessário atribuir pesos relativos para se determinar a importância relativa dos diferentes *inputs* e *outputs* e (2) a satisfação dos *inputs* e *outputs* é uma condição necessária para a otimização em qualquer sistema em que os pesos atribuídos sejam todos positivos.

Ragsdale (2001, p. 106) resume o modelo DEA de forma didática, conforme transcrito a seguir, de forma simplificada.

Se for empregado o modelo DEA, a eficiência de uma determinada DMU *i* pode ser assim definida:

$$\text{Eficiência da unidade } i = \frac{\text{Soma ponderada dos } \textit{outputs} \text{ da unidade } i}{\text{Soma ponderada dos } \textit{inputs} \text{ da unidade } i} = \frac{\sum_{j=1}^{n_o} O_{ij} W_j}{\sum_{j=1}^{n_i} I_{ij} V_j}$$

Em que:

O_{ij} = valor do *output* j para a DMU i ;

I_{ij} = valor do *input* j para a DMU i ;

W_j = peso não negativo atribuído ao *output* j ;

V_j = peso não negativo atribuído ao *input* j ;

n_o = número de *outputs*;

n_i = número de *inputs*;

O problema é determinar os valores dos pesos W_j e V_j . Essas são as variáveis de decisão num modelo DEA.

Por sua vez, o objetivo é maximizar a soma ponderada dos *outputs* de cada DMU. Assim, a função de objetivo para uma determinada DMU pode ser definida como:

$$\text{MAX: } \sum_{j=1}^{n_o} O_{ij} W_j$$

No modelo DEA, para cada DMU haverá a solução de um particular problema de programação linear. Assim, para cada DMU em análise, são calculados os pesos que maximizam a soma ponderada de seus *outputs*, sujeitos às restrições especificadas a seguir.

Em primeiro lugar, cabe destacar a impossibilidade de que cada DMU tenha um nível de eficiência maior que 100%. Desse modo, quando se soluciona o problema de programação

linear individual de cada DMU, os pesos devem ser calculados de tal forma que não originem um nível de eficiência maior que 100% para cada unidade.

Assim, para cada DMU, é necessário que a soma ponderada dos seus *outputs* seja menor ou igual à soma ponderada dos seus *inputs*. Ou então, em outros termos, a razão entre os *outputs* ponderados e os *inputs* ponderados deve ser menor ou igual a 100%. Dessa maneira, as restrições devem ser:

$$\sum_{j=1}^{n_o} O_{kj} W_j \leq \sum_{j=1}^{n_i} I_{kj} V_j, \text{ para } k = 1 \text{ para o número de unidades}$$

Ou, de forma equivalente:

$$\sum_{j=1}^{n_o} O_{kj} W_j - \sum_{j=1}^{n_i} I_{kj} V_j \leq 0, \text{ para } k = 1 \text{ para o número de unidades}$$

Para que não haja soluções irrestritas, deve-se assumir que os *inputs* ponderados para cada DMU em análise devam ser igual à unidade, a saber:

$$\sum_{j=1}^{n_i} I_{ij} V_j = 1$$

Como a soma dos *inputs* ponderados para a DMU em análise deve ser igual à unidade e a soma dos seus *outputs* ponderados (que estão sendo maximizados) não pode exceder a unidade, o valor máximo do escore de eficiência para a DMU em análise deve ser também igual à unidade. Dessa forma, as DMUs eficientes devem ter escore de eficiência igual a 100%.

Ragsdale (2001, p. 107) faz algumas observações importantes para a aplicação do DEA: os *outputs* devem ser variáveis cujo comportamento é do tipo “quanto maior, melhor” e os *inputs* devem ser variáveis do tipo “quanto menor, melhor”. Qualquer *input* ou *output* que não tenha esse comportamento deve sofrer um processo de transformação antes da aplicação do DEA.

Banker, Charnes e Cooper (1984, p. 1078) trazem novos desenvolvimentos ao trabalho de Charnes et al (1978), separando as ineficiências técnica e de escala. As ineficiências técnicas são identificadas com o insucesso na obtenção dos melhores níveis de *outputs* e/ou utilização de quantidades excessivas de *inputs*. Os autores introduzem uma nova variável em particular que torna possível identificar e corrigir as magnitudes dessas ineficiências para determinar se as operações estariam sendo conduzidas em regiões de incremento dos retornos de escala, de constante retorno de escala ou de redução de retornos de escala. O modelo de Banker *et al* (1984) é conhecido como modelo BCC.

Singh (2000, p. 236) explana as principais vantagens do DEA. A partir de suas explicações, a seguir se expõe um resumo de seus comentários.

- a) DEA permite trabalhar com um conjunto extenso de *inputs* e *outputs*, possibilitando também a maximização simultânea de um conjunto de *outputs*. Não requer que as variáveis sejam avaliadas monetariamente e admite a utilização de variáveis qualitativas;
- b) Como o procedimento é não-paramétrico, DEA não exige a definição de uma forma funcional. Isso pode ser visto como uma vantagem, pois a escolha inicial de uma forma funcional errada causa erros de especificação;
- c) Simulações realizadas e considerações teóricas indicam que o procedimento de estimação é estatisticamente robusto;
- d) Como o DEA utiliza rotinas individuais para cada *input*, pode indicar quais são os *inputs* específicos utilizados menos eficientemente no processo produtivo.

A desvantagem do DEA é que a elaboração de testes de hipóteses se torna muito mais difícil e para isso deve-se recorrer a um instrumental adicional muito mais sofisticado, como aqueles empregados, por exemplo, por Land *et al* (1990, apud SINGH (2000, p. 236) e Cooper *et al* (1995, apud SINGH (2000, p. 236).

No Brasil, existem muitas pesquisas com a utilização da Análise por Envoltória de Dados, DEA. Podem ser citadas muitas dissertações de mestrado, teses de doutorado, participações em congressos e publicações em periódicos. Algumas das principais pesquisas nesse sentido

são: Almeida (2000); Angulo Meza (1998); Angulo Meza (2002a); Angulo Meza (2002b); Belloni (2000); Casanova (2002); Lins *et al* (1999); Lins *et al* (2000); Lins *et al* (2004); Lins *et al* (2005); Macedo *et al* (2006); Macedo *et al* (2008); Mello *et al* (2005); Migon (2000); Moita (2002); Niederauer (1998); Niederauer (2002); Novaes (1998); Paiva (2000); Sant'anna (1999); Silva (2000); Senra (2004); Soares De Mello *et al* (2001); Soares De Mello *et al* (2002); Soares De Mello *et al* (2003a); e Soares De Mello *et al* (2003b).

Sobre DEA, é importante destacar os modelos mais largamente utilizados, segundo Casanova (2002, p. 72):

1. Modelo CCR (1978) – desenvolvido por Charnes, Cooper e Rhodes, permite uma avaliação objetiva da eficiência global e identifica as fontes e estimativas de montantes das ineficiências identificadas.
2. Modelo BCC (1984) – criado por Banker, Charnes e Cooper, distingue entre ineficiências técnicas e de escala, estimando a eficiência técnica pura, a uma dada escala de operações, e identificando se estão presentes ganhos de escala crescentes, decrescentes e constantes, para futura exploração.

Essa mesma autora (2002, p. 73-75) apresenta as formulações referentes a cada um desses dois modelos. Cada um deles pode ter orientação ao consumo (*input*) ou orientação ao produto (*output*). A seguir são demonstradas essas formulações, abordadas na tese de doutorado de Sílvia Pereira de Castro Casanova:

Modelo CCR (original) com orientação ao consumo (*input*)

$$\text{Maximizar } h_k = \sum_{r=1}^s u_r y_{rk}, \quad (2.5)$$

sujeito a

$$\sum_{r=1}^m u_r y_{rj} - \sum_{i=1}^n v_i x_{ij} \leq 0 \quad (2.6)$$

$$\sum_{i=1}^n v_i x_{ik} = 1 \quad (2.7)$$

$$u_r, v_i \geq 0 \quad (2.8)$$

$y = \text{produtos}$; $x = \text{insumos}$; $u, v = \text{pesos}$

$r = 1, \dots, m$; $i = 1, \dots, n$; $j = 1, \dots, N$

O modelo busca minimizar os consumos de insumos (*inputs*) de forma a produzir no mínimo o nível de produção dado, expresso pela maximização da somatória das quantidades produzidas y multiplicadas pelos pesos (preços) u.

A primeira restrição (eq. 2.6) pode ser definida como o resultado da empresa, pois é a subtração entre o somatório das quantidades produzidas multiplicadas pelos pesos (preços) dos produtos

$(\sum_{r=1}^m u_r y_{rj})$ e o somatório da multiplicação dos insumos consumidos pelos pesos $(\sum_{i=1}^n v_i x_{ij})$. Está

limitado a 0. Assim, as empresas eficientes obterão o resultado 0 para a primeira restrição.

Na segunda restrição (eq. 2.7), o somatório do produto das quantidades consumidas de recursos pelos pesos (preços) específicos para a empresa k $(\sum_{i=1}^n v_i x_{ik})$ é igual a 1. Portanto, o máximo resultado possível de se obter para h_k é 1. Se a empresa k for eficiente, h_k será igual a 1. Se não for, obterá um indicador sempre inferior a 1.

Após resolver o programa linear para cada uma das empresas, obtém-se o grupo de empresas eficientes, para as quais h_k é igual a 1.

Modelo CCR com orientação a produto (*output*)

$$\text{Minimizar } h_k = \sum_{i=1}^n v_i x_{ik}, \quad (2.9)$$

sujeito a

$$\sum_{r=1}^m u_r y_{rj} - \sum_{i=1}^n v_i x_{ij} \leq 0 \quad (2.10)$$

$$\sum_{r=1}^m u_r y_{rk} = 1 \quad (2.11)$$

$$u_r, v_i \geq 0 \quad (2.12)$$

$y = \text{produtos}$; $x = \text{insumos}$; $u, v = \text{pesos}$

$r = 1, \dots, m$; $i = 1, \dots, n$; $j = 1, \dots, N$

O objetivo é a maximização do nível de produção utilizando no máximo o consumo de insumos observados. Os modelos são equivalentes e pressupõem retornos constantes de escala.

O outro Modelo de DEA é chamado de BCC, abreviatura de Banker, Charnes e Cooper, que o desenvolveram e apresentaram em artigo publicado na *Management Science* em 1984. O Modelo BCC pressupõe que as unidades avaliadas apresentem retornos variáveis de escala. Segundo BELLONI (2000: 68), “ao possibilitar que a tecnologia exiba propriedades de retornos à escala diferentes ao longo de sua fronteira, esse modelo admite que a produtividade máxima varie em função da escala de produção”.

Modelo BCC com orientação aos insumos (*inputs*)

$$\text{Maximizar } \sum_{r=1}^m u_r y_{rk} - u_k, \quad (2.13)$$

sujeito a

$$\sum_{i=1}^n v_i x_{ik} = 1 \quad (2.14)$$

$$\sum_{r=1}^m u_r y_{rj} - \sum_{i=1}^n v_i x_{ij} - u_k \leq 0 \quad (2.15)$$

$$u_r, v_i \geq 0 \quad (2.16)$$

$y = \text{produtos}$; $x = \text{insumos}$; $u, v = \text{pesos}$

$r = 1, \dots, m$; $i = 1, \dots, n$; $j = 1, \dots, N$

Percebe-se que é introduzida uma variável u_k , representando os retornos variáveis de escala. Essa variável não deve atender à restrição de positividade; pode, portanto, assumir valores negativos.

Modelo BCC com orientação ao produto (*output*)

(2.17)

$$\text{Minimizar } \sum_{i=1}^n v_i x_{ki} + v_k,$$

sujeito a

$$\sum_{r=1}^m u_r y_{rk} = 1 \quad (2.18)$$

$$\sum_{r=1}^m u_r y_{jr} - \sum_{i=1}^n v_i x_{ji} - v_k \leq 0 \quad (2.19)$$

$$u_r, v_i \geq 0 \quad (2.20)$$

$y = \text{produtos}$; $x = \text{insumos}$; $u, v = \text{pesos}$

$r = 1, \dots, m$; $i = 1, \dots, n$; $j = 1, \dots, N$

Onde, novamente, o termo v_k representa a possibilidade de retornos de escala variáveis, podendo assumir valores negativos ou positivos.

Como ressaltado anteriormente, a possibilidade de retornos de escala variáveis do Modelo BCC admite que a produtividade máxima varie em função da escala de produção. O modelo permite, portanto, a utilização de unidades de portes distintos (BELLONI, 2000: 68).

Os Modelos CCR e BCC apresentam regiões de viabilidade distintas. A região viável do Modelo BCC é restringida a combinações convexas dos planos de produção observados, o que é caracterizado pelos retornos variáveis de escala. Como consequência, considerando orientação ao produto, o indicador de eficiência do Modelo BCC é menor ou igual ao indicador de eficiência do Modelo CCR (BELLONI, 2000: 69).

2.2 A condição da homogeneidade das DMUs

Saen e Memariani (2005, p. 313) comentam que o DEA pode ser aplicado em qualquer organização ou indústria onde um conjunto de DMUs razoavelmente homogêneas usem o mesmo conjunto de *inputs*, possivelmente em diferentes combinações, para produzir um conjunto de *outputs* identificáveis e, possivelmente, em diferentes combinações.

O pressuposto do modelo DEA clássico é baseado no princípio de que DMUs consomem *inputs* comuns para produzir *outputs* comuns. Em muitas aplicações de DEA esse pressuposto não é seguido.

É necessário saber quais são os fatores que afetam a avaliação da performance de DMUs sem uma certa função, comparadas com DMUs que possuem essa função.

As pesquisas que lidam com a não homogeneidade do DEA geralmente se têm concentrado na substituição dos dados faltantes (*missing data*), usando informações relativas a outras DMUs

e fazendo algum tipo de interpolação. Mas essa opção não proporciona uma boa estimativa dos dados faltantes. (SAEN e MEMARIANI, 2005, p. 313)

De acordo com Haas e Murphy (2003), uma suposição fundamental do DEA é que todas as DMUs em avaliação sejam homogêneas. No DEA essa suposição implica três condições: (1) as DMUs devem estar engajadas no mesmo processo; (2) as mesmas medidas de eficiência, definidas pelos *inputs* e *outputs* selecionados para o modelo DEA, devem ser aplicadas igualmente para todas as DMUs; (3) todas as DMUs devem ser operadas sob as mesmas condições.

Diversos autores têm feito aplicações do DEA nos casos em que as DMUs em análise são heterogêneas. Entre os autores podem ser citados:

Huang e McLaughlin (1989, apud HAAS e MURPHY (2003, p. 531) compararam os escores de eficiência gerados pelo DEA para diversos casos referentes a proporções de custos e obtiveram os resultados de uma regressão para avaliar os resultados do DEA.

Ray (1991, apud HAAS e MURPHY (2003, p. 531) calculou os escores de eficiência ajustados empregando um modelo de regressão e seus resíduos, com a utilização dos escores do DEA como variável dependente para identificar ineficiências administrativas não causadas por fatores extrínsecos.

Fizel e Nunnikhoven (1992, 1993, apud HAAS e MURPHY (2003, p. 531) também usaram a análise de regressão e os escores de eficiência gerados pelo DEA. Eles empregaram os resíduos da regressão como escores ajustados, tendo como objetivo determinar se as diferenças nos escores persistiam quando a análise era controlada por causas extrínsecas a essas diferenças.

Oum e Yu (1994, apud HAAS e MURPHY (2003, p. 531) e Nolan (1996, apud HAAS e MURPHY (2003, p. 531) utilizaram uma regressão com uma variável dependente limitada contra fatores exógenos para identificar variáveis controláveis pela administração que contribuíam para a eficiência.

Sexton *et al* (1994, apud HAAS e MURPHY (2003, p. 531) também usaram uma regressão sobre os escores de eficiência gerados pelo DEA para calcular os escores de eficiência ajustados. Seu procedimento é muito diferente, entretanto, eles aproveitaram o modelo de regressão e seus resíduos para calcular os valores ajustados dos *outputs* para o DEA e então aplicaram um segundo DEA para produzir os valores ajustados dos escores de eficiência para serem usados com os mesmos propósitos, como escores de eficiência para DMUs homogêneas.

Em um estudo abrangendo 144 agências bancárias, Soteriou e Zenios (1999, apud HAAS e MURPHY (2003, p. 531) lidaram com a não-homogeneidade alocando as agências em grupos homogêneos de acordo com o tipo de localização e tamanho. Daí, compararam a eficiência dentro dos grupos e também a eficiência entre os grupos.

O estudo de Haas e Murphy (2003, p. 531-532) explora e compara três diferentes métodos de compensação da homogeneidade. O primeiro é um método de dois estágios descrito por Sexton *et al* (1994, apud HAAS e MURPHY (2003, p. 531). Em seguida, são apresentados também dois métodos adicionais, demonstrando-se que estes possuíam as propriedades de consistência desejadas conforme apresentadas inicialmente por Sexton, Sleeper e Taggart (1994, apud HAAS e MURPHY (2003, p. 531) em seu método. Haas e Murphy avaliaram os três métodos usando DEA e empregaram parâmetros gerados aleatoriamente, simulando diferentes tamanhos de campos de avaliação, diferentes intervalos nos valores de *inputs* e *outputs* e diferentes graus de magnitude da não-homogeneidade.

Haas e Murphy (2003, p. 532) ainda compararam os resultados do DEA explorando a sensibilidade de cada um dos métodos para o tamanho, intervalos de valores dos *inputs* e *outputs* e magnitude da não homogeneidade.

Dadas as condições estabelecidas nos testes, esses autores (2003, p. 542) finalizaram, afirmando que não haveria evidências suficientes para estabelecer de forma conclusiva que qualquer dos métodos de compensação para a não-homogeneidade fosse superior ao do modelo CCR não ajustado para a avaliação da eficiência das DMUs sob diferentes condições de operação.

Em sua pesquisa, Thanassoulis (1996) escolheu como campo de estudo as unidades operadoras de recursos, como escolas, hospitais ou unidades que ofereçam algum tipo de serviço. No caso de recursos para escolas, a alocação poderia ser feita de acordo com uma espécie de “nível marginal de recursos”, isto é, taxas de recursos por unidade referentes a cada atividade de uma unidade operadora de recursos, considerada como uma DMU. Essas DMUs recebem recursos e prestam algum tipo de serviço. No caso de uma escola, por exemplo, a taxa de recurso por unidade poderia ser o valor do recurso por aluno comum, ou então o valor do recurso por aluno com necessidades especiais e assim por diante.

Esse estudo consistiu em (1) formar grupos homogêneos de unidades operadoras de recursos utilizando o instrumento *cluster analysis* e (2) determinar os níveis de eficiência de cada DMU através da utilização do DEA. Por meio de um exemplo hipotético, a pesquisa mostrou inicialmente como separar DMUs em grupos homogêneos em função de dois *outputs*. A partir dessa separação em DMUs com características homogêneas (*clusters*), determinaram-se quais eram as DMUs mais eficientes. Os escores de eficiência foram obtidos com a utilização do DEA. Assim, a alocação dos recursos para as DMUs seria feita em função dos níveis de eficiência da DMUs dentro de cada *cluster*.

A pesquisa de Saen e Memariani (2005, p. 314) sugere uma metodologia que leva em conta a substituição dos dados faltantes através de interpolação e também do potencial de determinada DMU naquela função em que ela se ausenta. Assim, sugerem uma estimação bidimensional considerada para reduzir o viés do escore de eficiência final.

De acordo com Saen e Memariani (2005, p. 315), o conceito e a importância da heterogeneidade no DEA foi discutida por Mester (1997), Dyson *et al* (2001), Sarrico e Dyson (2000), Athanassopoulos e Thanassoulis (1995), Sexton *et al* (1994) e Hass e Murphy (2003).

Mester (1997, apud SAEN e MEMARIANI (2005, p. 315) examinou o problema da heterogeneidade dos dados e concluiu que as estimativas de eficiência dos bancos podem ser estimadas se for ignorada a questão da heterogeneidade, apesar de não apresentar qualquer método para lidar com a heterogeneidade.

Dyson *et al* (2001, apud SAEN e MEMARIANI, 2005, p. 315-316) também trataram da heterogeneidade, argumentando que um possível enfoque seria testar comparadores externos e depois fazer comparações internas com relação às DMUs no seu próprio campo.

Em Sarrico e Dyson (2000, apud SAEN e MEMARIANI, 2005, p. 316), foi apresentado um exemplo da abordagem citada no parágrafo anterior, mas isso não causou a eficiência atual e geral das DMUs.

Ainda nesse sentido, Athanassopoulos e Thanassoulis (1995, apud SAEN e MEMARIANI, 2005, p. 316) sugeriram separar as DMUs em grupos homogêneos (*clusters*), embora isso também não tenha produzido a eficiência atual e geral das DMUs.

Sexton *et al* (1994 apud SAEN e MEMARIANI, 2005, p. 316) expuseram um método de dois estágios. Num primeiro estágio foi elaborado o DEA com a utilização de dados brutos, produzindo escores eficientes para todas as DMUs em seu campo de atuação. A seguir aplicou-se a análise de regressão múltipla e o método *stepwise* sobre o conjunto de escores eficientes usando um conjunto de DMUs características em que se esperava conter diferenças na eficiência não atribuíveis à administração. Depois, para levar em conta a não homogeneidade das DMUs características, os *outputs* das DMUs foram ajustados e se obteve um segundo DEA para produzir novos escores de eficiência com base nos dados ajustados. Esse método, porém, não considerou os pesos relativos das DMUs.

Hass e Murphy (2003 apud SAEN e MEMARIANI, 2005, p. 316) demonstraram dois métodos para compensar a não homogeneidade. O primeiro método utilizou a análise de regressão para determinar a magnitude dos erros (resíduos). *Inputs* e *outputs* foram estimados, usando-se análise de regressão em que as variáveis independentes são as causas da não homogeneidade. Aplicaram o DEA utilizando dados atuais e previstos dos *inputs* e *outputs*. O segundo método empregou a razão entre os valores atuais e os valores previstos, aplicando um DEA em que usaram a razão entre os valores atuais e previstos dos *inputs* e os valores atuais e previstos dos *outputs*. Nesses métodos não foram usados os pesos relativos das DMUs.

O método de Saen e Memariani, (2005, p. 316), ao contrário dos métodos de Hass e Murphy (2003), considerou os pesos relativos das DMUs. De acordo com a metodologia empregada,

os valores dos dados faltantes foram estimados levando-se em conta dois critérios: (1) os valores de outras DMUs e (2) o valor potencial da DMUs que apresentavam a falta de um determinado fator.

Quanto ao primeiro critério, foram feitas interpolações usando as médias de outras DMUs, embora pudesse ser usado algum outro parâmetro. Depois, foi usado o *Analytic Hierarchy Process* (AHP) (SAATY, 2005) para mensurar o valor potencial dos fatores faltantes. Os autores apresentaram uma metodologia para considerar os pesos relativos de cada DMU. A seguir, fez-se uma aplicação prática do método, para a qual foi utilizado o caso da *Iranian Research Organization for Science and Technology* (IROST), uma entidade de pesquisa e desenvolvimento com 16 DMUs em diferentes partes do Irã. (SAEN e MEMARIANI, 2005, p. 320)

Saen e Memariani (2005, p. 326), em suas conclusões, afirmaram: “Duas DMUs completamente diferentes não podem ser avaliadas pelo modelo DEA. Mas se elas são comuns em alguns fatores, então o DEA pode produzir uma avaliação confiável da eficiência para essas DMUs”

2.3 Cluster Analysis

Este trabalho se propõe a utilizar a *Cluster Analysis* para a separação das empresas (DMUs) em grupos com características homogêneas.

Cluster Analysis é uma técnica que, através de critérios de similaridade (ou dissimilaridade) e de algoritmos matemáticos, permite reunir objetos em grupos a partir de variáveis selecionadas pelo pesquisador.

Segundo Angelo e Sanvicente (1992, p. 113), o mercado pode sofrer influências externas comuns, mas a “conduta” das diversas empresas que compõem um determinado setor não é necessariamente a mesma. As empresas podem escolher tecnologias diferentes ou então podem financiar suas atividades com proporções diferentes entre capital de terceiros e capital próprio.

Como consequência da escolha entre essas alternativas, empresas do mesmo setor podem apresentar diversos graus de imobilização do ativo e de endividamento. Em função da escolha de diferentes alternativas de conduta, empresas que atuam no mesmo mercado acabam apresentando desempenho distinto (ANGELO e SANVICENTE, 1992, p. 114).

De acordo com Hair Junior e Black (2000, p. 147), *clusters analysis* é uma das técnicas de análise multivariada cujo propósito primário é reunir objetos, baseando-se nas características dos mesmos. Essa técnica classifica objetos (empresas, respondentes a questionários, produtos etc.) segundo a similaridade de cada objeto referente a outros dentro do grupo com respeito a um critério de seleção predeterminado. O grupo resultante deve então exibir um alto grau de homogeneidade interna e alta heterogeneidade externa. Desse modo, se a classificação é bem sucedida, os objetos dentro do grupo estarão juntos quando “plotados” geometricamente e diferentes grupos estarão distantes uns dos outros.

Assim, o valor primário da análise *cluster* repousa na classificação de dados, como sugerido pelo agrupamento natural de dados por si mesmo.

Bussab, Miazaki e Andrade (1990, p. I) referem-se a tal análise como sendo aquela que engloba uma variedade de técnicas e algoritmos, com o objetivo de encontrar e separar objetos em grupos similares. O problema a ser resolvido é: “dada uma amostra de n objetos (ou indivíduos), cada um deles medido segundo p variáveis, deve-se procurar um esquema de classificação que agrupe objetos em g grupos. Devem ser determinados também o número e as características desses grupos” (EVERITT *apud* BUSSAB, MIAZAKI e ANDRADE, 1990, p. I).

Cluster analysis é uma técnica em que nenhuma definição prévia é feita com relação ao número de grupos ou à sua estrutura. O processo de agrupamento é feito com base em similaridades ou distâncias (dissimilaridades). Os *inputs* exigidos são medidas de similaridade ou dados a partir dos quais as similaridades possam ser computadas (POHLMANN *apud* CORRAR, PAULO e DIAS FILHO, 2007, p. 326).

Segundo Green (*apud* A e SANVICENTE, 1990, p. 23), *cluster analysis* é uma técnica multivariada que permite formar grupos de objetos de acordo com suas semelhanças. Para ele, existem três questões básicas em *clusters analysis*:

- a) Como medir a semelhança entre objetos?
- b) Supondo que se possa medir a semelhança relativa entre objetos, como então colocar objetos semelhantes em *clusters*?
- c) Após ter efetuado o agrupamento, como descrever os *clusters* e saber se os *clusters* resultantes são reais, e não o produto de um simples artifício estatístico?

Sobre essas perguntas, Pohlmann *apud* Corrar, Paulo e Dias Filho (2007, p. 326) afirmam que:

No que se refere à primeira questão, a resposta está no conceito de “semelhança”: dois objetos são considerados semelhantes se seus perfis, em termos das variáveis utilizadas, são próximos, o que é determinado por medidas de distância, como a distância euclidiana entre pontos, ou através de um processo de *matching*, quando as escalas das variáveis são nominais.

A segunda pergunta é respondida através dos diversos algoritmos disponíveis para *clusters analysis*. Todos os algoritmos, porém, envolvem a determinação de ótimos locais, e não globais. De qualquer maneira, cabe à teoria a definição *a priori* do número de *clusters*, e não aos algoritmos.

De acordo com Figueiredo (2002, p. 3), pode-se descrever sinteticamente esse método do seguinte modo: dado um conjunto de n indivíduos para os quais existe informação sobre a forma de p variáveis, o método de análise de *clusters* procede ao agrupamento dos indivíduos em função da informação existente, de tal modo que os indivíduos pertencentes a um mesmo grupo sejam tão semelhantes quanto possível e sempre mais semelhantes aos elementos do mesmo grupo do que a elementos dos grupos restantes.

Ainda segundo esse referido autor, é fundamental ter particular cuidado na seleção das variáveis que vão caracterizar cada indivíduo. Nessa análise, não existe qualquer tipo de dependência entre as variáveis, isto é, os grupos configuram-se por si mesmos sem necessidade de ser definida uma relação causal entre as variáveis utilizadas.

Landim (2002, p. 1) acrescenta que a técnica classificatória multivariada da análise de agrupamentos é utilizada quando se deseja explorar as similaridades entre indivíduos (modo *Q*) ou entre variáveis (modo *R*), definindo-os em grupos, considerando simultaneamente, no

primeiro caso, todas as variáveis medidas em cada indivíduo e, no segundo, todos os indivíduos para os quais foram feitas as mesmas mensurações. Segundo esse método, desenvolvido inicialmente em Zoologia por taxonomistas numéricos, buscam-se agrupamentos homogêneos de itens representados por pontos num espaço n-dimensional onde se encontra um número conveniente de grupos, relacionando-os através de coeficientes de similaridade ou de semelhança.

Para Pohlmann apud Corrar, Paulo e Dias Filho (2007, p. 326),

Cluster analysis pode ser caracterizada como descritiva, ateorética e não-inferencial. A análise *cluster* não tem base estatística sobre a qual extrair inferências de uma população com base em uma amostra, e é usada precipuamente como técnica exploratória. As soluções não são únicas, porque os membros de qualquer número de soluções dependem dos elementos dos procedimentos; muitas soluções diferentes podem ser obtidas variando um ou mais elementos. Além disso, a análise *cluster* sempre cria grupos independentemente da verdadeira existência de qualquer estrutura nos dados.

Finalmente, a solução obtida é totalmente dependente das variáveis usadas como base de mensuração da similaridade. A adição ou exclusão de variáveis relevantes pode ter um substancial efeito na solução resultante. Desse modo, o pesquisador deve tomar cuidado na avaliação dos efeitos de cada decisão envolvida na realização da *cluster analysis* (HAIR, 2000, p. 149).

Adicionalmente, Johnson e Wichern (1998, p. 726) ressaltam que, na maioria das aplicações práticas da análise *cluster*, o pesquisador conhece suficientemente o problema a ponto de distinguir “bons” agrupamentos de “maus” agrupamentos. Busca-se então o porquê de não se enumerarem todas as possibilidades de agrupamento e de selecionar a “melhor” para posterior estudo. Os autores ilustram essa questão analisando as inúmeras possibilidades de agrupamento possíveis de cartas de um baralho e a impossibilidade de se escolher a melhor. Dessa maneira, a aplicação dessa análise, segundo Landim (2002, p. 1), é controversa.

Hair *et al* (1998, p. 475-476) explicam que:

O objetivo primário da análise *cluster* é definir a estrutura de dados colocando as observações mais similares dentro de grupos. Para executar esta tarefa, entretanto, o pesquisador deve tratar de três questões.

Primeiro, como medir a similaridade? Isso exige um método para comparar simultaneamente observações sobre duas variáveis de agrupamento (v_1 e v_2). Vários métodos são possíveis, incluindo a correlação entre objetos, uma medida de associação usada em outras técnicas

multivariadas; ou, possivelmente, uma medida de sua proximidade em um espaço bidimensional, tal que a distância entre as observações indique similaridade.

Segundo, como formar grupos? Não importa como a similaridade é medida, o procedimento deve agrupar, em um único grupo, aquelas observações que são mais similares. Este procedimento deve determinar, para cada observação, quais as outras observações que serão agrupadas com ela.

Terceiro, quantos grupos serão formados? A tarefa fundamental é medir o grau de similaridade dentro dos *clusters*, tal que crescendo a similaridade, menos similares serão os grupos. O pesquisador se depara então com uma questão: menos grupos versus menos homogeneidade.

Uma simples estrutura, num esforço em direção à parcimônia ou simplificação, é refletida em poucos grupos. Por outro lado, conforme o número de *clusters* diminui, a homogeneidade dentro dos *clusters* necessariamente diminui. Dessa forma, um equilíbrio deve ser encontrado na definição de uma estrutura mais básica (menos *clusters*) de forma a assegurar o necessário nível de similaridade dentro dos *clusters*.

A seleção da solução final para a análise *cluster* requer um substancial julgamento do pesquisador, e isso é considerado por muitos como extremamente subjetivo. Ainda que métodos mais sofisticados tenham sido desenvolvidos para a análise *cluster*, isso ainda não é suficiente para o pesquisador decidir qual o número de *clusters* deve ter a solução final.

A seleção das variáveis a serem incluídas no *cluster variate* deve ser feita tendo em vista aspectos tanto teóricos e conceituais como práticos. Qualquer aplicação da análise *cluster* deve ter alguma lógica em relação às variáveis selecionadas. Se a lógica é baseada em uma teoria explícita, em pesquisa anterior ou em suposições, o pesquisador deve ter noção quanto à importância de incluir somente aquelas variáveis que: (1) caracterizem os objetos agrupados, e (2) relacionem-se especificamente aos objetivos da análise *cluster* (HAIR, ANDERSON, TATHAM e BLACK, 1998, p.481 - 482) .

A técnica da análise *cluster* não permite diferenciar variáveis relevantes daquelas irrelevantes. A inclusão de uma variável irrelevante aumentará as chances de que *outliers* sejam criados nessas variáveis, podendo ter um substantivo efeito sobre os resultados. Dessa forma, não se devem incluir indiscriminadamente variáveis, mas ao contrário, cuidadosamente escolhê-las tendo por critério de seleção o objetivo da pesquisa (HAIR, ANDERSON, TATHAM e BLACK, 1998, p. 482).

Na prática, a análise *cluster* pode ser dramaticamente afetada pela inclusão de apenas uma ou duas variáveis não apropriadas ou indiferentes. O pesquisador é sempre encorajado a examinar os resultados e eliminar as variáveis que não são distintivas (isto é, que não diferem significativamente) dentre os grupos resultantes. Esse procedimento possibilita à análise *cluster* definir, com a maior segurança possível, os grupos baseados apenas naquelas variáveis que permitem diferenciação entre os objetos (HAIR, ANDERSON, TATHAM e BLACK, 1998, p. 482).

Bussab, Miazaki e Andrade (1990, p. 77) ressaltam que frequentemente o número de variáveis medidas é grande, dificultando a análise. Deve-se, então, respeitando o princípio da parcimônia, procurar diminuir essa quantidade, de forma que sua seleção contemple tanto a sua relevância como seu poder de discriminação face ao problema analisado. Em último caso, podem ser ainda utilizadas técnicas estatísticas para a redução da dimensionalidade da matriz de dados, tais como a Análise de Componentes Principais e a Análise Fatorial.

Conforme esses autores, variáveis que assumem praticamente o mesmo valor para todos os objetos são pouco discriminatórias e sua inclusão pouco contribuiria para a determinação da estrutura do agrupamento. Por outro lado, a inclusão de variáveis com grande poder de discriminação, mesmo que sejam irrelevantes ao problema, pode mascarar os grupos e levar a resultados equivocados. (BUSSAB, MIAZAKI E ANDRADE, 1990, p. 77)

Com os objetivos definidos e as variáveis selecionadas, o pesquisador deve tratar de três questões antes de iniciar o processo de agrupamento (HAIR, ANDERSON, TATHAM e BLACK, 1998, p. 482): (1) existem *outliers* e, em caso afirmativo, devem eles ser excluídos? (2) como deve ser medida a similaridade entre os objetos? (3) os dados devem ser padronizados?

Muitas abordagens diferentes podem ser usadas para solucionar essas questões, entretanto nenhuma delas tem sido considerada suficiente para oferecer uma resposta definitiva a qualquer dessas perguntas. Infelizmente, a maioria das abordagens implica diferentes resultados para o mesmo conjunto de dados. Assim, a análise *cluster* é muito mais uma arte do que uma ciência. Por essas razões, revisam-se tais questões de uma forma muito ampla, através da apresentação de exemplos das abordagens mais comumente empregadas e, quando

possível, da referência quanto a limitações práticas (POHLMANN apud CORRAR, PAULO e DIAS FILHO, 2007, p. 331).

Na sua busca por uma estrutura, a análise *cluster* é bastante sensível à inclusão de variáveis irrelevantes, bem como a *outliers*. Para Hair, Anderson, Tatham e Black (1998, p. 482), os *outliers* podem representar ou (1) observações que podem ser chamadas de verdadeiras “aberrações”, não representativas da população geral; ou (2) itens de um determinado grupo, obtidos de uma amostragem ruim de uma dada população, ocasionando uma má representação dos grupos através daquela amostra. (POHLMANN apud CORRAR, PAULO e DIAS FILHO, 2007, p. 333).

Em ambos casos, os *outliers* distorcem a real estrutura e tornam os grupos derivados não-representativos da verdadeira estrutura da população. Por essa razão, uma triagem preliminar em busca de *outliers* é necessária. (POHLMANN apud CORRAR, PAULO e DIAS FILHO, 2007, p. 332).

O conceito de similaridade é fundamental para a análise *cluster*. A similaridade entre objetos é uma medida de correspondência, ou semelhança, entre objetos a serem agrupados. Na análise fatorial, uma matriz de correlação entre variáveis é usada para agrupar variáveis dentro de fatores; um processo comparável a esse ocorre na análise *cluster* (HAIR, ANDERSON, TATHAM e BLACK, 1998, p. 483).

Nesse sentido, as características que definem similaridade são especificadas primeiramente. Em seguida, elas são combinadas segundo uma medida de similaridade e calculadas para todos os pares de objetos, assim como se usa a correlação entre variáveis na análise fatorial. Assim, qualquer objeto pode ser comparado a outro através da medida de similaridade. A análise *cluster* então reúne objetos similares dentro de um mesmo grupo (POHLMANN apud CORRAR, PAULO e DIAS FILHO, 2007, p. 333).

A similaridade entre objetos pode ser mensurada de várias maneiras, mas três métodos dominam as aplicações da análise *cluster*: medidas de correlação, medidas de distância e medidas de associação. Cada um desses métodos representa uma particular perspectiva de similaridade, dependendo dos objetivos e do tipo de dados. As medidas de correlação e de distância requerem dados quantitativos, enquanto as medidas de associação se constituem

para dados qualitativos (POHLMANN apud CORRAR, PAULO e DIAS FILHO, 2007, p. 333).

Bussab, Miazaki e Andrade (1990, p. 15) utilizam a expressão *coeficiente de parecença* para se referir ao critério que mede a distância entre dois objetos, ou quantifique o quanto eles são parecidos, dividindo-o em duas categorias: medidas de similaridade e de dissimilaridade.

Na primeira, quanto maior for o valor observado, mais parecidos são os objetos; já para a segunda, quanto maior o valor observado, menos parecidos (mais dissimilares) são os objetos. O coeficiente de correlação é um exemplo de medida de similaridade, enquanto a distância euclidiana é um exemplo de dissimilaridade. De um modo geral, é possível construir uma medida de dissimilaridade a partir de uma de similaridade, e vice-versa, razão pela qual preferem usar o termo *matriz de parecença* para indicar semelhança ou distância entre objetos. (BUSSAB, MIAZAKI E ANDRADE, 1990, p. 15)

Johnson e Wichern (1998, p. 728), por sua vez, afirmam que a maioria dos esforços para produzir uma melhor estrutura a partir de um conjunto de dados complexos exige uma medida de proximidade ou similaridade. Existe freqüentemente um considerável grau de subjetividade envolvido na escolha da medida de similaridade. Importantes considerações incluem a natureza das variáveis (discretas, contínuas, binárias), escalas de medida (nominal, ordinal, intervalar, proporcional) e o conhecimento da matéria, objeto da pesquisa ou análise. Quando itens são agrupados, a proximidade é normalmente indicada por algum tipo de medida de distância. Por outro lado, variáveis são normalmente agrupadas com base nos coeficientes de correlação ou semelhantes medidas de associação.

A medida de similaridade entre objetos que provavelmente se nota em primeiro lugar é o coeficiente de correlação entre um par de objetos mensurados a partir de certas variáveis. Alta correlação indica similaridade e baixa correlação denota ausência da mesma. Esse procedimento é seguido na aplicação da análise fatorial “*Q-type*” (HAIR, ANDERSON, TATHAM e BLACK, 1998, p. 484). Medidas correlacionais representam a similaridade pela correspondência de padrões entre as características (variáveis X).

As medidas de distância são, na verdade, medidas de dissimilaridade com altos valores, denotando menor similaridade. A distância é convertida dentro de uma medida de

similaridade através de utilização de uma relação inversa (POHLMANN apud CORRAR, PAULO e DIAS FILHO, 2007, p. 336).

A forma mais comum de padronização é a conversão de cada variável para escores padrões (também conhecidos como escores *Z*) pela subtração da média e divisão pelo desvio padrão de cada variável. Essa é uma opção disponível em todos os programas de computador e muitas vezes é mesmo diretamente inclusa nos procedimentos de análise *cluster* (POHLMANN apud CORRAR, PAULO e DIAS FILHO, 2007, p. 341).

Esse modo de padronização é a forma geral de uma função de distância normalizada, que utiliza uma medida de distância euclidiana, responsável pela transformação dos dados originais. Esse processo converte cada escore original de dados em um valor padronizado com média igual a zero e desvio padrão igual a um. Essa transformação, por sua vez, elimina a distorção introduzida pelas diferentes escalas de vários atributos ou pelas variáveis usadas na análise. (POHLMANN apud CORRAR, PAULO e DIAS FILHO, 2007, p. 341).

Um outro aspecto a ser considerado é a possibilidade de ponderação das variáveis, especialmente naquelas hipóteses em que as variáveis não têm a mesma importância para o problema. Nesses casos, os pesos são atribuídos de forma subjetiva e a sua escolha também depende do contexto e do grau de conhecimento do pesquisador sobre o problema. É importante ressaltar, entretanto, que, como a análise *cluster* é uma técnica exploratória em busca da formulação de hipóteses, a partir dos dados, o mais comum é se atribuir o mesmo peso para todas as variáveis (BUSSAB, MIAZAKI e ANDRADE, 1990, p. 79).

A análise *cluster* não é uma técnica de inferência estatística em que parâmetros de uma amostra são avaliados como sendo possivelmente representativos de uma população. Em vez disso, ela se constitui em uma metodologia objetiva para quantificar características estruturais de um conjunto de observações. Como tal, ela tem propriedades matemáticas fortes, mas não fundamentos estatísticos. Os requisitos de normalidade, linearidade e homocedasticidade (níveis iguais de variância da variável dependente ao longo de todas as variáveis preditoras), muito importantes em outras técnicas, realmente têm pouco significado na análise *cluster*. O pesquisador deve manter o foco, entretanto, em dois pontos críticos: representatividade da amostra e multicolinearidade (HAIR, ANDERSON, TATHAM e BLACK, 1998, p. 490).

Provavelmente, a questão mais complexa para o pesquisador é determinar o número final de grupos a serem formados, bem como saber a regra de parada. Infelizmente, não existe um procedimento padrão objetivo de seleção. Devido ao fato de não existir um critério estatístico interno usado para inferência, como testes estatísticos de significância de outros métodos multivariados, os pesquisadores devem desenvolver muitos critérios e guias para abordar o problema. O principal obstáculo é que existem muitos procedimentos de ocasião, os quais devem ser desenvolvidos pelo pesquisador, mesmo envolvendo muitas vezes aspectos bastante complexos (HAIR, ANDERSON, TATHAM e BLACK, 1998, p. 499).

Um tipo de regra de parada relativamente simples é examinar alguma medida de similaridade ou distância entre grupos a cada passo sucessivo, com a solução a ser definida quando a medida de similaridade exceder a um valor especificado ou quando os sucessivos valores entre os passos tiverem uma súbita elevação (POHLMANN apud CORRAR, PAULO e DIAS FILHO, 2007, p. 352).

Dentre os objetivos possíveis de uma análise *cluster*, pode-se determinar, para o presente estudo, primeiramente, o de identificar agrupamentos e descrever os grupos resultantes. Num segundo momento, tenta-se explorar mais profundamente as relações existentes entre os mesmos.

2.4 Indicadores contábeis

Um dos objetivos desta tese é provar que é possível avaliar a eficiência das empresas através da utilização de um pequeno número de indicadores contábeis. A questão inicial a ser respondida é: que indicadores contábeis devem ser utilizados para a avaliação de eficiência das empresas?

Os autores de textos sobre análise de balanços propõem uma lista infindável de indicadores obtidos do balanço, da demonstração de resultados, da demonstração das origens e aplicações de recursos, etc. A partir disso, a dúvida é justamente sobre como conseguir selecionar indicadores significativos.

Lyra (2008) realizou um estudo exploratório sobre livros de análise de demonstrações financeiras editados no Brasil a partir de 1980. O objetivo foi avaliar o conteúdo e ter uma visão dos assuntos abordados: que indicadores usavam, a utilização ou não de métodos quantitativos como instrumento adicional etc. Segundo Lyra (2008, p. 25), os aspectos principais são analisados nos seguintes livros:

- a) Apenas as obras de Iudícibus (2007), Marion (2005a) e Blatt (2001) tratam da análise de assuntos como Demonstração do Valor Adicionado (DVA) e *Economic Value Added* (EVA);
- b) Quatro livros (MATARAZZO, 1985, 1998; PEREIRA DA SILVA, 2004; MARION, 2005a; RIBEIRO, 1997) possuem um capítulo que trata dos índices-padrões;
- c) Apenas a obra de Pereira da Silva (2004) trata do uso de alguma técnica estatística, mais especificamente a aplicação da análise discriminante;
- d) Os livros de Matarazzo (1985), Savytzky (2007), Zdonowicz (1998) e Blatt (2001) tratam de previsões de falência. Contudo, limitam-se ao emprego de fórmulas já consagradas como as de Altmann e a de Kanitz, dentre outros;
- e) Dois livros, Zdonowicz (1998) e Marion (2005a) apresentam um capítulo específico para análise da Demonstração das Origens e Aplicações de Recursos (DOAR);
- f) Quatro livros possuem capítulo específico para análise da Demonstração do Fluxo de Caixa (DFC), Blatt (2001), Matarazzo (1985), Pereira da Silva (2004) e Marion (2005a).
- g) Apenas os livros de Matarazzo (1985 e 1998) e Iudícibus (2007) possuem um capítulo dedicado aos problemas de perda do poder aquisitivo da moeda para análise;
- h) O livro de Olinquevitch e Santi Filho (2004) diferencia-se dos demais pelo enfoque único na análise por fluxos, baseados na análise dinâmica do professor Fleuriet et al. (1978). Três outros livros, Matarazzo (1985), Pereira da Silva (2004) e Assaf Neto (2002), dedicam um capítulo a este assunto;
- i) Tópicos sobre objetivos e alcance da análise, estrutura das demonstrações, padronização, análise de liquidez, rentabilidade, endividamento, índices de prazo, análise vertical e horizontal são comuns em todos os livros, sendo as diferenças entre um livro e outro mais um problema de estilo literário dos autores do que de conteúdo.

2.4.1 Seleção de indicadores contábeis através da técnica Delphi

O professor Ricardo Luiz Wüst Corrêa de Lyra (2008) respondeu a essa questão de seleção de indicadores em sua tese de doutorado apresentada ao Departamento de Contabilidade e Atuária da Universidade de São Paulo. Ele utilizou a técnica Delphi e a participação de especialistas. Resumidamente, destacam-se os seguintes aspectos do procedimento utilizado pelo Professor Lyra (2008, p. 85- 94):

- (a) elaboração de questionário com um conjunto de perguntas e uma relação de indicadores contábeis definidos com base na literatura existente;
- (b) pré-teste realizado junto a dois professores especialistas em metodologia da pesquisa com o objetivo de aperfeiçoar o questionário. Segundo Lyra, os professores que participaram do pré-teste são:

- Dra. Amélia Silveira: doutora em comunicação pela Universidade de São Paulo, ex-professora titular da Universidade Federal de Santa Catarina (aposentada) e professora do mestrado de Administração da Universidade Regional de Blumenau (FURB), autora de livros de metodologia da pesquisa.

- Dra. Ilse Maria Beuren: doutora em Controladoria e Contabilidade pela Universidade de São Paulo, ex-professora titular da Universidade Federal de Santa Catarina, professora e coordenadora do mestrado de Ciências Contábeis da FURB, autora de livros de metodologia da pesquisa.

- (c) o questionário ajustado com base na opinião dos dois especialistas mencionados foi enviado por e-mail acompanhado de uma carta de apresentação a um conjunto de especialistas convidados.

A partir disso, Lyra (2008, p. 86) convidou 45 professores especialistas, sendo 15 norte-americanos e 30 brasileiros. Entretanto, apenas 19 especialistas participaram efetivamente da pesquisa com a técnica *Delphi*. A seguir, conforme Lyra, encontra-se a relação dos profissionais participantes, em ordem alfabética:

- Alexandre Assaf Neto: doutor em Controladoria e Contabilidade, autor de livros e artigos diversos sobre a matéria e professor titular da Universidade de São Paulo (USP).
- Ariovaldo dos Santos: doutor em Controladoria e Contabilidade, coordenador técnico do anuário “Melhores e Maiores” da revista Exame e professor titular da Universidade de São Paulo (USP).
- Eliseu Martins: doutor em Controladoria e Contabilidade, autor de artigos diversos sobre a matéria e professor titular da Universidade de São Paulo (USP).
- Eric Press: *associate professor of Temple University – Department of Accounting*.
- Frank Heflin: *associate professor of College of Business Florida State University*.
- Ilse Maria Beuren: doutora em Controladoria e Contabilidade, é professora titular da Universidade Regional de Blumenau (FURB).
- José Carlos Marion: doutor em Controladoria e Contabilidade pela Universidade de São Paulo, pós-doutor pela *University of Kansas*, autor de livros e artigos diversos sobre a matéria e professor titular da Pontifícia Universidade Católica de São Paulo (PUC-SP).
- Luiz Alberton: doutor em Engenharia de Produção e professor adjunto II da Universidade Federal de Santa Catarina (UFSC).
- Maria Ivanice Vendruscolo: mestre em Contabilidade e professora da Pontifícia Universidade Católica do Rio Grande do Sul (PUC-RS).
- Maria Thereza Pompa Antunes: doutora em Controladoria e Contabilidade pela Universidade de São Paulo e professora assistente II da Universidade Presbiteriana Mackenzie.
- Moises Prates Silveira: mestre em Administração e professor titular da Universidade Federal do Paraná (UFPR).

- Octávio Ribeiro de Mendonça: doutor em Controladoria e Contabilidade pela Universidade de São Paulo e professor assistente da Universidade Presbiteriana Mackenzie.
- Odair Gonçalves: mestre em Ciências Contábeis e professor da Universidade do Vale do Rio dos Sinos (UNISINOS).
- Otávio Ribeiro de Medeiros: doutor em Economia e professor titular da Universidade de Brasília (UNB).
- Sérgio de Iudicibus: doutor em Controladoria e Contabilidade pela Universidade de São Paulo, autor de livros e artigos diversos sobre a matéria e professor titular da Pontifícia Universidade Católica de São Paulo (PUC-SP).
- Stephen R. Moehrle: *assistant professor of University of Missouri at St. Louis - Accounting Area.*
- Vera Marleide Loureiro dos Anjos: professora da Universidade de Brasília (UNB).
- Wagner Moura Lamounier: doutor em Economia e professor titular da Universidade Federal de Minas Gerais (UFMG).
- Yumara Lúcia Vasconcelos: doutora em Administração autora de livros e artigos sobre a matéria e professora da Fundação Escola Álvares Penteado (FECAP).

(d) De acordo com Lyra (2008), a técnica Delphi se processou em três etapas ou “rodadas”. Cada etapa consistiu basicamente no envio de questionários aos especialistas, na obtenção das respectivas respostas, no tratamento quantitativo dessas respostas e na emissão de novo questionário com as respostas consolidadas e tratadas, obtidas da etapa anterior. O objetivo dessas três etapas da técnica Delphi era identificar indicadores contábeis e estabelecer uma estrutura hierárquica entre eles, considerando a opinião e o consenso dos especialistas que participaram do processo. O Professor Lyra (2008) elaborou uma lista inicial de indicadores tomando por base o anuário Melhores e Maiores, publicado pela Revista Exame da Editora Abril. A seguir incluiu outros indicadores considerados importantes na literatura existente sobre o assunto. As variáveis inicialmente listadas foram filtradas e selecionadas com base nas respostas dos especialistas e no tratamento estatístico dos dados. Ao final de três etapas, foram selecionados e hierarquizados os seguintes indicadores (LYRA, 2008, p. 88-93):

- Rentabilidade sobre o Patrimônio Líquido;
- Rentabilidade sobre o Ativo;
- Crescimento das Vendas;
- Liquidez Corrente;
- Composição do Endividamento;

- Margem Líquida;
- Giro do Ativo

2.4.2 Indicadores contábeis utilizados neste estudo

Uma das questões de pesquisa em artigos e teses relacionados com avaliação de eficiência em empresas é a seleção de indicadores significativos. A definição desses indicadores foi uma das preocupações iniciais deste trabalho. Entretanto, após consulta da tese do professor Lyra, esse problema foi resolvido, pois se decidiu adotar os indicadores contábeis definidos na referida tese. Tal escolha ocorreu devido à metodologia utilizada pelo pesquisador, à eficiência da aplicação da técnica Delphi e ao notório saber dos especialistas que participaram dessa pesquisa.

Na tese do professor Lyra (2008), ficou comprovado que os indicadores selecionados mostram-se eficientes na avaliação de desempenho de empresas. Nessa tese foi utilizada a técnica *Analytic Hierarchy Process* (AHP) para hierarquizar o desempenho econômico-financeiro de empresas através da utilização dos referidos indicadores.

Dessa forma, os indicadores contábeis utilizados na presente tese são aqueles definidos na tese do professor Lyra, exceto a Rentabilidade sobre o Ativo. Essa exclusão foi feita porque a Rentabilidade sobre o Ativo resulta da divisão do lucro operacional antes das despesas financeiras pelo ativo operacional. O presente estudo se restringiu ao uso exclusivo de indicadores obtidos da Edição de Melhores & Maiores (M&M) da Revista Exame. Nessa Edição esse indicador não pôde ser obtido da forma como estabelecida acima, ou seja, tendo como numerador o lucro operacional e como denominador da fórmula o ativo operacional. Os dados utilizados pela M&M provêm dos balanços das empresas e não segregam os lucros operacionais e os ativos operacionais. Um analista interno da empresa conseguiria obter essa informação, mas o externo não. Assim, nesta tese, foi adotada a posição de um analista externo e, em consequência disso, tal indicador foi excluído da análise.

Desse modo, foram obtidos do anuário Melhores & Maiores, edição de 2008, os seguintes indicadores utilizados neste trabalho:

- (a) Rentabilidade sobre o Patrimônio Líquido – resulta da divisão do lucro líquido ajustado pelo patrimônio líquido ajustado. Os valores são ajustados para reconhecer os efeitos inflacionários que as empresas deixaram, por imposição legal, de considerar nas demonstrações contábeis;

- (b) Crescimento das Vendas – mostra a evolução da receita bruta de vendas em reais, descontada a inflação média apontada pela variação do IGP-M. As empresas que não publicaram demonstrações contábeis com correção monetária integral tiveram suas vendas atualizadas por Melhores & Maiores. Os valores foram convertidos para moeda de poder aquisitivo de 31/12/2007;
- (c) Liquidez Corrente – é o ativo circulante dividido pelo passivo circulante;
- (d) Composição do Endividamento – é a divisão do passivo circulante pela soma do passivo circulante com o exigível de longo prazo;
- (e) Margem Líquida – é o lucro líquido ajustado dividido pela receita bruta de vendas;
- (f) Giro do Ativo – é a receita bruta de vendas dividida pelo ativo total ajustado.

2.4.3 O banco de dados do anuário Melhores & Maiores

Os dados foram fornecidos pelo professor doutor Ariovaldo dos Santos, da Fundação Instituto de Pesquisas Contábeis Atuariais e Financeiras – FIPECAFI, e coordenador técnico do anuário Melhores & Maiores da Revista Exame.

O anuário contém dados relativos às 500 maiores empresas e uma lista complementar com outras 500 empresas, totalizando 1000. O critério de classificação utilizado é o da receita de vendas (faturamento bruto). Os valores referem-se ao exercício social findo em 2007. E todos os valores estão ajustados – considerando a variação da inflação – para o dia 31 de dezembro de 2007.

Além dos indicadores citados no item anterior, existe nesse anuário um conjunto de outros indicadores a serem utilizados numa fase posterior desta pesquisa, quando será explicado o motivo de sua utilização.

Cada empresa listada no banco de dados apresenta as seguintes informações:

- Cidade (sede);
- Estado (sede);
- Setor (energia, atacado, varejo, etc.);
- Grupo ao qual pertence;
- Ramo de atividade (indústria, comércio ou serviços);
- Controle acionário (estatal, país de origem);
- Ações na bolsa;

- Empresa responsável pela auditoria;
- Nome do principal executivo;
- Cargo do principal executivo.

As variáveis obtidas a partir dos dados fornecidos são as seguintes:

- Valor das vendas;
- Crescimento das vendas;
- Lucro líquido ajustado;
- Lucro líquido legal;
- Patrimônio líquido ajustado;
- Patrimônio líquido legal;
- Rentabilidade ajustada;
- Rentabilidade legal;
- Valor do Capital Circulante Líquido – CCL;
- Liquidez geral;
- Endividamento geral;
- Endividamento de longo prazo;
- Valor da riqueza criada;
- Número de empregados;
- Riqueza criada por empregado;
- EBITDA;
- Salários;
- Impostos;
- Margem das vendas;
- Giro;
- Liquidez corrente;
- Investimentos no imobilizado;
- Vendas (exportação);
- Percentual das vendas (exportação sobre vendas).

A seguir, será então aplicado todo esse arcabouço teórico na metodologia e análise dos resultados.

3 Metodologia e Análise dos Resultados

O objetivo deste capítulo é apresentar a abordagem metodológica utilizada no desenvolvimento da pesquisa na medida em que sejam expostos e analisados seus resultados. Optou-se por tal processo de explanação simultânea da metodologia e da análise dos resultados pelo fato de o procedimento metodológico abranger muitas fases, sendo que sua demonstração torna-se mais didática e prática no decorrer do cumprimento das diversas etapas.

3.1 Características da pesquisa

O presente estudo trata-se eminentemente de uma pesquisa empírica em que são utilizados instrumentos não paramétricos e também procedimentos de estatística paramétrica, caracterizando-se, quanto ao aspecto epistemológico, como empírico-analítico. Pode ser considerado também como uma pesquisa exploratória e apresenta uma abordagem dedutiva.

O instrumento não paramétrico empregado é o *Data envelopment analysis* (DEA), amplamente utilizado no mundo acadêmico e empresarial para a avaliação de eficiência de empresas públicas e privadas e de subdivisões de diversos tipos de organização.

Para usar o referido instrumento, é necessário que essas organizações ou subunidades possuam características homogêneas e, para separar as empresas em grupos homogêneos, utiliza-se a *Cluster Analysis*.

A seguir expõe-se a avaliação da eficiência das empresas dentro de cada grupo homogêneo. Então, emprega-se o instrumental estatístico relativo a teste de hipótese e análise de regressão para avaliar se os indicadores contábeis utilizados têm influência significativa na eficiência das empresas, considerando cada grupo de empresas.

Para consecução da pesquisa são utilizados dados secundários. A avaliação da eficiência é realizada “ex post facto”.

3.2 Revisão bibliográfica

O primeiro passo na execução desse projeto foi a realização de um amplo levantamento bibliográfico, o qual consistiu principalmente em obter ampla documentação literária sobre avaliação de eficiência de empresas. O instrumento escolhido para a avaliação de eficiência foi a Análise Por Envoltória de Dados – DEA, um dos instrumentos mais utilizados no mundo inteiro para a avaliação de eficiência de entidades públicas e privadas. Constatou-se que o volume de artigos publicados sobre essa técnica e a grande diversidade de aplicações acarretou uma quantidade de referências muito acima do esperado. O estudo efetuado gerou como produto inicial uma relação dos artigos publicados em ordem alfabética de autores, co-autores, com os respectivos títulos, local e data de publicação. Esta pesquisa inicial pode gerar um artigo proveitoso para os futuros pesquisadores, pois contempla o referencial teórico bem como os diversos tipos de aplicações, nos diversos ramos de atividade e nos diferentes países.

Um dos objetivos desta tese é provar a possibilidade de avaliar a eficiência das empresas através da utilização de um pequeno número de indicadores contábeis. A questão inicial é: que indicadores contábeis significativos devem ser utilizados para a avaliação de eficiência das empresas?

Os autores de livros sobre análise de balanços propõem uma lista infindável de indicadores obtidos do balanço, da demonstração de resultados, da demonstração das origens e aplicações de recursos etc.

O professor Ricardo Luiz Wüst Corrêa de Lyra (2008) responde essa questão de seleção de indicadores em sua tese de doutorado apresentada ao Departamento de Contabilidade e Atuária da Universidade de São Paulo. Ele utiliza a técnica Delphi e a participação de especialistas. O procedimento empregado pelo Professor Lyra (2008, p. 85- 94) foi resumido e se encontra no Capítulo 2 deste estudo.

3.3 Indicadores contábeis utilizados

Uma das questões de pesquisa em artigos e teses relacionados com avaliação de eficiência em empresas é a seleção de indicadores significativos. A definição desses indicadores impôs-se como uma das preocupações iniciais deste trabalho. Entretanto, após a conclusão da tese do professor Lyra (2008) esse problema se resolveu, pois se decidiu adotar os indicadores

contábeis definidos nesse estudo. Tal escolha ocorreu devido à metodologia utilizada pelo referido pesquisador, à eficiência da aplicação da técnica Delphi e ao notório saber dos especialistas que participaram dessa pesquisa.

Na tese do professor Lyra (2008) está comprovado que os indicadores selecionados mostram-se eficientes na avaliação de desempenho de empresas. Nessa tese foi empregada a técnica *Analytic Hierarchy Process* (AHP) para hierarquizar o desempenho econômico-financeiro de empresas através da utilização dos referidos indicadores.

Portanto, os indicadores contábeis utilizados no presente trabalho são aqueles definidos na tese do professor Lyra, exceto a Rentabilidade sobre o Ativo. Essa exclusão acontece porque a Rentabilidade sobre o Ativo resulta da divisão do lucro operacional antes das despesas financeiras pelo ativo operacional. O estudo aqui efetuado se restringiu ao uso exclusivo de indicadores obtidos da Edição de Melhores & Maiores (M&M) da Revista Exame. De acordo com esta publicação, esse indicador não pode ser obtido da forma como estabelecida acima, ou seja, tendo como numerador o lucro operacional e como denominador da fórmula o ativo operacional. Os dados empregados pela M&M provêm dos balanços das empresas e estas não segregam os lucros operacionais e os ativos operacionais. Um analista interno da empresa consegue obter essa informação, mas o externo não. Assim, esta tese considera a posição de um analista externo e, conseqüentemente, esse indicador foi eliminado da análise.

Assim, os indicadores utilizados nesta tese são obtidos do anuário Melhores & Maiores, edição de 2008 e estão discriminados abaixo. Cabe observar que, para cada variável, atribui-se uma sigla que aparece em letras maiúsculas logo após o seu respectivo nome:

(a) Rentabilidade sobre o Patrimônio Líquido (RENTAJU7) resulta da divisão do lucro líquido ajustado pelo patrimônio líquido ajustado. Os valores são ajustados para reconhecer os efeitos inflacionários que as empresas deixaram, por imposição legal, de considerar nas demonstrações contábeis;

(b) Crescimento das Vendas (CRESVEN7) mostra a evolução da receita bruta de vendas em reais, descontada a inflação média apontada pela variação do IGP-M. As empresas que não publicaram demonstrações contábeis com correção monetária integral tiveram suas vendas

atualizadas por Melhores & Maiores. Os valores foram convertidos para moeda de poder aquisitivo de 31/12/2007;

(c) Liquidez Corrente (LIQCOR7) é o ativo circulante dividido pelo passivo circulante;

(d) Composição do Endividamento (COMPEND7) é a divisão do passivo circulante pela soma do passivo circulante com o exigível de longo prazo;

(e) Margem Líquida (MARVEN7) é o lucro líquido ajustado dividido pela receita bruta de vendas;

(f) Giro do Ativo (GIRO7) é a receita bruta de vendas dividida pelo ativo total ajustado.

3.4 O banco de dados do anuário Melhores & Maiores

Os dados expostos foram fornecidos pelo professor Ariovaldo dos Santos da Fundação Instituto de Pesquisas Contábeis Atuariais e Financeiras – FIPECAFI – e coordenador técnico do anuário Melhores & Maiores da Revista Exame.

O anuário contém dados relativos às 500 maiores empresas e uma lista complementar com outras 500 empresas, totalizando 1000. O critério de classificação utilizado é o da receita de vendas (faturamento bruto). Os valores referem-se ao exercício social findo em 2007 e todos os valores estão ajustados – considerando a variação da inflação – para o dia 31 de dezembro de 2007.

Além dos indicadores citados no item anterior, há, nesse anuário, um conjunto de outros indicadores a serem utilizados numa fase posterior desta pesquisa, momento em que se explicará o motivo de sua utilização.

Cada empresa listada no banco de dados apresenta as seguintes informações:

- Cidade (sede);
- Estado (sede);
- Setor (energia, atacado, varejo, etc.);

- Grupo ao qual pertence;
- Ramo de atividade (indústria, comércio ou serviços);
- Controle acionário (estatal, país de origem);
- Ações na bolsa;
- Empresa responsável pela auditoria;
- Nome do principal executivo;
- Cargo do principal executivo.

As variáveis obtidas a partir dos dados fornecidos são as seguintes:

- Valor das vendas;
- Crescimento das vendas;
- Lucro líquido ajustado;
- Lucro líquido legal;
- Patrimônio líquido ajustado;
- Patrimônio líquido legal;
- Rentabilidade ajustada;
- Rentabilidade legal;
- Valor do Capital Circulante Líquido – CCL;
- Liquidez geral;
- Endividamento geral;
- Endividamento de longo prazo;
- Valor da riqueza criada;
- Número de empregados;
- Riqueza criada por empregado;
- EBITDA;
- Salários;
- Impostos;
- Margem das vendas;
- Giro;
- Liquidez corrente;

- Investimentos no imobilizado;
- Vendas (exportação);
- Percentual das vendas (exportação sobre vendas).

3.5 O tratamento dos *outliers*

Quanto aos desvios padrões, decidiu-se por trabalhar com uma massa de dados cujos valores estivessem entre mais ou menos três desvios padrões da média. Assim, de acordo com o critério adotado, todos os dados que estivessem fora desse intervalo seriam considerados como *outliers* e conseqüentemente eliminados da amostra.

O processo de eliminação dos *outliers* se iniciou com a padronização de todos os dados relativos às seguintes variáveis:

- Rentabilidade sobre o Patrimônio Líquido;
- Crescimento das Vendas;
- Liquidez Corrente;
- Composição do Endividamento;
- Margem Líquida;
- Giro do Ativo.

A seguir procedeu-se à eliminação de todos os valores que estavam fora do intervalo considerado de mais ou menos três desvios padrões. Após essa primeira eliminação, os valores das médias das variáveis se tornaram menores e, portanto, foram padronizados uma segunda vez e assim surgiram novos casos cujos valores se situavam fora dos três desvios padrões considerados como limítrofes. Esses casos foram também eliminados.

O mesmo processo foi repetido até que não fosse encontrado mais nenhum caso fora dos três desvios padrões. Para que o processo de eliminação dos *outliers* fosse finalizado, foi necessário repetir o mesmo procedimento por quatro vezes. Ao final do processo restavam 758 empresas das 1.000 integrantes do banco de dados inicial. Essas 758 empresas apresentavam valores que se situavam entre mais ou menos três desvios padrões em relação à média para cada uma das seis variáveis consideradas na pesquisa.

3.6 Utilização da *Cluster Analysis* para separar empresas em grupos homogêneos

Para a técnica DEA, as empresas, objetos da pesquisa, são denominadas DMUs, *Decision Making Units*. Essas DMUS podem ser também grupos de empresas ou subunidades da empresa como departamentos, por exemplo.

Há certos pressupostos para a comparação de eficiência entre empresas:

- Haver condições de comparabilidade entre as DMUs;
- As condições em que atuam devem ser semelhantes;
- Para cada DMU devem ser considerados os mesmos *inputs* e *outputs*.

As 1.000 empresas componentes da amostra são classificadas por setor e por ramo de atividade. Entretanto, através de análise efetuada, verifica-se que empresas componentes de um mesmo setor apresentam características bem diferentes de outras participantes do mesmo setor. Um dos grandes problemas desta pesquisa é descobrir como obter condições semelhantes de comparabilidade e de atuação dessas empresas. Devido a esse fato, a maioria dos pesquisadores que utilizam o DEA optam pela escolha de um setor específico como: bancos, empresas de energia elétrica, hospitais, empresas petroquímicas, empresas de saneamento, escolas etc.

Um trabalho inicial é, portanto, separar as empresas em grupos homogêneos. Dessa forma, foram consideradas as seis variáveis mencionadas anteriormente obtidas da pesquisa Delphi e, a partir dos dados de empresas que compõem a Edição de Maiores e Melhores foi possível investigar a possível formação de agrupamentos naturais.

Optou-se pela aplicação da técnica *Cluster Analysis*, ou, em português, Análise de Agrupamento ou também Análise de Conglomerados, cujo objetivo é agrupar objetos segundo características que os mesmos possuam, levando em conta diversas variáveis escolhidas pelo pesquisador.

Tomando-se por base a amostra obtida da Edição de 2007 da Revista Exame, Edição Melhores e Maiores, a pesquisa testa a seguinte hipótese:

Hipótese nula (H_0): não há padrões de comportamento similares significativos que evidenciem a existência de agrupamentos naturais entre as empresas listadas pela Edição de Melhores e Maiores;

Hipótese alternativa: há evidências que permitem rejeitar a hipótese nula.

Assim, no presente caso, o objetivo passa a ser o de identificar agrupamentos naturais eventualmente existentes entre as empresas componentes da edição de Melhores e Maiores da Revista Exame. Conforme os resultados encontrados, a hipótese nula será aceita ou rejeitada. A partir disso, são utilizados os seis índices já referidos. Tais indicadores são considerados suficientes e aptos a propiciar uma análise de conglomerados.

Cabe ressaltar, por fim, que para a viabilização dos cálculos da análise de agrupamento, empregou-se o programa SPSS.

Seguindo os passos indicados pela literatura, há a necessidade de se certificar, primeiramente, quanto à representatividade da amostra e verificar se a amostra selecionada representa a população para a qual se pretende generalizar os resultados da pesquisa.

Nesse particular, é importante destacar que o método de amostragem selecionado é da espécie não-casual, denominado amostragem por julgamento ou conveniência. Segundo Levin (1985, p. 120), a idéia básica envolvida é a de que a lógica, o senso comum ou um julgamento equilibrado podem ser usados na seleção de uma amostra representativa de um grupo maior (população).

O universo em estudo abrange todas as empresas que atuam no Brasil, organizadas em setores segundo a natureza das respectivas atividades. Seria, então, preciso dispor das informações de cada variável relativas a todas as empresas, divididas em setores.

Tendo em vista a indisponibilidade desses dados, e qualquer esforço no sentido da sua obtenção seria impraticável dentro dos limites temporais deste estudo, optou-se por recorrer aos dados disponibilizados pela Revista Exame.

Dessa maneira, voltando ao problema básico do presente trabalho, consubstanciado nas hipóteses da pesquisa enunciada anteriormente, conclui-se pela rejeição da hipótese nula (H_0)

e, consequentemente, pela aceitação da hipótese alternativa (H1), uma vez que está demonstrada a existência de agrupamentos naturais distintos.

A *Cluster Analysis*, técnica multivariada que tem por finalidade permitir a classificação de objetos segundo suas similaridades, mostrou-se bastante útil e apropriada para a solução do problema proposto, consubstanciado nas hipóteses definidas.

3.7 *Cluster Analysis* e as opções do programa SPSS

Para a operacionalização da *cluster analysis* utilizou-se o programa SPSS. Entre as várias opções oferecidas por esse programa foi escolhida a função *TwoStep Cluster Analysis*, uma ferramenta exploratória usada para detectar agrupamentos naturais dentro de um conjunto de dados.

O algoritmo empregado nessa função possibilita a seleção automática do número de agrupamentos, sendo vantajoso em relação aos outros procedimentos porque uma das grandes dificuldades da *cluster analysis* é justamente determinar esse número de forma objetiva. Além disso, é possível definir o número máximo de *clusters* a serem selecionados, considerando-se um limite de 20 *clusters*. O anuário M&M seleciona 24 setores e o algoritmo determina o melhor número de agrupamentos. O critério utilizado para formar o número de agrupamentos é o *Bayesian Information Criterion* (BIC).

Foi escolhida também a opção *likelihood* que é uma das formas de medir a similaridade entre os agrupamentos e assume uma distribuição de probabilidades sobre as variáveis: variáveis contínuas apresentam distribuição normal; e categóricas têm distribuição multinomial. Todas as variáveis são assumidas como sendo independentes. Testes empíricos realizados pelos criadores do procedimento indicam que este é robusto em relação às violações de ambos os pressupostos.

3.8 Análise dos resultados da *Cluster Analysis*

Inicialmente obteve-se uma estatística descritiva das seis variáveis obtidas através da técnica Delphi, com a utilização dos dados da M&M referentes ao exercício de 2007.

A Tabela 1 (*Auto-Clustering*) sintetiza o processo pelo qual é escolhido o número ideal de *clusters*.

Tabela 1: Critério de Seleção do Número de *Clusters* (*Auto-Clustering*)

<i>Number of Clusters</i>	<i>Schwarz's Bayesian Criterion (BIC)</i>	<i>BIC Change^a</i>	<i>Ratio of BIC Changes^b</i>	<i>Ratio of Distance Measures^c</i>
1	2464,609			
2	2256,646	-207,962	1,000	1,258
3	2107,005	-149,642	,720	1,402
4	2022,161	-84,843	,408	1,412
5	1984,330	-37,831	,182	1,102
6	1957,083	-27,248	,131	1,022
7	1932,080	-25,002	,120	1,679
8	1948,026	15,946	-,077	1,201
9	1974,069	26,043	-,125	1,030
10	2001,559	27,490	-,132	1,160
11	2035,791	34,232	-,165	1,067
12	2072,645	36,854	-,177	1,044
13	2111,177	38,532	-,185	1,035
14	2150,988	39,811	-,191	1,241
15	2197,872	46,884	-,225	1,035
16	2245,762	47,890	-,230	1,082
17	2295,807	50,045	-,241	1,050
18	2347,099	51,292	-,247	1,048
19	2399,525	52,426	-,252	1,024
20	2452,508	52,983	-,255	1,123

a. The changes are from the previous number of clusters in the table.

b. The ratios of changes are relative to the change for the two cluster solution.

c. The ratios of distance measures are based on the current number of clusters against the previous number of clusters.

No caso exposto na tabela, foi selecionado o *Schwarz's Bayesian Criterion* (BIC). O critério de agrupamento (BIC) é computado para cada potencial número de *clusters*. Os menores

valores do BIC indicam os melhores modelos e, nessa situação, a melhor solução apresenta o menor valor do BIC (1932,080). Assim, o melhor modelo indica sete *clusters*.

A Tabela 2 a seguir mostra as frequências absoluta e relativa de cada cluster. Do total de 1.000 empresas, 425 foram excluídas da análise por algumas razões, como a eliminação dos *outliers* (fator já mencionado anteriormente) e *missing values* (valores faltantes) em uma ou mais variáveis.

Dos 575 casos válidos, 148 casos foram computados para o primeiro (1) *cluster*, 72 para o segundo (2) *cluster*, 74 para o terceiro (3), 69 para o quarto (4), 55 para o quinto (5), 70 para o sexto (6) e 87 para o sétimo(7). Abaixo se visualiza a figura onde estão representados os tamanhos dos *clusters* (*Cluster Size*).

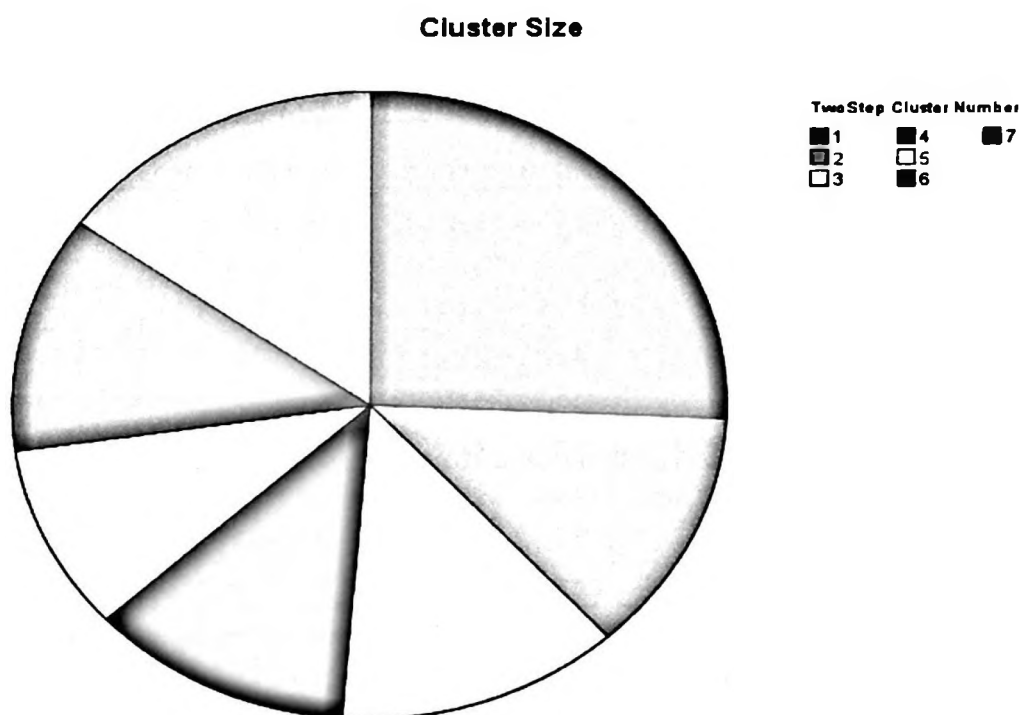


Figura 1: Representação dos tamanhos dos *clusters* (*Cluster Size*)

Tabela 2: Distribuição dos *Clusters* (Cluster Distribution)

		N	% of Combined	% of Total
Cluster	1	148	25,7%	14,8%
	2	72	12,5%	7,2%
	3	74	12,9%	7,4%
	4	69	12,0%	6,9%
	5	55	9,6%	5,5%
	6	70	12,2%	7,0%
	7	87	15,1%	8,7%
	Combined	575	100,0%	57,5%
	Excluded Cases	425		42,5%
	Total	1000		100,0%

A tabela 3, Perfil dos *Clusters* (*Centroids*), apresenta os valores dos centróides. Os centróides são os valores médios das empresas contidas em um determinado *cluster* e específicos para cada variável. Com base na análise comparativa dos valores contidos nessa tabela é possível traçar um possível perfil dos *clusters*. Avaliando-se os resultados, destacam-se os seguintes aspectos:

- (a) *Cluster 1*: As empresas desse *cluster* apresentam valores elevados para a composição do endividamento, ou seja, valores elevados de exigíveis a curto prazo sobre o total dos exigíveis ou também alto grau de endividamento a curto prazo (COMPEND7 = 77,88%) e alto nível do giro do ativo (GIRO7 = 2,58);
- (b) *Cluster 2*: Nesse caso, as empresas demonstram os piores desempenhos: queda no crescimento das vendas (CRESVEN7 = -0,59); rentabilidade sobre o patrimônio líquido negativa (RENTAJU7 = -2,59); margem líquida negativa (MARVEN7 = -1,09%);
- (c) *Cluster 3*: Nessa situação, as empresas evidenciam alta liquidez corrente (LIQCOR7 = 2,78) e valores moderados para todos os outros indicadores;
- (d) *Cluster 4*: As empresas desse item se destacam por apresentarem a maior rentabilidade sobre o patrimônio líquido (RENTAJU7 = 31,92%); além disso, apresentam índices

significativos para a margem líquida (MARVEN7 = 7,92%) e o giro do ativo (GIRO7 = 2,06);

(e) *Cluster 5*: Nesse *cluster*, as empresas expõem a maior margem líquida (MARVEN7 = 15,54%); em compensação, exibem o menor giro do ativo (GIRO7 = 0,67) e o menor índice de liquidez corrente (LIQCOR77 = 1,20); apresentam ainda baixo crescimento das vendas (CRESVEN7 = 1,07);

(f) *Cluster 6*: As empresas, nessa situação, denotam a maior taxa de crescimento das vendas (CRESVEN7 = 30,14%; nota-se que talvez esse forte crescimento se deva à baixa margem líquida (MARVEN7 = 3,12%);

(g) *Cluster 7*: As empresas desse *cluster* apresentam a menor composição do endividamento ou endividamento a curto prazo (COMPEND7 = 38,42%); evidenciam também redução do crescimento das vendas (CRESVEN7 = -0,15).

Tabela 3: Perfil dos *Clusters* (*Centroids*)

		Cluster							
		1	2	3	4	5	6	7	Combined
CRESVEN7	Mean	8,57	-0,59	4,80	9,57	1,07	30,14	-0,15	7,65
	Std. Deviation	8,29	13,46	11,56	10,45	12,04	10,37	10,13	14,04
RENTAJU7	Mean	11,30	-2,59	10,95	31,92	16,88	10,24	7,31	11,79
	Std. Deviation	7,82	7,77	7,10	7,98	7,62	8,97	5,04	11,69
COMPEND7	Mean	77,88	72,10	54,73	64,67	50,29	69,30	38,42	62,94
	Std. Deviation	14,09	18,12	20,64	17,58	24,37	16,67	11,73	21,92
MARVEN7	Mean	2,27	-1,09	6,63	7,92	15,54	3,12	4,52	4,80
	Std. Deviation	1,91	3,57	4,19	2,66	3,15	2,95	3,00	5,28
GIRO7	Mean	2,58	1,21	1,18	2,06	0,67	1,42	0,80	1,57
	Std. Deviation	0,74	0,53	0,54	0,85	0,33	0,57	0,38	0,93
LIQCOR7	Mean	1,54	1,35	2,78	1,35	1,20	1,56	1,23	1,57
	Std. Deviation	0,56	0,45	0,51	0,51	0,51	0,56	0,43	0,70

(1) Valores em negrito com letras em tamanho maior apontam o maior valor para aquela variável;

(2) Valores em negrito com letras menores indicam o menor valor para aquela variável.

Concluindo, pode-se constatar que cada *cluster* oferece sempre um indicador de destaque na sua formação, o que ocorre devido ao fato de o *cluster* apresentar o maior valor para aquele tipo de indicador comparativamente aos valores desse indicador para os demais *clusters*. Tal destaque também pode acontecer pelo motivo de o *cluster* evidenciar o menor valor para aquele indicador comparativamente aos valores desse indicador para os demais *clusters*. Isso pode ser visualizado através das figuras apresentadas a seguir.

Em cada figura aparece uma variável no eixo vertical e os respectivos valores dos centróides de cada *cluster*, no eixo horizontal. É possível verificar que para cada variável há um *cluster* apresentado com maior valor (ou menor valor) e, portanto, mais identificado com essa variável. Isso quer dizer que essa variável é a principal, embora não seja a única responsável pela formação do *cluster*, como será abordado numa fase mais adiante.

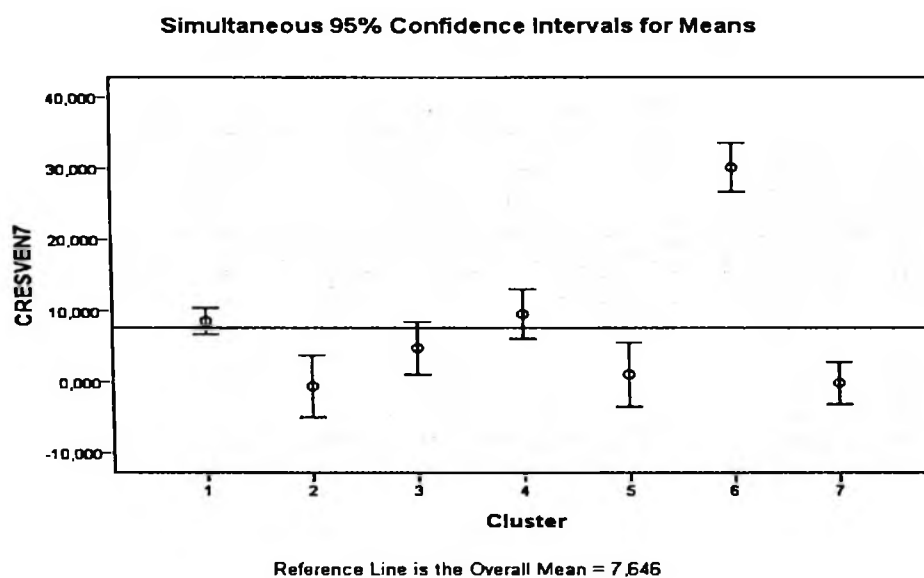


Figura 2: Variável Crescimento das Vendas para os 7 clusters

A linha de referência se refere à média da variável considerando todos os *clusters*. Verifica-se que o *cluster* 6 evidencia o maior valor para o crescimento das vendas, sendo tal crescimento uma das principais características desse *cluster*.

Por outro lado, nota-se que uma das principais características do *cluster* 2 é o baixo crescimento das vendas.

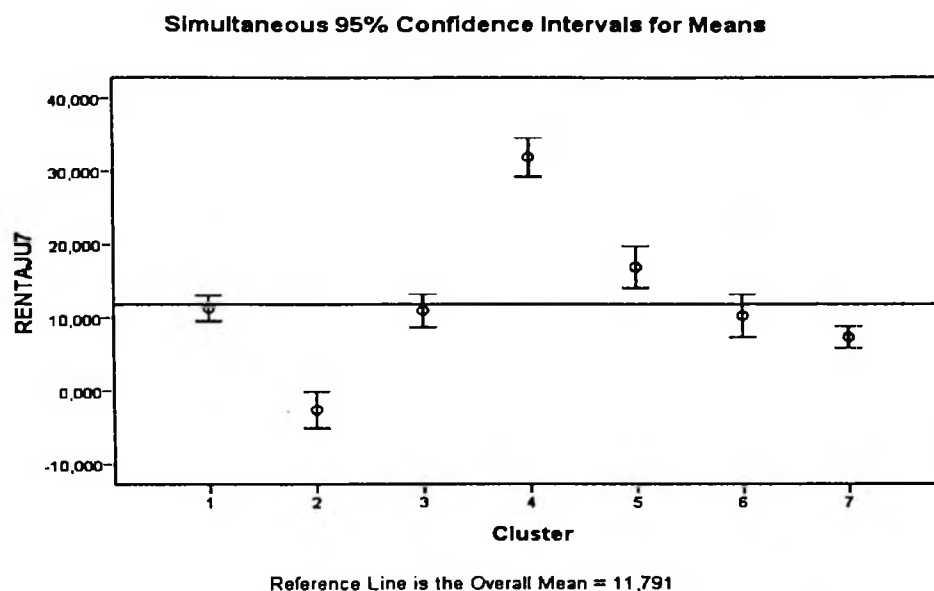


Figura 3: Variável Rentabilidade sobre o Patrimônio para os 7 clusters

De acordo com a figura 3, constata-se que o *cluster* 4 apresenta o maior valor para a rentabilidade sobre o patrimônio líquido. Portanto, uma das principais características desse *cluster* é a rentabilidade sobre o patrimônio líquido. Por outro lado, verifica-se o baixo desempenho do *cluster* 2 com relação à rentabilidade sobre o patrimônio líquido, sendo essa mais uma das características do *cluster* 2.

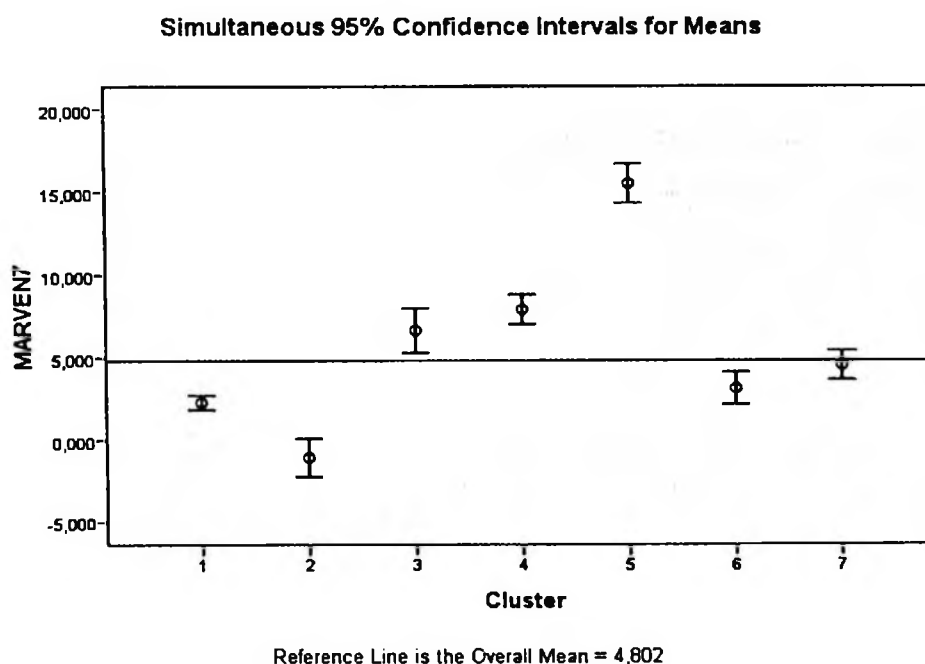


Figura 4: Variável Margem Líquida para os 7 clusters

Analisando a figura 4, constata-se que o *cluster* 5 demonstra o maior valor para a margem líquida sobre o patrimônio líquido. Dessa maneira, uma das principais características desse *cluster* é a margem líquida sobre o patrimônio líquido. Apesar disso, nota-se mais uma das particularidades do *cluster* 2: seu baixo desempenho com relação à margem líquida.

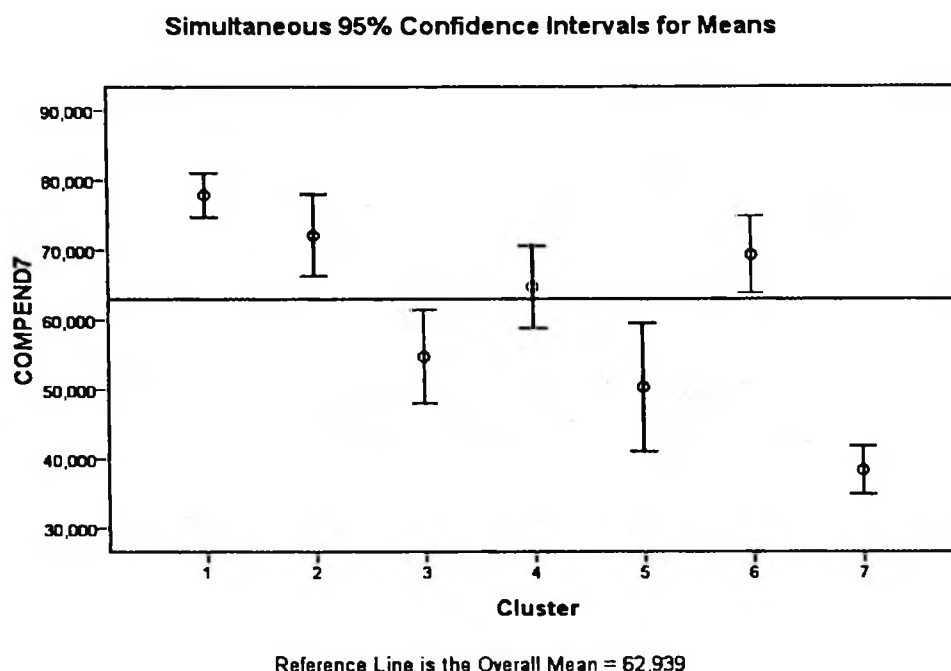


Figura 5: Variável Composição do Endividamento para os 7 clusters

Conforme a figura 5, o *cluster* 1 expõe o maior valor para a composição do endividamento, o que significa elevado endividamento a curto prazo. Assim, uma das principais características desse *cluster* é a composição do endividamento. Por outro lado, percebe-se que o *cluster* 7 demonstra o menor valor da composição do endividamento, sendo essa mais uma de suas características.

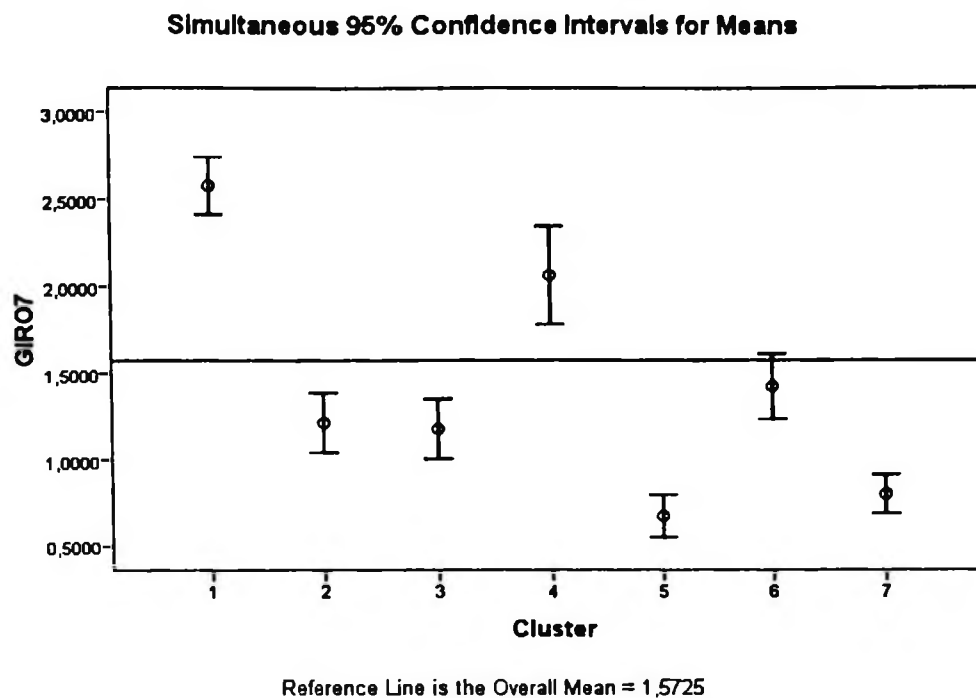


Figura 6: Variável Giro do Ativo para os 7 clusters

A partir da análise da figura 6, verifica-se que o *cluster* 1 apresenta o maior valor para o giro do ativo. Dessa forma, uma das principais características desse *cluster* é o giro do ativo. Já para o *cluster* 5, é evidente uma de suas particularidades, o baixo nível de giro do ativo.

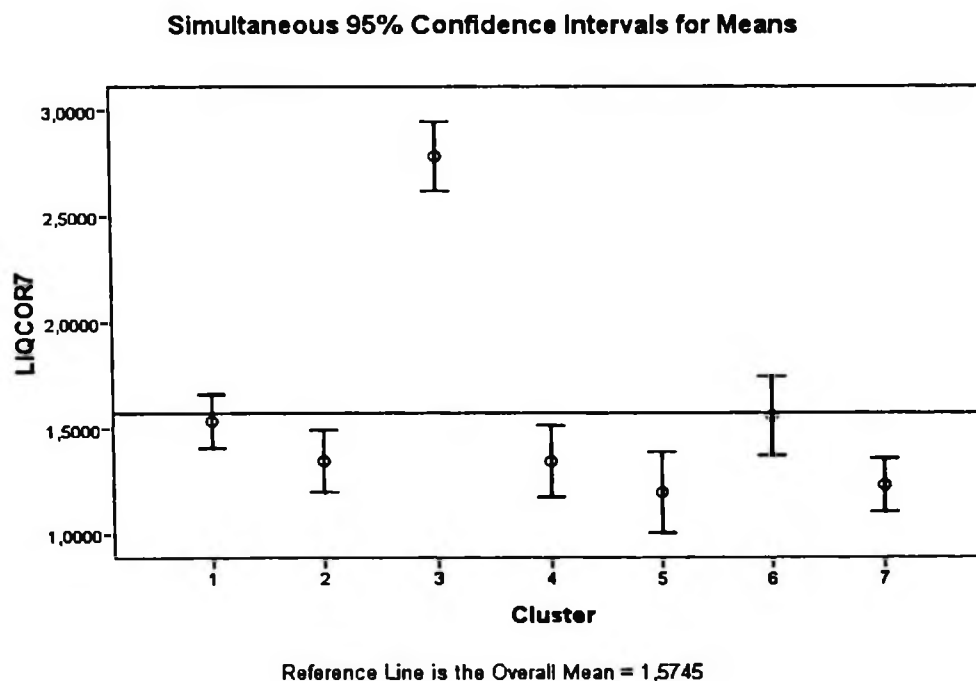


Figura 7: Variável Liquidez Corrente para os 7 clusters

Segundo a figura 7, o *cluster* 3 apresenta o maior valor para a liquidez corrente, sendo uma das principais características desse *cluster* a liquidez corrente.

Conforme já exposto anteriormente, diversas variáveis contribuem, em geral, para a formação de um determinado *cluster*. As figuras a seguir indicam a importância das variáveis na formação de cada *cluster*, havendo uma figura para cada *cluster*. No eixo vertical são listadas as variáveis em ordem decrescente de importância para a formação do *cluster*. A linha tracejada vertical indica os valores críticos para o nível de significância de 5% (ou nível de confiança de 95%). Para uma variável ser considerada significativa, seus valores, representados pelas barras horizontais, devem exceder a linha tracejada.

No caso do *cluster* 1, de acordo com a figura 8, as variáveis significativas para sua formação são: giro do ativo (GIRO7), margem líquida (MARVEN7) e composição do endividamento (COMPEND7)

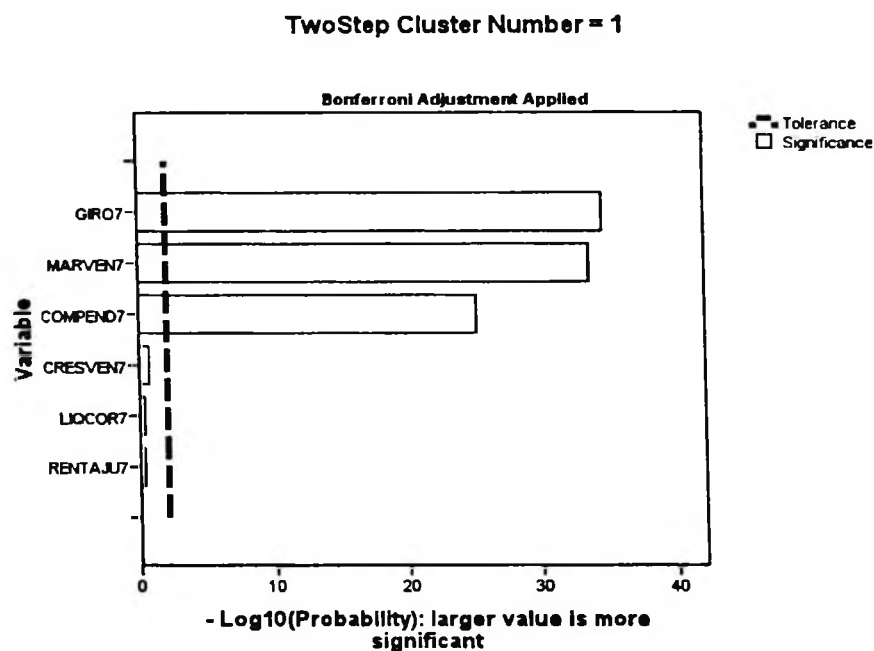


Figura 8: Variáveis de importância na formação do *cluster* 1

Considerando a figura 9, no caso do *cluster* 2, todas as 6 variáveis são significativas para sua formação. Em ordem decrescente de importância: rentabilidade do patrimônio líquido (RENTAJU7), margem líquida (MARVEN7), giro do ativo (GIRO7), crescimento das vendas (CRESVEN7), liquidez corrente (LIQCOR7) e composição do endividamento (COMPEND7).

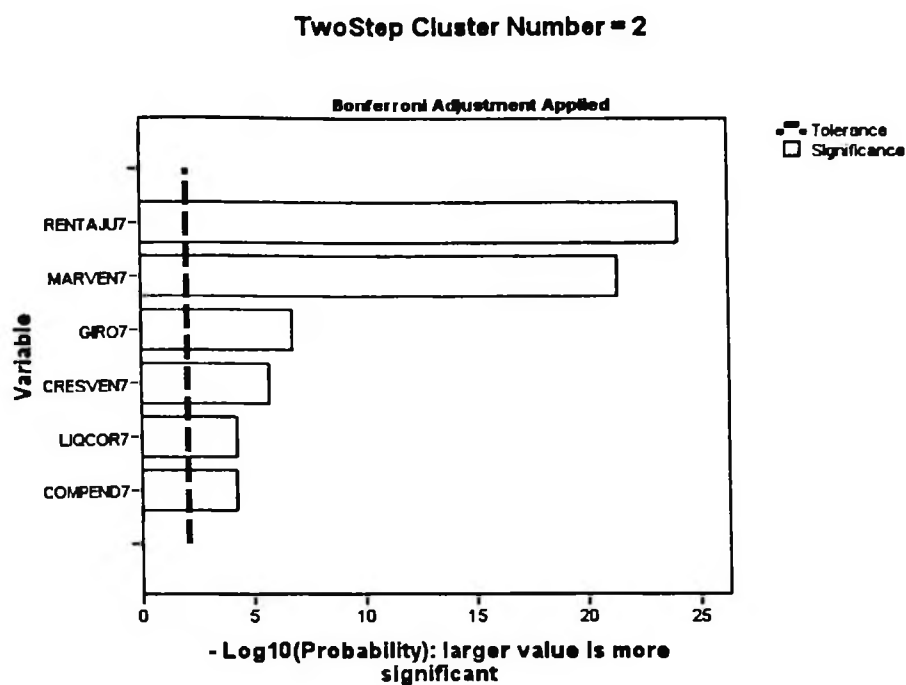


Figura 9: Variáveis de importância na formação do *cluster 2*

Segundo a figura 10, no *cluster 3*, as variáveis significativas, em ordem decrescente de importância, são: liquidez corrente (LIQCOR7), giro do ativo (GIRO7), margem líquida (MARVEN7), e composição do endividamento (COMPEND7).

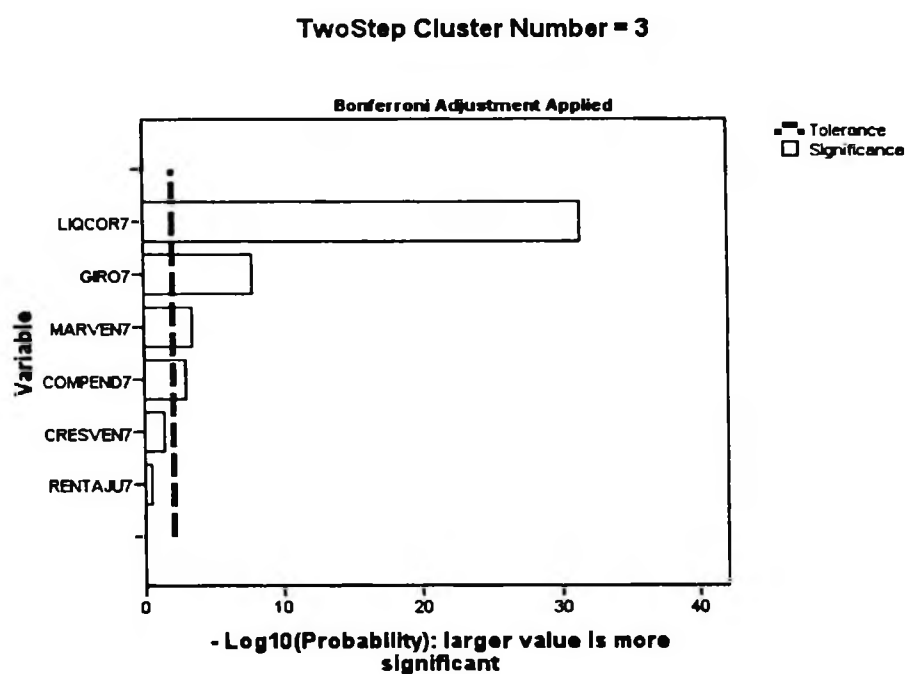


Figura 10: Variáveis de importância na formação do *cluster 3*

Avaliando o *cluster* 4, na figura 11 a seguir, as variáveis significativas são, em ordem decrescente de importância: rentabilidade do patrimônio líquido (RENTAJU7), margem líquida (MARVEN7), giro do ativo (GIRO7) e liquidez corrente (LIQCOR7).

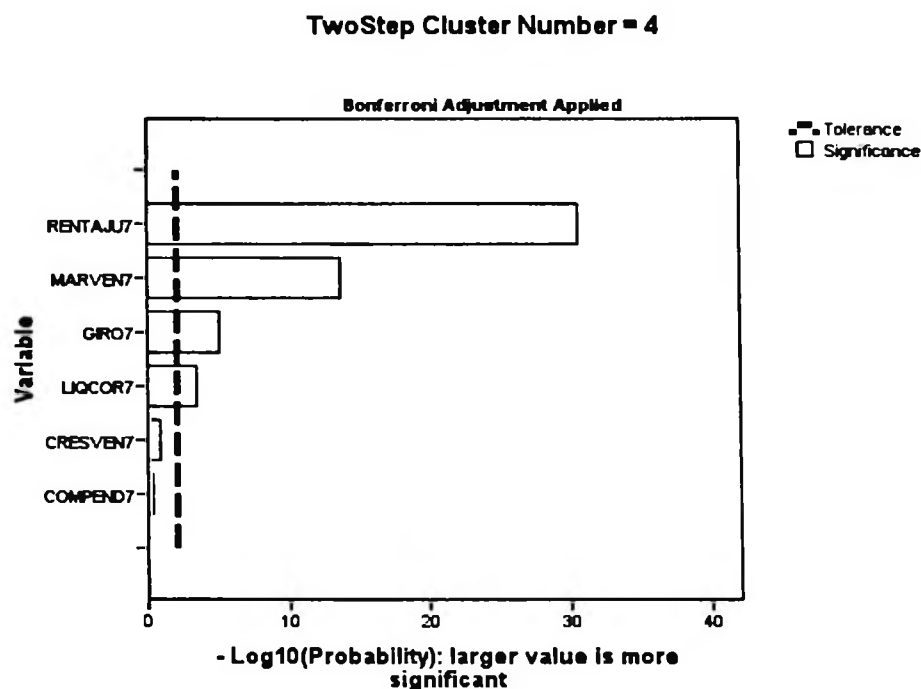


Figura 11: Variáveis de importância na formação do *cluster* 4

Já no caso do *cluster* 5, conforme a figura 12, as variáveis significativas são, em ordem decrescente de importância: margem líquida (MARVEN7), giro do ativo (GIRO7), liquidez corrente (LIQCOR7), rentabilidade do patrimônio líquido (RENTAJU7), crescimento das vendas (CRESVEN7) e composição do endividamento (COMPEND7).

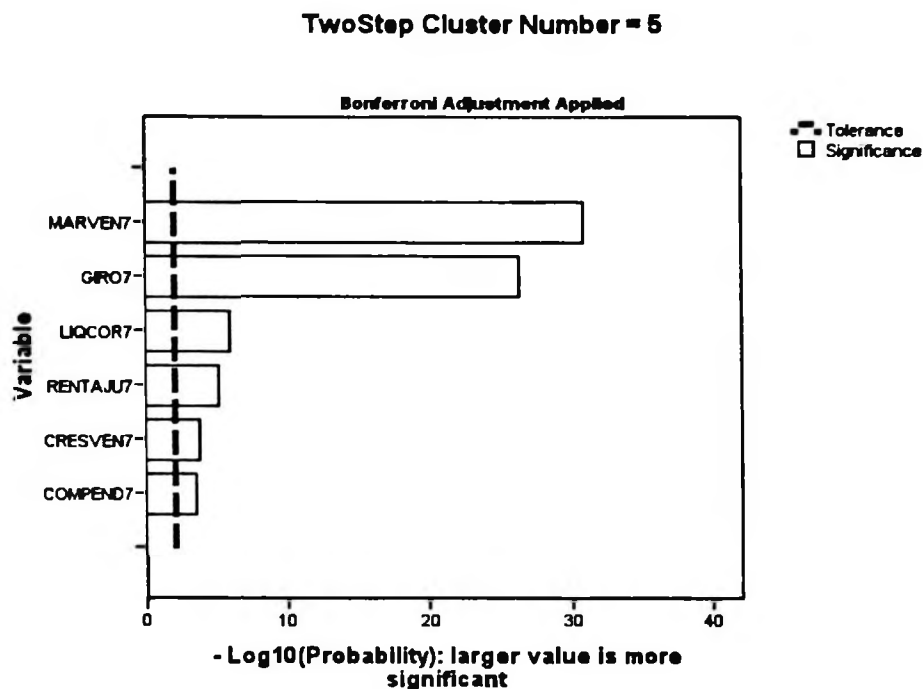


Figura 12: Variáveis de importância na formação do *cluster* 5

Ao avaliar a figura 13, quanto ao *cluster* 6, as variáveis significativas são, em ordem decrescente de importância: crescimento das vendas (CRESVEN7), margem líquida (MARVEN7) e composição do endividamento (COMPEND7).

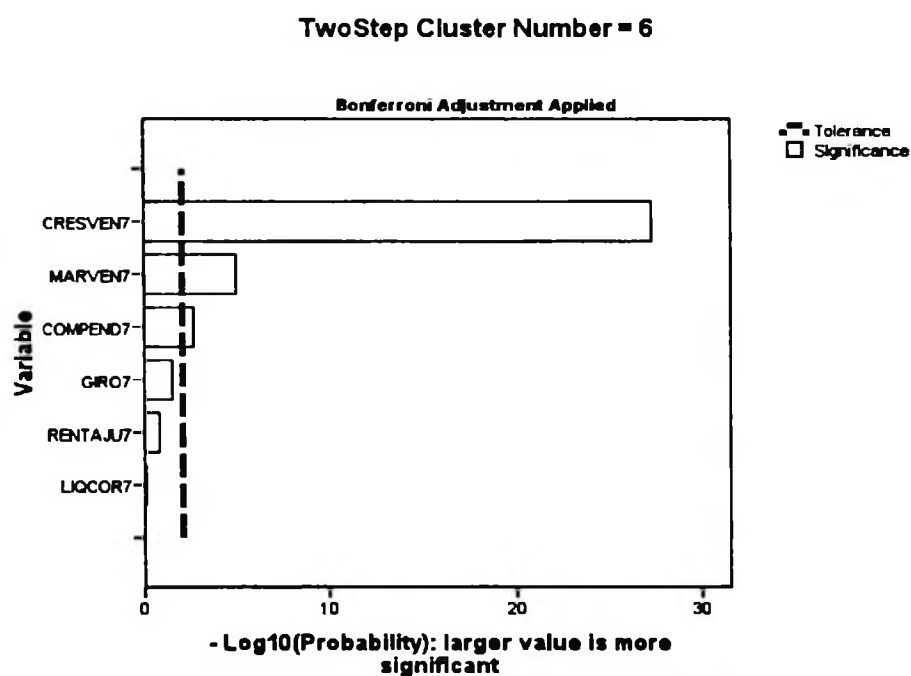


Figura 13: Variáveis de importância na formação do *cluster* 6

Por fim, em relação ao *cluster* 7, conforme a figura 14, as variáveis significativas são, em ordem decrescente de importância: composição do endividamento (COMPEND7), giro do ativo (GIRO7), rentabilidade do patrimônio líquido (RENTAJU7), liquidez corrente (LIQCOR7) e crescimento das vendas (CRESVEN7).

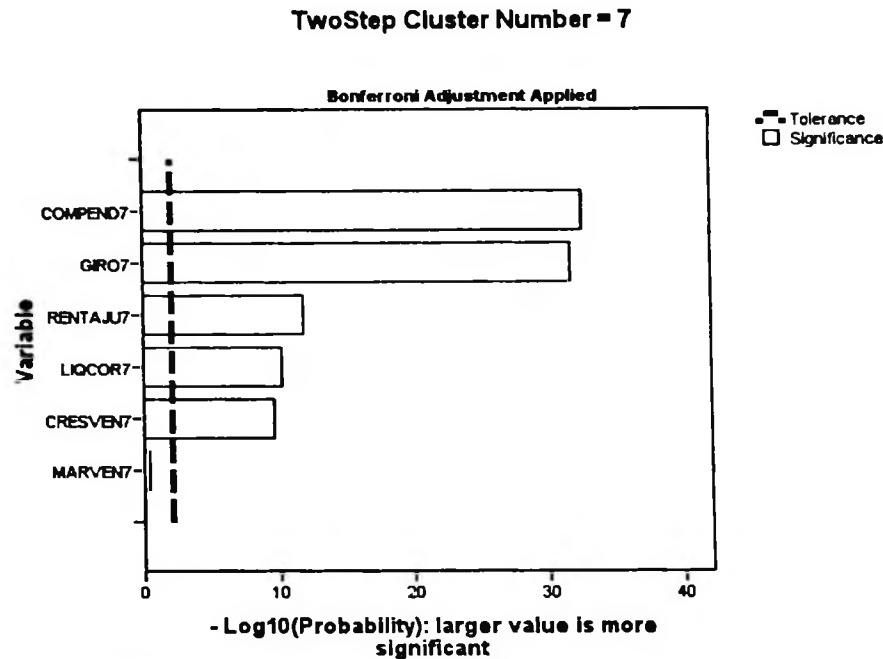


Figura 14: Variáveis de importância na formação do *cluster* 7

3.9 Avaliação da eficiência das empresas

Através da aplicação da *cluster analysis*, as empresas foram separadas em sete grupos com características homogêneas. Conforme explanado anteriormente, essa separação tem por objetivo estabelecer condições de comparabilidade entre as DMUs, *Decision Making Units*. É importante salientar que em cada *cluster* estão inclusas empresas que podem ter produtos diferentes, podem dispor de fatores de produção diferentes, embora a conduta das mesmas em relação aos 6 indicadores apresente características homogêneas. Essas empresas demonstram similaridade de comportamento quanto aos indicadores utilizados, os quais se constituem nas principais variáveis a serem consideradas na comparação para se apurar a eficiência relativa dessas empresas dentro de cada *cluster*.

As empresas serão então comparadas dentro de cada *cluster* através dos mesmos indicadores. A medida de eficiência comumente utilizada assume a forma da seguinte razão:

$Efficiency = Output / Input$. Compara-se aquilo produzido por elas (*outputs*) com aquilo que poderia ser produzido empregando-se os mesmos recursos (*inputs*).

Ao avaliar as características dos seis indicadores utilizados, cinco deles podem considerados como *outputs*, a saber:

- Rentabilidade sobre o Patrimônio Líquido;
- Crescimento das Vendas;
- Liquidez Corrente;
- Margem Líquida;
- Giro do Ativo.

Apenas um desses indicadores pode ser considerado como *input*, a Composição do Endividamento. Esse *input* é estimado como controlado, significando que o gestor da DMU tem o controle e em consequência pode alterar a quantidade utilizada desse fator.

A mensuração da eficiência deve ser feita de forma “parcial”, de acordo com Cooper *et al* (2007, p. 1) porque será usado apenas um indicador como *input* e 5 indicadores como *outputs*. Pretende-se investigar quais foram as empresas que alcançaram os maiores níveis de produtos medidos pelos cinco indicadores citados (*outputs*), dada a composição do endividamento atual (*input*), referente a dezembro de 2007.

Para tal investigação, é utilizado o programa *Frontier Analyst*, que, por sua vez, emprega a técnica DEA – *Data Envelopment Analysis*. Também se utiliza o modelo CCR, desenvolvido por Charnes, Cooper e Rhodes (1978). Esse modelo conduz o processo de mensuração considerando retorno constante de escala. Se o aumento em uma unidade do *input* estimado resultar em aumento proporcional nos *outputs*, a DMU apresenta um retorno constante de escala. O modo de análise usado é o *output maximisation*, o qual procura maximizar os *outputs*, dados os *inputs* considerados.

A seguir são apresentados e analisados os resultados referentes a cada *cluster*.

3.10 Avaliação dos resultados do *cluster 1*

Tabela 4: escores das empresas componentes do *cluster 1*

Unit name	Score	Unit name	Score	Unit name	Score
Atlas Transportes	100,00%	DPaschoal	70,10%	Supermix	58,40%
Brasilata	100,00%	Unimed Nordeste RS	69,60%	Aurora	58,30%
Mili	100,00%	Spaipa Coca-Cola	68,50%	Merck	58,20%
Modelo	100,00%	GRSA	68,40%	CTA - Continental	57,90%
Profarma	100,00%	Politec	68,40%	CCV	57,70%
Sada	100,00%	Unimed Vitória	68,40%	Copérdia	57,60%
Kraft Foods	97,90%	Zona Sul	68,20%	Lar	57,60%
Ultragaz	97,00%	Hermes	67,90%	Magazine Luiza	57,40%
Liquigás	95,60%	Makro	67,70%	Techint	57,40%
Coop	95,00%	B. Braun	67,50%	Yoki Alimentos	57,10%
Sulamed	85,40%	Grupo Delga	67,50%	Whirlpool	57,00%
G&P	85,30%	Acro	67,40%	Aunde	56,80%
Coca-Cola Ribeirão Preto	84,20%	Ipiranga Distribuidora	67,30%	Casas Pernambucanas-SP	56,80%
Compar	84,10%	Nutron Alimentos	66,90%	Brasmetal	56,50%
Shell	83,30%	Fratelli Vita Bebidas	66,40%	Acrinor	55,40%
Cocatrel	83,10%	Mahle	66,10%	Magneti Marelli Cofap	54,90%
Unimed Porto Alegre	82,70%	Servimed	65,80%	Elgin	54,70%
Unimed Rio	82,50%	Aethra	65,10%	Romagnole	54,50%
Rio Doce Café	81,60%	DHB Componentes	64,70%	Bristol-Myers Squibb	54,40%
Frigol	80,30%	Arcom	64,50%	Ficap	54,20%
Fras-Le	80,10%	Giroflex	64,50%	Dufer	54,10%
Clariant	78,50%	Panvel Farmácias	64,50%	BBA	54,00%
RR Donnelley	78,40%	Drogasil	64,00%	Gol	54,00%
SHV Gás	78,20%	Central Nacional Unimed	63,90%	Norsa	54,00%
Crios	77,80%	Farmacias Sant'Ana	63,90%	Stemac	54,00%
South Service	75,40%	Josapar	63,30%	Scopus	53,70%
Batávia	75,30%	Dow Brasil Sudeste	63,10%	Henkel	52,70%
Mabe Campinas	75,30%	Clarion	63,00%	Móveis Gazin	52,70%
CTIS	75,10%	Manchester	62,30%	Renner	52,40%
Zaraplast	75,00%	Prysmian Telecom	62,30%	Alcatel	52,30%
IBF	74,90%	Integrada	62,20%	KSB	52,10%
Unimed Cuiabá	74,70%	Coperguaçu	61,30%	Brasif	52,00%
Drogarias Pacheco	74,10%	Pesa	61,20%	FGV	52,00%
Marilan	74,10%	Refap	61,20%	Basf	51,70%
Angeloni	74,00%	TAM - Linhas Aéreas	60,90%	Frigelar Moto Refrigeração	50,60%
Nexans	73,60%	Armco	60,10%	Tower Automotive do Brasil	50,50%
Novartis	73,40%	GAM	60,00%	Lojas Yamada	49,90%
Morlan	72,60%	Rio Negro	60,00%	Schneider Electric	49,90%
Contax	72,40%	Tambasa	60,00%	Iharabras	49,30%
Pettenati	72,40%	Intermédica	59,90%	Giassi Supermercados	49,00%
Grupo Martins	71,90%	Transpetro	59,90%	Lojas Colombo	48,60%
Painço	71,80%	Selmi	59,80%	Acument	48,20%
G. Barbosa	71,70%	Dacarto	59,60%	Unimed Fortaleza	47,90%
Galvão	71,70%	Droga Raia	59,40%	Cosuel	46,70%
Pif Paf	71,40%	Frimesa	59,30%	Orlândia	46,40%
Gerdau Aços Especiais	71,10%	Crown	59,20%	Hexion Química	44,40%
Farmácia Pague Menos	71,00%	Placibras	58,80%	Lojas Quero-Quero	43,70%
Dori Alimentos	70,70%	Comau do Brasil	58,60%	SLC Alimentos	42,80%
Frisa	70,70%	Sandvik	58,50%	Lojas Riachuelo	41,20%
				Cognis	39,60%

A tabela 4 apresenta os escores, em ordem decrescente, referentes às 148 empresas componentes do *cluster* 1.

Na primeira coluna da tabela 4 encontram-se as seis empresas que demonstram maiores escores no *cluster*: Atlas Transportes, Brasilata, Mili, Modelo, Profarma e Sada. Todas elas têm o mesmo escore 100%, sendo, dessa forma, 100% eficientes.

Na penúltima coluna, última linha da tabela 4, aparece a empresa Cognis com o escore de 39,60%, ou seja, essa empresa é 39,60% eficiente.

As empresas com escore de 100% são as que poderiam ser tomadas como referências para as demais componentes do *cluster* 1. Elas estão na denominada fronteira eficiente, constituída pelas empresas mais eficientes de um determinado conjunto de empresas, no caso, as que compõem o *cluster* 1. Essa fronteira se constitui num padrão de melhores performances e pode ser usada como parâmetro para avaliar o comportamento de todas as demais empresas do grupo.

As empresas que se encontram na fronteira eficiente “envelopam” ou “dominam” todas as demais empresas daquele *cluster*. É importante ressaltar que as empresas da fronteira são 100% eficientes. Qualquer empresa fora da fronteira tem nível de eficiência menor que 100%. Quando se tem um *input* e dois *outputs* ou dois *inputs* e um *output*, o programa *Frontier Analyst* faz a representação gráfica. Entretanto, no presente caso, tem-se um *input* e cinco *outputs* e a representação gráfica não é efetuada pelo programa, pois se trata de gráfico multidimensional. O programa faz a plotagem dos casos, o que se pode ver na Figura 15, onde cada empresa é representada por um ponto.

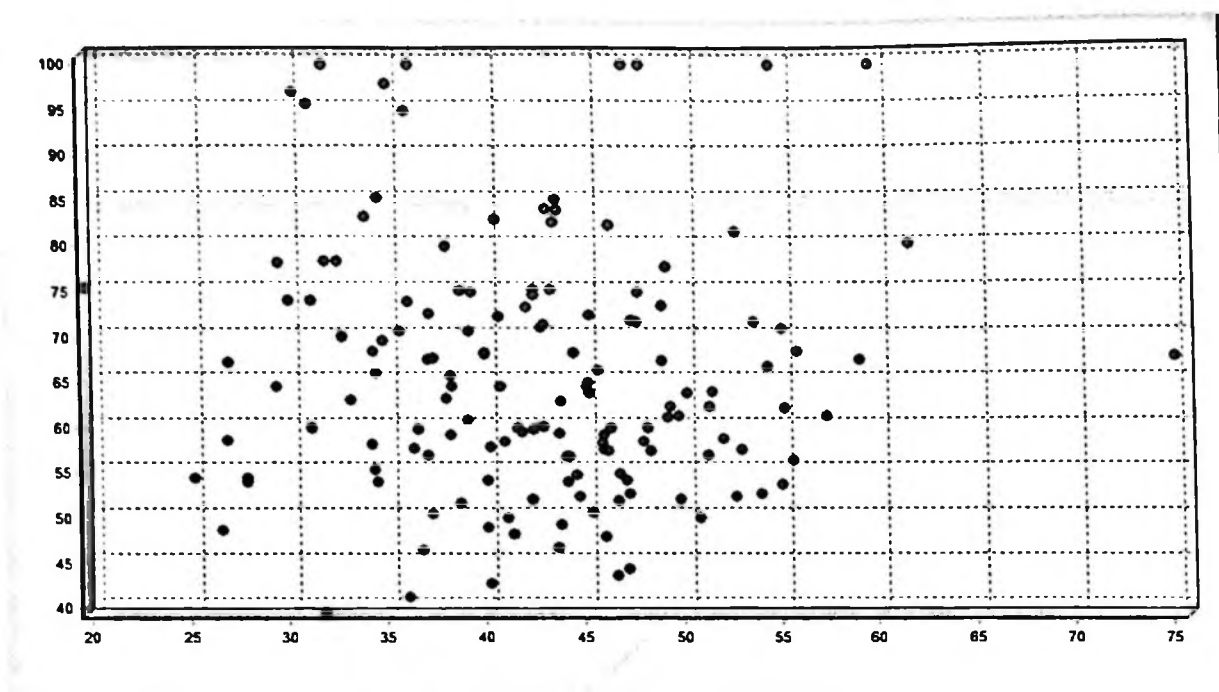


Figura 15: Plotagem das empresas do *cluster 1*

Pode-se verificar nessa figura 15 que as seis empresas 100% eficientes encontram-se na parte superior e todas as outras estão abaixo delas. A visualização pode ser feita olhando-se no eixo vertical cuja escala é de 100%.

A seguir, ao analisar a figura 16, pode-se ter uma visão geral da distribuição dos escores. Da esquerda para a direita encontram-se barras representativas das diversas empresas ineficientes (barras em cor vermelha), daquelas próximas da eficiência (barras em cor amarela) e das eficientes (barras em cor verde). O eixo vertical indica o número de empresas e o eixo horizontal mostra o percentual de eficiência.

No caso do *cluster 1* verifica-se que não existe caso de empresas com nível de eficiência entre 0% a 30%. Para o nível de 31% a 40%, há uma única empresa, Cognis, com o escore de 39,60%. Entre 51% e 60% de eficiência, surgem 49 empresas, representando a maior frequência.

O programa *Frontier Analyst* aponta empresas com o nível de 81% até 99,9% como aproximadamente eficientes (barras em cor amarela). No *cluster 1*, são apontadas 13 empresas nessas condições. E, por fim, há as 6 empresas 100% eficientes (barra em cor verde).

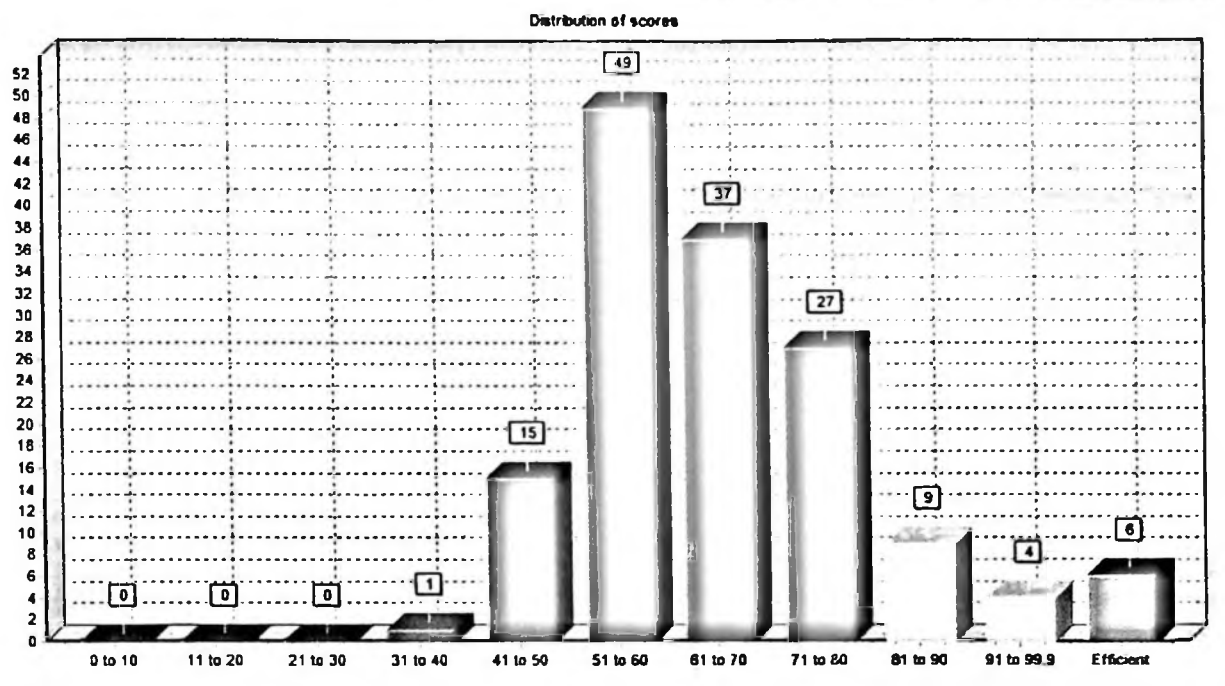


Figura 16: Distribuição dos escores do *cluster 1*

A figura acima apresenta as empresas eficientes, com base no número de vezes que servem como referência para a atuação das empresas ineficientes.

Quanto maior a frequência com que uma unidade eficiente aparece no conjunto de referência mais ela é considerada como um exemplo de boa performance. As unidades eficientes com pequena frequência são consideradas como tendo uma não usual combinação de *inputs* e *outputs* e, por esse motivo, não oferecem as melhores práticas operacionais para que as ineficientes as imitem. (*Frontier Analyst Manual*, 2008, p. 97).

A unidade eficiente que aparece com a maior frequência pode ser considerada como a líder geral e como tal poderia servir como exemplo de boas práticas operacionais para as unidades ineficientes.

Na figura 17, a empresa Mili é a líder geral com 93 referências. A Profarma, com 18 referências, apresenta menor frequência.

Embora as seis empresas sejam consideradas 100% eficientes, ainda assim aparecem ranqueadas de acordo com esse critério de número de referências para as empresas ineficientes, conforme exposto na figura 17.

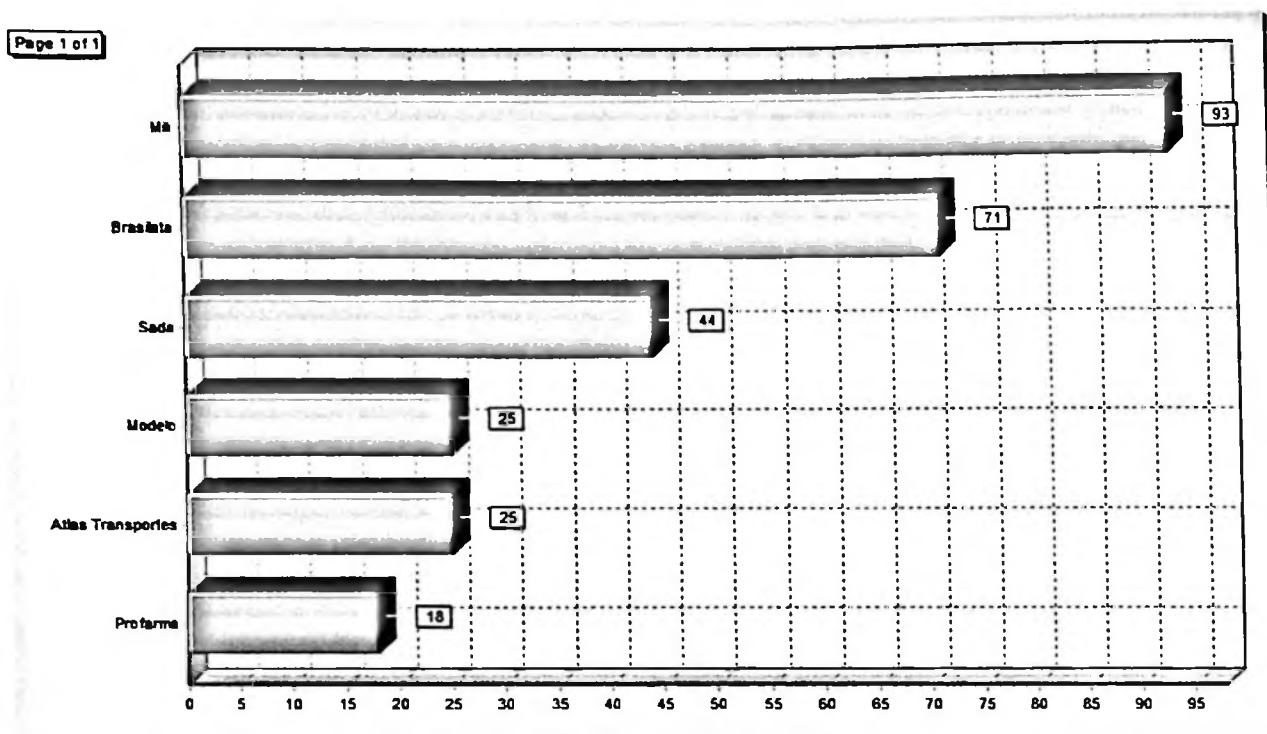


Figura 17: Empresas de referência do *cluster I*

A seguir, a figura 18 demonstra as possíveis melhorias indicadas pela análise, apresentando também uma breve visão das variáveis nas quais poderiam ser implementadas as melhorias. Trata-se de um gráfico em formato de pizza no qual as fatias representam as participações relativas das melhorias potenciais para cada *input* ou *output*.

Como se nota na figura 18, o *input* composição do endividamento aparece com o valor de zero por cento, significando que, dada uma determinada composição do endividamento, deve haver melhorias nos *outputs*. Como explicitado anteriormente, a opção escolhida foi a maximização dos *outputs* e não a minimização dos *inputs*. Assim, embora essa composição do endividamento possa ser alterada pelo gestor, considerando que tal posição permaneça como está, são indicadas as potenciais melhorias nos *outputs*.

A melhoria potencial indicada com maior ênfase é no indicador liquidez corrente (30,35%).

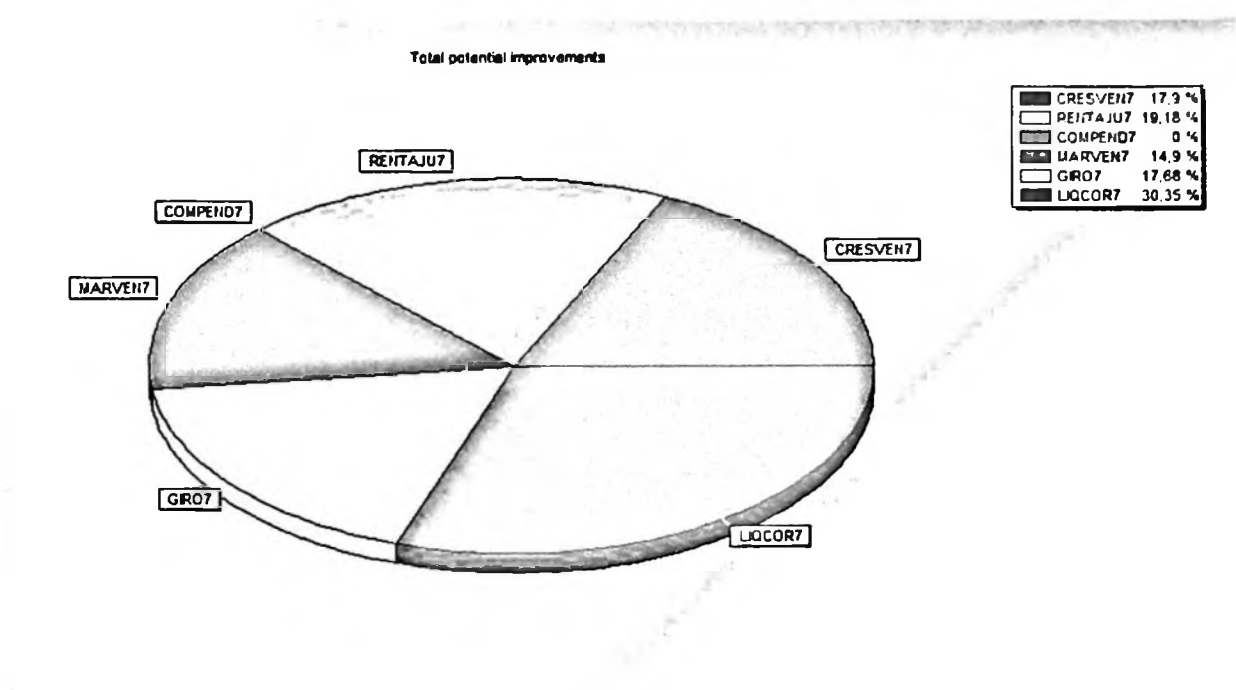


Figura 18: Melhorias potenciais para o *cluster 1*

O programa *Frontier Analyst* demonstra inúmeros outros detalhes no nível de cada empresa não incluídos neste trabalho.

3.11 Avaliação dos resultados do *cluster 2*

A tabela 5 apresenta os escores, em ordem decrescente, referentes às 72 empresas componentes do *cluster 2*.

Tabela 5: Escores das empresas componentes do *cluster 2*

Unit name	Score	Unit name	Score
Biscoitos Mabel	100,00%	Ponto Frio	64,80%
Evonik Degussa	100,00%	CSU	64,50%
Holcim	100,00%	Alcicla	64,40%
Ceva	99,90%	CETESB	63,30%
Parati	99,80%	Accor Hotels	63,00%
Caraíba	96,30%	Cimento Tupi	62,20%
Enersul	89,20%	Corol	62,00%
Coteminas	89,10%	Vitapelli	61,50%
Usina Nova América	88,60%	Armazém Paraíba	61,00%
Lanxess	88,30%	Engenpack-SP	61,00%
Rigesa	87,50%	Medial Saúde	60,50%
Ceal	83,40%	Iesa	60,40%
Leco	82,20%	Rexam	59,00%
Prada	80,90%	Coopercitrus	58,60%
Quero	80,70%	Camda	58,40%
Equipav	80,60%	Santa Amália	58,20%
Ipiranga Asfaltos	79,70%	Eliane	57,90%
Penha	79,00%	Milenia	57,50%
Paramount Têxteis	77,10%	Cooagri	55,70%
JBS	76,30%	Cacique	55,20%
Socic - Armazém Paraíba	73,30%	Medidata	53,50%
ThyssenKrupp	73,10%	Epcos do Brasil	52,90%
Eluma	72,10%	Vulcabrás-NE	52,50%
Doux	71,60%	Fiagril Agromercantil	52,10%
Carol	71,40%	Millennium	50,90%
Mercúrio	71,00%	Seara	50,70%
União Farmacêutica	70,80%	Conab	49,90%
Interagropecuária	70,60%	TIM-NE	49,40%
Mabe Itu	70,20%	Peixoto	48,80%
Cia. Iguaçu	69,00%	Gomes da Costa	48,20%
Egesa	67,20%	Siemens	48,00%
Laginha	67,00%	Bardella	47,00%
Weg Exportadora	66,80%	Galvani	46,30%
Trombini	66,40%	Fluxo	44,90%
Usina Bazan	66,30%	Cedro e Cachoeira	43,80%
CCE-AM	65,00%	Infraero	41,50%

Na primeira coluna da tabela 5 surgem as três empresas com maiores escores no *cluster*: Biscoitos Mabel, Evonik Degussa e Holcim. Todas elas apresentam o mesmo escore, 100%, sendo, portanto, 100% eficientes. Essas empresas poderiam ser tomadas como referências para as demais que compõem o *cluster 2*, por estarem na fronteira eficiente.

Na segunda coluna, última linha da tabela, aparece a empresa Infraero com o escore de 41,50%, isto é, essa empresa é 41,50% eficiente.

A figura 19 apresenta a plotagem de todas as empresas do *cluster 2*. Cada empresa é representada por um ponto da figura.

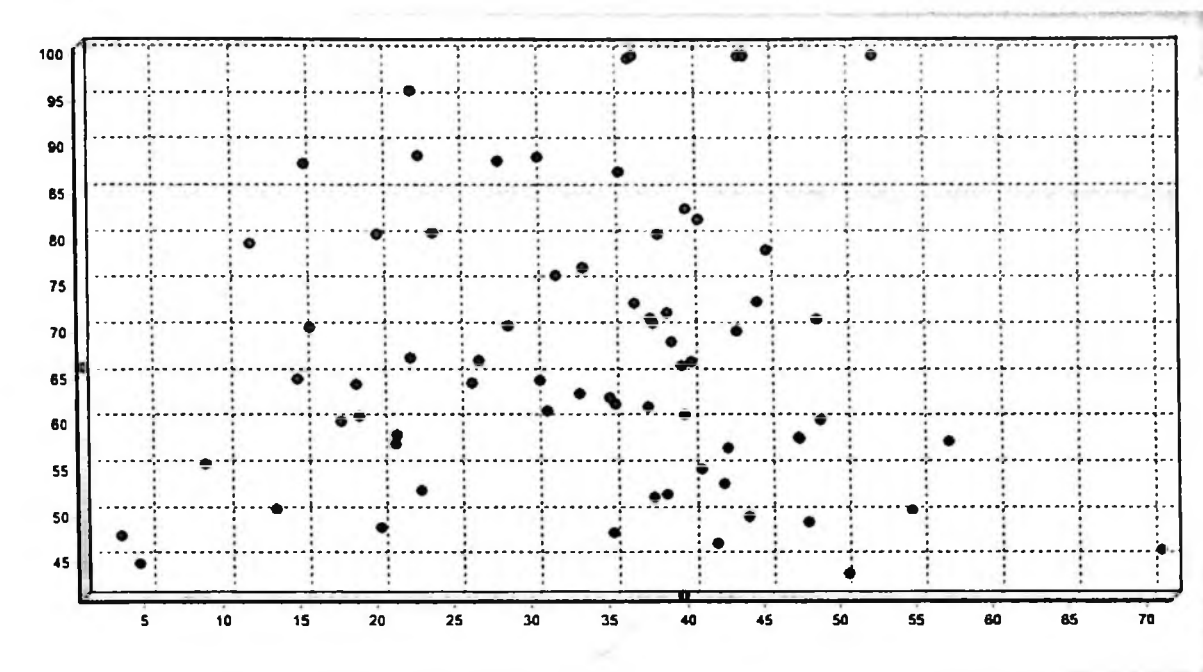


Figura 19: Plotagem das empresas do *cluster 2*

Pode-se verificar nessa figura 19 as três empresas 100% eficientes na parte superior e todas as outras abaixo delas. A visualização pode ser feita olhando-se no eixo vertical cuja escala é de 100%. Duas empresas, Ceva (99,90%) e Parati (99,80%), praticamente se sobrepõem às que estão na escala dos 100%.

A seguir apresenta-se a figura 20 na qual se pode ter uma visão geral da distribuição dos escores. Analisando a figura 20 da esquerda para a direita, observa-se que não existe nenhum caso de empresas com nível de eficiência entre 0% a 40%. Para o nível de 41% a 50%, constam 12 empresas. Entre 61% e 70% de eficiência encontram-se 19 empresas, representando a maior frequência.

As empresas com o nível de 81% até 99,9% são apontadas como aproximadamente eficientes (barras em cor amarela). No *cluster 2*, são apontadas 10 empresas nessas condições. Finalmente, encontram-se as 3 empresas 100% eficientes (barra em cor verde).

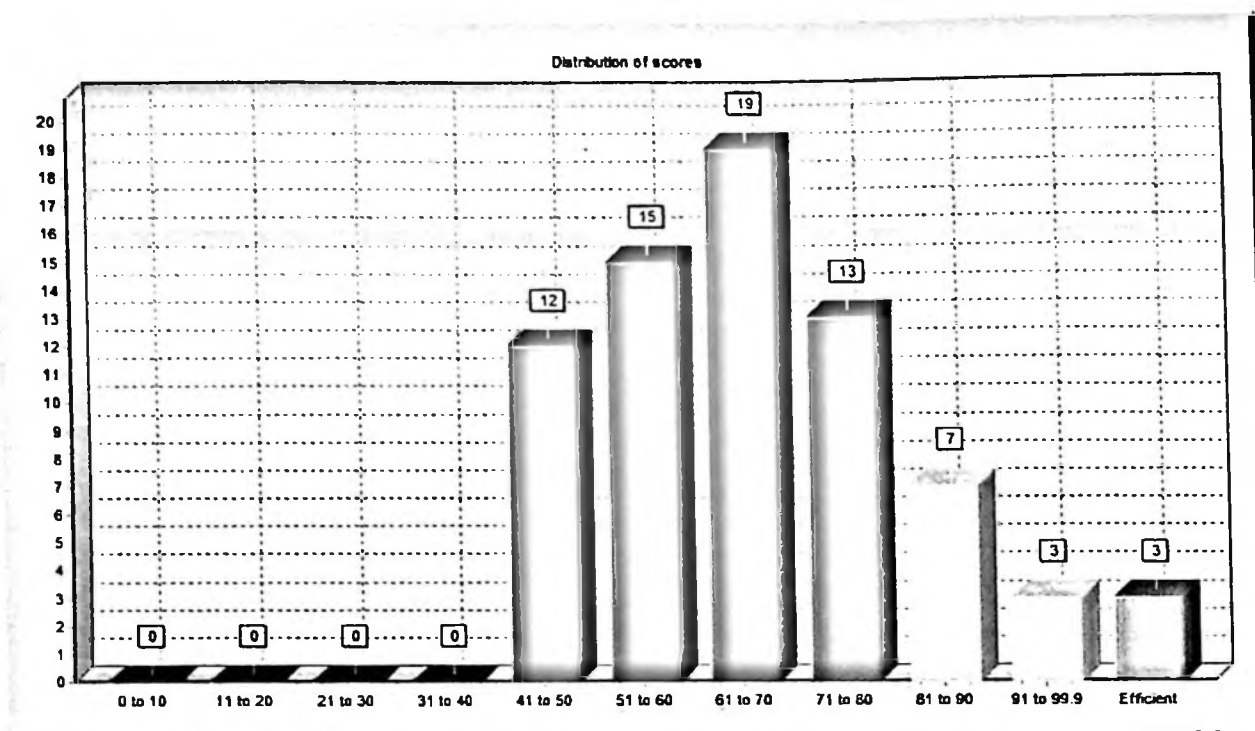


Figura 20: Distribuição dos escores do *cluster 2*

A figura 21 a seguir apresenta quantas vezes as empresas eficientes são referência para a atuação das empresas ineficientes.

Na figura 21, a empresa Evonik Degussa é a líder geral com 63 referências. A Biscoitos Mabel segue com 33 referências. Holcim demonstra menor frequência (20), ressaltando que as três empresas são 100% eficientes.

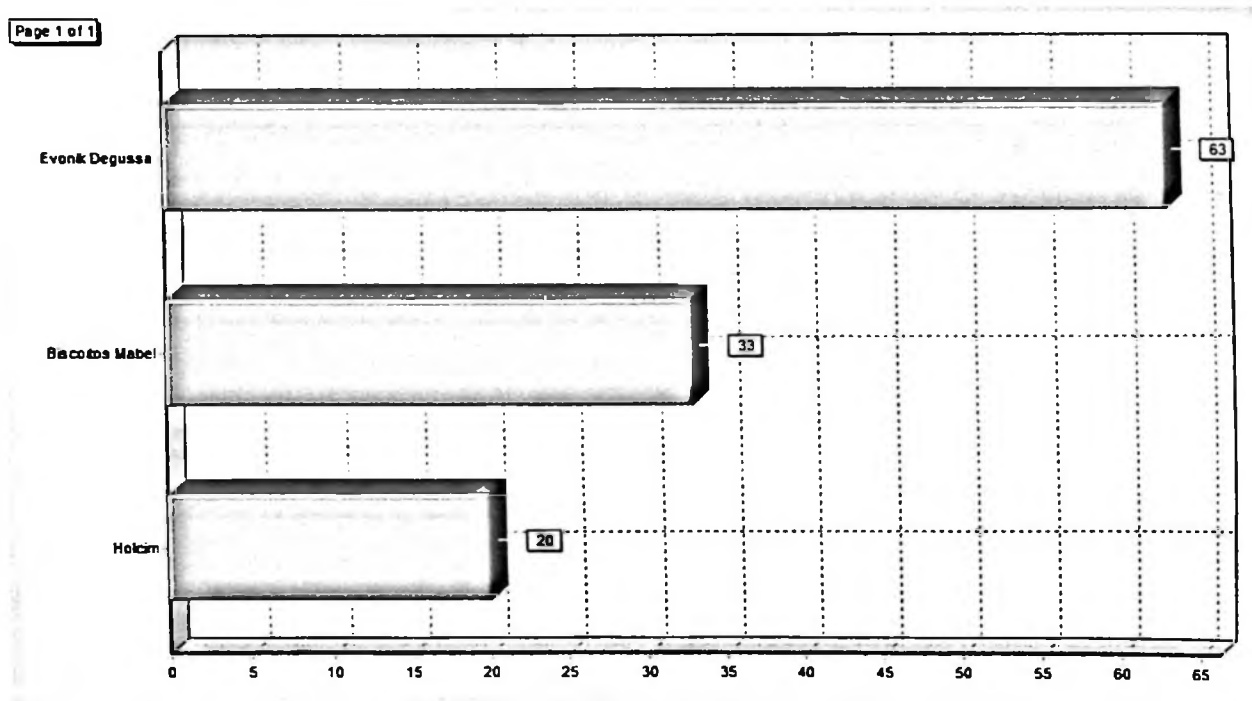


Figura 21: Empresas de referência do *cluster 2*

A figura 22 proporciona uma breve visão das variáveis nas quais poderiam ser implementadas as melhorias indicadas pela análise. Conforme se constata na figura 22, o *input* composição do endividamento aparece com o valor de zero por cento.

Entre os *outputs*, cabe um grande destaque para o indicador crescimento das vendas, no qual deveria se concentrar o maior esforço de melhoria. Nesse caso, a melhoria potencial indicada é de 38,54%.

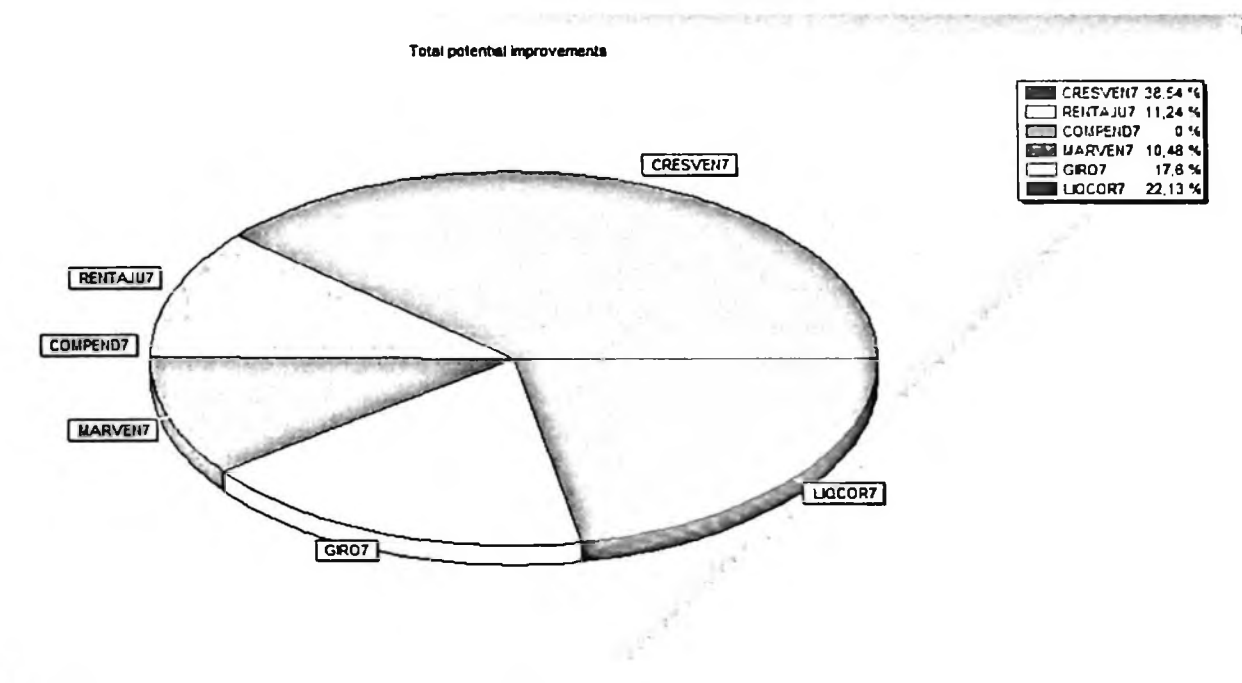


Figura 22: Melhorias potenciais para o *cluster 2*

3.12 Avaliação dos resultados do *cluster 3*

A tabela 6 apresenta os escores, em ordem decrescente, referentes às 74 empresas componentes do *cluster 3*.

Na primeira coluna da tabela 6 estão as quatro empresas com os maiores escores no cluster: Cagece, Farmasa, Usina Colombo, Weg Indústrias. Todas elas exibem o mesmo escore, 100% e são, assim, 100% eficientes. Essas empresas poderiam ser consideradas como referências para as demais componentes do *cluster 3* e estão na fronteira eficiente.

Na segunda coluna, última linha da Tabela 6, aparece a empresa Semp Toshiba-AM com o escore de 20,80%, significando que essa empresa é 20,80% eficiente. É válido salientar que no *cluster 3* aparecem 14 empresas eficientes nos níveis entre 20% e 30%.

Tabela 6: Escores das empresas componentes do *cluster 3*

Unit name	Score	Unit name	Score
Cagece	100,00%	Tejofran	44,00%
Farmasa	100,00%	Karsten	43,40%
Usina Colombo	100,00%	Romi	43,40%
Weg Indústrias	100,00%	Gonvarri	42,10%
Berneck Painéis e Serrados	88,80%	Arcelor	41,90%
Dataprev	83,70%	Santos-Brasil	40,40%
Celtins	83,40%	Tramontina Cutelaria	40,20%
Copasa	82,80%	CIV	39,90%
Petrocoque	79,70%	Vilma Alimentos	39,50%
Zamproгна	74,50%	Santanense	39,40%
Frigorífico Minerva	73,70%	Abnote	37,60%
NEC	65,40%	Casa da Moeda do Brasil	35,90%
Dass-Sport & Style	63,30%	Tortuga	35,70%
Coca-Cola Sorocaba	62,90%	Nadir Figueiredo	34,90%
Satipel	62,60%	Compesa	34,50%
Takata-Petri	62,40%	Gasmig	33,70%
Vipal	60,60%	Via Engenharia	32,90%
SPAL	58,70%	Viena	32,30%
Lupo	57,90%	Bic-AM	31,80%
Veja do Sul	55,70%	Administração Central	31,70%
Marfrig	55,00%	Videolar	31,00%
Solvay Indupa	54,90%	Oi	30,90%
Elekeiroz	54,60%	Editora Saraiva	30,20%
Andrade Gutierrez	52,30%	Construcap	29,30%
O Boticário	52,30%	Lojas Koerich	29,30%
Beira Rio	51,80%	Eternit	28,60%
Azaléia-NE	51,50%	Lorenzetti	28,40%
Mackenzie	51,50%	Magnesita	27,80%
São Paulo Alpargatas	51,10%	White Martins-N	27,80%
FCC	50,60%	CBC	27,10%
Suzano Petroquímica	50,20%	Lojas CEM	26,20%
Construtora Odebrecht	49,20%	Prodesp	25,60%
Duralex	48,80%	FTD	24,90%
Máster	47,70%	Mendes Junior	22,80%
Itautec	47,30%	Monsanto-NE	22,50%
SEM	47,20%	Hospital Santa Catarina	22,30%
Rossi	45,60%	Semp Toshiba-AM	20,80%

A figura 23 expõe a plotagem de todas as empresas do cluster 3. Cada empresa é representada por um ponto da figura.

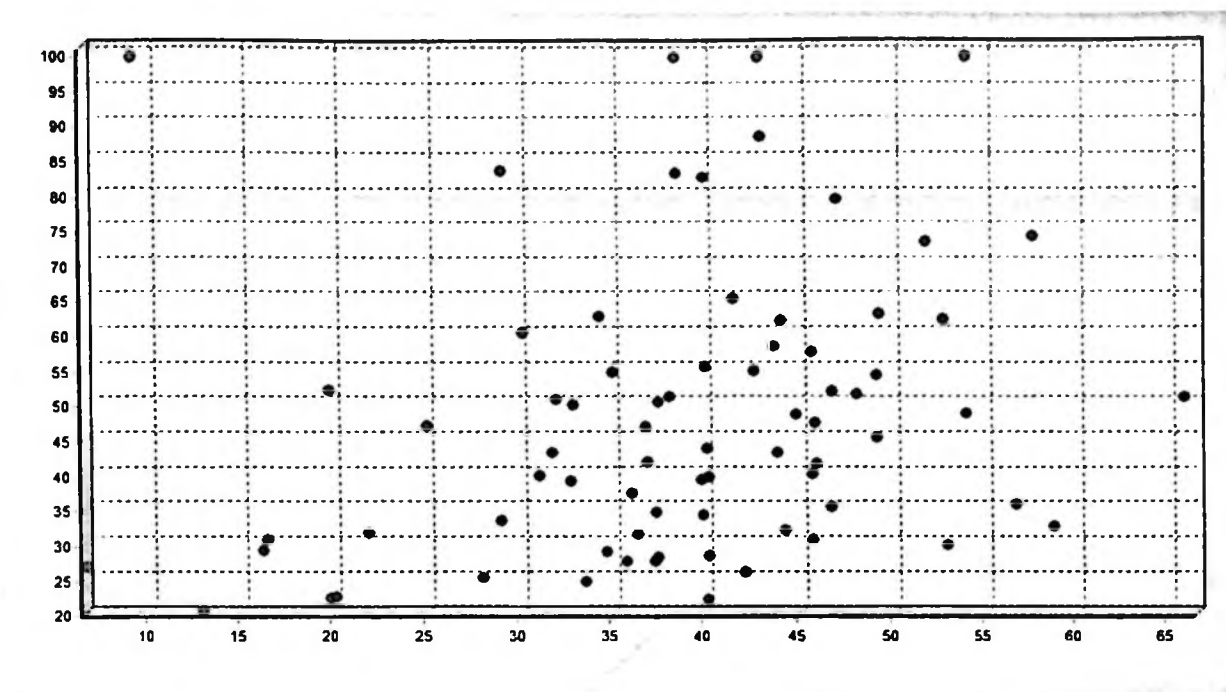


Figura 23: Plotagem das empresas do *cluster 3*

Pode-se perceber na figura 23 que as quatro empresas 100% eficientes encontram-se na parte superior e todas as outras se situam abaixo delas. A visualização pode ser feita olhando-se no eixo vertical cuja escala vai até 100%.

A seguir é apresentada a figura 24, na qual se pode ter uma visão geral da distribuição dos escores.

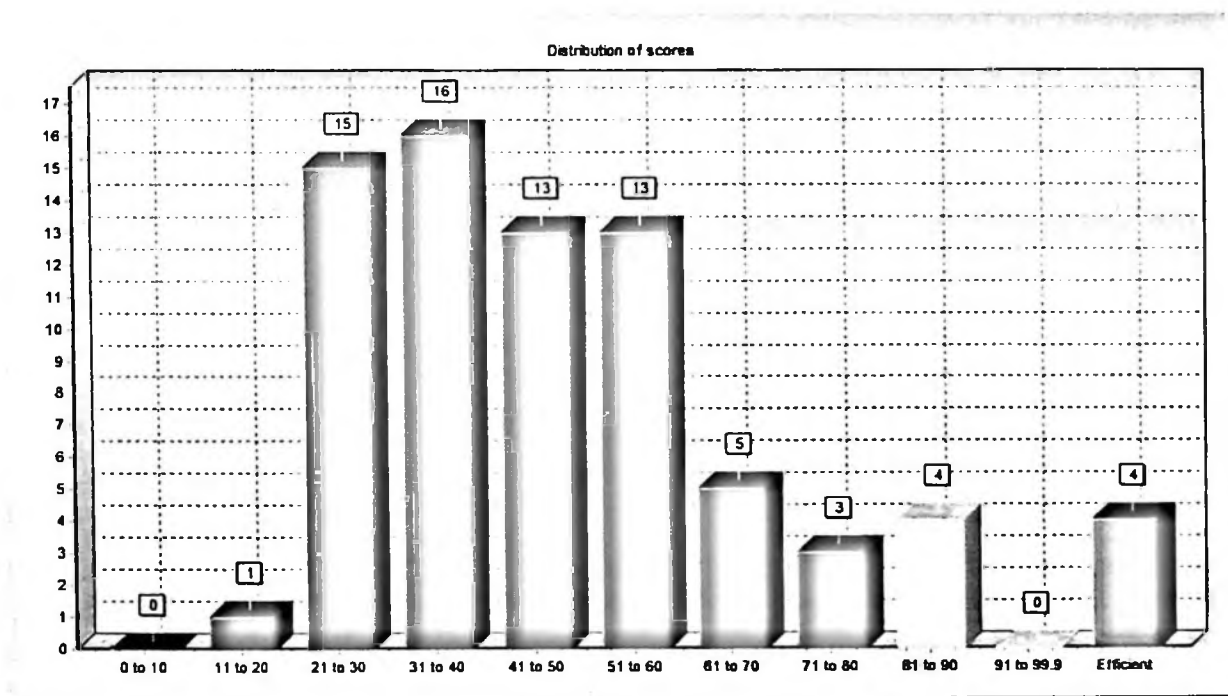


Figura 24: Distribuição dos escores do *cluster 3*

Analisando a figura 24, da esquerda para a direita, há apenas uma empresa com nível de eficiência entre 0% a 20%. Para o nível de 21% a 30%, constam 15 empresas. Entre 31% e 40% de eficiência, existem 16 empresas, representando a maior frequência.

As empresas com o nível de 81% até 99,9% são apontadas como aproximadamente eficientes (barras em cor amarela). No *cluster* 3, são apontadas 4 empresas nessas condições e elas estão no nível ente 81% a 90%.

Finalmente, encontram-se as 4 empresas 100% eficientes (barra em cor verde).

A figura 25 expõe a quantidade de vezes que as empresas eficientes servem como referência para a atuação das empresas ineficientes.

Na figura 25, a empresa Weg Indústrias é a líder geral com 67 referências. Em seguida vem a Cagece com 46 referências e logo após vem a Usina Colombo com 31 citações. A empresa com menor frequência (24) é a Farmasa. Cabe ressaltar que as 4 empresas são 100% eficientes.

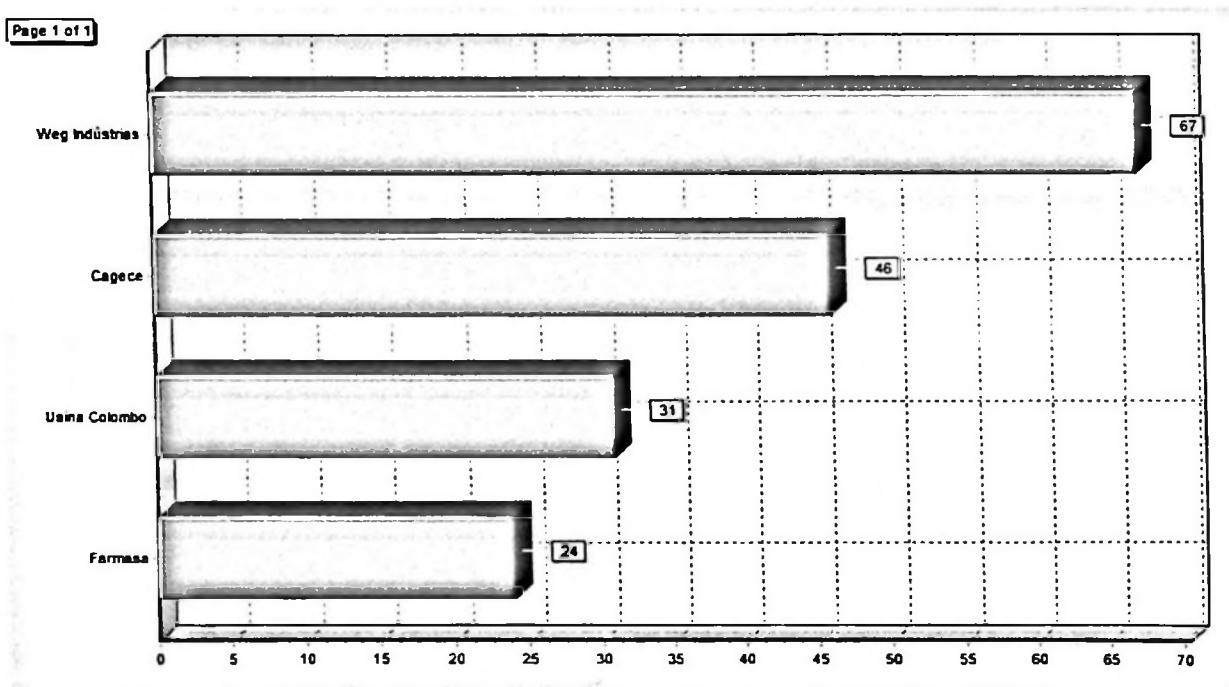


Figura 25: Empresas de referência do *cluster* 3

A figura 26 traz uma breve visão das variáveis nas quais poderiam ser implementadas as melhorias indicadas pela análise.

Como se percebe na figura 26, o *input* composição do endividamento aparece com o valor de zero por cento.

Entre os *outputs*, vale destacar os indicadores: crescimento das vendas, em que se sugere melhoria de 27,48% e o indicador liquidez corrente com melhoria potencial de 22,81%.

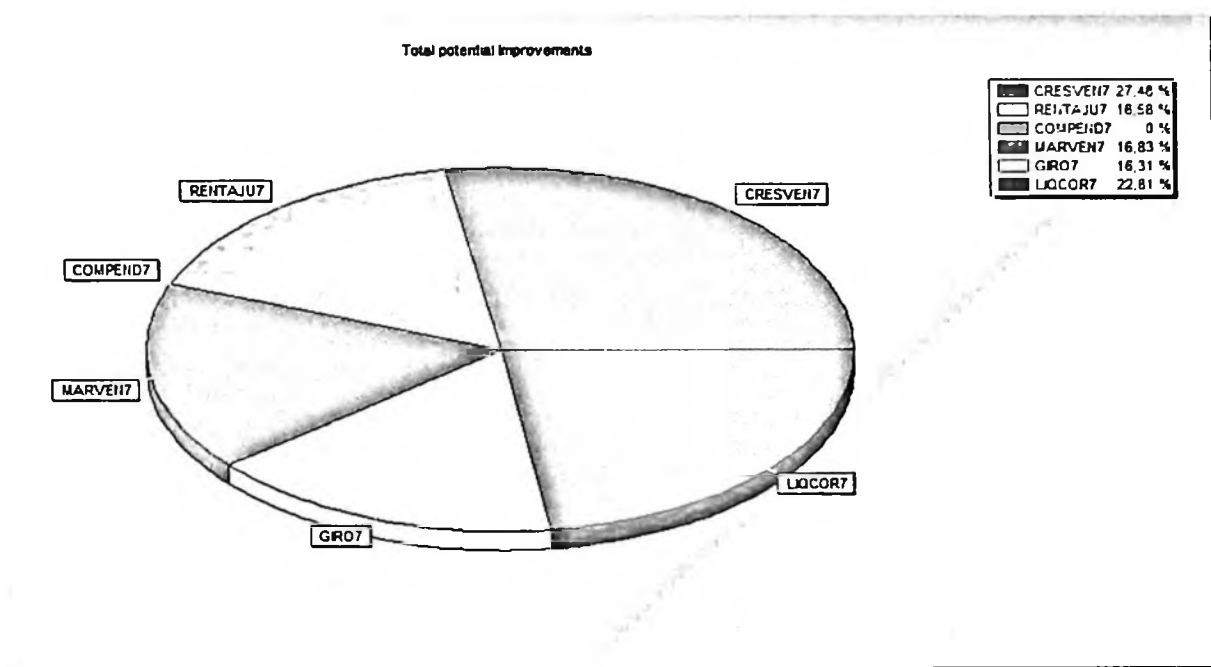


Figura 26: Melhorias potenciais para o *cluster* 3

3.13 Avaliação dos resultados do *cluster* 4

A tabela 7 apresenta os escores, em ordem decrescente, referentes às 69 empresas componentes do *cluster* 4.

Na primeira coluna da tabela 7, há as três empresas com os maiores escores no *cluster*: Cia Petrolífera Marlim, Rio de Janeiro Refrescos e Tenneco. Essas empresas são consideradas 100% eficientes e portanto, poderiam ser consideradas como referências para as demais que compõem o *cluster* 4. Elas estão na fronteira eficiente.

Na segunda coluna, última linha da Tabela 7, aparece a empresa Nibrasco com o escore de 28,60%, ou seja, essa empresa é 28,60% eficiente.

Cabe salientar que no *cluster* 4 aparecem 28 empresas eficientes nos níveis que se situam abaixo de 50%.

Tabela 7: Escores das empresas componentes do *cluster* 4

Unit name	Score	Unit name	Score
Cia. Petrolífera Marlim	100,00%	Wirex	53,90%
Rio de Janeiro Refrescos	100,00%	Autotrac	53,70%
Tenneco	100,00%	CBC Cartuchos	53,50%
Copesul	93,60%	IPQ	53,20%
Omint	79,50%	Schulz	52,70%
Fleury Medicina e Saúde	79,40%	Prysmian Energia	52,60%
Telemont	75,30%	Lojas Americanas	50,90%
Aços Villares	74,00%	Sundown	49,30%
PPE Fios	74,00%	Vix	49,20%
Medabil S/A	72,10%	Estadão	48,00%
Suspensys	70,80%	E.C.T.	47,90%
Petroquímica União	66,50%	Unimed BH	47,40%
Disbrave	66,30%	Cemig Distribuição	47,10%
Garoto	64,60%	Borlem	46,40%
Amil	63,00%	Pósitron	46,40%
Comgás	63,00%	Natura	45,80%
FL Brasil AS	62,20%	UTC Engenharia	44,60%
Ceg Rio	61,90%	Kobrasco	44,10%
3M	61,80%	Atlas Schindler	43,60%
Infoglobo	61,60%	Autometal	42,20%
CPFL – Paulista	60,90%	Tigre	42,00%
Souza Cruz	59,40%	Rip	41,90%
Esab	59,30%	Atento	40,90%
Prosegur	58,90%	Belgo Bekaert-NE	40,70%
Santher	58,90%	Gerdau Aços Longos	39,70%
CEG	58,70%	Ultrafertil	35,80%
CPFL - Piratininga	58,30%	Weg Equipamentos	35,10%
AstraZeneca	57,10%	Bahiagás	35,00%
Jorlan	57,10%	ATP	34,60%
Triunfo	56,70%	V & M do Brasil	33,60%
Maxion	56,50%	Itabrasco	33,40%
Cinemark	56,40%	Unicafé	33,20%
Embaré	55,30%	Hispanobras	32,80%
Copebrás	54,30%	Furukawa	30,40%
		Nibrasco	28,60%

A figura 27 apresenta a plotagem de todas as empresas do cluster 4. Cada empresa é representada por um ponto da figura.

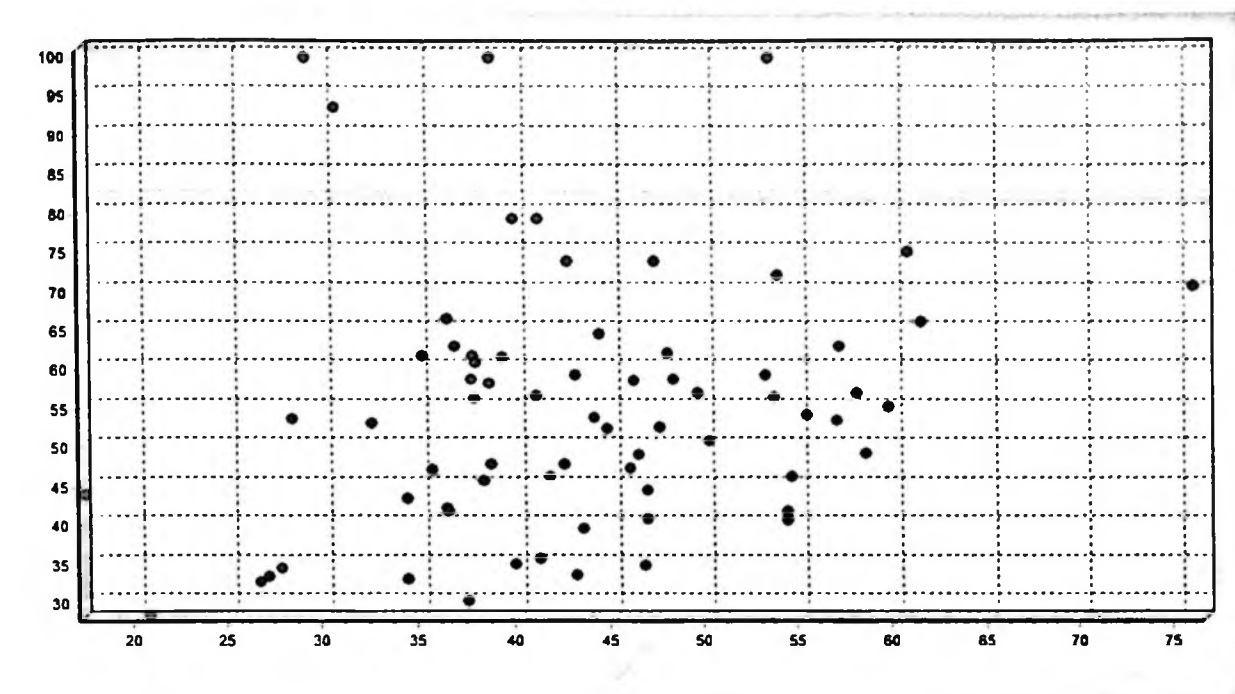


Figura 27: Plotagem das empresas do *cluster* 4

Pode-se notar na figura 27 que as três empresas 100% eficientes aparecem na parte superior e todas as outras estão situadas abaixo delas. A visualização pode ser feita no eixo vertical cuja escala vai até 100%.

A seguir se apresenta a figura 28 em que se pode ter uma visão geral da distribuição dos escores.

Ao verificar a figura 28, da esquerda para a direita, não existe caso de empresas com nível de eficiência entre 0% a 20%. Somando aquelas que estão entre os níveis de 21% a 50%, há um total de 29 empresas. Entre 51% e 60% de eficiência encontram-se 20 empresas, que representam a maior frequência.

As empresas com o nível de 81% até 99,9% são apontadas como aproximadamente eficientes (barras em cor amarela). No *cluster* 4, apenas 1 empresa está nessa condição, a Copesul, sendo 93,60% eficiente. Finalmente, encontram-se as 3 empresas 100% eficientes (barra em cor verde).

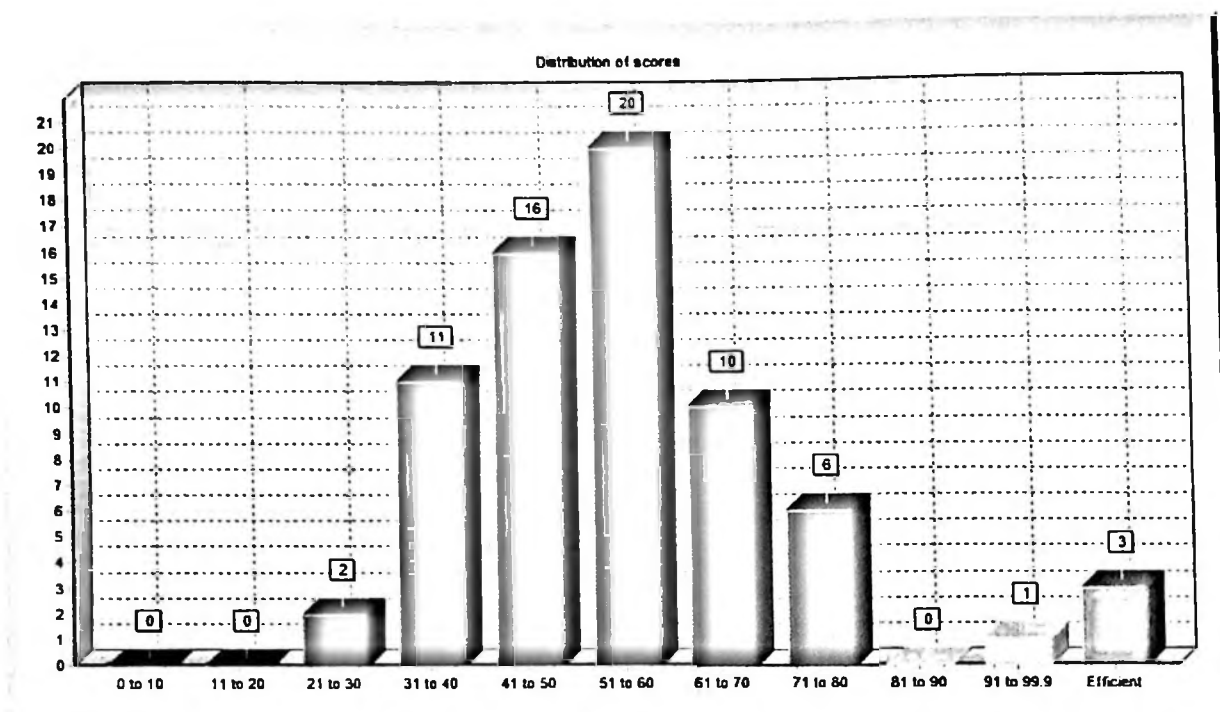


Figura 28: Distribuição dos escores do *cluster 4*

A figura 29 mostra quantas vezes as empresas eficientes servem como referência para a atuação das empresas ineficientes.

Na figura 29, a empresa Rio de Janeiro Refrescos é a líder geral com 67 referências. A seguir, vem a Cia Petrolifera Marlim com 66 referências e, dessa forma, quase com o mesmo número de referências da líder. Aquela com menor número de referências é a Tenneco, com 14. É válido ressaltar que as 3 empresas são 100% eficientes.

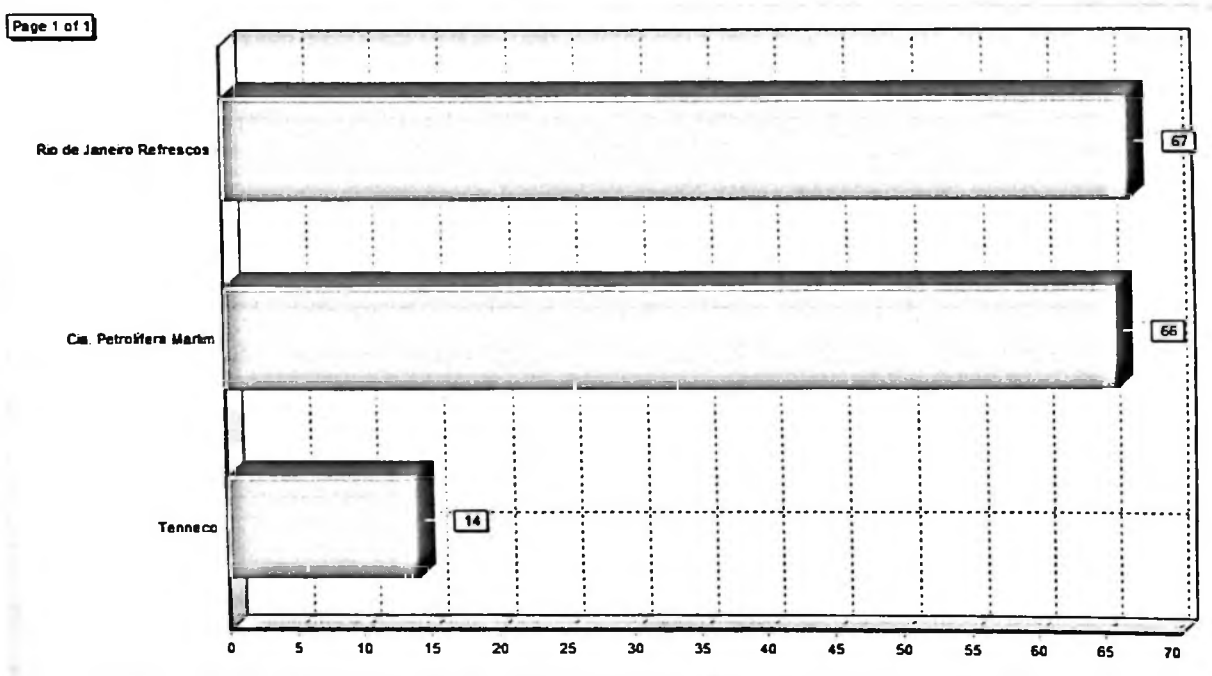


Figura 29: Empresas de referência do *cluster 4*

Na figura 30 há uma breve visão das variáveis nas quais poderiam ser implementadas as melhorias indicadas pela análise.

Conforme se verifica na figura 30, o *input* composição do endividamento aparece com o valor de zero por cento.

Entre os *outputs*, cabe destacar os indicadores: liquidez corrente com indicação de 27,06% de melhoria potencial e o indicador rentabilidade sobre o patrimônio líquido com melhoria potencial de 24,41%.

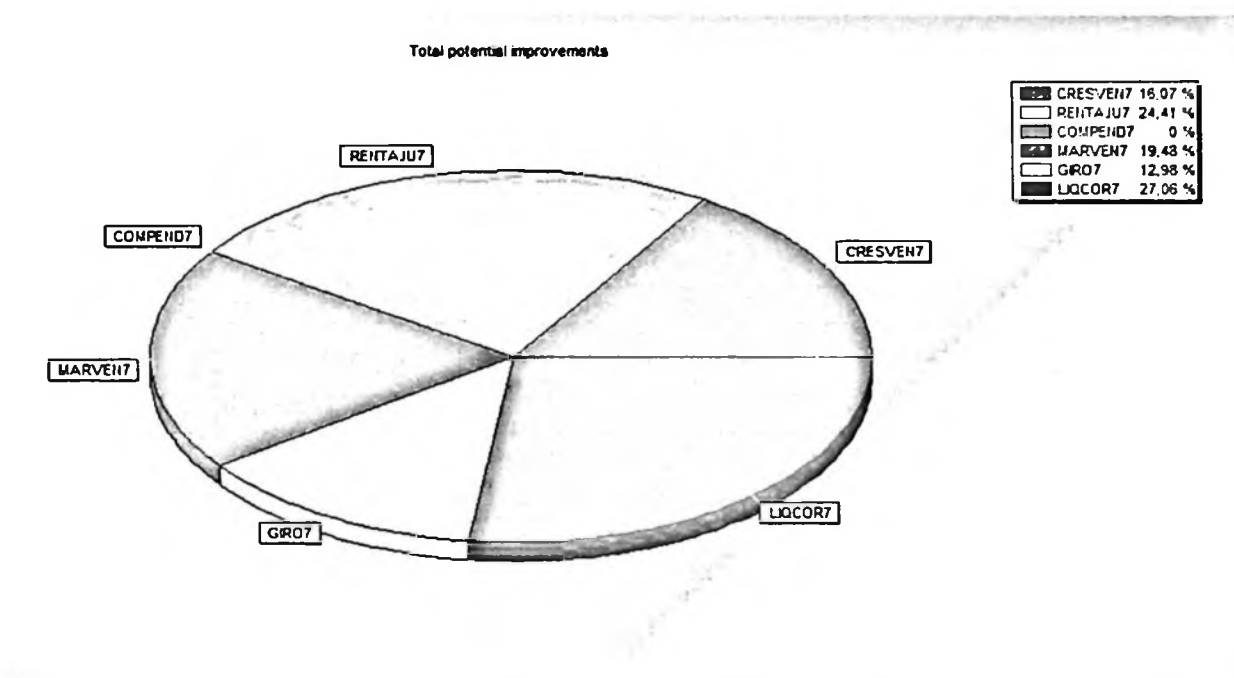


Figura 30: Melhorias potenciais para o *cluster 4*

3.14 Avaliação dos resultados do *cluster 5*

A tabela 8 mostra os escores, em ordem decrescente, referentes às 55 empresas componentes do *cluster 5*.

Na primeira coluna da tabela 8 estão as três empresas com maiores escores no *cluster*: Carbocloro, Eletronuclear e Suzano. Essas empresas são consideradas 100% eficientes e portanto, poderiam ser consideradas como referências para as demais que compõem o *cluster 5*. Elas estão na fronteira eficiente.

Na segunda coluna, última linha da Tabela 8, localiza-se a empresa Rio Capim Caulim, com o escore de 20,80%, ou seja, essa empresa é 20,80% eficiente.

Cabe salientar que no *cluster* 5 aparecem 20 empresas eficientes nos níveis que se situam abaixo de 50%.

Tabela 8: Escores das empresas componentes do *cluster* 5

Unit name	Score	Unit name	Score
Carbocloro	100,00%	ArcelorMittal Brasil	60,70%
Eletronuclear	100,00%	Localiza	59,70%
Suzano	100,00%	Cemar	59,60%
Usina Batatais	99,80%	Telefônica	59,20%
White Martins-NE	98,30%	Coelba	59,10%
Elektro	96,20%	Copergás	57,50%
Usina São Martinho	93,60%	SCGás	54,00%
Corsan	92,90%	Compagas	53,20%
Klabin	91,50%	Alcoa	49,40%
Cosipa	90,30%	Primesys	47,00%
UTE Norte Fluminense	90,00%	Petrobras	46,70%
Clealco	89,40%	Total Fleet	45,40%
Sanepar	88,70%	Guarani	45,20%
Rodonorte	83,60%	Coelce	45,00%
Alunorte	79,10%	Santa Cruz	45,00%
Duke Energy	77,10%	Saint-Gobain Vidros	44,70%
Usina da Pedra	77,10%	Arcelormittal Tubarão	42,00%
Sabesp	76,00%	Nitro Química	41,90%
Chesf	74,90%	Gerdau Açominas	39,30%
Termope	71,10%	Casan	39,00%
VCP	68,00%	CBA	37,70%
Pisa	67,60%	Orbitall	37,10%
Albrás	66,90%	Serasa	36,20%
Cosern	66,00%	Furnas	34,30%
Azaléia	65,90%	Ericsson Telecomunicações	29,80%
Usaçúcar	63,20%	Ipiranga Química	28,60%
CGTF	62,70%	Ticket	28,00%
		Rio Capim Caulim	20,80%

A figura 31 expõe a plotagem de todas as empresas do *cluster* 5, sendo cada uma representada por um ponto da figura.

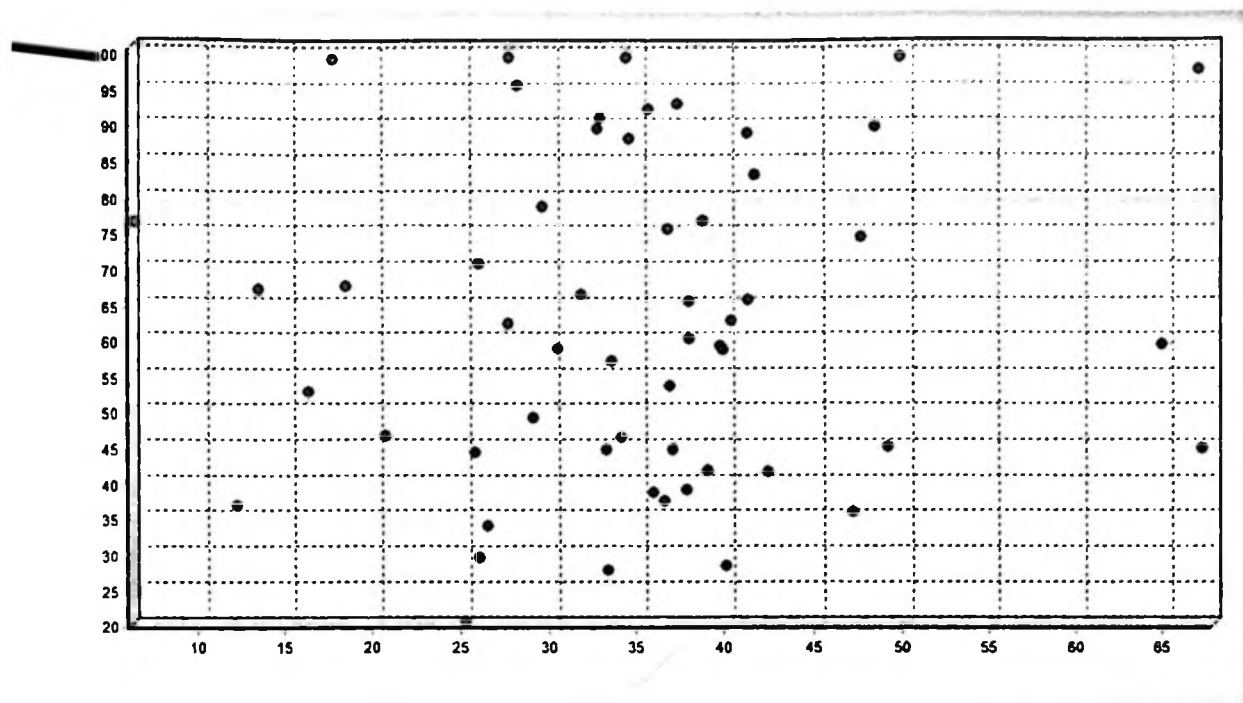


Figura 31: Plotagem das empresas do *cluster 5*

Pode-se constatar na figura 31 que as três empresas 100% eficientes encontram-se na parte superior e todas as outras estão situadas abaixo delas. A visualização pode ser feita no eixo vertical cuja escala vai até 100%.

Cabe salientar que numa visualização rápida dá a impressão de serem 5 as empresas que estariam no topo da figura, porém das 5 empresas, 3 de fato estão na escala dos 100%, a quarta está na escala de 99,80% e a quinta está na escala de 98,30%, o qual se localiza na extrema direita da figura e está ligeiramente abaixo em relação aos outros 4.

Em seguida, expõe-se a figura 32 em que se pode ter uma visão geral da distribuição dos escores.

De acordo com a figura 32, da esquerda para a direita, não se verifica qualquer caso de empresas com nível de eficiência entre 0% a 10%. Juntando as empresas que estão entre os níveis de 11% a 50% há um total de 20 empresas. Entre 41% e 50% de eficiência encontram-se 10 empresas, representantes da maior frequência.

As empresas com o nível de 81% até 99,9% são apontadas como aproximadamente eficientes (barras em cor amarela). No *cluster 5*, verificam-se 11 empresas nessa condição. Também se visualizam as 3 empresas 100% eficientes (barra em cor verde).

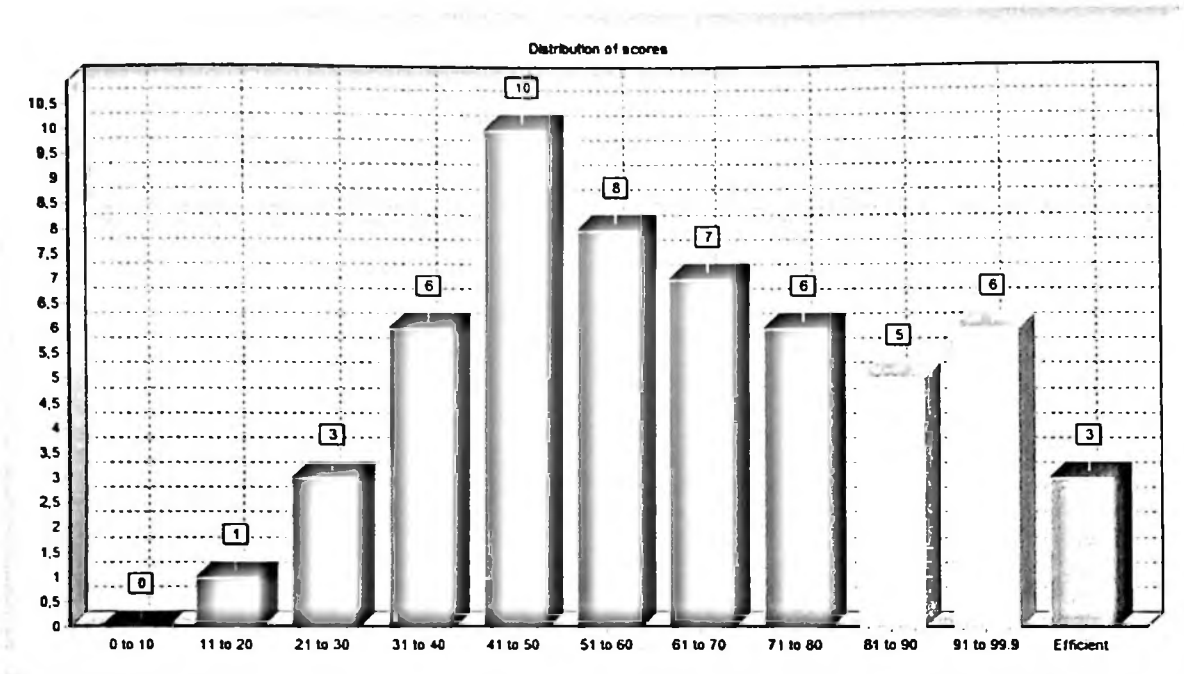


Figura 32: Distribuição dos escores do cluster 5

A figura 33 expõe quantas vezes as empresas eficientes servem como referência para a atuação das empresas ineficientes.

Na figura 33, a empresa Suzano é a líder geral com 53 referências, acompanhada pela Carbocloro, com 42 referências. Aquela com o menor número de referências é a Eletronuclear, com 10. Cabe destacar que as três empresas são 100% eficientes.

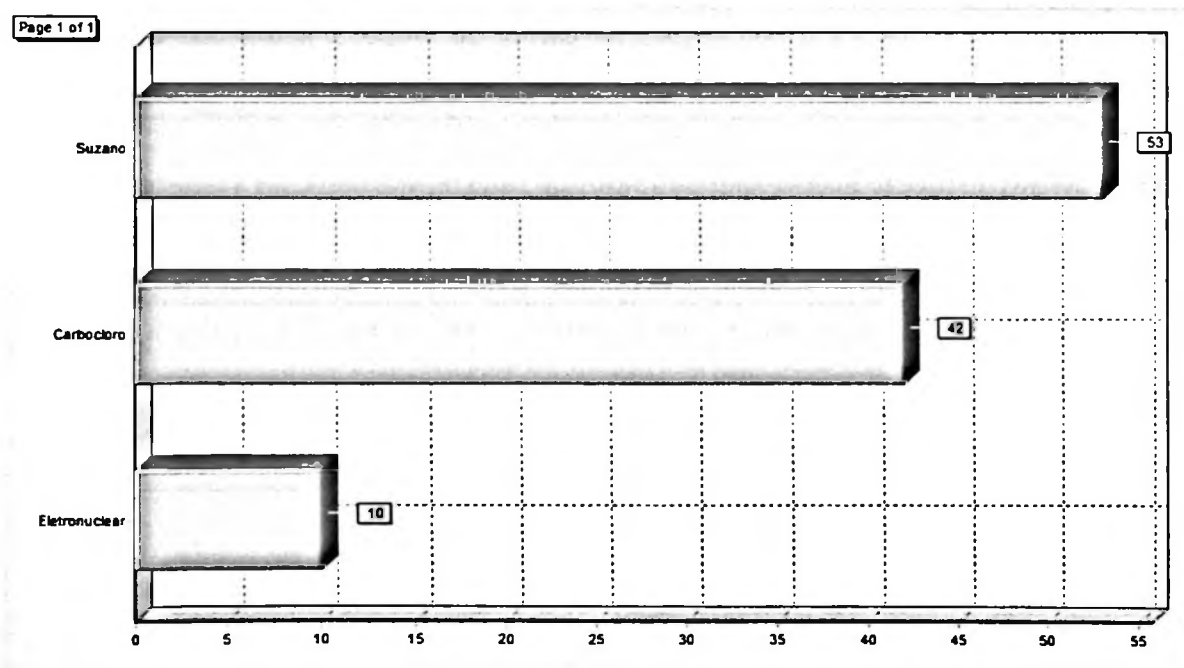


Figura 33: Empresas de referência do cluster 5

Na figura 34, percebe-se uma breve visão das variáveis nas quais poderiam ser implementadas as melhorias indicadas pela análise.

Como se analisa na figura 34, o *input* composição do endividamento aparece com o valor de zero por cento.

Entre os *outputs*, é importante destacar os indicadores: liquidez corrente com indicação de 39,33% de melhoria potencial e o indicador crescimento das vendas com melhoria potencial de 22,79%.

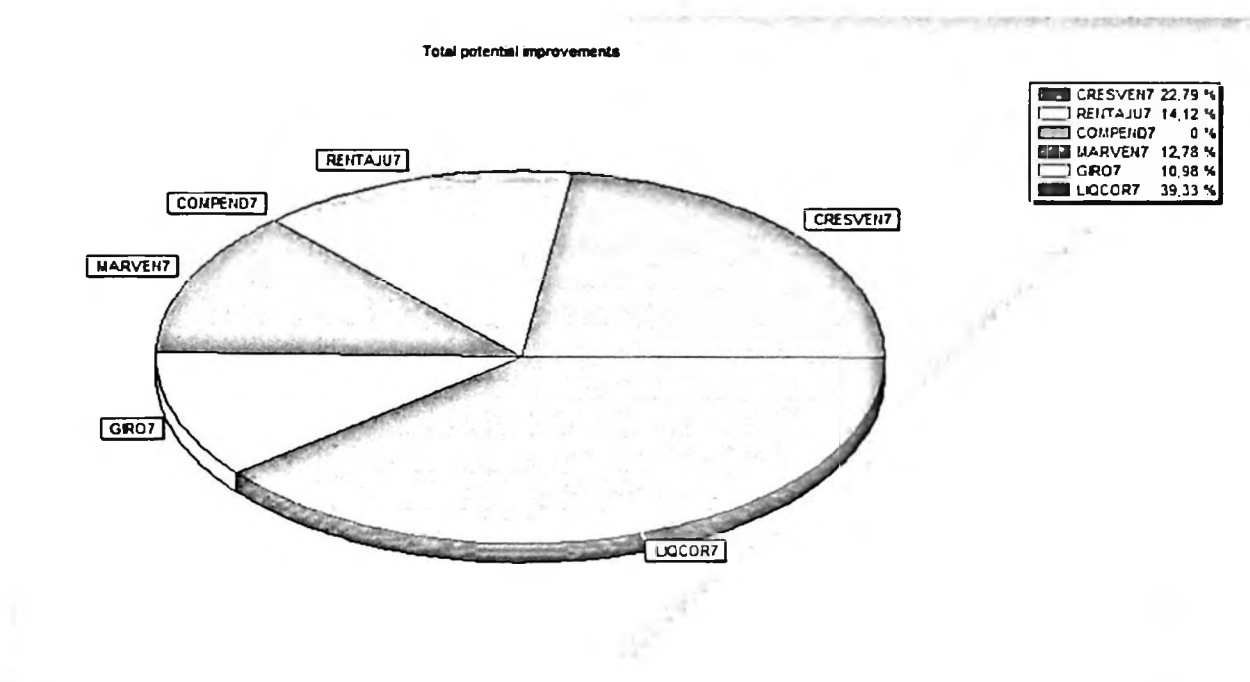


Figura 34: Melhorias potenciais para o *cluster* 5

3.15 Avaliação dos resultados do *cluster* 6

A tabela 9 apresenta os escores, em ordem decrescente, referentes às 70 empresas componentes do *cluster* 6.

As quatro empresas que demonstraram maiores escores no *cluster*, Cisa Trading, CNEC, Polietilenos e Schincariol N-NE, surgem na primeira coluna da tabela 9. Essas empresas são analisadas como 100% eficientes e, portanto, poderiam ser consideradas como referências para as demais que compõem o *cluster* 6. Elas estão na fronteira eficiente.

A empresa Montecitrus Trading está na segunda coluna, última linha da tabela 9, com o escore de 29,50%, sendo, dessa maneira, 29,50% eficiente.

Cabe salientar que no *cluster* 6 aparecem 15 empresas eficientes nos níveis situados abaixo de 50%.

Tabela 9: Escores das empresas componentes do *cluster* 6

Unit name	Score	Unit name	Score
Cisa Trading	100,00%	Cocari	57,20%
CNEC	100,00%	Telemig Celular	56,30%
Polietilenos	100,00%	Positivo	55,80%
Schincariol N-NE	100,00%	RDM	55,40%
Rações Total	87,90%	LDC Brasil	55,20%
Cooperalfa	83,30%	Agrale	54,90%
Facchini	78,50%	Cooperativa Batavo	54,80%
Gafisa	78,40%	Brascan	54,30%
Eleva	76,60%	Amaggi	53,70%
Languiru	76,60%	Cooper AI	53,50%
Castrolanda	76,00%	C. Vale	53,40%
Pires do Rio – Citep	76,00%	Copercampos	53,30%
Iesa Óleo & Gás	74,20%	Coopavel	53,20%
Imcopa	74,00%	Cocamar	53,10%
Livraria Saraiva	72,40%	Bunge Alimentos	52,20%
Cotriel	70,40%	Voith Paper	51,60%
Tuper	69,30%	Baterias Moura	50,80%
Confepar	68,90%	Coamo	50,80%
EIT	68,20%	GPC	50,60%
Bom Gosto	67,00%	Brastemp da Amazônia	50,10%
Sandvik MGS	66,70%	Ciser	49,90%
Itambé	66,30%	Perdigão Agroindustrial	49,50%
Sadia	65,90%	Embraer	48,20%
Cotripal	64,60%	Copacol	47,30%
CRA	64,50%	Cotrisal	47,10%
Cooparaíso	63,60%	Marcopolo	47,10%
Granol	63,60%	Cooxupé	46,50%
Copagril	62,80%	Cooperativa Agrária	45,60%
Randon	61,40%	Thomson Multimidia-AM	45,40%
Metalfrio	59,50%	Gevisa	45,20%
Teuto	59,20%	Dias Branco	43,20%
Vivara	59,00%	Villares Metals	42,10%
Bianchini	58,50%	Usiminas Mecânica	40,60%
Comigo	58,50%	Bunge Fertilizantes	35,40%
Fertipar	58,30%	Montecitrus Trading	29,50%

A plotagem de todas as empresas do *cluster* 6 é exibida na figura 35, estando cada empresa representada por um ponto da figura.

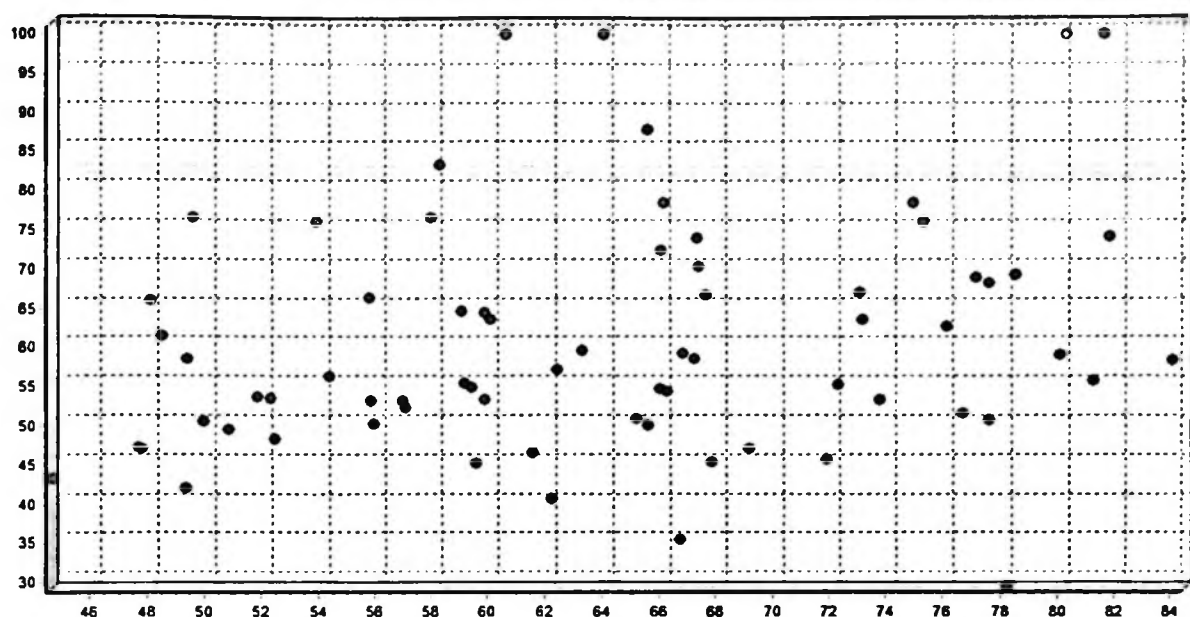


Figura 35: Plotagem das empresas do *cluster 6*

Verifica-se, na figura 35, que as quatro empresas 100% eficientes encontram-se na parte superior e todas as outras estão situadas abaixo delas. A visualização pode ser feita no eixo vertical cuja escala vai até 100%.

Na figura 36, é possível ter uma visão geral da distribuição dos escores.

Observando a figura 36, da esquerda para a direita, verifica-se que não existe caso de empresas com nível de eficiência entre 0% a 20%. Ao somar as empresas localizadas entre os níveis de 21% a 50% há um total de 19 empresas. Entre 51% e 60% de eficiência encontram-se 22 empresas, que representam a maior frequência.

As empresas com o nível de 81% até 99,9% são apontadas como aproximadamente eficientes (barras em cor amarela). No *cluster 6*, apenas 2 empresas estão na faixa entre 81% e 90% e nenhuma empresa entre 91% e 99,9%. Por fim, encontram-se as 4 empresas 100% eficientes (barra em cor verde).

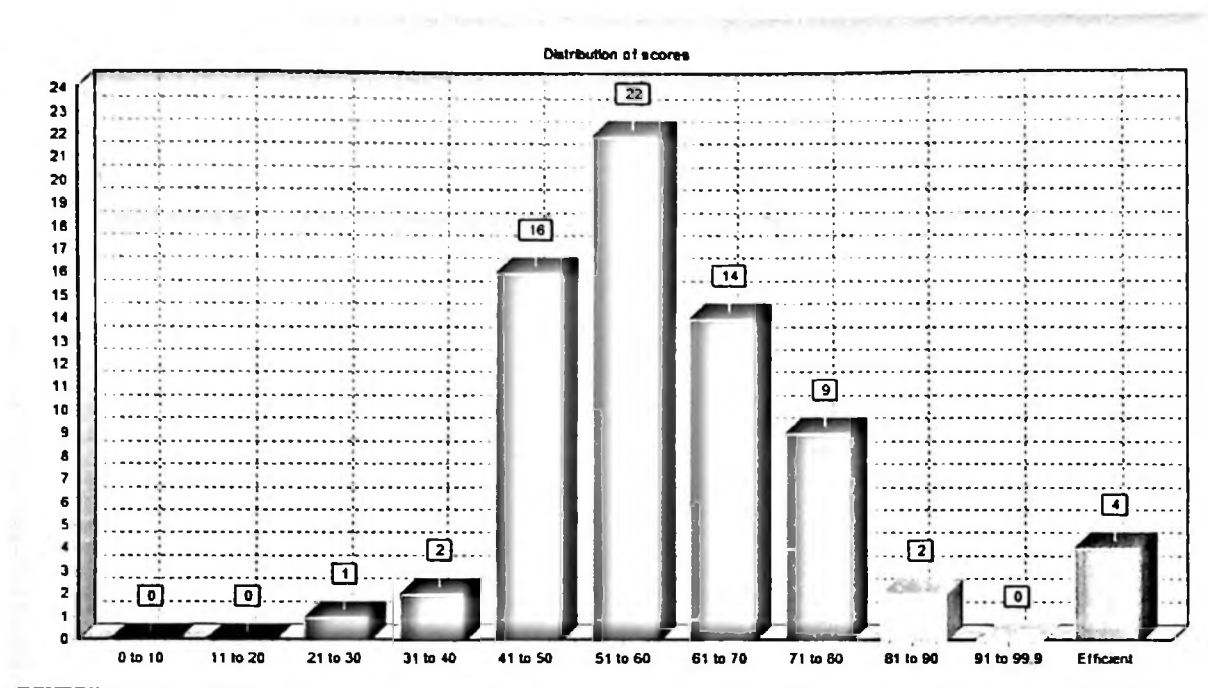


Figura 36: Distribuição dos escores do *cluster 6*

A figura 37 demonstra a quantidade de vezes que as empresas eficientes servem como referência para a atuação das empresas ineficientes.

Segundo a figura 37, a empresa Schincariol N-NE é a líder geral com 50 referências. A seguir, vem a Polietilenos com 44 referências. Em terceiro lugar está a Cisa Trading com 26 referências e aquela que apresenta menor número de referências é a CNEC, com 24. Vale destacar que as 4 empresas são 100% eficientes.

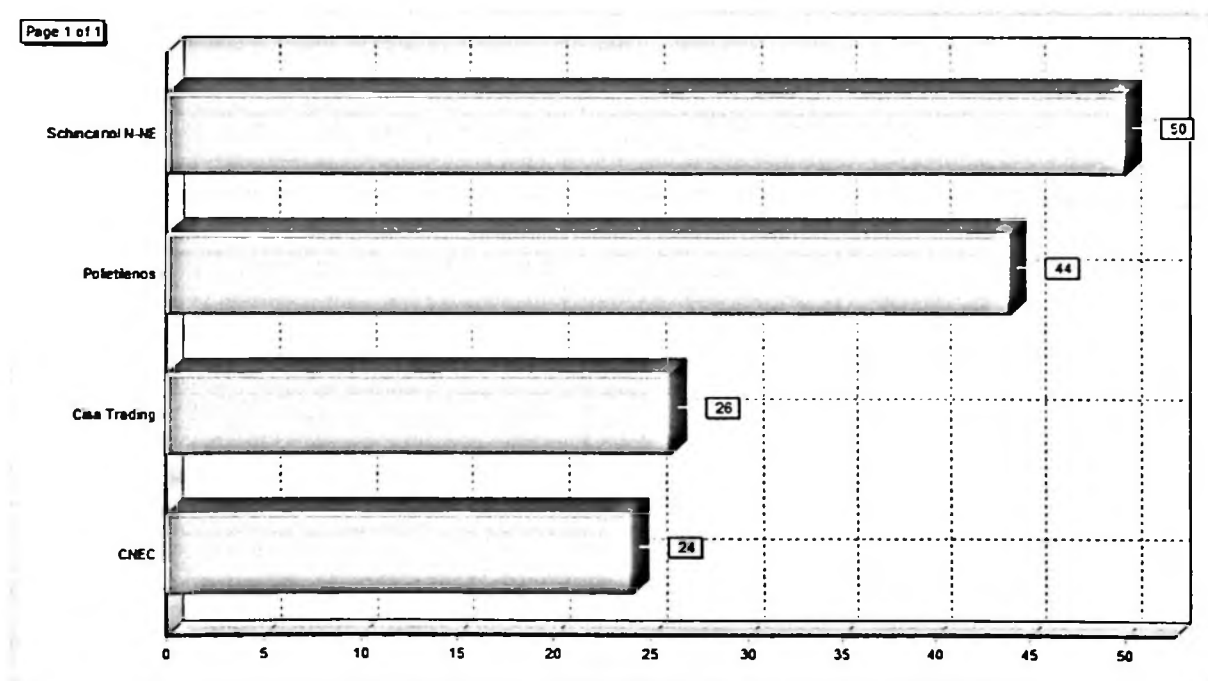


Figura 37: Empresas de referência do *cluster 6*

Uma breve visão das variáveis nas quais poderiam ser implementadas as melhorias indicadas pela análise é exposta na figura 38.

Conforme a figura 38, o *input* composição do endividamento aparece com o valor de zero por cento.

Entre os *outputs*, cabe destacar os indicadores: crescimento das vendas com melhoria potencial de 27,93% e liquidez corrente com indicação de 22,44% de melhoria potencial.

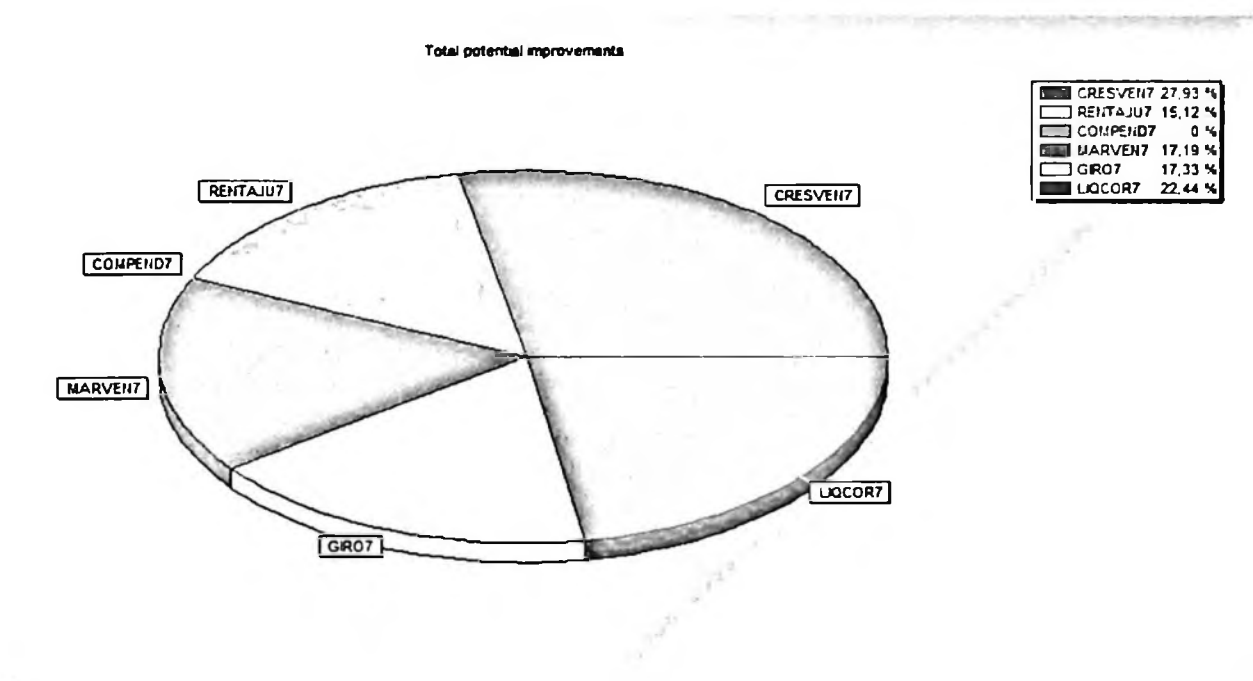


Figura 38: Melhorias potenciais para o *cluster* 6

3.16 Avaliação dos resultados do *cluster* 7

A tabela 10 expõe os escores, em ordem decrescente, referentes às 87 empresas componentes do *cluster* 7.

As empresas Eletronorte, Embasa, Eucatex, Innova, Melhoramentos Papéis e Virgolino de Oliveira apresentaram maiores escores no *cluster*, localizando-se na primeira coluna da tabela 10. Essas seis empresas são avaliadas como 100% eficientes e, assim, poderiam ser consideradas como referências para as demais que compõem o *cluster* 7. Elas estão na fronteira eficiente.

A empresa CBTU aparece na segunda coluna, última linha da Tabela 10, com o escore de 24,70%, ou seja, essa empresa é 24,70% eficiente. Cabe salientar que no *cluster* 7 há 17 empresas eficientes nos níveis que se situam abaixo de 50%.

Tabela 10: Escores das empresas componentes do *cluster* 7

Unit name	Score	Unit name	Score
Eletronorte	100,00%	Cauê	62,10%
Embasa	100,00%	Placas	61,90%
Eucatex	100,00%	Brasil Telecom	61,40%
Innova	100,00%	Unipar	61,40%
Melhoramentos Papéis	100,00%	Braskem	61,20%
Virgolino de Oliveira	100,00%	Escelsa	61,00%
Vigor	99,50%	Sanasa Campinas	60,40%
Usina Coruripe	93,00%	Santista Brasil	59,60%
Vicunha	92,40%	Telemar	59,50%
Caiuá	92,00%	Sifco	59,00%
Serpro	88,90%	Celpe	58,80%
Light Sesa	86,40%	AES Sul	58,10%
Zero Hora	85,40%	Comercial de Alimentos Carrefour	58,00%
Açucareira Zillo Lorenzetti	83,10%	Energipe	57,90%
Petroflex	83,10%	Ampla	57,50%
Usina Caeté	82,10%	AmBev	56,80%
Cesan	81,60%	Caesb	55,70%
CTBC Telecom	80,90%	Bandeirante	55,40%
AES Eletropaulo	79,60%	Copel DIS	55,30%
Barra Grande	79,00%	Saelpa	55,10%
Cesp	78,80%	Dixie Toga	51,90%
Celulose Irani	77,90%	Cecrisa	51,60%
Orsa	77,80%	Cagepa	51,30%
Hering	77,60%	Globo	51,30%
Schincariol-NE	77,20%	Saneago	51,10%
Rima	77,00%	ALL Brasil	50,90%
Amazônia Celular	76,80%	Ripasa	50,10%
Golden Cross	76,30%	Gás Natural São Paulo Sul	49,70%
Copersucar	75,50%	RGE	49,40%
Mundial	75,10%	CFLCL	47,10%
CTBC Celular	74,50%	Posigraf	46,70%
Caramuru Alimentos	73,80%	Embratel	46,60%
Cemat	70,80%	SPVias	46,30%
Veracel	70,40%	Oxiteno	45,90%
Adami	69,50%	PUC-PR	43,90%
Gontijo	66,50%	CCB	43,70%
SBT	66,50%	Santa Casa de São Paulo	41,90%
Celpa	66,30%	Santelisa	38,00%
Pão de Açúcar	66,10%	Valesul	38,00%
Dasa	65,10%	Ubea	35,00%
Ache	64,30%	Usina São João	33,50%
Tecban	63,10%	Usina Cerradinho	29,40%
Plascar	62,30%	Hospital Beneficência Portuguesa	26,80%
		CBTU	24,70%

Na figura 39, verifica-se a plotagem de todas as empresas do *cluster 7*, estando cada empresa representada por um ponto da Figura.

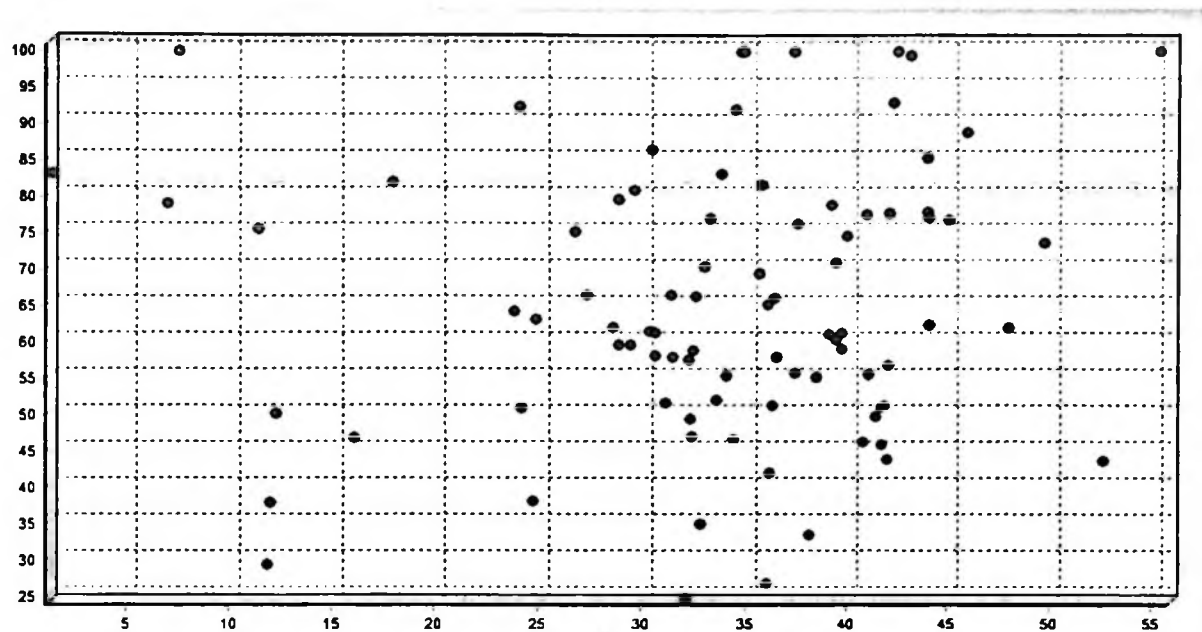


Figura 39: Plotagem das empresas do *cluster 7*

Na figura 39, constata-se que as seis empresas 100% eficientes surgem na parte superior e todas as outras estão localizadas abaixo delas. Isso pode ser analisado no eixo vertical cuja escala vai até 100%.

Em seguida, a figura 40 possibilita uma visão geral da distribuição dos escores. Analisando a figura 39, da esquerda para a direita, não se observa caso de empresas com nível de eficiência entre 0% a 20%. Há um total de 19 empresas entre os níveis de 21% a 50%. Representando a maior frequência, estão 20 empresas entre 51% e 60% de eficiência.

Apontadas como aproximadamente eficientes (barras em cor amarela) são as empresas com o nível de 81% até 99,9%. No *cluster 7*, encontram-se 11 empresas com essa classificação. Ainda verificam-se as 6 empresas 100% eficientes (barra em cor verde).

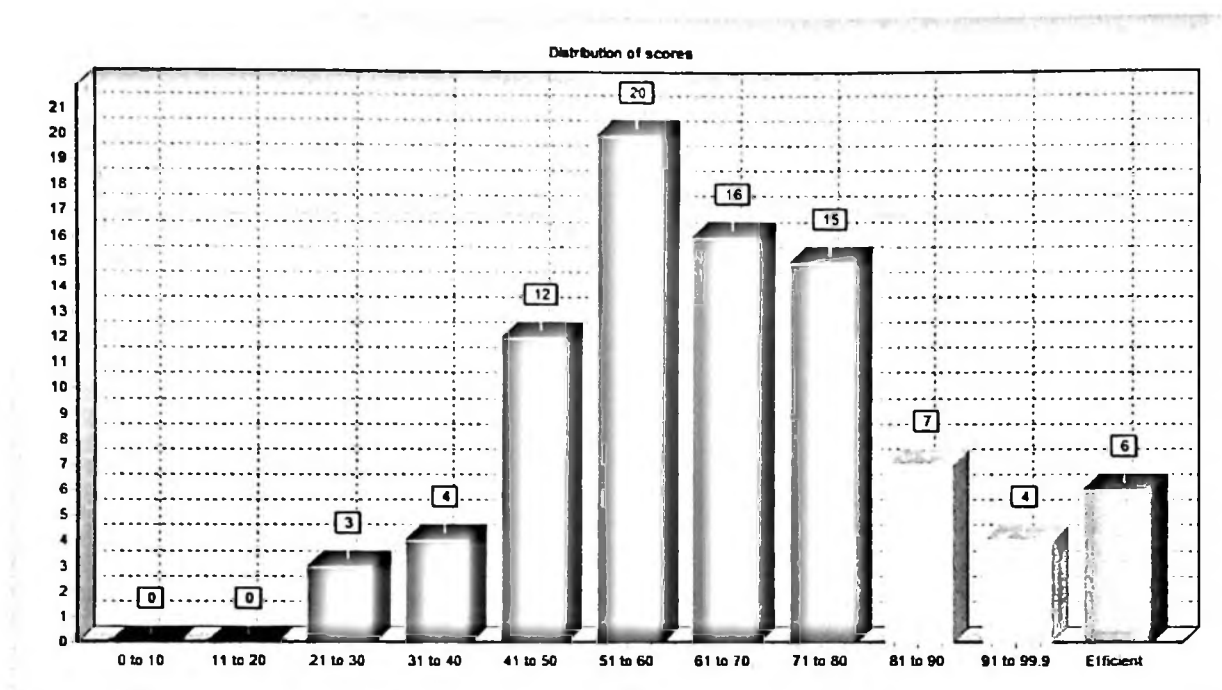


Figura 40: Distribuição dos escores do *cluster 7*

A figura 41 apresenta o número de vezes que as empresas eficientes aparecem como referência para a atuação das empresas ineficientes.

Conforme a figura 41, a empresa Eucatex é a líder geral com 78 referências. A partir disso, vêm as seguintes empresas e respectivos números de vezes em que foram referenciadas: Innova, 38; Melhoramentos Papéis, 32; Eletronorte, 24; Embasa, 22; Virgolino de Oliveira, 10. Vale ressaltar que as 6 empresas são 100% eficientes.

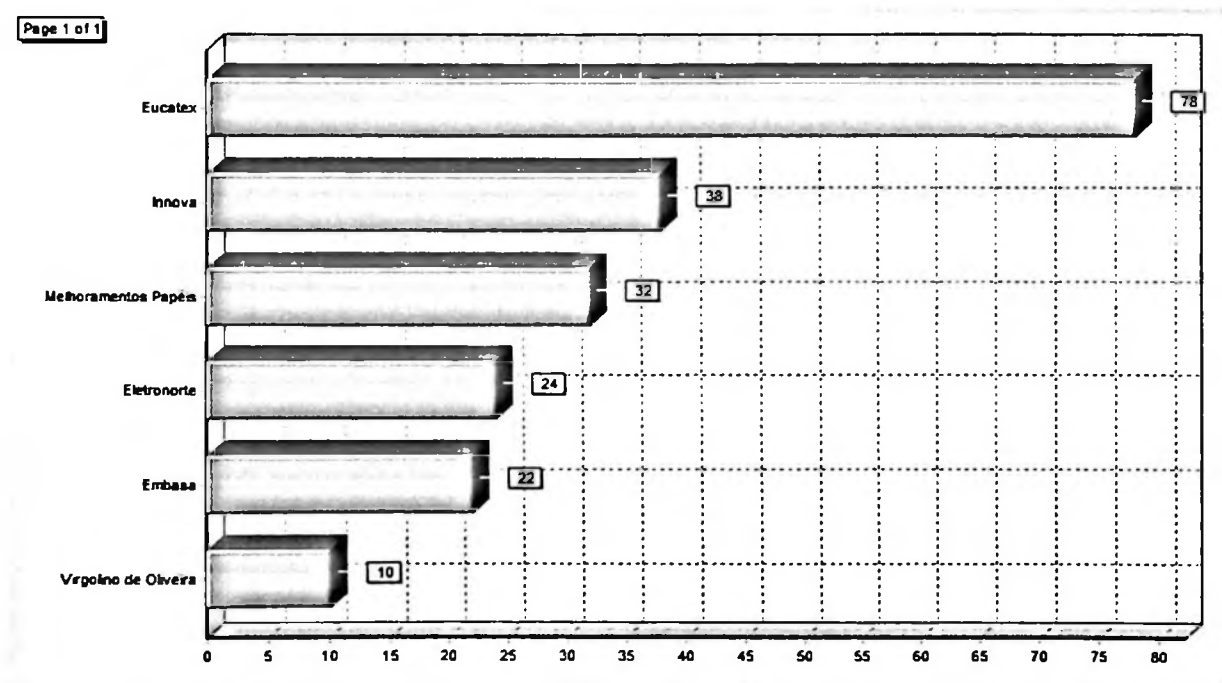


Figura 41: Empresas de referência do *cluster 7*

A figura 42 demonstra uma breve visão das variáveis nas quais poderiam ser implementadas as melhorias indicadas pela análise.

Como se percebe na figura 42, o *input* composição do endividamento aparece com o valor de zero%.

Entre os *outputs*, cabe destacar os indicadores: liquidez corrente com indicação de 28,07% de melhoria potencial e crescimento das vendas com melhoria potencial de 23,22%.

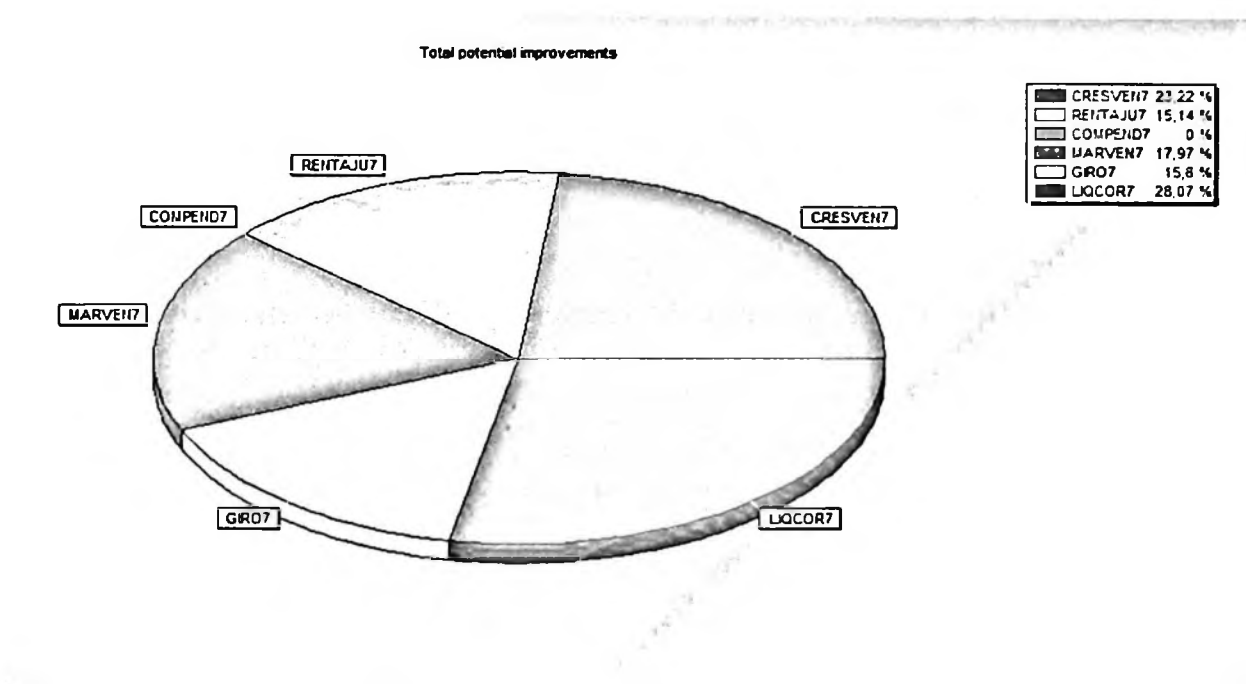


Figura 42: Melhorias potenciais para o *cluster* 7

3.17 Hipóteses a serem testadas

A técnica DEA *Data Envelopment Analysis* é não paramétrica e não se presta a testes de hipóteses. Com sua utilização, é possível determinar os níveis de eficiência das empresas componentes da amostra dentro de cada *cluster*. Nesta fase do trabalho, será empregada a técnica de regressão linear múltipla para realização dos testes de hipóteses, um dos objetivos deste estudo.

Inicialmente serão testadas as seguintes hipóteses:

H_0 : os indicadores obtidos através da técnica Delphi não têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA;

H_1 : os indicadores obtidos através da técnica Delphi têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA.

As variáveis a serem utilizadas para testar essas hipóteses são:

Variável dependente: escores de eficiência definidos para cada empresa dentro de cada *cluster*;

Variáveis independentes: os 6 indicadores definidos através da técnica Delphi.

3.18 Regressão e análise dos resultados do *cluster* 1: variáveis estabelecidas pela Delphi

Inicialmente, foram usados o programa SPSS e o método *stepwise* para fazer uma pré-seleção das variáveis do modelo. O método *stepwise* permite selecionar automaticamente as variáveis significativas com base no nível de significância desejado pelo pesquisador. Dessa forma, estabeleceu-se que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

Assim, a variável rentabilidade sobre o patrimônio líquido foi rejeitada com um nível de significância superior a 5%.

Desse modo, o programa *Eviews* foi utilizado para continuar com a análise de regressão. Esse programa tem opções que possibilitam efetuar automaticamente os testes dos pressupostos básicos da regressão.

A tabela 11 mostra os resultados da regressão em que a variável dependente é constituída pelos escores e as independentes são os indicadores estabelecidos pela Delphi, exceto a variável rentabilidade sobre o patrimônio líquido.

Tabela 11: Resultados da regressão para o cluster 1

Dependent Variable: SCORE

Method: Least Squares

Date: 12/18/08 Time: 17:28

Sample: 1 148

Included observations: 148

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEN7	-0.008150	0.000318	-25.61578	0.0000
CRESVEN7	0.001581	0.000533	2.969144	0.0035
GIRO7	0.093647	0.006583	14.22540	0.0000
LIQCOR7	0.030963	0.007884	3.927380	0.0001
MARVEN7	0.016907	0.002554	6.618816	0.0000
C	0.693217	0.052732	13.14604	0.0000
R-squared	0.851555	Mean dependent var		0.656054
Adjusted R-squared	0.846328	S.D. dependent var		0.134980
S.E. of regression	0.052914	Akaike info criterion		-3.000612
Sum squared resid	0.397580	Schwarz criterion		-2.879104
Log likelihood	228.0453	F-statistic		162.9166
Durbin-Watson stat	1.786658	Prob(F-statistic)		0.000000

Ao analisar os resultados, verifica-se que as cinco variáveis independentes incluídas no modelo são significativas em um nível de significância menor que 1%. Pode-se constatar também que os valores Prob para as cinco variáveis são todos menores que 0,01.

Os sinais dos coeficientes angulares de todas as variáveis são coerentes. A variável composição do endividamento tem sinal negativo, o que é coerente, pois o aumento da composição do endividamento, isto é, o aumento das dívidas de curto prazo, acarreta redução no nível de eficiência.

O modelo de regressão como um todo é significativo (*Prob(F-statistic)*) e menor que 0,01. O poder explicativo da regressão (*Adjusted R-squared*) é de 0,846, podendo ser considerado bastante razoável.

A figura 43 exibe os resultados do teste de normalidade dos resíduos, no qual foi empregado o teste Jarque-Bera. O valor para a Probability = 0,075 sendo maior que 0,05 indica que não há evidências para a rejeição da hipótese de normalidade dos resíduos.

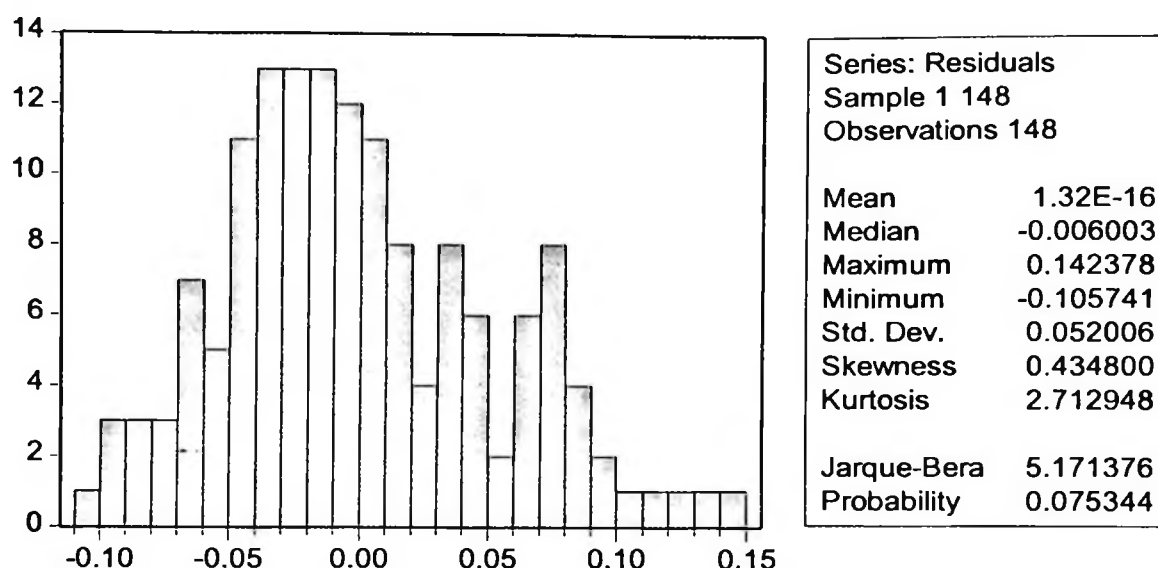


Figura 43: Teste da normalidade dos resíduos da regressão para o *cluster 1*

Na tabela 12, o teste Breusch_Godfrey para correlação serial indica Probability = 0,37, significando que não se pode rejeitar a hipótese de ausência de correlação serial.

Tabela 12: Teste de correlação serial dos resíduos da regressão para o *cluster 1*

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.991594	Probability	0.373575
Obs*R-squared	2.067229	Probability	0.355719

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/18/08 Time: 17:31

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEN7	-8.75E-06	0.000325	-0.026899	0.9786
CRESVEN7	-5.36E-05	0.000535	-0.100181	0.9203
GIRO7	0.000218	0.006689	0.032635	0.9740
LIQCOR7	0.000396	0.007891	0.050191	0.9600
MARVEN7	-6.62E-05	0.002573	-0.025737	0.9795
C	0.002722	0.052779	0.051574	0.9589
RESID(-1)	0.112454	0.085601	1.313696	0.1911
RESID(-2)	-0.051972	0.086792	-0.598812	0.5503
R-squared	0.013968	Mean dependent var	1.32E-16	
Adjusted R-squared	-0.035334	S.D. dependent var	0.052006	
S.E. of regression	0.052917	Akaike info criterion	-2.987652	
Sum squared resid	0.392027	Schwarz criterion	-2.825640	
Log likelihood	229.0862	F-statistic	0.283313	
Durbin-Watson stat	2.007050	Prob(F-statistic)	0.959607	

A tabela 13 mostra os resultados do teste de heterocedasticidade de White. O valor Probability = 0,000001 sugere a existência de evidências para se rejeitar a hipótese de homocedasticidade dos resíduos, em um nível de significância menor que 0,01.

Tabela 13: Teste de heterocedasticidade de White para os resíduos do *cluster* 1

White Heteroskedasticity Test:

F-statistic	3.860196	Probability	0.000001
Obs*R-squared	55.95476	Probability	0.000030

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 12/18/08 Time: 17:32

Sample: 1 148

Included observations: 148

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.066118	0.028351	2.332100	0.0213
COMPEND7	-0.000523	0.000256	-2.039953	0.0434
COMPEND7^2	1.77E-06	1.24E-06	1.427378	0.1559
COMPEND7*CRESVEN7	5.61E-07	2.40E-06	0.233802	0.8155
COMPEND7*GIRO7	-1.85E-05	3.13E-05	-0.589644	0.5565
COMPEND7*LIQCOR7	7.25E-05	3.86E-05	1.876294	0.0629
COMPEND7*MARVEN7	8.79E-06	1.11E-05	0.793951	0.4287
CRESVEN7	-0.000378	0.000480	-0.787291	0.4326
CRESVEN7^2	3.30E-06	2.41E-06	1.368991	0.1734
CRESVEN7*GIRO7	1.56E-05	5.75E-05	0.270938	0.7869
CRESVEN7*LIQCOR7	0.000135	5.56E-05	2.432254	0.0164
CRESVEN7*MARVEN7	-1.48E-05	2.20E-05	-0.673659	0.5018
GIRO7	-0.000271	0.005953	-0.045601	0.9637
GIRO7^2	0.000725	0.000455	1.591012	0.1141
GIRO7*LIQCOR7	-0.000294	0.000778	-0.377894	0.7061
GIRO7*MARVEN7	-0.000168	0.000268	-0.626488	0.5321
LIQCOR7	-0.007294	0.006665	-1.094365	0.2759
LIQCOR7^2	0.001206	0.000643	1.874440	0.0632
LIQCOR7*MARVEN7	-0.000486	0.000309	-1.570156	0.1189
MARVEN7	-0.004197	0.002527	-1.660399	0.0993
MARVEN7^2	0.000190	6.79E-05	2.789104	0.0061
R-squared	0.378073	Mean dependent var	0.002686	
Adjusted R-squared	0.280131	S.D. dependent var	0.003528	
S.E. of regression	0.002993	Akaike info criterion	-8.654196	
Sum squared resid	0.001138	Schwarz criterion	-8.228916	
Log likelihood	661.4105	F-statistic	3.860196	
Durbin-Watson stat	1.802300	Prob(F-statistic)	0.000001	

Devido ao problema detectado quanto à homocedasticidade dos resíduos, recorreu-se à correção, conforme processo de White. Comparando os resultados obtidos a partir desse procedimento com o da regressão da Tabela 14, verifica-se que os erros padrões dos coeficientes angulares aumentaram e os valores dos testes t reduziram-se, mas, mesmo assim, os valores de Prob relativos a todas as variáveis são inferiores a 0,01, indicando que todas as variáveis são significativas.

Tabela 14: Processo de White para heterocedasticidade referente ao *cluster 1*

Dependent Variable: SCORE

Method: Least Squares

Date: 12/18/08 Time: 17:33

Sample: 1 148

Included observations: 148

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEN7	-0.008150	0.000356	-22.91896	0.0000
CRESVEN7	0.001581	0.000601	2.632920	0.0094
GIRO7	0.093647	0.006730	13.91390	0.0000
LIQCOR7	0.030963	0.009307	3.326909	0.0011
MARVEN7	0.016907	0.003234	5.228524	0.0000
C	0.693217	0.061948	11.19033	0.0000
R-squared	0.851555	Mean dependent var		0.656054
Adjusted R-squared	0.846328	S.D. dependent var		0.134980
S.E. of regression	0.052914	Akaike info criterion		-3.000612
Sum squared resid	0.397580	Schwarz criterion		-2.879104
Log likelihood	228.0453	F-statistic		162.9166
Durbin-Watson stat	1.786658	Prob(F-statistic)		0.000000

Comparando os resultados obtidos, é possível obter as seguintes evidências:

Rejeita-se H_0 , de que os indicadores obtidos através da técnica Delphi não tem influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA. Portanto, todas as variáveis obtidas pelo modelo Delphi, com exceção da variável retorno sobre o patrimônio líquido são significativas em um nível de significância inferior a 0,01.

3.19 Regressão e análise dos resultados do *cluster 2*: variáveis estabelecidas pela Delphi

O programa SPSS e o método *stepwise* foram empregados, primeiramente, para fazer uma pré-seleção das variáveis do modelo. Conforme já afirmado, o método *stepwise* aceita uma seleção automática de variáveis significativas com base no nível de significância desejado pelo pesquisador. Foi estabelecido que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

Desse modo, as variáveis crescimento das vendas e liquidez corrente foram rejeitadas, considerando-se um nível de significância superior a 5%.

A seguir, empregou-se o programa *Eviews* para continuar com a análise de regressão, pois o mesmo tem opções que possibilitam efetuar automaticamente os testes dos pressupostos básicos da regressão.

A tabela 15 expõe os resultados da regressão em que a variável dependente é constituída pelos escores e as independentes são os indicadores estabelecidos pela Delphi, exceto as variáveis crescimento das vendas e liquidez corrente.

Tabela 15: Resultados da regressão para o cluster 2

Dependent Variable: SCORE

Method: Least Squares

Date: 12/18/08 Time: 17:40

Sample: 1 72

Included observations: 72

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.009032	0.000532	-16.97266	0.0000
MARVEN7	0.020056	0.004800	4.178637	0.0001
GIRO7	0.103466	0.016977	6.094499	0.0000
RENTAJU7	0.005288	0.002160	2.447542	0.0170
C	0.868584	0.038375	22.63425	0.0000
R-squared	0.822059	Mean dependent var		0.674986
Adjusted R-squared	0.811436	S.D. dependent var		0.154073
S.E. of regression	0.066905	Akaike info criterion		-2.504185
Sum squared resid	0.299907	Schwarz criterion		-2.346083
Log likelihood	95.15064	F-statistic		77.38245
Durbin-Watson stat	2.107619	Prob(F-statistic)		0.000000

Analisando os resultados, constata-se que as quatro variáveis independentes incluídas no modelo são significativas. Três delas, sendo a composição do endividamento, margem líquida e giro do ativo, constituem-se significativas a um nível de significância menor que 1%. Pode-se verificar que os valores Prob para essas três variáveis são todos menores que 0,01. A variável rentabilidade sobre o patrimônio líquido é significativa ao nível de 1,7% e é possível avaliar que o valor Prob para essa variável é de 0,017.

Os sinais dos coeficientes angulares de todas as variáveis são coerentes. A variável composição do endividamento tem sinal negativo e isso faz sentido, pois o aumento da composição do endividamento, ou seja, o aumento das dívidas de curto prazo acarreta redução no nível de eficiência.

O modelo de regressão como um todo é significativo (Prob(F-statistic) e menor que 0,01. O poder explicativo da regressão (*Adjusted R-squared*) é de 0,811, considerado bastante razoável.

A figura 44 demonstra os resultados do teste de normalidade dos resíduos, a partir da utilização do teste Jarque-Bera. O valor para a *Probability* = 0,0901, sendo maior que 0,05, indica não haver evidências para a rejeição da hipótese de normalidade dos resíduos.

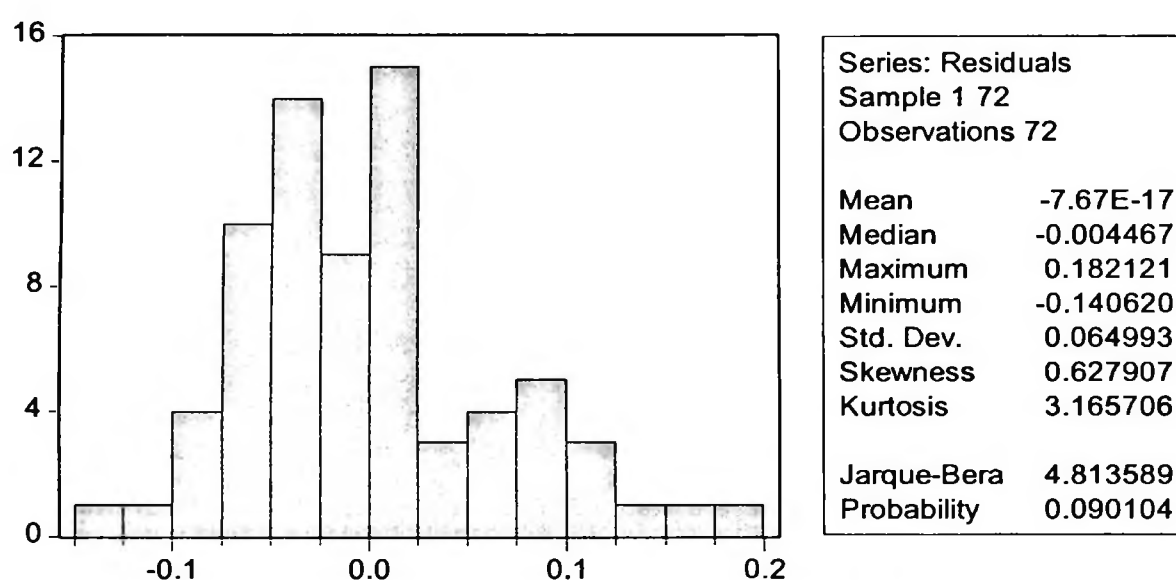


Figura 44: Teste da normalidade dos resíduos da regressão para o *cluster 2*

O teste Breusch_Godfrey para a correlação serial, presente na tabela 16, sugere *Probability* = 0,694, expressando que não se pode rejeitar a hipótese de ausência de correlação serial.

Tabela 16: Teste de correlação serial dos resíduos da regressão para o *cluster 2*

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.366772	Probability	0.694393
Obs*R-squared	0.803473	Probability	0.669157

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/18/08 Time: 17:45

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-4.74E-05	0.000542	-0.087467	0.9306
MARVEN7	0.000761	0.004951	0.153771	0.8783
GIRO7	-0.000298	0.017189	-0.017335	0.9862
RENTAJU7	-0.000290	0.002208	-0.131183	0.8960
C	0.001708	0.038848	0.043965	0.9651
RESID(-1)	-0.065526	0.127472	-0.514041	0.6090
RESID(-2)	0.081814	0.128427	0.637046	0.5263
R-squared	0.011159	Mean dependent var	-7.67E-17	
Adjusted R-squared	-0.080118	S.D. dependent var	0.064993	
S.E. of regression	0.067546	Akaike info criterion	-2.459851	
Sum squared resid	0.296560	Schwarz criterion	-2.238509	
Log likelihood	95.55464	F-statistic	0.122257	
Durbin-Watson stat	1.991713	Prob(F-statistic)	0.993308	

A tabela 17 expõe os resultados do teste de heterocedasticidade de White. O valor *Probability* = 0,00000 indica a presença de evidências para se rejeitar a hipótese de homocedasticidade dos resíduos, a um nível de significância menor que 0,01.

Tabela 17: Teste de heterocedasticidade de White para os resíduos do *cluster 2*

White Heteroskedasticity Test:

F-statistic	14.05923	Probability	0.000000
Obs*R-squared	55.83165	Probability	0.000001

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 12/18/08 Time: 17:46

Sample: 1 72

Included observations: 72

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.051413	0.009256	5.554606	0.0000
COMPEN7	-0.000565	0.000190	-2.970472	0.0043
COMPEN7^2	1.10E-06	1.92E-06	0.575771	0.5670
COMPEN7*MARVEN7	-5.62E-05	2.64E-05	-2.128968	0.0376
COMPEN7*GIRO7	0.000125	6.46E-05	1.940550	0.0573
COMPEN7*RENTAJU7	3.61E-05	1.01E-05	3.569530	0.0007
MARVEN7	0.003281	0.001439	2.280771	0.0263
MARVEN7^2	5.43E-05	0.000109	0.498190	0.6203
MARVEN7*GIRO7	0.001873	0.000978	1.915243	0.0605
MARVEN7*RENTAJU7	-4.67E-05	0.000114	-0.407653	0.6851
GIRO7	-0.029486	0.008755	-3.367693	0.0014
GIRO7^2	0.004092	0.001871	2.186837	0.0329
GIRO7*RENTAJU7	-0.000614	0.000362	-1.696915	0.0952
RENTAJU7	-0.002507	0.000537	-4.668634	0.0000
RENTAJU7^2	1.23E-05	2.80E-05	0.438358	0.6628
R-squared	0.775440	Mean dependent var	0.004165	
Adjusted R-squared	0.720284	S.D. dependent var	0.006173	
S.E. of regression	0.003265	Akaike info criterion	-8.428221	
Sum squared resid	0.000608	Schwarz criterion	-7.953916	
Log likelihood	318.4160	F-statistic	14.05923	
Durbin-Watson stat	1.930495	Prob(F-statistic)	0.000000	

Em função do problema detectado quanto à homocedasticidade dos resíduos, fez-se a correção conforme o processo de White. Comparando os resultados obtidos nesse procedimento da Tabela 18 com o processo da regressão inicial (tabela 15), verifica-se que os erros padrões dos coeficientes angulares aumentam e os valores dos testes t diminuem, embora, mesmo assim, os valores de Prob relativos a três variáveis – composição do endividamento, margem líquida e giro do ativo – sejam inferiores a 0,01, demonstrando sua significação.

Entretanto, após o processo de White, a variável rentabilidade sobre o patrimônio líquido apresentou um Prob = 0,0608 (ou 6,08%), indicando que essa variável não é significativa no nível de significância de 5%.

Tabela 18: Processo de White para heterocedasticidade referente ao *cluster 2*

Dependent Variable: SCORE

Method: Least Squares

Date: 12/18/08 Time: 17:47

Sample: 1 72

Included observations: 72

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.009032	0.000540	-16.72165	0.0000
MARVEN7	0.020056	0.005766	3.478317	0.0009
GIRO7	0.103466	0.024475	4.227418	0.0001
RENTAJU7	0.005288	0.002773	1.906810	0.0608
C	0.868584	0.056743	15.30742	0.0000
R-squared	0.822059	Mean dependent var		0.674986
Adjusted R-squared	0.811436	S.D. dependent var		0.154073
S.E. of regression	0.066905	Akaike info criterion		-2.504185
Sum squared resid	0.299907	Schwarz criterion		-2.346083
Log likelihood	95.15064	F-statistic		77.38245
Durbin-Watson stat	2.107619	Prob(F-statistic)		0.000000

Comparando os resultados obtidos, é possível obter as seguintes evidências:

Rejeita-se H_0 , de que três indicadores obtidos através da técnica Delphi não têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA. Portanto, os indicadores – composição do endividamento, margem líquida e giro do ativo – obtidos pelo modelo Delphi são significativos em um nível de significância inferior a 0,01.

Já a variável rentabilidade sobre o patrimônio não se mostrou significativa nem mesmo no nível de 5% após a aplicação do processo de White.

Finalmente, não há evidências para se rejeitar a hipótese nula de que variável liquidez corrente não é significativa, considerando o nível de significância de 5%.

3.20 Regressão e análise dos resultados do *cluster 3*: variáveis estabelecidas pela Delphi

Inicialmente, da mesma forma como foram analisados os resultados do *cluster 2*, para a avaliação dos dados finais do *cluster 3*, foram empregados o programa SPSS e o método *stepwise* para fazer uma pré-seleção das variáveis do modelo. Também, para a análise desse

cluster, foi estabelecido que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

Foi constatado que as variáveis crescimento das vendas e rentabilidade sobre o patrimônio líquido foram rejeitadas, considerando-se um nível de significância superior a 5%.

A seguir, utilizou-se o programa *Eviews* para continuar com a análise de regressão. Esse programa tem opções que possibilitam efetuar automaticamente os testes dos pressupostos básicos da regressão.

A tabela 19 apresenta os resultados da regressão em que a variável dependente é constituída pelos escores e as independentes são os indicadores estabelecidos pela Delphi, exceto as variáveis crescimento das vendas e rentabilidade sobre o patrimônio líquido.

Tabela 19: Resultados da regressão para o *cluster* 3

Dependent Variable: SCORE
Method: Least Squares
Date: 12/18/08 Time: 17:58
Sample: 1 74
Included observations: 74

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.009471	0.000505	-18.76615	0.0000
GIRO7	0.135267	0.019927	6.787977	0.0000
MARVEN7	0.009355	0.002375	3.938089	0.0002
LIQCOR7	0.067647	0.019151	3.532309	0.0007
C	0.481680	0.078220	6.158014	0.0000
R-squared	0.849544	Mean dependent var		0.484919
Adjusted R-squared	0.840822	S.D. dependent var		0.204940
S.E. of regression	0.081765	Akaike info criterion		-2.104758
Sum squared resid	0.461301	Schwarz criterion		-1.949078
Log likelihood	82.87603	F-statistic		97.40157
Durbin-Watson stat	2.331943	Prob(F-statistic)		0.000000

Na verificação dos resultados, as quatro variáveis independentes inclusas no modelo são significativas em um nível de significância menor que 1%. Pode-se também perceber que os valores Prob para essas três variáveis são todos menores que 0,01.

Os sinais dos coeficientes angulares de todas as variáveis são coerentes, como a variável composição do endividamento, que tem sinal negativo, pois o aumento da composição do

endividamento, ou seja, o aumento das dívidas de curto prazo, acarreta redução no nível de eficiência.

O modelo de regressão como um todo é significativo ($Prob(F\text{-statistic})$) e menor que 0,01. O poder explicativo da regressão ($Adjusted R\text{-squared}$), de 0,8408, pode ser considerado bastante razoável.

A figura 45 mostra os resultados do teste de normalidade dos resíduos. Foi usado o teste Jarque-Bera, com o valor para a $Probability = 0,027$ sendo menor que 0,05, o que sugere a existência de evidências para a rejeição da hipótese de normalidade dos resíduos.

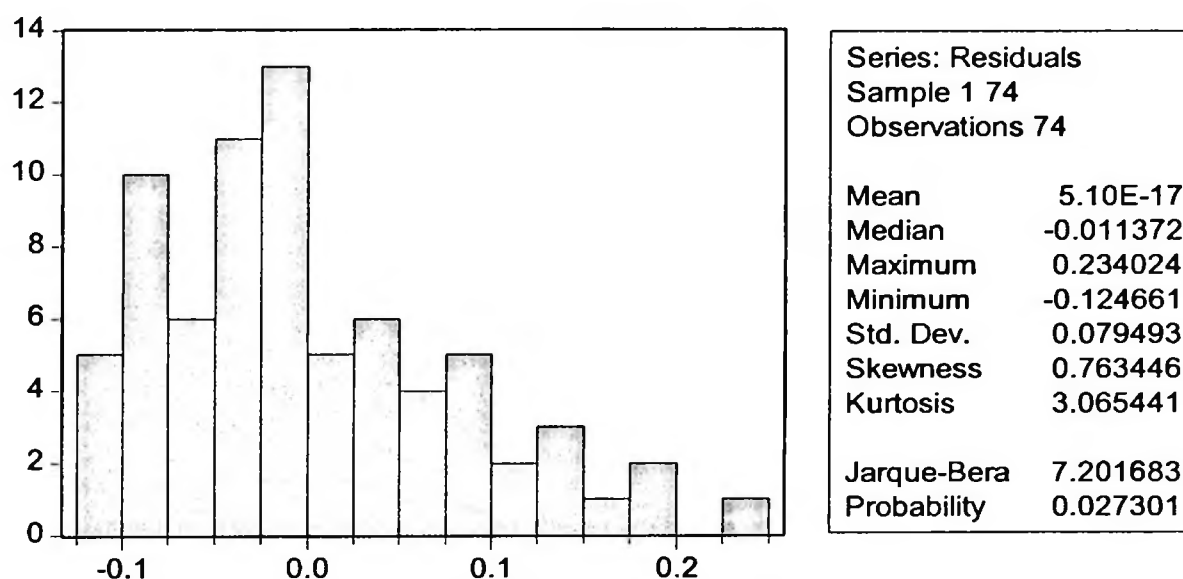


Figura 45: Teste da normalidade dos resíduos da regressão para o cluster 3

Aproveitando o uso do programa SPSS, foram efetuados outros testes de normalidade dos resíduos, como Kolmogorov-Smirnov e Shapiro- Wilk. Os resultados estão na Tabela 20.

Tabela 20: Teste da normalidade dos resíduos da regressão para o *cluster 3*

Tests of Normality						
	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	0,109	74	0,031	0,949	74	0,005
a. Lilliefors Significance Correction						

O teste Kolmogorov-Smirnov apresenta um nível de significância Sig = 0,031 e o teste Shapiro- Wilk um nível Sig = 0,005. Portanto, ambos confirmam a rejeição da hipótese de normalidade dos resíduos em um nível de significância menor do que 0,05.

O teste Breusch_Godfrey de correlação serial da tabela 21 indica *Probability* = 0,2418, implicando que não se pode rejeitar a hipótese de ausência de correlação serial.

Tabela 21: Teste de correlação serial dos resíduos da regressão para o *cluster 3*

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.449779	Probability	0.241887
Obs*R-squared	3.069652	Probability	0.215493

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/18/08 Time: 18:00

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	0.000176	0.000513	0.343367	0.7324
GIRO7	0.001861	0.019837	0.093792	0.9256
MARVEN7	2.70E-05	0.002360	0.011428	0.9909
LIQCOR7	1.15E-06	0.019210	5.99E-05	1.0000
C	-0.012517	0.078534	-0.159390	0.8738
RESID(-1)	-0.198731	0.124716	-1.593468	0.1158
RESID(-2)	-0.110972	0.124399	-0.892063	0.3756
R-squared	0.041482	Mean dependent var	5.10E-17	
Adjusted R-squared	-0.044356	S.D. dependent var	0.079493	
S.E. of regression	0.081237	Akaike info criterion	-2.093070	
Sum squared resid	0.442165	Schwarz criterion	-1.875118	
Log likelihood	84.44360	F-statistic	0.483260	
Durbin-Watson stat	1.981226	Prob(F-statistic)	0.818548	

Os resultados do teste de heterocedasticidade de White podem ser visualizados na tabela 22. O valor *Probability* = 0,00003 indica que há evidências para se rejeitar a hipótese de homocedasticidade dos resíduos, em um nível de significância menor que 0,01.

Tabela 22: Teste de heterocedasticidade de White para os resíduos do *cluster* 3

White Heteroskedasticity Test:

F-statistic	4.357059	Probability	0.000030
Obs*R-squared	37.61631	Probability	0.000595

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 12/18/08 Time: 18:01

Sample: 1 74

Included observations: 74

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.002789	0.048842	-0.057109	0.9547
COMPEND7	-0.001438	0.000510	-2.820506	0.0065
COMPEND7^2	1.18E-05	2.49E-06	4.762722	0.0000
COMPEND7*GIRO7	-3.84E-05	0.000134	-0.285797	0.7760
COMPEND7*MARVEN7	2.49E-06	1.20E-05	0.206618	0.8370
COMPEND7*LIQCOR7	-1.14E-06	0.000109	-0.010404	0.9917
GIRO7	0.001817	0.013880	0.130897	0.8963
GIRO7^2	-0.001887	0.003598	-0.524586	0.6018
GIRO7*MARVEN7	-0.000356	0.000509	-0.699282	0.4871
GIRO7*LIQCOR7	0.003615	0.004097	0.882251	0.3812
MARVEN7	0.004517	0.002260	1.998604	0.0503
MARVEN7^2	-9.94E-05	5.10E-05	-1.948112	0.0562
MARVEN7*LIQCOR7	-0.000119	0.000510	-0.233148	0.8165
LIQCOR7	0.002140	0.026141	0.081872	0.9350
LIQCOR7^2	-0.000439	0.003678	-0.119405	0.9054
R-squared	0.508329	Mean dependent var	0.006234	
Adjusted R-squared	0.391661	S.D. dependent var	0.009020	
S.E. of regression	0.007035	Akaike info criterion	-6.896858	
Sum squared resid	0.002920	Schwarz criterion	-6.429818	
Log likelihood	270.1837	F-statistic	4.357059	
Durbin-Watson stat	1.764799	Prob(F-statistic)	0.000030	

Devido ao problema detectado quanto à homocedasticidade dos resíduos, efetuou-se a correção conforme o processo de White. Comparando os resultados obtidos nesse procedimento da tabela 23 com o da regressão da tabela 19, constata-se que os erros padrões dos coeficientes angulares aumentam e os valores dos testes t diminuem, mas, mesmo assim, os valores de Prob relativos às quatro variáveis continuam inferiores a 0,01, confirmando que essas variáveis são significativas.

Tabela 23: Processo de White para heterocedasticidade referente ao *cluster* 3

Dependent Variable: SCORE

Method: Least Squares

Date: 12/18/08 Time: 18:02

Sample: 1 74

Included observations: 74

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.009471	0.000659	-14.38109	0.0000
GIRO7	0.135267	0.015004	9.015434	0.0000
MARVEN7	0.009355	0.002144	4.362580	0.0000
LIQCOR7	0.067647	0.018961	3.567642	0.0007
C	0.481680	0.061837	7.789516	0.0000
R-squared	0.849544	Mean dependent var		0.484919
Adjusted R-squared	0.840822	S.D. dependent var		0.204940
S.E. of regression	0.081765	Akaike info criterion		-2.104758
Sum squared resid	0.461301	Schwarz criterion		-1.949078
Log likelihood	82.87603	F-statistic		97.40157
Durbin-Watson stat	2.331943	Prob(F-statistic)		0.000000

Comparando os resultados obtidos, é possível obter as seguintes evidências:

Rejeita-se H_0 , de que quatro indicadores obtidos através da técnica Delphi não têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA. Portanto, as variáveis composição do endividamento, giro do ativo, margem líquida e liquidez corrente, obtidas pelo modelo Delphi são significativas em um nível de significância inferior a 0,01.

Entretanto, cabe salientar que foram encontradas evidências para rejeitar a hipótese de da normalidade dos resíduos em um nível de significância menor que 5%.

3.20.1 *Cluster* 3: Mais testes para a normalidade dos resíduos da regressão

Ao se avaliar a regressão referente ao *cluster* 3, foi feita uma ressalva quanto à normalidade dos resíduos da regressão. A hipótese nula de normalidade dos resíduos foi rejeitada.

Assim, procurou-se examinar melhor o assunto. Foi obtida nova regressão com a inclusão de todas as variáveis que haviam sido selecionadas pela técnica Delphi. Os resultados estão na tabela 24. Através destes, verifica-se que as variáveis crescimento das vendas e rentabilidade sobre o patrimônio líquido não são significativas, conforme já explanado anteriormente. O poder explicativo da regressão é de Adjusted *R-squared* = 0,839.

Tabela 24: Cluster 3: regressão com todas as variáveis da técnica DELPHI

Dependent Variable: SCORE

Method: Least Squares

Date: 01/18/09 Time: 10:17

Sample: 1 74

Included observations: 74

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.509589	0.085431	5.964918	0.0000
COMPEND7	-0.009658	0.000546	-17.67365	0.0000
CRESVEN7	-0.001061	0.000899	-1.180389	0.2420
GIRO7	0.142063	0.031341	4.532850	0.0000
LIQCOR7	0.069095	0.019697	3.507891	0.0008
MARVEN7	0.010236	0.003982	2.570893	0.0124
RENTAJU7	-0.000142	0.002526	-0.056055	0.9555
R-squared	0.852626	Mean dependent var	0.484919	
Adjusted R-squared	0.839429	S.D. dependent var	0.204940	
S.E. of regression	0.082122	Akaike info criterion	-2.071402	
Sum squared resid	0.451851	Schwarz criterion	-1.853449	
Log likelihood	83.64186	F-statistic	64.60441	
Durbin-Watson stat	2.350149	Prob(F-statistic)	0.000000	

A figura 46 apresenta teste Jarque-Bera para normalidade dos resíduos. Como o valor $Probability = 0,0688$ é maior que 0,05, não se pode rejeitar a hipótese da normalidade dos resíduos.

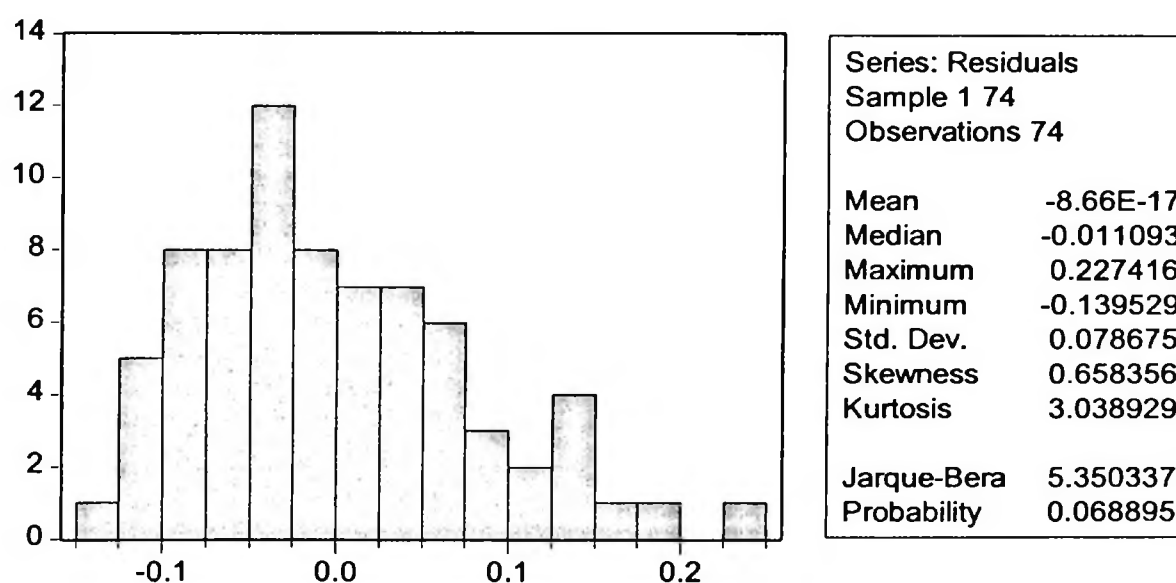


Figura 46: Normalidade dos resíduos da regressão

A tabela 25 expõe o teste de correlação serial de Breusch-Godfrey . Como o valor *Probability* = 0,21 é maior que 0,05, não se pode rejeitar a hipótese nula de ausência de correlação serial.

Tabela 25: Teste de correlação serial

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.579615	Probability	0.213861
Obs*R-squared	3.429954	Probability	0.179968

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 01/18/09 Time: 10:18

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.021687	0.086243	-0.251461	0.8023
COMPEN7	0.000226	0.000558	0.404763	0.6870
CRESVEN7	6.65E-05	0.000892	0.074569	0.9408
GIRO7	-0.000579	0.031169	-0.018590	0.9852
LIQCOR7	0.000715	0.019751	0.036190	0.9712
MARVEN7	-0.000335	0.003958	-0.084587	0.9329
RENTAJU7	0.000326	0.002519	0.129420	0.8974
RESID(-1)	-0.217688	0.128152	-1.698668	0.0942
RESID(-2)	-0.110932	0.127398	-0.870745	0.3871
R-squared	0.046351	Mean dependent var	-8.66E-17	
Adjusted R-squared	-0.071021	S.D. dependent var	0.078675	
S.E. of regression	0.081421	Akaike info criterion	-2.064807	
Sum squared resid	0.430907	Schwarz criterion	-1.784583	
Log likelihood	85.39785	F-statistic	0.394904	
Durbin-Watson stat	1.978442	Prob(F-statistic)	0.919460	

A tabela 26 demonstra o teste de White para a heterocedasticidade dos resíduos. Como o valor *probability* = 0,00068 é menor que 0,05, rejeita-se a hipótese nula de homocedasticidade dos resíduos.

Tabela 26: Teste de heterocedasticidade dos resíduos

White Heteroskedasticity Test:

F-statistic	2.914333	Probability	0.000684
Obs*R-squared	46.69964	Probability	0.010685

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 01/18/09 Time: 10:19

Sample: 1 74

Included observations: 74

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.074660	0.073089	1.021494	0.3124
COMPEND7	-0.002378	0.000600	-3.962263	0.0003
COMPEND7^2	1.27E-05	2.75E-06	4.638432	0.0000
COMPEND7*CRESVEN7	5.54E-06	5.70E-06	0.970727	0.3368
COMPEND7*GIRO7	-0.000273	0.000238	-1.149871	0.2561
COMPEND7*LIQCOR7	0.000113	0.000123	0.920929	0.3619
COMPEND7*MARVEN7	-2.09E-05	2.69E-05	-0.776842	0.4412
COMPEND7*RENTAJU7	2.71E-05	2.09E-05	1.299531	0.2002
CRESVEN7	-0.001668	0.001020	-1.635664	0.1087
CRESVEN7^2	-2.97E-06	5.96E-06	-0.497705	0.6211
CRESVEN7*GIRO7	-5.04E-06	0.000345	-0.014577	0.9884
CRESVEN7*LIQCOR7	0.000212	0.000212	0.998891	0.3231
CRESVEN7*MARVEN7	-3.97E-06	5.03E-05	-0.078938	0.9374
CRESVEN7*RENTAJU7	2.74E-05	3.24E-05	0.843946	0.4031
GIRO7	0.059060	0.032486	1.818027	0.0756
GIRO7^2	-0.004390	0.008337	-0.526534	0.6010
GIRO7*LIQCOR7	-0.002213	0.007664	-0.288729	0.7741
GIRO7*MARVEN7	-0.000672	0.002159	-0.311336	0.7570
GIRO7*RENTAJU7	-0.000285	0.001044	-0.273141	0.7860
LIQCOR7	-0.008104	0.032482	-0.249482	0.8041
LIQCOR7^2	-0.001245	0.004377	-0.284445	0.7773
LIQCOR7*MARVEN7	-0.000699	0.000846	-0.826106	0.4130
LIQCOR7*RENTAJU7	0.000508	0.000590	0.861546	0.3934
MARVEN7	0.014765	0.004334	3.406278	0.0014
MARVEN7^2	-0.000288	0.000138	-2.087584	0.0424
MARVEN7*RENTAJU7	4.11E-05	0.000161	0.254814	0.8000
RENTAJU7	-0.007313	0.003488	-2.096780	0.0415
RENTAJU7^2	3.02E-05	5.21E-05	0.580076	0.5647
R-squared	0.631076	Mean dependent var	0.006106	
Adjusted R-squared	0.414534	S.D. dependent var	0.008778	
S.E. of regression	0.006717	Akaike info criterion	-6.887042	
Sum squared resid	0.002075	Schwarz criterion	-6.015234	
Log likelihood	282.8206	F-statistic	2.914333	
Durbin-Watson stat	1.900836	Prob(F-statistic)	0.000684	

Devido ao problema detectado quanto à homocedasticidade dos resíduos procedeu-se à correção, de acordo com o processo de White. Comparando os resultados obtidos no processo da tabela 27 com o da regressão da tabela 24, nota-se que os erros padrões dos coeficientes angulares aumentam e os valores dos testes t reduzem, no entanto, mesmo assim, os valores de Prob relativos a quatro variáveis são inferiores a 0,01, admitindo que essas variáveis –

composição do endividamento, giro do ativo, liquidez corrente e margem líquida – são significativas. Por outro lado, duas variáveis não são significativas: crescimento das vendas e rentabilidade do patrimônio líquido.

Tabela 27: Teste de heterocedasticidade de White para os resíduos

Dependent Variable: SCORE

Method: Least Squares

Date: 01/18/09 Time: 10:24

Sample: 1 74

Included observations: 74

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.509589	0.075192	6.777210	0.0000
COMPEN7	-0.009658	0.000762	-12.67623	0.0000
CRESVEN7	-0.001061	0.000981	-1.082144	0.2831
GIRO7	0.142063	0.029408	4.830779	0.0000
LIQCOR7	0.069095	0.019420	3.557923	0.0007
MARVEN7	0.010236	0.004333	2.362432	0.0211
RENTAJU7	-0.000142	0.002646	-0.053509	0.9575
R-squared	0.852626	Mean dependent var		0.484919
Adjusted R-squared	0.839429	S.D. dependent var		0.204940
S.E. of regression	0.082122	Akaike info criterion		-2.071402
Sum squared resid	0.451851	Schwarz criterion		-1.853449
Log likelihood	83.64186	F-statistic		64.60441
Durbin-Watson stat	2.350149	Prob(F-statistic)		0.000000

Comparando os resultados obtidos, é possível obter as seguintes evidências:

Incluir todas as variáveis obtidas através da técnica Delphi foi suficiente para confirmar que as quatro variáveis – composição do endividamento, giro do ativo, liquidez corrente e margem líquida – continuam significativas e não há rejeição da normalidade dos resíduos.

3.21 Regressão e análise dos resultados do *cluster* 4: variáveis estabelecidas pela Delphi

Inicialmente, da mesma forma como foram analisados os resultados dos outros *clusters*, para a avaliação dos dados finais do *cluster* 4, foram usados o programa SPSS e o método *stepwise* para fazer uma pré-seleção das variáveis do modelo. Novamente, para a análise desse *cluster*, foi instituído que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

A variável margem líquida foi rejeitada, considerando-se um nível de significância superior a 5%.

A seguir, empregou-se o programa *Eviews* para continuar com a análise de regressão. Esse programa tem opções que possibilitam efetuar automaticamente os testes dos pressupostos básicos da regressão.

A tabela 28 apresenta os resultados da regressão em que a variável dependente é constituída pelos escores e as independentes são os indicadores estabelecidos pela Delphi, exceto a variável margem líquida.

Tabela 28: Resultados da regressão para o cluster 4

Dependent Variable: SCORE

Method: Least Squares

Date: 12/18/08 Time: 18:08

Sample: 1 69

Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.008463	0.000405	-20.90337	0.0000
GIRO7	0.086893	0.009844	8.826648	0.0000
LIQCOR7	0.065279	0.013708	4.762090	0.0000
CRESVEN7	0.002573	0.000689	3.732061	0.0004
RENTAJU7	0.002930	0.000844	3.469730	0.0009
C	0.556086	0.059490	9.347486	0.0000
R-squared	0.903813	Mean dependent var		0.551826
Adjusted R-squared	0.896179	S.D. dependent var		0.163044
S.E. of regression	0.052535	Akaike info criterion		-2.971733
Sum squared resid	0.173875	Schwarz criterion		-2.777463
Log likelihood	108.5248	F-statistic		118.3944
Durbin-Watson stat	2.577450	Prob(F-statistic)		0.000000

Ao analisar os resultados, nota-se que as cinco variáveis independentes inclusas no modelo são significativas em um nível de significância menor que 1%. Percebe-se ainda que os valores Prob para essas cinco variáveis são todos menores que 0,01.

Os sinais dos coeficientes angulares de todas as variáveis são coerentes. A variável composição do endividamento tem sinal negativo e isso é coerente, pois o aumento da composição do endividamento, isto é, o aumento das dívidas de curto prazo, acarreta redução no nível de eficiência.

O modelo de regressão como um todo é significativo ($Prob(F\text{-statistic})$ e menor que 0,01. O poder explicativo da regressão ($Adjusted R\text{-squared}$) é de 0,896, podendo ser considerado muito bom.

A figura 47 mostra os resultados do teste de normalidade dos resíduos, através do uso do teste Jarque-Bera. O valor para a $Probability = 0,000026$, menor que 0,05, indica que há evidências para a rejeição da hipótese de normalidade dos resíduos.

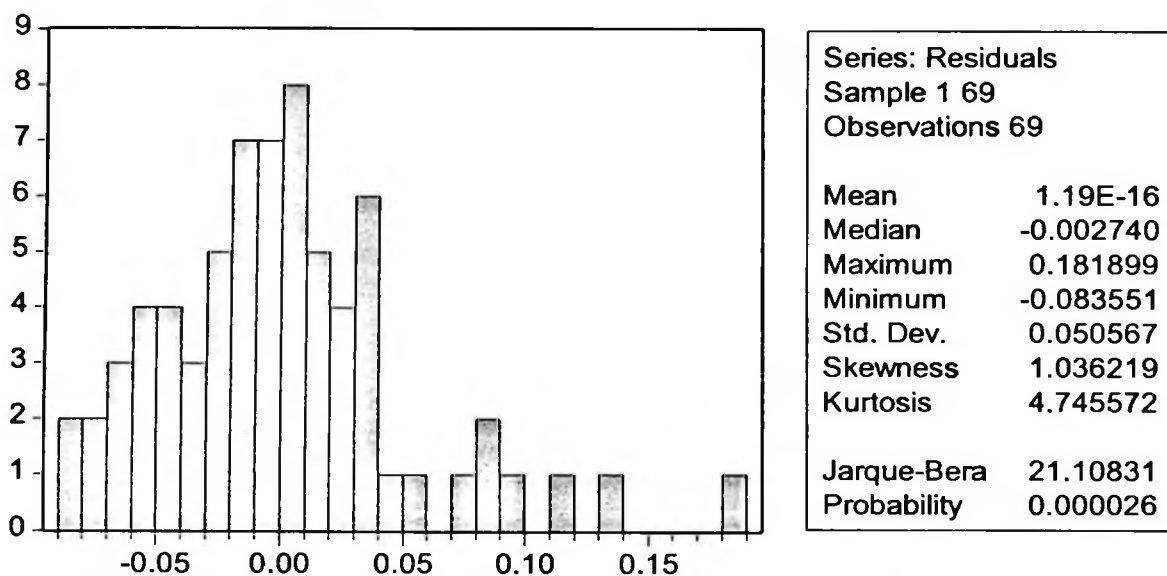


Figura 47: Teste da normalidade dos resíduos da regressão para o cluster 4

O teste de correlação serial de Breusch_Godfrey da tabela 29 indica $Probability = 0,000568$, o que significa haver evidências para se rejeitar a hipótese de ausência de correlação serial, pois esse valor é menor que o nível de significância de 0,05 e mesmo de 0,01.

Tabela 29: Teste de correlação serial dos resíduos da regressão para o *cluster* 4**Breusch-Godfrey Serial Correlation LM Test:**

F-statistic	8.468273	Probability	0.000568
Obs*R-squared	14.99453	Probability	0.000555

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/18/08 Time: 18:10

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.000164	0.000370	-0.442597	0.6596
GIRO7	0.002108	0.008894	0.237008	0.8134
LIQCOR7	0.002969	0.012346	0.240495	0.8108
CRESVEN7	-0.000521	0.000637	-0.818604	0.4162
RENTAJU7	0.000281	0.000765	0.367075	0.7148
C	0.009715	0.053571	0.181346	0.8567
RESID(-1)	-0.409780	0.119449	-3.430592	0.0011
RESID(-2)	-0.408938	0.127014	-3.219639	0.0021
R-squared	0.217312	Mean dependent var	1.19E-16	
Adjusted R-squared	0.127495	S.D. dependent var	0.050567	
S.E. of regression	0.047233	Akaike info criterion	-3.158783	
Sum squared resid	0.136090	Schwarz criterion	-2.899757	
Log likelihood	116.9780	F-statistic	2.419506	
Durbin-Watson stat	2.038010	Prob(F-statistic)	0.029677	

A tabela 30 retrata os resultados do teste de heterocedasticidade de White. O valor *Probability* = 0,00000 expressa a existência de evidências para se rejeitar a hipótese de homocedasticidade dos resíduos, em um nível de significância menor que 0,01.

Tabela 30: Teste de heterocedasticidade de White para os resíduos do *cluster 4*

White Heteroskedasticity Test:

F-statistic	9.492440	Probability	0.000000
Obs*R-squared	55.07519	Probability	0.000040

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 12/18/08 Time: 18:11

Sample: 1 69

Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.007193	0.025642	-0.280526	0.7803
COMPEND7	-7.17E-05	0.000227	-0.316110	0.7533
COMPEND7^2	4.53E-06	1.34E-06	3.371954	0.0015
COMPEND7*GIRO7	-5.10E-06	3.84E-05	-0.132773	0.8949
COMPEND7*LIQCOR7	-0.000167	6.01E-05	-2.772295	0.0079
COMPEND7*CRESVEN7	3.22E-06	3.52E-06	0.915552	0.3645
COMPEND7*RENTAJU7	-8.07E-06	2.98E-06	-2.708600	0.0093
GIRO7	-0.003929	0.006056	-0.648758	0.5196
GIRO7^2	0.000697	0.000730	0.954527	0.3446
GIRO7*LIQCOR7	0.003140	0.001950	1.610306	0.1139
GIRO7*CRESVEN7	8.13E-05	7.55E-05	1.076697	0.2870
GIRO7*RENTAJU7	-0.000131	8.91E-05	-1.474727	0.1468
LIQCOR7	0.007712	0.009032	0.853902	0.3974
LIQCOR7^2	-0.001352	0.001366	-0.989490	0.3274
LIQCOR7*CRESVEN7	-0.000154	0.000104	-1.480022	0.1454
LIQCOR7*RENTAJU7	0.000160	0.000156	1.022650	0.3116
CRESVEN7	-4.85E-05	0.000463	-0.104823	0.9170
CRESVEN7^2	6.35E-07	3.55E-06	0.178754	0.8589
CRESVEN7*RENTAJU7	-4.63E-06	5.06E-06	-0.915410	0.3646
RENTAJU7	0.000431	0.000763	0.564199	0.5752
RENTAJU7^2	3.15E-06	6.03E-06	0.522658	0.6036
R-squared	0.798191	Mean dependent var	0.002520	
Adjusted R-squared	0.714104	S.D. dependent var	0.004913	
S.E. of regression	0.002627	Akaike info criterion	-8.800334	
Sum squared resid	0.000331	Schwarz criterion	-8.120388	
Log likelihood	324.6115	F-statistic	9.492440	
Durbin-Watson stat	1.635197	Prob(F-statistic)	0.000000	

Em função do problema detectado quanto à correlação serial e também quanto à homocedasticidade dos resíduos, houve a correção conforme o processo de Newey-West. Comparando os resultados obtidos no procedimento da tabela 31 com o da regressão da tabela 28, verifica-se que os erros padrões dos coeficientes angulares aumentam e os valores dos testes t reduzem; no entanto, mesmo assim, os valores de Prob relativos às cinco variáveis são inferiores a 0,05, indicando que essas variáveis são significativas.

Tabela 31: Processo de Newey-West para correlação serial e heterocedasticidade referente ao *cluster 4*

Dependent Variable: SCORE

Method: Least Squares

Date: 12/18/08 Time: 18:13

Sample: 1 69

Included observations: 69

Newey-West HAC Standard Errors & Covariance (lag truncation=3)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.008463	0.000462	-18.31026	0.0000
GIRO7	0.086893	0.010342	8.402088	0.0000
LIQCOR7	0.065279	0.014606	4.469212	0.0000
CRESVEN7	0.002573	0.000588	4.374713	0.0000
RENTAJU7	0.002930	0.001245	2.354326	0.0217
C	0.556086	0.050348	11.04490	0.0000
R-squared	0.903813	Mean dependent var	0.551826	
Adjusted R-squared	0.896179	S.D. dependent var	0.163044	
S.E. of regression	0.052535	Akaike info criterion	-2.971733	
Sum squared resid	0.173875	Schwarz criterion	-2.777463	
Log likelihood	108.5248	F-statistic	118.3944	
Durbin-Watson stat	2.577450	Prob(F-statistic)	0.000000	

Comparando os resultados obtidos, é possível obter as seguintes evidências:

Rejeita-se H_0 , de que cinco indicadores obtidos através da técnica Delphi não têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA. Portanto, as variáveis composição do endividamento, giro do ativo, liquidez corrente, crescimento das vendas e rentabilidade sobre o patrimônio líquido e obtidas pelo modelo Delphi são significativas em um nível de significância inferior a 0,05. Cabe salientar que os quatro primeiros indicadores citados demonstram nível de significância inferior a 0,01. Apenas o indicador rentabilidade sobre o patrimônio apresenta um nível de significância de Prob = 0,0217, mas ainda bastante inferior aos 0,05.

Entretanto, é importante ressaltar que foram encontradas evidências para rejeitar a hipótese da normalidade dos resíduos em um nível de significância menor que 5%.

3.21.1 Cluster 4: Mais testes para a normalidade dos resíduos da regressão

No momento da análise da regressão referente ao *cluster* 4, foi feita uma ressalva quanto à normalidade dos resíduos da regressão. A hipótese nula de normalidade dos resíduos foi rejeitada.

Assim, procurou-se esclarecer melhor o assunto. Foi obtida nova regressão com a inclusão de todas as variáveis selecionadas pela técnica Delphi. Os resultados mostram que a hipótese nula da normalidade dos resíduos continua a ser rejeitada considerando o nível de significância de 5%.

Mais uma tentativa de resolver o problema da normalidade foi realizada. Agora cada uma das variáveis foi transformada de acordo com a seguinte expressão geral: $\log(x + 1)$. Os resultados localizam-se na tabela 32.

Verifica-se que a variável: $\log(\text{margem líquida} + 1)$ não é significativa, pois o valor Prob = 0,262 é maior que 0,05. O poder explicativo da regressão é de Adjusted R-squared = 0,944.

Tabela 32: Cluster 4: regressão com todas as variáveis da técnica Delphi

Dependent Variable: LOG(SCORE+1)

Method: Least Squares

Date: 01/18/09 Time: 10:53

Sample: 1 69

Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.934414	0.120591	7.748624	0.0000
LOG(COMPEND7+1)	-0.327184	0.011865	-27.57499	0.0000
LOG(CRESVEN7+1)	0.088930	0.013079	6.799230	0.0000
LOG(GIRO7+1)	0.180278	0.018414	9.790254	0.0000
LOG(LIQCOR7+1)	0.072188	0.016517	4.370478	0.0000
LOG(MARVEN7+1)	-0.034351	0.030377	-1.130836	0.2625
LOG(RENTAJU7+1)	0.091634	0.025565	3.584321	0.0007
R-squared	0.949617	Mean dependent var		0.434227
Adjusted R-squared	0.944741	S.D. dependent var		0.101786
S.E. of regression	0.023927	Akaike info criterion		-4.531688
Sum squared resid	0.035495	Schwarz criterion		-4.305039
Log likelihood	163.3432	F-statistic		194.7617
Durbin-Watson stat	1.921695	Prob(F-statistic)		0.000000

A figura 48 expõe o teste Jarque-Bera para normalidade dos resíduos. Como o valor Probability = 0,0356 é menor que 0,05, rejeita-se a hipótese da normalidade dos resíduos.

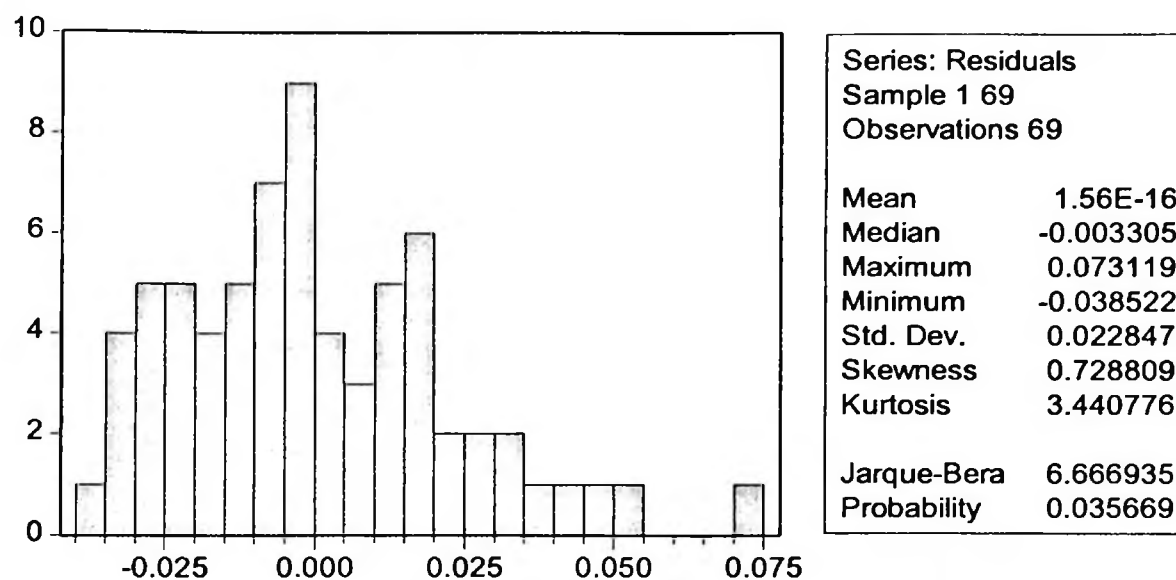


Figura 48: Normalidade dos resíduos da regressão

A tabela 33 apresenta o teste de correlação serial de Breusch-Godfrey . Como o valor Probability = 0,50 é maior que 0,05, não se pode rejeitar a hipótese nula de ausência de correlação serial.

Tabela 33: Teste de correlação serial

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.691347	Probability	0.504846
Obs*R-squared	1.554280	Probability	0.459719

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 01/18/09 Time: 10:54

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.007993	0.121389	0.065849	0.9477
LOG(COMPEN7+1)	0.000445	0.011937	0.037256	0.9704
LOG(CRESVEN7+1)	-0.003496	0.013490	-0.259176	0.7964
LOG(GIRO7+1)	-0.000817	0.018532	-0.044102	0.9650
LOG(LIQCOR7+1)	0.003142	0.016819	0.186814	0.8524
LOG(MARVEN7+1)	-0.007923	0.031267	-0.253380	0.8008
LOG(RENTAJU7+1)	0.006373	0.026259	0.242686	0.8091
RESID(-1)	0.029893	0.128409	0.232795	0.8167
RESID(-2)	-0.158949	0.137246	-1.158135	0.2514
R-squared	0.022526	Mean dependent var	1.56E-16	
Adjusted R-squared	-0.107804	S.D. dependent var	0.022847	
S.E. of regression	0.024047	Akaike info criterion	-4.496500	
Sum squared resid	0.034696	Schwarz criterion	-4.205095	
Log likelihood	164.1293	F-statistic	0.172837	
Durbin-Watson stat	1.961090	Prob(F-statistic)	0.993746	

Por sua vez, a tabela 34 mostra o teste de White para a heterocedasticidade dos resíduos. Como o valor probability = 0,000003 é menor que 0,05, rejeita-se a hipótese nula de homocedasticidade dos resíduos.

Tabela 34: Teste de heterocedasticidade dos resíduos

White Heteroskedasticity Test:

F-statistic	4.826477	Probability	0.000003
Obs*R-squared	52.48655	Probability	0.002320

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 01/18/09 Time: 10:54

Sample: 1 69

Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.019746	0.100673	-0.196145	0.8455
LOG(COMPEND7+1)	0.000134	0.012234	0.010925	0.9913
(LOG(COMPEND7+1))^2	-0.002162	0.001132	-1.910354	0.0631
(LOG(COMPEND7+1))*(LOG(CRESVEN7+1))	-0.001477	0.002308	-0.639858	0.5258
(LOG(COMPEND7+1))*(LOG(GIRO7+1))	0.001813	0.002212	0.819391	0.4173
(LOG(COMPEND7+1))*(LOG(LIQCOR7+1))	-0.004064	0.002165	-1.877010	0.0676
(LOG(COMPEND7+1))*(LOG(MARVEN7+1))	0.003498	0.003671	0.952951	0.3462
(LOG(COMPEND7+1))*(LOG(RENTAJU7+1))	0.003425	0.002551	1.342986	0.1867
LOG(CRESVEN7+1)	0.004355	0.019772	0.220258	0.8268
(LOG(CRESVEN7+1))^2	0.000941	0.001312	0.717333	0.4772
(LOG(CRESVEN7+1))*(LOG(GIRO7+1))	0.005079	0.002514	2.020198	0.0499
(LOG(CRESVEN7+1))*(LOG(LIQCOR7+1))	-0.003495	0.002260	-1.546660	0.1296
(LOG(CRESVEN7+1))*(LOG(MARVEN7+1))	0.009865	0.004216	2.339599	0.0243
(LOG(CRESVEN7+1))*(LOG(RENTAJU7+1))	-0.009434	0.003308	-2.851784	0.0068
LOG(GIRO7+1)	-0.013116	0.020501	-0.639773	0.5259
(LOG(GIRO7+1))^2	0.000388	0.002420	0.160299	0.8734
(LOG(GIRO7+1))*(LOG(LIQCOR7+1))	-0.006454	0.004296	-1.502296	0.1407
(LOG(GIRO7+1))*(LOG(MARVEN7+1))	0.004159	0.004926	0.844224	0.4034
(LOG(GIRO7+1))*(LOG(RENTAJU7+1))	-0.005664	0.005147	-1.100449	0.2776
LOG(LIQCOR7+1)	0.051230	0.019912	2.572744	0.0138
(LOG(LIQCOR7+1))^2	0.008426	0.001919	4.389652	0.0001
(LOG(LIQCOR7+1))*(LOG(MARVEN7+1))	-0.022272	0.005536	-4.023003	0.0002
(LOG(LIQCOR7+1))*(LOG(RENTAJU7+1))	0.009868	0.005262	1.875251	0.0679
LOG(MARVEN7+1)	-0.003491	0.031864	-0.109547	0.9133
(LOG(MARVEN7+1))^2	0.002677	0.005010	0.534315	0.5960
(LOG(MARVEN7+1))*(LOG(RENTAJU7+1))	-0.013037	0.008878	-1.468461	0.1496
LOG(RENTAJU7+1)	0.002023	0.034515	0.058606	0.9536
(LOG(RENTAJU7+1))^2	0.007256	0.005146	1.409831	0.1661
R-squared	0.760675	Mean dependent var	0.000514	
Adjusted R-squared	0.603070	S.D. dependent var	0.000810	
S.E. of regression	0.000510	Akaike info criterion	-12.03307	
Sum squared resid	1.07E-05	Schwarz criterion	-11.12648	
Log likelihood	443.1409	F-statistic	4.826477	
Durbin-Watson stat	1.820817	Prob(F-statistic)	0.000003	

Devido ao problema detectado quanto à homocedasticidade dos resíduos, recorreu-se à correção, conforme o processo de White. Comparando os resultados obtidos nesse procedimento da tabela 35 com o processo da regressão da tabela 32, verifica-se que os erros padrões dos coeficientes angulares aumentam e os valores dos testes t diminuem, embora, mesmo assim, os valores de Prob relativos a quatro variáveis continuem inferiores a 0,01, ratificando que essas variáveis – composição do endividamento, giro do ativo, liquidez corrente e margem líquida – são significativas. Diferentemente, duas variáveis não são significativas: crescimento das vendas e rentabilidade do patrimônio líquido.

Tabela 35: Teste de heterocedasticidade de White para os resíduos

Dependent Variable: LOG(SCORE+1)

Method: Least Squares

Date: 01/18/09 Time: 11:15

Sample: 1 69

Included observations: 69

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.934414	0.111358	8.391042	0.0000
LOG(COMPEN7+1)	-0.327184	0.011339	-28.85367	0.0000
LOG(CRESVEN7+1)	0.088930	0.017556	5.065467	0.0000
LOG(GIRO7+1)	0.180278	0.016810	10.72466	0.0000
LOG(LIQCOR7+1)	0.072188	0.024071	2.998952	0.0039
LOG(MARVEN7+1)	-0.034351	0.029968	-1.146273	0.2561
LOG(RENTAJU7+1)	0.091634	0.026290	3.485452	0.0009
R-squared	0.949617	Mean dependent var		0.434227
Adjusted R-squared	0.944741	S.D. dependent var		0.101786
S.E. of regression	0.023927	Akaike info criterion		-4.531688
Sum squared resid	0.035495	Schwarz criterion		-4.305039
Log likelihood	163.3432	F-statistic		194.7617
Durbin-Watson stat	1.921695	Prob(F-statistic)		0.000000

Comparando os resultados obtidos, é possível obter as seguintes evidências:

A única variável não significativa é log (margem líquida + 1), pois seu valor Prob = 0,256 é maior que 0,05. As outras cinco variáveis transformadas e correspondentes às variáveis obtidas pela técnica Delphi são significativas em um nível de significância de 1%.

Entretanto, apesar das transformações, a hipótese nula da normalidade dos resíduos foi rejeitada através do teste de Jarque-Bera. Por outro lado, se for efetuado o teste da normalidade dos resíduos com o teste de Kolmogorov-Smirnov, conclui-se que não se pode rejeitar a hipótese nula da normalidade uma vez que o valor Sig. = 0,175 é maior que 0,05. Tais resultados estão expostos na tabela 36.

Tabela 36: Teste de normalidade dos resíduos

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	DF	Sig.	Statistic	df	Sig.
RESÍDUOS	,097	69	,175	,962	69	,034

a. Lilliefors Significance Correction

3.22 Regressão e análise dos resultados do *cluster* 5: variáveis estabelecidas pela Delphi

Da mesma maneira como foram analisados os resultados dos outros *clusters*, para a avaliação dos dados finais do *cluster* 5, foram utilizados o programa SPSS e o método *stepwise* para fazer uma pré-seleção das variáveis do modelo. Novamente, para a análise desse *cluster*, foi estabelecido que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

Constata-se que as variáveis crescimento das vendas e margem líquida foram rejeitadas, considerando-se um nível de significância superior a 5%.

A seguir, para continuar com a análise de regressão, foi empregado o programa *Eviews*, o qual tem opções que possibilitam efetuar automaticamente os testes dos pressupostos básicos da regressão.

A tabela 37 demonstra os resultados da regressão em que a variável dependente é constituída pelos escores e as independentes são os indicadores estabelecidos pela Delphi, exceto crescimento das vendas e margem líquida.

Tabela 37: Resultados da regressão para o cluster 5

Dependent Variable: SCORE

Method: Least Squares

Date: 12/18/08 Time: 18:21

Sample: 1 55

Included observations: 55

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.009357	0.000583	-16.04280	0.0000
GIRO7	0.226265	0.056331	4.016705	0.0002
LIQCOR7	0.072767	0.024104	3.018859	0.0040
RENTAJU7	0.005136	0.002198	2.336710	0.0235
C	0.653080	0.083673	7.805128	0.0000
R-squared	0.859717	Mean dependent var		0.632000
Adjusted R-squared	0.848494	S.D. dependent var		0.225943
S.E. of regression	0.087945	Akaike info criterion		-1.937696
Sum squared resid	0.386719	Schwarz criterion		-1.755211
Log likelihood	58.28664	F-statistic		76.60564
Durbin-Watson stat	1.485020	Prob(F-statistic)		0.000000

Analisando os resultados, compreende-se que as quatro variáveis independentes incluídas no modelo são significativas. Três delas, composição do endividamento, giro do ativo e liquidez corrente, são significativas a um nível de significância menor que 1%. Pode-se afirmar que os valores Prob para essas três variáveis são todos menores que 0,01. A variável rentabilidade sobre o patrimônio líquido é significativa no nível de 2,35%. Ressalta-se ainda que o valor Prob para essa variável é de 0,0235.

Os sinais dos coeficientes angulares de todas as variáveis são coerentes, como a variável composição do endividamento, que tem sinal negativo, pois o aumento da composição do endividamento, ou, em outras palavras, o aumento das dívidas de curto prazo, implica redução no nível de eficiência.

O modelo de regressão como um todo é significativo ($Prob(F\text{-statistic})$) e menor que 0,01. O poder explicativo da regressão ($Adjusted\ R\text{-squared}$) é de 0,848, considerado bastante razoável.

A figura 49 mostra os resultados do teste de normalidade dos resíduos. Utilizou-se o teste Jarque-Bera, sendo o valor para a Probability = 0,4389 maior que 0,05. Isso indica que não há evidências para a rejeição da hipótese de normalidade dos resíduos.

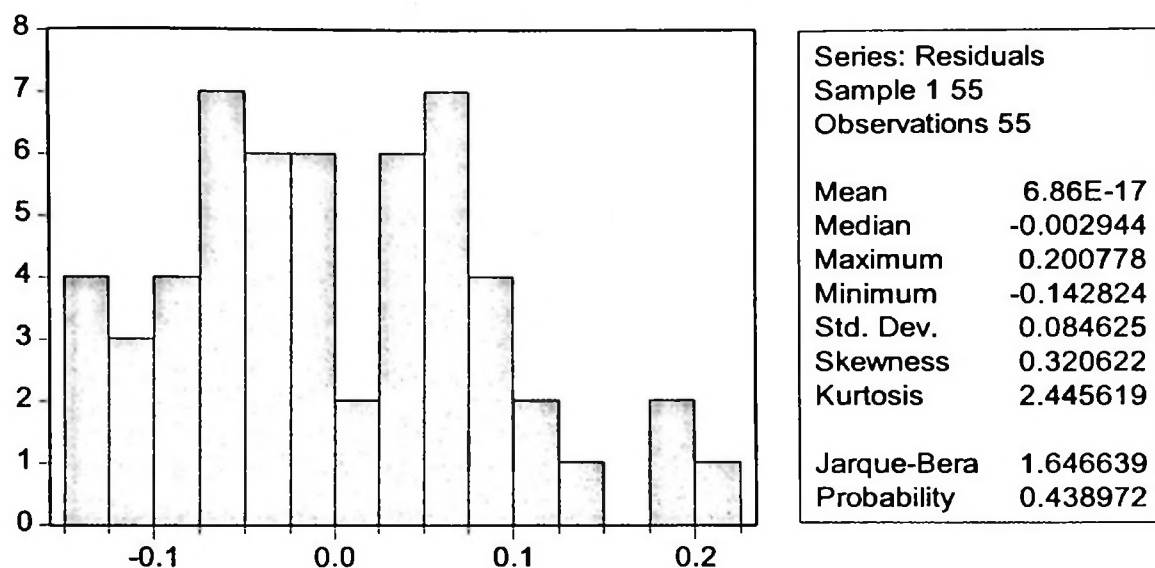


Figura 49: Teste da normalidade dos resíduos da regressão para o cluster 5

A tabela 38 a seguir exibe o teste de correlação serial dos resíduos da regressão. O teste de correlação serial de Breusch_Godfrey da Tabela 39 indica *Probability* = 0,1009, o que significa que não se pode rejeitar a hipótese de ausência de correlação serial pois esse valor é superior a 0,05.

Tabela 38: Teste de correlação serial dos resíduos da regressão para o cluster 5

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	2.406628	Probability	0.100916
Obs*R-squared	5.012550	Probability	0.081572

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/18/08 Time: 18:22

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-1.36E-05	0.000568	-0.024040	0.9809
GIRO7	-0.011219	0.055435	-0.202374	0.8405
LIQCOR7	0.009502	0.023859	0.398270	0.6922
RENTAJU7	0.000324	0.002154	0.150506	0.8810
C	-0.016011	0.081926	-0.195437	0.8459
RESID(-1)	0.295664	0.150737	1.961458	0.0556
RESID(-2)	-0.225144	0.152632	-1.475073	0.1467
R-squared	0.091137	Mean dependent var	6.86E-17	
Adjusted R-squared	-0.022471	S.D. dependent var	0.084625	
S.E. of regression	0.085571	Akaike info criterion	-1.960530	
Sum squared resid	0.351474	Schwarz criterion	-1.705051	
Log likelihood	60.91458	F-statistic	0.802209	
Durbin-Watson stat	1.794599	Prob(F-statistic)	0.573110	

A tabela 39 corrobora os resultados do teste de heterocedasticidade de White. O valor Probability = 0,169 indica que não há evidências para se rejeitar a hipótese de homocedasticidade dos resíduos, em um nível de significância menor que 0,01.

Tabela 39: Teste de heterocedasticidade de White para os resíduos do *cluster* 5

White Heteroskedasticity Test:

F-statistic	1.464176	Probability	0.169810
Obs*R-squared	18.63544	Probability	0.179352

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 12/18/08 Time: 18:23

Sample: 1 55

Included observations: 55

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.005126	0.060438	0.084821	0.9328
COMPEND7	0.000200	0.000560	0.357401	0.7227
COMPEND7^2	1.57E-06	2.93E-06	0.538289	0.5934
COMPEND7*GIRO7	-0.000358	0.000234	-1.527055	0.1346
COMPEND7*LIQCOR7	-7.92E-05	0.000130	-0.607407	0.5470
COMPEND7*RENTAJU7	7.51E-07	1.08E-05	0.069240	0.9451
GIRO7	-0.063062	0.048749	-1.293606	0.2032
GIRO7^2	-0.003204	0.020102	-0.159396	0.8742
GIRO7*LIQCOR7	0.005791	0.016645	0.347931	0.7297
GIRO7*RENTAJU7	0.001678	0.001058	1.585421	0.1207
LIQCOR7	0.020205	0.025647	0.787814	0.4354
LIQCOR7^2	-0.001731	0.005155	-0.335856	0.7387
LIQCOR7*RENTAJU7	-0.000374	0.000580	-0.645169	0.5225
RENTAJU7	0.000595	0.002895	0.205399	0.8383
RENTAJU7^2	-1.93E-05	3.30E-05	-0.585269	0.5617
R-squared	0.338826	Mean dependent var	0.007031	
Adjusted R-squared	0.107415	S.D. dependent var	0.008532	
S.E. of regression	0.008061	Akaike info criterion	-6.576652	
Sum squared resid	0.002599	Schwarz criterion	-6.029198	
Log likelihood	195.8579	F-statistic	1.464176	
Durbin-Watson stat	1.861579	Prob(F-statistic)	0.169810	

Comparando os resultados obtidos, é possível obter as seguintes evidências:

Rejeita-se H_0 , de que quatro indicadores obtidos através da técnica Delphi não têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA. Os indicadores composição do endividamento, giro do ativo, liquidez corrente e obtidos pelo modelo Delphi são significativos em um nível de significância inferior a 0,01. Já o indicador rentabilidade sobre o patrimônio oferece um nível de significância de Prob = 0,0235, mas ainda bastante inferior aos 0,05.

Os indicadores crescimento das vendas e margem líquida não se mostram significativos no nível de significância de 5%.

3.23 Regressão e análise dos resultados do cluster 6: variáveis estabelecidas pela Delphi

Do mesmo modo como foram examinados os resultados dos outros *clusters*, para a avaliação dos dados finais do *cluster* 6, foram utilizados o programa SPSS e o método *stepwise* para fazer uma pré-seleção das variáveis do modelo. Novamente, para a análise desse *cluster*, instituiu-se que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

Verificou-se que a variável margem líquida foi rejeitada, considerando-se um nível de significância superior a 5%.

Para continuar com a análise de regressão, a seguir, foi utilizado o programa *Eviews*, pois este tem opções que possibilitam efetuar automaticamente os testes dos pressupostos básicos da regressão.

A tabela 40 apresenta os resultados da regressão em que a variável dependente é constituída pelos escores e as independentes são os indicadores estabelecidos pela Delphi, exceto a margem líquida.

Tabela 40: Resultados da regressão para o cluster 6

Dependent Variable: SCORE

Method: Least Squares

Date: 12/18/08 Time: 18:27

Sample: 1 70

Included observations: 70

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.008394	0.000566	-14.83991	0.0000
GIRO7	0.131126	0.016308	8.040465	0.0000
CRESVEN7	0.003230	0.000868	3.723804	0.0004
RENTAJU7	0.003920	0.001049	3.737694	0.0004
LIQCOR7	0.055764	0.015964	3.493128	0.0009
C	0.576032	0.063403	9.085230	0.0000
R-squared	0.799459	Mean dependent var		0.608986
Adjusted R-squared	0.783791	S.D. dependent var		0.150641
S.E. of regression	0.070046	Akaike info criterion		-2.397524
Sum squared resid	0.314009	Schwarz criterion		-2.204796
Log likelihood	89.91333	F-statistic		51.02720
Durbin-Watson stat	1.737646	Prob(F-statistic)		0.000000

Averiguando os resultados, verifica-se que as cinco variáveis independentes incluídas no modelo são significativas em um nível de significância menor que 1%. Pode-se destacar que os valores Prob para essas cinco variáveis são todos menores que 0,01.

Os sinais dos coeficientes angulares de todas as variáveis são coerentes, como a variável composição do endividamento, a qual tem sinal negativo, pois o aumento da composição do endividamento, ou seja, o aumento das dívidas de curto prazo, ocasiona redução no nível de eficiência.

O modelo de regressão como um todo é significativo ($Prob(F\text{-statistic})$) e menor que 0,01. O poder explicativo da regressão ($Adjusted\ R\text{-squared}$) é de 0,78379 e pode ser considerado bastante razoável.

A figura 50 demonstra os resultados do teste de normalidade dos resíduos, a partir da utilização do teste Jarque-Bera. O valor para a $Probability = 0,043$, menor que 0,05, sugere que existem evidências para a rejeição da hipótese de normalidade dos resíduos.

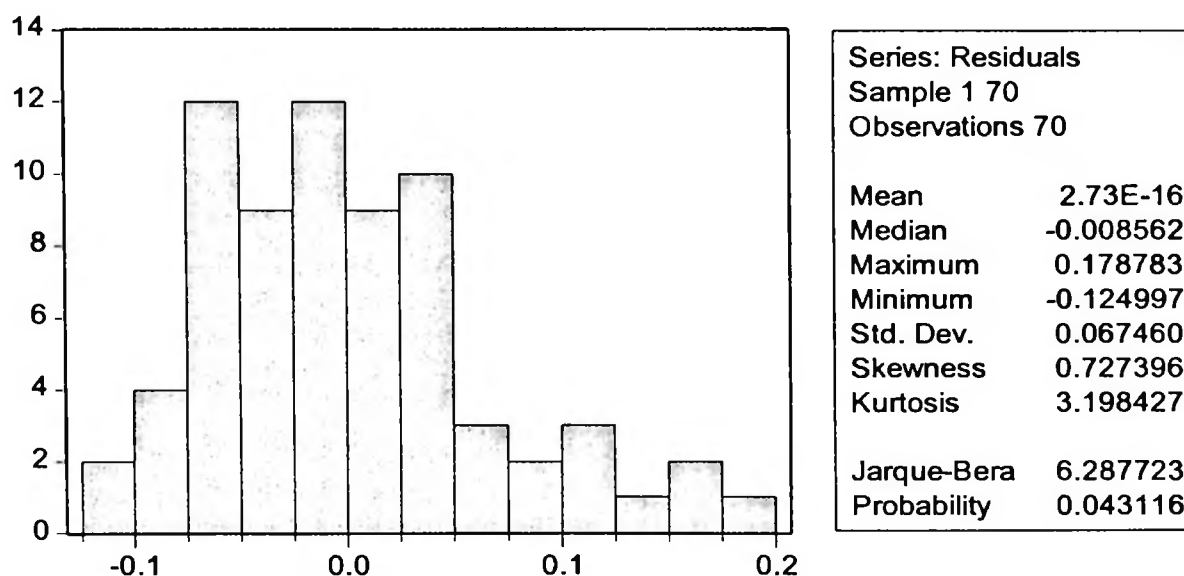


Figura 50: Teste da normalidade dos resíduos da regressão para o *cluster* 6

A tabela 41 expõe o teste de correlação serial dos resíduos da regressão.

Conforme a tabela abaixo, o teste Breusch_Godfrey indica *Probability* = 0,629, significando que não se pode rejeitar a hipótese de ausência de correlação serial em um nível de significância menor que 0,01.

Tabela 41: Teste de correlação serial dos resíduos da regressão para o *cluster* 6

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.466426	Probability	0.629423
Obs*R-squared	1.037609	Probability	0.595232

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/18/08 Time: 18:30

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	8.73E-05	0.000578	0.151113	0.8804
GIRO7	0.000930	0.016561	0.056146	0.9554
CRESVEN7	-2.91E-05	0.000879	-0.033160	0.9737
RENTAJU7	-0.000158	0.001086	-0.145539	0.8848
LIQCOR7	-0.000296	0.016109	-0.018380	0.9854
C	0.000476	0.064256	0.007408	0.9941
RESID(-1)	0.126167	0.130798	0.964591	0.3385
RESID(-2)	-0.023467	0.133296	-0.176055	0.8608
R-squared	0.014823	Mean dependent var	2.73E-16	
Adjusted R-squared	-0.096407	S.D. dependent var	0.067460	
S.E. of regression	0.070637	Akaike info criterion	-2.355315	
Sum squared resid	0.309354	Schwarz criterion	-2.098344	
Log likelihood	90.43602	F-statistic	0.133265	
Durbin-Watson stat	1.967635	Prob(F-statistic)	0.995407	

Por sua vez, a tabela 42 mostra os resultados do teste de heterocedasticidade de White. O valor *Probability* = 0,00002 sugere que há evidências para se rejeitar a hipótese de homocedasticidade dos resíduos, em um nível de significância menor que 0,01.

Tabela 42: Teste de heterocedasticidade de White para os resíduos do *cluster 6*

White Heteroskedasticity Test:

F-statistic	4.197107	Probability	0.000022
Obs*R-squared	44.19930	Probability	0.001415

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 12/18/08 Time: 18:30

Sample: 1 70

Included observations: 70

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.019408	0.038888	-0.499090	0.6199
COMPEND7	-0.001046	0.000381	-2.748635	0.0084
COMPEND7^2	8.05E-06	2.22E-06	3.623989	0.0007
COMPEND7*GIRO7	-1.23E-05	9.62E-05	-0.127320	0.8992
COMPEND7*CRESVEN7	1.67E-06	4.45E-06	0.375184	0.7091
COMPEND7*RENTAJU7	-8.16E-06	6.26E-06	-1.303186	0.1986
COMPEND7*LIQCOR7	4.42E-05	8.64E-05	0.511768	0.6111
GIRO7	-0.002182	0.010678	-0.204306	0.8390
GIRO7^2	0.002163	0.002439	0.886569	0.3796
GIRO7*CRESVEN7	0.000298	0.000168	1.777132	0.0818
GIRO7*RENTAJU7	-0.000428	0.000180	-2.376178	0.0214
GIRO7*LIQCOR7	-0.007136	0.002375	-3.004483	0.0042
CRESVEN7	0.001273	0.000827	1.538748	0.1303
CRESVEN7^2	-1.28E-05	6.99E-06	-1.825053	0.0741
CRESVEN7*RENTAJU7	-3.79E-06	1.03E-05	-0.368924	0.7138
CRESVEN7*LIQCOR7	4.39E-05	0.000172	0.255359	0.7995
RENTAJU7	0.000723	0.000774	0.933656	0.3551
RENTAJU7^2	1.29E-05	8.02E-06	1.604200	0.1151
RENTAJU7*LIQCOR7	-0.000113	0.000236	-0.479699	0.6336
LIQCOR7	0.009134	0.012547	0.727996	0.4701
LIQCOR7^2	-0.000413	0.003251	-0.126949	0.8995
R-squared	0.631419	Mean dependent var	0.004486	
Adjusted R-squared	0.480977	S.D. dependent var	0.006699	
S.E. of regression	0.004826	Akaike info criterion	-7.586133	
Sum squared resid	0.001141	Schwarz criterion	-6.911585	
Log likelihood	286.5147	F-statistic	4.197107	
Durbin-Watson stat	2.063132	Prob(F-statistic)	0.000022	

Em função do problema detectado quanto à homocedasticidade dos resíduos procedeu-se à correção conforme o processo de White. Comparando os resultados obtidos no procedimento da tabela 43 com o da regressão da tabela 40, verifica-se que os erros padrões dos coeficientes angulares aumentam e os valores dos testes t reduzem, mas, mesmo assim, os valores de Prob. relativos às cinco variáveis são inferiores a 0,01, o que indica que essas variáveis são significativas.

Tabela 43: Processo de White para heterocedasticidade referente ao *cluster* 6

Dependent Variable: SCORE

Method: Least Squares

Date: 12/18/08 Time: 18:31

Sample: 1 70

Included observations: 70

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.008394	0.000781	-10.74783	0.0000
GIRO7	0.131126	0.020944	6.260862	0.0000
CRESVEN7	0.003230	0.000718	4.499072	0.0000
RENTAJU7	0.003920	0.001209	3.242656	0.0019
LIQCOR7	0.055764	0.018450	3.022367	0.0036
C	0.576032	0.062269	9.250671	0.0000
R-squared	0.799459	Mean dependent var		0.608986
Adjusted R-squared	0.783791	S.D. dependent var		0.150641
S.E. of regression	0.070046	Akaike info criterion		-2.397524
Sum squared resid	0.314009	Schwarz criterion		-2.204796
Log likelihood	89.91333	F-statistic		51.02720
Durbin-Watson stat	1.737646	Prob(F-statistic)		0.000000

Comparando os resultados obtidos, é possível obter as seguintes evidências:

Rejeita-se H_0 , de que os cinco indicadores obtidos através da técnica Delphi não têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA. Portanto, os indicadores composição do endividamento, giro do ativo, crescimento das vendas, rentabilidade sobre o patrimônio líquido e liquidez corrente obtidos pelo modelo Delphi são significativos em um nível de significância inferior a 0,01.

Entretanto, cabe salientar que foram encontradas evidências para rejeitar a hipótese de da normalidade dos resíduos em um nível de significância menor que 5%.

3.23.1 *Cluster* 6: Mais testes para a normalidade dos resíduos da regressão

Quando foi feita a análise da regressão referente ao *cluster* 6, houve uma ressalva quanto à normalidade dos resíduos da regressão. A hipótese nula de normalidade dos resíduos foi rejeitada.

Assim, buscou-se investigar melhor o assunto. Foi obtida nova regressão com a inclusão de todas as variáveis selecionadas pela técnica Delphi. Os resultados estão na Tabela 44.

Constata-se que a variável margem líquida não é significativa, conforme já afirmado anteriormente. O poder explicativo da regressão é de *Adjusted R-squared* = 0,78.

Tabela 44: Cluster 6: regressão com todas as variáveis da técnica Delphi

Dependent Variable: SCORE

Method: Least Squares

Date: 01/18/09 Time: 11:55

Sample: 1 70

Included observations: 70

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.582049	0.071857	8.100071	0.0000
COMPEN7	-0.008358	0.000603	-13.85929	0.0000
CRESVEN7	0.003207	0.000884	3.629544	0.0006
GIRO7	0.129345	0.019103	6.770856	0.0000
LIQCOR7	0.056512	0.016598	3.404809	0.0012
MARVEN7	-0.000936	0.005117	-0.182922	0.8554
RENTAJU7	0.004168	0.001721	2.421240	0.0184
R-squared	0.799565	Mean dependent var		0.608986
Adjusted R-squared	0.780476	S.D. dependent var		0.150641
S.E. of regression	0.070581	Akaike info criterion		-2.369483
Sum squared resid	0.313842	Schwarz criterion		-2.144634
Log likelihood	89.93192	F-statistic		41.88606
Durbin-Watson stat	1.727050	Prob(F-statistic)		0.000000

A figura 51 exibe teste Jarque-Bera para normalidade dos resíduos. Como o valor Probability = 0,0503 é maior que 0,05, não se pode rejeitar a hipótese da normalidade dos resíduos.

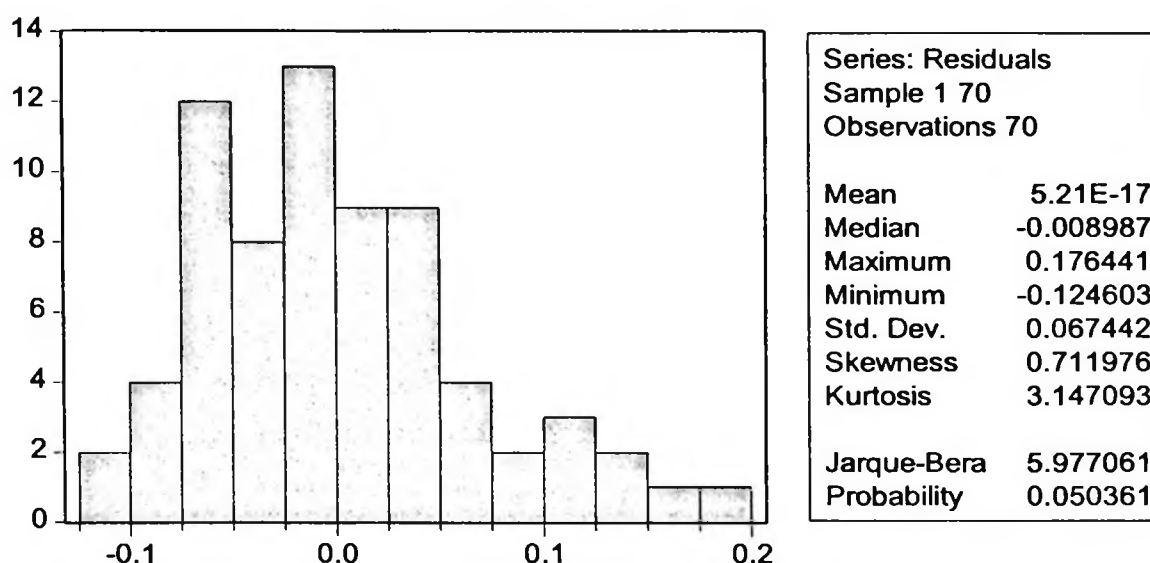


Figura 51: Normalidade do resíduos da regressão

A tabela 45 apresenta o teste de correlação serial de Breusch-Godfrey. Como o valor *Probability* = 0,58 é maior que 0,05, não se pode rejeitar a hipótese nula de ausência de correlação serial.

Tabela 45: Teste de correlação serial

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.534537	Probability	0.588660	
Obs*R-squared	1.205676	Probability	0.547256	
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Date: 01/18/09 Time: 11:57				
Presample missing value lagged residuals set to zero.				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.008201	0.072932	0.112446	0.9108
COMPEND7	0.000144	0.000625	0.230817	0.8182
CRESVEN7	-6.44E-05	0.000896	-0.071838	0.9430
GIRO7	-0.001263	0.019314	-0.065377	0.9481
LIQCOR7	0.000590	0.016751	0.035235	0.9720
MARVEN7	-0.001171	0.005292	-0.221358	0.8256
RENTAJU7	0.000136	0.001763	0.076930	0.9389
RESID(-1)	0.139361	0.135144	1.031204	0.3065
RESID(-2)	-0.027648	0.134343	-0.205801	0.8376
R-squared	0.017224	Mean dependent var	5.21E-17	
Adjusted R-squared	-0.111665	S.D. dependent var	0.067442	
S.E. of regression	0.071108	Akaike info criterion	-2.329714	
Sum squared resid	0.308437	Schwarz criterion	-2.040622	
Log likelihood	90.54001	F-statistic	0.133634	
Durbin-Watson stat	1.965613	Prob(F-statistic)	0.997444	

A tabela 46 evidencia o teste de White para a heterocedasticidade dos resíduos. Como o valor *probability* = 0,000019 é menor que 0,05, rejeita-se a hipótese nula de homocedasticidade dos resíduos.

Tabela 46: Teste de heterocedasticidade dos resíduos

White Heteroskedasticity Test:

F-statistic	4.147313	Probability	0.000019
Obs*R-squared	50.90630	Probability	0.003570

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 01/18/09 Time: 11:57

Sample: 1 70

Included observations: 70

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.030009	0.042230	0.710618	0.4812
COMPEND7	-0.000903	0.000618	-1.461844	0.1512
COMPEND7^2	9.15E-06	2.65E-06	3.453542	0.0013
COMPEND7*CRESVEN7	4.70E-07	5.71E-06	0.082414	0.9347
COMPEND7*GIRO7	-0.000113	0.000127	-0.893288	0.3768
COMPEND7*LIQCOR7	-3.86E-05	8.81E-05	-0.438071	0.6636
COMPEND7*MARVEN7	-2.80E-05	4.76E-05	-0.588799	0.5591
COMPEND7*RENTAJU7	5.53E-06	1.60E-05	0.346284	0.7309
CRESVEN7	0.000133	0.000929	0.142902	0.8871
CRESVEN7^2	-7.98E-06	7.83E-06	-1.019129	0.3140
CRESVEN7*GIRO7	0.000326	0.000233	1.397874	0.1695
CRESVEN7*LIQCOR7	0.000115	0.000210	0.548476	0.5863
CRESVEN7*MARVEN7	5.76E-06	5.12E-05	0.112479	0.9110
CRESVEN7*RENTAJU7	8.47E-06	2.34E-05	0.362732	0.7186
GIRO7	-0.028567	0.017674	-1.616342	0.1135
GIRO7^2	0.005552	0.002855	1.944313	0.0586
GIRO7*LIQCOR7	-0.012649	0.004165	-3.037111	0.0041
GIRO7*MARVEN7	0.004430	0.001419	3.122854	0.0032
GIRO7*RENTAJU7	-0.001466	0.000427	-3.431800	0.0014
LIQCOR7	0.010260	0.015199	0.675002	0.5034
LIQCOR7^2	0.001240	0.003978	0.311773	0.7568
LIQCOR7*MARVEN7	-0.001246	0.001626	-0.766259	0.4478
LIQCOR7*RENTAJU7	0.000556	0.000764	0.727550	0.4709
MARVEN7	-0.003392	0.004104	-0.826697	0.4131
MARVEN7^2	0.000283	0.000198	1.425625	0.1614
MARVEN7*RENTAJU7	-0.000233	0.000135	-1.724150	0.0920
RENTAJU7	0.002073	0.001424	1.456125	0.1528
RENTAJU7^2	2.63E-05	2.22E-05	1.182999	0.2435
R-squared	0.727233	Mean dependent var	0.004483	
Adjusted R-squared	0.551882	S.D. dependent var	0.006617	
S.E. of regression	0.004430	Akaike info criterion	-7.711866	
Sum squared resid	0.000824	Schwarz criterion	-6.812467	
Log likelihood	297.9153	F-statistic	4.147313	
Durbin-Watson stat	1.919641	Prob(F-statistic)	0.000019	

Em razão do problema detectado quanto à homocedasticidade dos resíduos, fez-se a correção conforme o processo de White. Na comparação dos resultados obtidos no procedimento da tabela 47 com o da regressão da tabela 44, constata-se que os erros padrões dos coeficientes

angulares aumentam e os valores dos testes t reduzem, no entanto, mesmo assim, os valores de Prob relativos a quatro variáveis são inferiores a 0,01, demonstrando que essas variáveis – composição do endividamento, crescimento das vendas, giro do ativo e liquidez corrente – são significativas. Por outro lado, duas variáveis não são significativas: margem líquida e rentabilidade do patrimônio líquido.

Tabela 47: Teste de heterocedasticidade de White para os resíduos

Dependent Variable: SCORE

Method: Least Squares

Date: 01/18/09 Time: 11:59

Sample: 1 70

Included observations: 70

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.582049	0.066157	8.798028	0.0000
COMPEN7	-0.008358	0.000884	-9.457932	0.0000
CRESVEN7	0.003207	0.000766	4.188315	0.0001
GIRO7	0.129345	0.028700	4.506728	0.0000
LIQCOR7	0.056512	0.019706	2.867741	0.0056
MARVEN7	-0.000936	0.006220	-0.150468	0.8809
RENTAJU7	0.004168	0.002429	1.715863	0.0911
R-squared	0.799565	Mean dependent var	0.608986	
Adjusted R-squared	0.780476	S.D. dependent var	0.150641	
S.E. of regression	0.070581	Akaike info criterion	-2.369483	
Sum squared resid	0.313842	Schwarz criterion	-2.144634	
Log likelihood	89.93192	F-statistic	41.88606	
Durbin-Watson stat	1.727050	Prob(F-statistic)	0.000000	

Comparando os resultados obtidos, é possível obter as seguintes evidências:

É possível incluir todas as variáveis obtidas através da técnica Delphi para confirmar que as quatro variáveis citadas – composição do endividamento, crescimento das vendas, giro do ativo e liquidez corrente – continuam significativas, não havendo rejeição da normalidade dos resíduos.

3.24 Regressão e análise dos resultados do *cluster 7*: variáveis estabelecidas pela Delphi

Da mesma forma como foram examinados os resultados dos outros *clusters*, para a avaliação dos dados finais do *cluster 7*, foram utilizados o programa SPSS e o método *stepwise* para fazer uma pré-seleção das variáveis do modelo. Igualmente, para a análise desse *cluster*,

estabeleceu-se que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

Constata-se que as variáveis margem líquida e rentabilidade sobre o patrimônio líquido foram rejeitadas, considerando-se um nível de significância superior a 5%.

A seguir, foi utilizado o programa *Eviews* para continuar com a análise de regressão. Esse programa tem opções que possibilitam efetuar automaticamente os testes dos pressupostos básicos da regressão.

A tabela 48 destaca os resultados da regressão em que a variável dependente é constituída pelos escores e as independentes são os indicadores estabelecidos pela Delphi, exceto as variáveis margem líquida e rentabilidade sobre o patrimônio líquido.

Tabela 48: Resultados da regressão para o cluster 7

Dependent Variable: SCORE

Method: Least Squares

Date: 12/18/08 Time: 18:36

Sample: 1 87

Included observations: 87

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.015289	0.000704	-21.70516	0.0000
GIRO7	0.304387	0.020885	14.57437	0.0000
LIQCOR7	0.054444	0.017686	3.078324	0.0028
CRESVEN7	0.001584	0.000711	2.228952	0.0286
C	0.875887	0.046943	18.65854	0.0000
R-squared	0.882280	Mean dependent var		0.652540
Adjusted R-squared	0.876538	S.D. dependent var		0.184833
S.E. of regression	0.064945	Akaike info criterion		-2.574795
Sum squared resid	0.345865	Schwarz criterion		-2.433076
Log likelihood	117.0036	F-statistic		153.6425
Durbin-Watson stat	1.604932	Prob(F-statistic)		0.000000

Ao se verificarem os resultados, as quatro variáveis independentes contidas no modelo são significativas em um nível de significância menor que 5%. Pode-se afirmar que os valores Prob para essas três variáveis são todos menores que 0,05. Vale salientar que três dessas variáveis, composição do endividamento, giro do ativo e liquidez corrente, são significativas a um nível de significância menor que 1%. Já a variável crescimento das vendas é significativa em um nível de 0,0286 e, portanto, menor que 0,05.

Os sinais dos coeficientes angulares de todas as variáveis são coerentes, como a variável composição do endividamento, que tem sinal negativo, pois o aumento da composição do endividamento, isto é, o aumento das dívidas de curto prazo, causa redução no nível de eficiência.

O modelo de regressão como um todo é significativo ($Prob(F\text{-statistic})$ e menor que 0,01. O poder explicativo da regressão ($Adjusted R\text{-squared}$) é de 0,8765, sendo considerado bom.

Na figura 52, expõem-se os resultados do teste de normalidade dos resíduos. Foi empregado o teste Jarque-Bera e o valor para a Probability = 0,1795 é maior que 0,05, indicando não haver evidências para a rejeição da hipótese de normalidade dos resíduos.

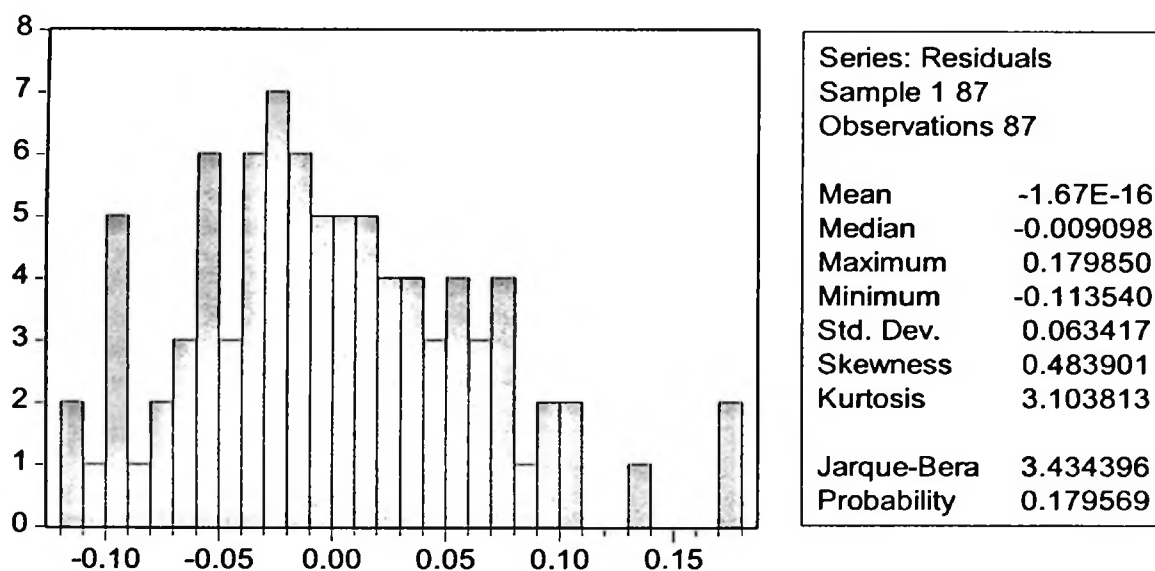


Figura 52: Teste da normalidade dos resíduos da regressão para o cluster 7

A tabela 49 mostra os resultados do teste de correlação serial dos resíduos da regressão.

De acordo com a tabela 49, o teste Breusch_Godfrey sugere $Probability = 0,091$, denotando que não se pode rejeitar a hipótese de ausência de correlação serial em um nível de significância de 0,05.

Tabela 49: Teste de correlação serial dos resíduos da regressão para o *cluster 7*

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	2.465472	Probability	0.091403
Obs*R-squared	5.051070	Probability	0.080015

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/18/08 Time: 18:38

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	3.78E-05	0.000692	0.054623	0.9566
GIRO7	0.000975	0.020654	0.047193	0.9625
LIQCOR7	0.002954	0.017429	0.169461	0.8659
CRESVEN7	-5.06E-05	0.000701	-0.072156	0.9427
C	-0.004465	0.046324	-0.096389	0.9235
RESID(-1)	0.221017	0.111290	1.985959	0.0505
RESID(-2)	-0.156945	0.117373	-1.337147	0.1850
R-squared	0.058058	Mean dependent var	-1.67E-16	
Adjusted R-squared	-0.012587	S.D. dependent var	0.063417	
S.E. of regression	0.063815	Akaike info criterion	-2.588630	
Sum squared resid	0.325785	Schwarz criterion	-2.390223	
Log likelihood	119.6054	F-statistic	0.821824	
Durbin-Watson stat	1.970584	Prob(F-statistic)	0.556313	

A tabela 50 demonstra os resultados do teste de heterocedasticidade de White. O valor *Probability* = 0,000001 sugere que há evidências para se rejeitar a hipótese de homocedasticidade dos resíduos, em um nível de significância menor que 0,01.

Tabela 50: Teste de heterocedasticidade de White para os resíduos do *cluster 7*

White Heteroskedasticity Test:

F-statistic	5.190073	Probability	0.000001
Obs*R-squared	43.69877	Probability	0.000066

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 12/18/08 Time: 18:38

Sample: 1 87

Included observations: 87

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.043780	0.016669	-2.626455	0.0105
COMPEND7	0.000177	0.000392	0.451174	0.6532
COMPEND7^2	6.94E-06	4.09E-06	1.697968	0.0938
COMPEND7*GIRO7	-0.000141	0.000163	-0.869714	0.3873
COMPEND7*LIQCOR7	-0.000127	0.000131	-0.964366	0.3381
COMPEND7*CRESVEN7	-1.53E-05	4.87E-06	-3.138568	0.0025
GIRO7	0.005341	0.011757	0.454275	0.6510
GIRO7^2	0.003444	0.003107	1.108143	0.2715
GIRO7*LIQCOR7	-0.000351	0.003958	-0.088790	0.9295
GIRO7*CRESVEN7	-0.000216	0.000192	-1.129041	0.2626
LIQCOR7	0.032373	0.012705	2.547948	0.0130
LIQCOR7^2	-0.001682	0.002858	-0.588594	0.5580
LIQCOR7*CRESVEN7	-0.000652	0.000141	-4.640602	0.0000
CRESVEN7	0.001196	0.000355	3.368366	0.0012
CRESVEN7^2	6.84E-06	3.52E-06	1.943172	0.0559
R-squared	0.502285	Mean dependent var	0.003975	
Adjusted R-squared	0.405507	S.D. dependent var	0.005800	
S.E. of regression	0.004472	Akaike info criterion	-7.826500	
Sum squared resid	0.001440	Schwarz criterion	-7.401344	
Log likelihood	355.4528	F-statistic	5.190073	
Durbin-Watson stat	2.036646	Prob(F-statistic)	0.000001	

Devido ao problema detectado quanto à homocedasticidade dos resíduos, fez-se a correção de acordo com o processo de White. Ao se compararem os resultados obtidos no procedimento da tabela 51 e com o processo da regressão da tabela 48, verifica-se que os erros padrões dos coeficientes angulares aumentam e os valores dos testes t reduzem mas, mesmo assim, os valores de Prob relativos às variáveis composição do endividamento, giro do ativo e liquidez corrente são inferiores a 0,01, evidenciando que essas variáveis são significativas.

A variável crescimento das vendas apresentou o valor Prob = 0,0595 indicando a presença de evidências para se rejeitar a hipótese nula de que essa variável é significativa no nível de 0,05.

Tabela 51: Processo de White para heterocedasticidade referente ao *cluster* 7

Dependent Variable: SCORE

Method: Least Squares

Date: 12/18/08 Time: 18:39

Sample: 1 87

Included observations: 87

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.015289	0.000784	-19.50376	0.0000
GIRO7	0.304387	0.020742	14.67486	0.0000
LIQCOR7	0.054444	0.018058	3.014993	0.0034
CRESVEN7	0.001584	0.000829	1.910746	0.0595
C	0.875887	0.039373	22.24605	0.0000
R-squared	0.882280	Mean dependent var		0.652540
Adjusted R-squared	0.876538	S.D. dependent var		0.184833
S.E. of regression	0.064945	Akaike info criterion		-2.574795
Sum squared resid	0.345865	Schwarz criterion		-2.433076
Log likelihood	117.0036	F-statistic		153.6425
Durbin-Watson stat	1.604932	Prob(F-statistic)		0.000000

Comparando os resultados obtidos, é possível obter as seguintes evidências:

Rejeita-se H_0 , de que três indicadores obtidos através da técnica Delphi não têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA. Portanto, os indicadores composição do endividamento, giro do ativo e liquidez corrente obtidos pelo modelo Delphi são significativos em um nível de significância inferior a 0,01.

Conforme as evidências expostas, o indicador crescimento das vendas não se mostrou significativo no nível de significância de 0,05, mesmo após a aplicação do processo de correção de White.

Não foram encontrados dados para rejeitar a hipótese nula de que os indicadores rentabilidade sobre o patrimônio líquido e margem líquida não sejam significativos no nível de significância de 5%.

3.25 Incluindo todas variáveis para regressão e análise dos resultados dos *clusters*

Na presente etapa do trabalho serão testados mais uma vez os modelos de regressão linear múltipla para cada agrupamento. Nesse caso, são incluídas como variáveis independentes não só os seis indicadores definidos pela técnica Delphi, mas todas as outras variáveis publicadas pela Edição de Melhores & Maiores.

Além dessas variáveis, são também incluídas variáveis *dummies* para testar diferença de comportamento devido a diversas situações específicas, para cada um dos seis indicadores e para cada um dos sete agrupamentos. Dependendo do tipo de empresa, pode haver diferença de comportamento em um ou mais indicadores em função do fato, por exemplo, de algumas empresas de um dado agrupamento possuírem ações na bolsa de valores ou não, ou então por serem auditadas pelas grandes empresas de auditoria consideradas como “*big four*” ou não. Sendo assim, foram criadas variáveis *dummies* para diversas situações e para cada um dos seis indicadores. A lista de todas as variáveis usadas na regressão encontra-se no início desse estudo.

3.25.1 Regressão e análise dos resultados do *cluster* 1: incluindo todas as variáveis

Inicialmente, para a análise dos resultados do *cluster* 1, utilizaram-se o programa SPSS e o método *stepwise* para fazer uma pré-seleção das variáveis do modelo. O método *stepwise* permite selecionar automaticamente as variáveis significativas com base no nível de significância desejado pelo pesquisador. Foi estabelecido que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

Dessa forma, usando o método *stepwise* e incluindo todas as variáveis anteriormente mencionadas, foram selecionadas as variáveis significativas em um nível de significância de 5%, conforme quadro 1.

Quadro 1: Variáveis selecionadas pelo método *stepwise*

Variável	Sigla
Composição do endividamento	COMPEND7
Crescimento das vendas	CRESVEN7
<i>Dummy</i> liquidez corrente das estatais	DESTLIQ
<i>Dummy</i> giro do ativo da indústria	DINGIR
<i>Dummy</i> margem líquida da indústria	DINMAR
Giro do ativo	GIRO7
Liquidez corrente	IQCOR7
Margem líquida	MARVEN7

A seguir, a fim de prosseguir com a análise de regressão, foi empregado o programa *Eviews*, o qual tem opções que possibilitam efetuar automaticamente os testes dos pressupostos básicos da regressão.

A tabela 52 apresenta os resultados da regressão em que a variável dependente é constituída pelos escores e as independentes são as variáveis que constam do Quadro 1. Analisando os

resultados, constata-se que as oito variáveis independentes incluídas no modelo são significativas em um nível de significância menor que 1%. Pode-se também notar que os valores Prob para as cinco variáveis são todos menores que 0,01.

Os sinais dos coeficientes angulares de todas as variáveis são coerentes como a variável composição do endividamento que tem sinal negativo, pois o aumento da composição do endividamento, isto é, o aumento das dívidas de curto prazo, acarreta redução no nível de eficiência.

Mais adiante serão feitos testes para avaliação da coerência dos sinais das variáveis *dummies* introduzidas.

O modelo de regressão como um todo é significativo (*Prob(F-statistic)*) e menor que 0,01. O poder explicativo da regressão (*Adjusted R-squared*) é de 0,8787 e pode ser considerado bom.

Tabela 52: Resultados da regressão para o *cluster* 1 com todas as variáveis

Dependent Variable: SCORE
Method: Least Squares
Date: 12/15/08 Time: 11:09
Sample: 1 148
Included observations: 148

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.734432	0.048227	15.22855	0.0000
COMPEN7	-0.008226	0.000283	-29.03800	0.0000
CRESVEN7	0.001855	0.000477	3.891757	0.0002
DESTLIQ	0.060272	0.021310	2.828409	0.0054
DINGIR	-0.040730	0.009668	-4.212983	0.0000
DINMAR	0.005065	0.001873	2.704402	0.0077
GIRO7	0.100836	0.007100	14.20218	0.0000
LIQCOR7	0.037002	0.007069	5.234387	0.0000
MARVEN7	0.012360	0.002992	4.130768	0.0001
R-squared	0.885341	Mean dependent var		0.656054
Adjusted R-squared	0.878741	S.D. dependent var		0.134980
S.E. of regression	0.047003	Akaike info criterion		-3.218321
Sum squared resid	0.307092	Schwarz criterion		-3.036058
Log likelihood	247.1557	F-statistic		134.1608
Durbin-Watson stat	1.911072	Prob(F-statistic)		0.000000

A figura 53 expõe os resultados do teste de normalidade dos resíduos, empregando o teste Jarque-Bera. O valor para a Probability = 0,0425 menor que 0,05 demonstra a presença de evidências para a rejeição da hipótese de normalidade dos resíduos.

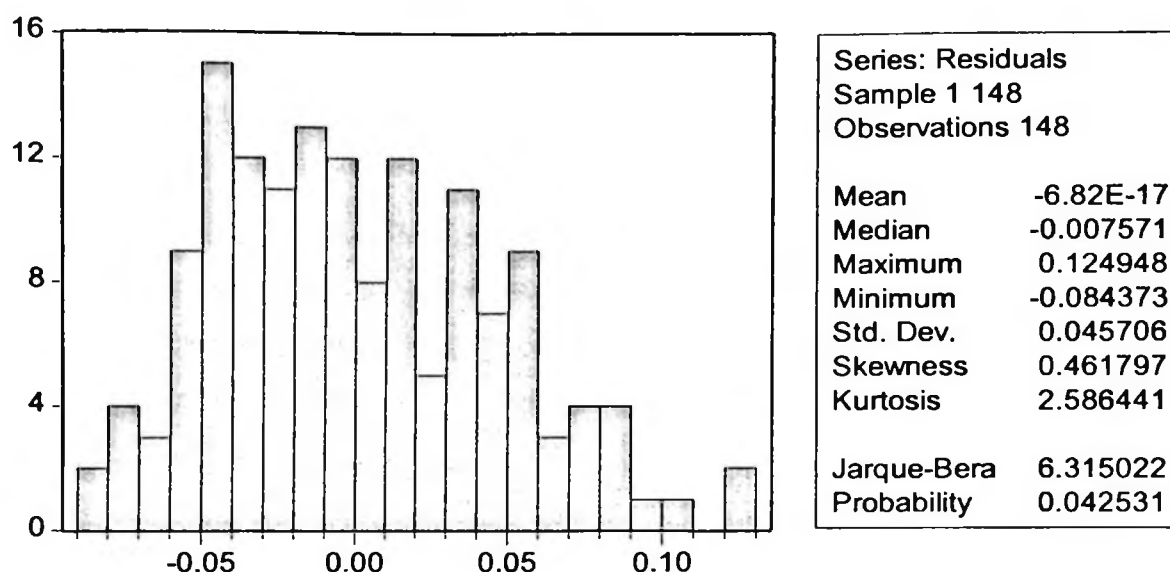


Figura 53: Teste da normalidade dos resíduos da regressão para o *cluster* 1 com todas as variáveis

A tabela 53 mostra os resultados da aplicação do teste de correlação serial dos resíduos da regressão. Segundo essa tabela, o teste Breusch_Godfrey indica *Probability* = 0,87, denotando que não se pode rejeitar a hipótese de ausência de correlação serial.

Tabela 53: Teste de correlação serial dos resíduos da regressão para o *cluster* 1

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.135104	Probability	0.873741
Obs*R-squared	0.291330	Probability	0.864447

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/15/08 Time: 21:57

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000348	0.048542	0.007159	0.9943
COMPEN7	-1.88E-05	0.000289	-0.065066	0.9482
CRESVEN7	-1.13E-05	0.000481	-0.023413	0.9814
DESTLIQ	-1.07E-05	0.021488	-0.000496	0.9996
DINGIR	0.000351	0.009786	0.035825	0.9715
DINMAR	-4.33E-05	0.001889	-0.022930	0.9817
GIRO7	0.000109	0.007211	0.015149	0.9879
LIQCOR7	5.30E-05	0.007116	0.007448	0.9941
MARVEN7	7.88E-05	0.003020	0.026084	0.9792
RESID(-1)	0.045159	0.087026	0.518917	0.6047
RESID(-2)	-0.004549	0.089080	-0.051070	0.9593
R-squared	0.001968	Mean dependent var	-6.82E-17	
Adjusted R-squared	-0.070881	S.D. dependent var	0.045706	
S.E. of regression	0.047298	Akaike info criterion	-3.193264	
Sum squared resid	0.306488	Schwarz criterion	-2.970498	
Log likelihood	247.3015	F-statistic	0.027021	
Durbin-Watson stat	1.998587	Prob(F-statistic)	1.000000	

A tabela 54 comprova os resultados do teste de heterocedasticidade de White. O valor *Probability* = 0,00367 indica que há evidências para se rejeitar a hipótese de homocedasticidade dos resíduos, em um nível de significância menor que 0,01.

Tabela 54: Teste de heterocedasticidade de White para os resíduos do *cluster* 1 com todas as variáveis
White Heteroskedasticity Test:

F-statistic	1.996230	Probability	0.003674
Obs*R-squared	55.53672	Probability	0.011291

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 12/15/08 Time: 21:57

Sample: 1 148

Included observations: 148

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.032471	0.024557	1.322280	0.1887
COMPEND7	-0.000460	0.000223	-2.060365	0.0417
COMPEND7^2	2.31E-06	1.03E-06	2.245800	0.0267
COMPEND7*CRESVEN7	-2.98E-07	2.07E-06	-0.143717	0.8860
COMPEND7*DESTLIQ	-8.72E-06	0.000162	-0.053959	0.9571
COMPEND7*DINGIR	-4.95E-06	4.58E-05	-0.108176	0.9140
COMPEND7*DINMAR	-2.77E-06	8.64E-06	-0.321144	0.7487
COMPEND7*GIRO7	-9.19E-06	3.41E-05	-0.269197	0.7883
COMPEND7*LIQCOR7	6.45E-05	3.27E-05	1.972104	0.0510
COMPEND7*MARVEN7	2.84E-06	1.22E-05	0.232602	0.8165
CRESVEN7	-0.000348	0.000428	-0.812214	0.4184
CRESVEN7^2	2.86E-06	1.98E-06	1.446292	0.1509
CRESVEN7*DESTLIQ	5.41E-05	0.000260	0.208571	0.8352
CRESVEN7*DINGIR	5.94E-05	7.31E-05	0.811763	0.4186
CRESVEN7*DINMAR	-1.45E-06	1.47E-05	-0.098894	0.9214
CRESVEN7*GIRO7	3.34E-05	5.26E-05	0.633971	0.5274
CRESVEN7*LIQCOR7	7.77E-05	4.73E-05	1.641385	0.1035
CRESVEN7*MARVEN7	-1.12E-05	2.45E-05	-0.456128	0.6492
DESTLIQ	-0.001244	0.012992	-0.095740	0.9239
DINGIR	0.007050	0.008266	0.852890	0.3955
DINGIR^2	0.000471	0.000943	0.499729	0.6182
DINGIR*DINMAR	-0.000800	0.000332	-2.410158	0.0176
DINGIR*LIQCOR7	0.000930	0.001281	0.726139	0.4693
DINMAR	-0.001559	0.001584	-0.983664	0.3274
DINMAR^2	0.000134	7.22E-05	1.854471	0.0663
DINMAR*LIQCOR7	-0.000116	0.000259	-0.450091	0.6535
GIRO7	-0.004226	0.006652	-0.635271	0.5265
GIRO7^2	0.000223	0.000618	0.361321	0.7185
GIRO7*LIQCOR7	-0.000194	0.000791	-0.244786	0.8071
GIRO7*MARVEN7	0.000147	0.000342	0.430404	0.6677
LIQCOR7	-0.006400	0.006312	-1.013935	0.3128
LIQCOR7^2	0.000705	0.000580	1.215960	0.2265
LIQCOR7*MARVEN7	-0.000270	0.000419	-0.644997	0.5202
MARVEN7	0.000970	0.002600	0.373147	0.7097
MARVEN7^2	-2.40E-05	9.16E-05	-0.262371	0.7935
R-squared	0.375248	Mean dependent var	0.002075	
Adjusted R-squared	0.187270	S.D. dependent var	0.002622	
S.E. of regression	0.002364	Akaike info criterion	-9.053696	
Sum squared resid	0.000632	Schwarz criterion	-8.344896	
Log likelihood	704.9735	F-statistic	1.996230	
Durbin-Watson stat	1.680157	Prob(F-statistic)	0.003674	

Em razão do problema detectado quanto à homocedasticidade dos resíduos, foi efetuada a correção segundo o processo de White. Comparando os resultados obtidos no procedimento da tabela 55 com o da regressão da tabela 52, verifica-se que os erros padrões dos coeficientes angulares aumentam e os valores dos testes t diminuem, porém, mesmo assim, os valores de Prob relativos a sete das oito variáveis do modelo são inferiores a 0,01, indicando que essas variáveis são significativas nesse nível de significância. A oitava variável, *dummy* da margem líquida da indústria, obteve um valor de Prob = 0,147, demonstrando-se que ela também é significativa, mas em um nível de 0,05.

Tabela 55: Processo de White para heterocedasticidade referente ao *cluster* 1

Dependent Variable: SCORE

Method: Least Squares

Date: 12/15/08 Time: 22:12

Sample: 1 148

Included observations: 148

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.008226	0.000323	-25.46228	0.0000
CRESVEN7	0.001855	0.000515	3.597657	0.0004
GIRO7	0.100836	0.006829	14.76502	0.0000
LIQCOR7	0.037002	0.008133	4.549852	0.0000
MARVEN7	0.012360	0.002848	4.339510	0.0000
DESTLIQ	0.060272	0.020473	2.944018	0.0038
DINGIR	-0.040730	0.010896	-3.738167	0.0003
DINMAR	0.005065	0.002051	2.469354	0.0147
C	0.734432	0.049842	14.73524	0.0000
R-squared	0.885341	Mean dependent var		0.656054
Adjusted R-squared	0.878741	S.D. dependent var		0.134980
S.E. of regression	0.047003	Akaike info criterion		-3.218321
Sum squared resid	0.307092	Schwarz criterion		-3.036058
Log likelihood	247.1557	F-statistic		134.1608
Durbin-Watson stat	1.911072	Prob(F-statistic)		0.000000

Comparando os dados obtidos, é possível conceber as seguintes evidências:

Rejeita-se H_0 , de que os indicadores obtidos da através da técnica Delphi não têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA. Portanto, os indicadores composição do endividamento, crescimento das vendas, giro do ativo, liquidez corrente e margem líquida são significativos em um nível de significância inferior a 0,01.

O indicador retorno sobre o patrimônio líquido não é significativo a um nível de significância de 0,05. Dos demais indicadores testados, mostram-se significativos em um nível de

significância de 0,01, os seguintes: *dummy* liquidez corrente das estatais, *dummy* giro do ativo da indústria, *dummy* margem líquida da indústria.

Entretanto, cabe salientar que foram encontradas evidências para rejeitar a hipótese de da normalidade dos resíduos em um nível de significância menor que 5%.

3.25.1.1 Cluster 1: Análise mais detalhada de outras variáveis significativas

3.25.1.1.1 Dummy Giro do ativo da indústria

A variável *Dummy* Giro do ativo da indústria (DINGIR) mostrou-se significativa na regressão. Esse fato sugere que o giro do ativo do ramo industrial é significativamente diferente dos ramos comercial e de serviços. A hipótese a ser testada é:

H_0 : não há diferença significativa no giro do ativo entre os ramos comercial (1), industrial (2) e de serviços.

Para avaliar essa hipótese, emprega-se inicialmente o teste ANOVA – Análise de variância, porém, para a aplicação desse teste, é necessário que sejam atendidos dois pressupostos: o da normalidade das variáveis e homogeneidade das variâncias.

Conforme resultados apresentados na Tabela 56 não há evidências para se rejeitar a hipótese da normalidade da variável para os três ramos de atividades, pois o valor *Sig* = 0,20, apresentado pelo teste Kolmogorov-Smirnov é maior que 0,05.

Tabela 56: Teste de normalidade

Tests of Normality

CODRAMO		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
GIRO7	1	,109	39	,200*	,953	39	,102
	2	,053	84	,200*	,991	84	,861
	3	,082	25	,200*	,959	25	,401

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

De acordo com resultados apresentados na tabela 57, rejeita-se a hipótese de homogeneidade das variâncias, pois o valor $Sig = 0,02$, exposto pelo teste de Levene, é menor que 0,05.

Tabela 57: Teste de Homogeneidade das variâncias

Test of Homogeneity of Variances			
GIRO7			
Levene Statistic	df1	Df2	Sig.
4,019	2	145	,020

A tabela 58 mostra o resultado do teste ANOVA. O valor $Sig = 0,000$ menor que 0,01 comprova que há evidências para rejeitar a hipótese nula de que não há diferença significativa no giro do ativo entre os três ramos de atividade.

Tabela 58: Teste ANOVA

ANOVA					
GIRO7					
	Sum of Squares	DF	Mean Square	F	Sig.
Between Groups	11,645	2	5,823	12,252	,000
Within Groups	68,911	145	,475		
Total	80,557	147			

Como o pressuposto da homogeneidade das variâncias não foi atendido, aplicaram-se os testes de Welch e de Brown-Forsythe, ambos robustos quanto à violação desse pressuposto. De acordo com a tabela 59, os dois testes originam um valor $Sig = 0,000$, indicando que há evidências para rejeitar a hipótese nula.

Tabela 59: Testes de Welch e de Brown-Forsythe

Robust Tests of Equality of Means				
GIRO7				
	Statistic ^a	df1	df2	Sig.
Welch	11,096	2	51,753	,000
Brown-Forsythe	9,818	2	68,189	,000

a. Asymptotically F distributed.

Além disso, foi aplicado também o teste de Kruskal-Wallis, o correspondente teste não paramétrico usado com a mesma finalidade. A tabela 60 mostra o valor *Asymp Sig* = 0,000, o que também leva à rejeição da hipótese nula.

Tabela 60: Teste de Kruskal-Wallis

Ranks			
CODRAMO		N	Mean Rank
GIRO7	1	39	90,26
	2	84	61,58
	3	25	93,32
	Total	148	

Test Statistics^{a,b}

	GIRO7
Chi-Square	17,714
DF	2
Asymp. Sig.	,000

a. Kruskal Wallis Test

b. Grouping Variable: CODRAMO

A tabela 61, com o teste de Tamhane, faz múltiplas comparações entre os giros médios dos três ramos. Os resultados demonstram que não há diferença significativa entre os ramos 1 e 3, apesar de haver diferenças significativas entre os ramos 1 e 2 e também entre os ramos 2 e 3, em um nível de significância menor que 0,01.

Tabela 61: Teste de Tamhane

Post Hoc Tests

Multiple Comparisons

Dependent Variable:GIRO7

	(I)	(J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tamhane 1.		2	,5324668*	,1368755	,001	,196460	,868474
		3	-,0802521	,2100267	,974	-,600565	,440061
2		1	-,5324668*	,1368755	,001	-,868474	-,196460
		3	-,6127190*	,1840892	,007	-1,077132	-,148306
3		1	,0802521	,2100267	,974	-,440061	,600565
		2	,6127190*	,1840892	,007	,148306	1,077132

*. The mean difference is significant at the 0.05 level.

Homogeneous Subsets

GIRO7

	CODRAMO	N	Subset for alpha = 0.05	
			1	2
Tukey B ^a	2	84	2,333869	
	1	39		2,866336
	3	25		2,946588

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 38,687.

Analizando os dados obtidos, é possível considerar as seguintes evidências:

Existem indícios para se rejeitar a hipótese nula de que não há diferença significativa entre o giro do ativo do ramo industrial e o dos outros dois ramos comercial e de serviços em um nível de significância de 1%. O giro do ativo do ramo industrial é significativamente menor que o dos outros dois ramos e por isso o coeficiente angular da regressão apresenta sinal negativo. Isso demonstra que giro do ativo do ramo industrial influi negativamente os escores de eficiência do *cluster* 1.

3.25.1.1.2 *Dummy* Margem líquida da indústria

A variável *Dummy* margem líquida da indústria (DINMAR) mostrou-se significativa na regressão. De acordo com tal fato, a margem líquida do ramo industrial é significativamente diferente dos ramos comercial e de serviços. Sendo assim, a hipótese a ser testada é:

H_0 : não há diferença significativa na margem líquida entre os ramos comercial (1), industrial (2) e de serviços.

Para testar essa hipótese, será utilizado o teste ANOVA – Análise de variância. Entretanto, para a aplicação desse teste, é necessário que sejam atendidos dois pressupostos: o da normalidade das variáveis e da homogeneidade das variâncias.

Conforme resultados apresentados na tabela 62, não há evidências para se rejeitar a hipótese da normalidade da variável para os três ramos de atividades, pois o valor Sig é maior que 0,05, de acordo com o teste Kolmogorov-Smirnov.

Tabela 62: Teste de normalidade

Tests of Normality						
CODRAMO		Kolmogorov-Smirnov ^a			Shapiro-Wilk	
		Statistic	df	Sig.	Statistic	df
MARVEN7	1	,064	39	,200*	,976	39
	2	,087	84	,175	,970	84
	3	,161	25	,095	,833	25

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Segundo os resultados apresentados na tabela 63, rejeita-se a hipótese de homogeneidade das variâncias, pois o valor Sig = 0,02, apresentado pelo teste de Levene, é menor que 0,05.

Tabela 63: Teste de Homogeneidade das variâncias

Test of Homogeneity of Variances			
MARVEN7			
Levene Statistic	df1	Df2	Sig.
4,045	2	145	,020

A tabela 64 expõe o resultado do teste ANOVA. O valor Sig = 0,000 menor que 0,01 indica a existência de evidências para rejeitar a hipótese nula de que não há diferença significativa no giro do ativo entre os três ramos de atividade.

Tabela 64: Teste ANOVA

MARVEN7	ANOVA				
	Sum of Squares	DF	Mean Square	F	Sig.
Between Groups	81,211	2	40,606	12,965	,000
Within Groups	454,113	145	3,132		
Total	535,324	147			

Como o pressuposto da homogeneidade das variâncias não foi atendido, aplicaram-se os testes de Welch e de Brown-Forsythe, ambos robustos quanto à violação desse pressuposto. De acordo com a tabela 65, os testes resultaram em um valor Sig = 0,000, sugerindo que há evidências para rejeitar a hipótese nula.

Tabela 65: Testes de Welch e de Brown-Forsythe

MARVEN7	Robust Tests of Equality of Means			
	Statistic ^a	df1	df2	Sig.
Welch	17,429	2	63,612	,000
Brown-Forsythe	15,318	2	91,053	,000

a. Asymptotically F distributed.

Além disso, foi aplicado também o teste de Kruskal-Wallis, o teste correspondente não paramétrico, usado com a mesma finalidade. A tabela 66 mostra o valor *Asymp Sig* = 0,000, o que também leva à rejeição da hipótese nula.

Tabela 66: Teste de Kruskal-Wallis

Ranks			
CODRAMO		N	Mean Rank
MARVEN7	1	39	45,97
	2	84	86,64
	3	25	78,22
	Total	148	

Test Statistics^{a,b}

	MARVEN7
Chi-Square	24,191
DF	2
Asymp. Sig.	,000

a. Kruskal Wallis Test

b. Grouping Variable: CODRAMO

A tabela 67, com o teste de Tamhane, faz múltiplas comparações entre os giros médios dos três ramos. Os resultados confirmam que não há diferença significativa entre os ramos 1 e 3, apesar de haver diferenças significativas entre os ramos 1 e 2 e também entre os ramos 2 e 3, em um nível de significância menor que 0,01.

Tabela 67: Teste de Tamhane

Post Hoc Tests

Multiple Comparisons

MARVEN7

Tamhane

(I) CODRAMO (J) CODRAMO		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-1,743049*	,300222	,000	-2,47143	-1,01467
	3	-1,308673*	,397410	,006	-2,29668	-,32066
2	1	1,743049*	,300222	,000	1,01467	2,47143
	3	,434376	,399634	,631	-,55652	1,42527
3	1	1,308673*	,397410	,006	,32066	2,29668
	2	-,434376	,399634	,631	-1,42527	,55652

*. The mean difference is significant at the 0.05 level.

Avaliando os resultados, é possível obter as seguintes evidências:

Existem indícios para se rejeitar a hipótese nula de não há diferença significativa na margem líquida entre os ramos comercial, industrial e de serviços em um nível de significância de 1%. A margem líquida do ramo industrial é significativamente maior que o dos outros dois ramos. O coeficiente angular da regressão apresenta sinal negativo, indicando que margem líquida do ramo industrial influi positivamente nos escores de eficiência do *cluster* 1.

3.25.1.1.3 *Dummy* Liquidez corrente das empresas estatais

Na regressão, a variável *Dummy* liquidez corrente das empresas estatais mostrou-se significativa. Conforme a tabela 68, a liquidez corrente média das estatais é 1,28 (*dummy* = 0) enquanto as demais empresas possuem índice médio igual a 1,54 (*dummy* = 1). Entretanto, o número de estatais que fazem parte do *cluster* 1 é muito pequeno (apenas 3 empresas) e portanto, não se pode fazer qualquer generalização.

Tabela 68: Liquidez corrente *cluster* 1

Group Statistics				
EST	N	Mean	Std. Deviation	Std. Error Mean
LIQCOR7 0	145	1,541314	,5643682	,0468682
1	3	1,283167	,6081964	,3511423

3.25.1.2 *Cluster* 1: Mais testes para a normalidade dos resíduos da regressão com variáveis adicionais

Na análise da regressão referente ao *cluster* 1, foi feita uma ressalva quanto à normalidade dos resíduos da regressão. A hipótese nula de normalidade dos resíduos foi rejeitada, pois o valor Probability = 0,042 é menor que 0,05.

Assim, o assunto foi melhor investigado, obtendo-se, dessa forma nova regressão com a inclusão de todas as variáveis selecionadas pela técnica Delphi e também as variáveis adicionais selecionadas pelo método *stepwise*. Os resultados mostram que a hipótese nula da normalidade dos resíduos continuou a ser rejeitada, considerando o nível de significância de 5%. O teste de normalidade utilizado foi o de Jarque-Bera.

Mais uma tentativa de resolver o problema da normalidade foi realizada. Nesse sentido, cada uma das variáveis mencionadas no parágrafo anterior foi transformada de acordo com a seguinte expressão geral: $\log(x + 1)$. O teste de normalidade empregado foi o de Jarque-Bera. Os resultados apontam que a hipótese nula da normalidade dos resíduos continuou a ser rejeitada, com o nível de significância de 5%.

Testou-se mais uma vez a normalidade dos resíduos, usando-se, nesse caso, o teste de Kolmogorov-Smirnov. Os resultados da tabela 69 expõem que a hipótese nula da normalidade dos resíduos não pode ser rejeitada, uma vez que o valor Sig = 0,06 é maior que 0,05.

Tabela 69: Teste de normalidade dos resíduos da regressão

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	DF	Sig.	Statistic	DF	Sig.
Resíduos	,071	148	,063	,974	148	,006

a. Lilliefors Significance Correction

Para Gujarati (2006, p.88), de acordo com o teorema central do limite, se houver um grande número de variáveis aleatórias independentes e com distribuição idêntica, a distribuição de sua soma tende à distribuição normal à medida que o número de variáveis aumenta indefinidamente.

Mesmo que o número de variáveis não seja muito grande ou que essas variáveis não sejam estritamente independentes, sua soma ainda pode ser normalmente distribuída, segundo Gujarati (2006, p.89).

Essa autora ainda salienta que, no caso de uma amostra de tamanho pequeno, ou finito, de menos de cem observações, a premissa da normalidade assume um papel fundamental porque permite recorrer aos testes t, F. Como em amostras pequenas esses testes exigem a normalidade, é necessário que essa premissa seja testada principalmente se o objetivo for não só a estimação, mas também a inferência (Gujarati, 2006, P.2001).

De acordo com Gujarati (2006, p.273), se os resíduos não forem distribuídos normalmente, é possível contar com o apoio do teorema central do limite e usar os testes t, F em grandes amostras.

Como observado nos parágrafos anteriores, para Gujarati, amostra grande é aquela com mais de cem observações. Entretanto, não há unanimidade entre os autores quanto à definição do que é uma amostra grande. Assim, por exemplo, segundo McClave *et al* (1998, p. 254), para a maior parte das populações amostradas, uma amostra de tamanho igual ou maior que 30 será suficiente para uma aproximação da normal ser considerada razoável. Conforme Hildebrand *et al* (1998, p. 241), um bom guia para o uso do teorema central do limite seria quando o tamanho da amostra excede 30 dados. Porém, esse autor afirma ainda que uma regra melhor seria a análise dos efeitos da assimetria. Se um histograma indicar uma curva com pequeno grau de assimetria pode ser aproximada de uma normal, mesmo com amostras de tamanho 15 ou 20 (Hildebrand, 1998, p. 242).

De acordo com Evans *et al* (1998, p.77), coeficientes de assimetria entre 0,5 e -0,5 podem ser considerados aproximadamente simétricos.

Nos testes dos resíduos da regressão, quando foram utilizadas todas as variáveis, obteve-se um grau de assimetria de 0,46, conforme visto anteriormente na figura 53, o que se enquadraria na regra de Evans. Além disso, deve-se analisar que o tamanho da amostra é de 148 casos, e, dessa forma, ela pode ser considerada grande, conforme Gujarati.

Dessa maneira, avaliando os dados, é possível obter as seguintes evidências:

No modelo do *cluster* 1 tudo foi testado e o único problema era não atender o pressuposto da normalidade. Entretanto, dadas todas as considerações e justificativas explanadas e, em

função dos testes efetuados, é possível considerar que as conclusões obtidas ao longo das avaliações sobre o *cluster* 1 aproximam-se razoavelmente da realidade.

3.25.2 Regressão e análise dos resultados do *cluster* 2: incluindo todas as variáveis

Como na análise dos resultados do *cluster* 1, foram empregados o programa SPSS e o método *stepwise* para fazer uma pré-seleção das variáveis do modelo. Instituiu-se que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

Utilizando o método *stepwise* e incluindo todas as variáveis – citadas na lista de variáveis –, verifica-se que nenhuma outra variável mostrou-se significativa além daquelas que já haviam sido consideradas como tal na análise da regressão, com a inclusão apenas das variáveis estabelecidas pela técnica Delphi.

Ao avaliar os dados obtidos, é possível destacar as seguintes evidências:

Os resultados do *cluster* 2, introduzindo-se todas as variáveis no modelo, demonstram que apenas os indicadores composição do endividamento, margem líquida e giro do ativo – obtidos pelo modelo Delphi – são significativos em um nível de significância inferior a 0,01.

3.25.3 Regressão e análise dos resultados do *cluster* 3: incluindo todas as variáveis

Como na análise dos resultados dos outros *clusters*, empregaram-se o programa SPSS e o método *stepwise* para fazer uma pré-seleção das variáveis do modelo. Estabeleceu-se que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

Utilizando o método *stepwise* e incluindo todas as variáveis – mencionadas na lista de variáveis, presente no início do trabalho –, são selecionadas aquelas significativas em um nível de significância de 5%, conforme quadro 2.

Quadro 2: Variáveis selecionadas pelo método *stepwise*

Variável	Sigla
Composição do endividamento	COMPEND7
Giro do ativo	GIRO7
Margem líquida	MARVEN7
Endividamento em longo prazo	ENDLP7
Dummy margem líquida do ramo indústria	DINMAR
Dummy giro do ativo do ramo serviços	DSEGIR
Dummy de liquidez corrente de empresas com ações	DACLIQ
Endividamento geral	ENDGE7
Rentabilidade sobre o patrimônio líquido	RENTAJU7

A seguir, a fim de prosseguir com a análise de regressão, utilizou-se o programa *Eviews*, o qual tem opções que possibilitam efetuar automaticamente os testes dos pressupostos básicos da regressão.

A tabela 70 explana os resultados da regressão em que a variável dependente é constituída pelos escores e as independentes são as variáveis que do Quadro 2 .

Ao verificar os resultados, nota-se que as nove variáveis independentes incluídas no modelo são significativas em um nível de significância menor que 1%. Pode-se constatar que os valores Prob para as cinco variáveis são todos menores que 0,01.

Os sinais dos coeficientes angulares de todas as variáveis são coerentes. A variável composição do endividamento tem sinal negativo, sendo, dessa forma, coerente, pois o aumento da composição do endividamento, isto é, o aumento das dívidas de curto prazo, ocasiona redução no nível de eficiência.

Mais adiante serão feitos testes para avaliação da coerência dos sinais das variáveis *dummies* introduzidas.

O modelo de regressão como um todo é significativo (*Prob(F-statistic)*) e menor que 0,01. O poder explicativo da regressão (*Adjusted R-squared*) é de 0,916, que pode ser considerado muito bom.

Tabela 70: Resultados da regressão para o *cluster* 3 com todas as variáveis

Dependent Variable: SCORE

Method: Least Squares

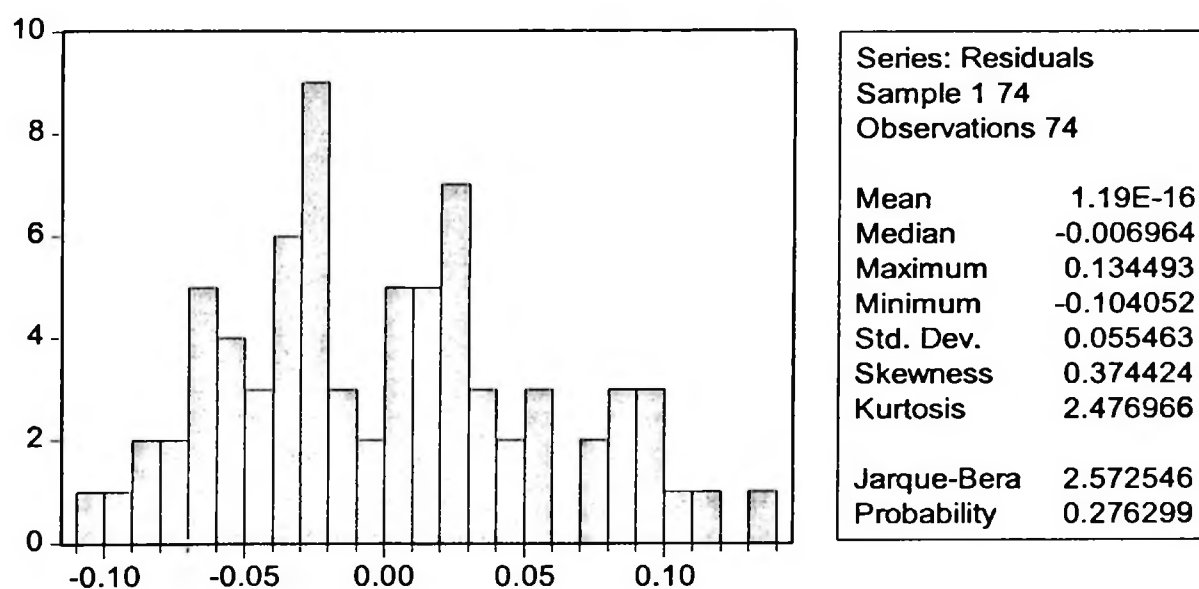
Date: 12/16/08 Time: 07:39

Sample: 1 74

Included observations: 74

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.003966	0.001068	-3.713523	0.0004
GIRO7	0.219863	0.022969	9.572251	0.0000
MARVEN7	0.027046	0.003505	7.715489	0.0000
ENDLP7	0.015426	0.003045	5.066127	0.0000
DINMAR	-0.006594	0.001382	-4.772900	0.0000
DSEGIR	-0.075749	0.024463	-3.096495	0.0029
DACLIQ	-0.029433	0.006672	-4.411532	0.0000
ENDGE7	-0.005650	0.001952	-2.893868	0.0052
RENTAJU7	-0.007360	0.002237	-3.290488	0.0016
C	0.247728	0.078748	3.145824	0.0025
R-squared	0.926760	Mean dependent var	0.484919	
Adjusted R-squared	0.916461	S.D. dependent var	0.204940	
S.E. of regression	0.059234	Akaike info criterion	-2.689553	
Sum squared resid	0.224555	Schwarz criterion	-2.378193	
Log likelihood	109.5135	F-statistic	89.98238	
Durbin-Watson stat	2.244051	Prob(F-statistic)	0.000000	

A figura 54 mostra os resultados do teste de normalidade dos resíduos. Foi empregado o teste Jarque-Bera, com o valor para a Probability = 0,276, sendo maior que 0,05, e indicando não haver evidências para a rejeição da hipótese de normalidade dos resíduos.

**Figura 54: Teste da normalidade dos resíduos da regressão para o *cluster* 3 com todas as variáveis**

A tabela 71 expõe os resultados da aplicação do teste de correlação serial dos resíduos da regressão. Conforme esta tabela, o teste Breusch Godfrey indica $Probability = 0,33$, e assim não se pode rejeitar a hipótese de ausência de correlação serial.

Tabela 71: Teste de correlação serial dos resíduos da regressão para o *cluster* 3

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.120575	Probability	0.332606
Obs*R-squared	2.581602	Probability	0.275050

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/16/08 Time: 07:42

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.000283	0.001094	-0.259026	0.7965
GIRO7	-0.001048	0.022948	-0.045652	0.9637
MARVEN7	-2.38E-05	0.003501	-0.006786	0.9946
ENDLP7	-0.001055	0.003177	-0.332070	0.7410
DINMAR	0.000194	0.001388	0.139865	0.8892
DSEGIR	0.002568	0.025012	0.102668	0.9186
DACLIQ	-0.000871	0.006736	-0.129339	0.8975
ENDGE7	0.000615	0.002050	0.300055	0.7651
RENTAJU7	-0.000134	0.002236	-0.060148	0.9522
C	0.015659	0.079431	0.197138	0.8444
RESID(-1)	-0.166275	0.132444	-1.255435	0.2140
RESID(-2)	-0.129351	0.140515	-0.920555	0.3609
R-squared	0.034887	Mean dependent var	1.19E-16	
Adjusted R-squared	-0.136343	S.D. dependent var	0.055463	
S.E. of regression	0.059123	Akaike info criterion	-2.671008	
Sum squared resid	0.216721	Schwarz criterion	-2.297376	
Log likelihood	110.8273	F-statistic	0.203741	
Durbin-Watson stat	1.945290	Prob(F-statistic)	0.996789	

A tabela 72 demonstra os resultados do teste de heterocedasticidade de White. O valor $Probability = 0,00417$ sugere a existência de evidências para se rejeitar a hipótese de homocedasticidade dos resíduos, a um nível de significância menor que 0,01.

Tabela 72: Teste de heterocedasticidade de White para os resíduos do *cluster 3* com todas as variáveis

White Heteroskedasticity Test:

F-statistic	3.849964	Probability	0.000417
Obs*R-squared	66.10202	Probability	0.063123

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 12/16/08 Time: 07:42

Sample: 1 74

Included observations: 74

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.032288	0.063814	0.505961	0.6177
COMPEND7	-0.001341	0.001590	-0.843390	0.4077
COMPEND7^2	1.44E-05	1.29E-05	1.116650	0.2757
COMPEND7*GIRO7	-0.000813	0.000562	-1.446049	0.1617
COMPEND7*MARVEN7	-0.000300	0.000112	-2.672132	0.0136
COMPEND7*ENDLP7	0.000119	8.16E-05	1.461304	0.1575
COMPEND7*DINMAR	5.45E-05	5.16E-05	1.056632	0.3017
COMPEND7*DSEGIR	-0.001762	0.001448	-1.217293	0.2358
COMPEND7*DACLQ	0.000505	0.000148	3.405676	0.0024
COMPEND7*ENDGE7	-6.21E-05	4.14E-05	-1.499691	0.1473
COMPEND7*RENTAJU7	0.000215	8.89E-05	2.419101	0.0239
GIRO7	0.055806	0.030753	1.814667	0.0826
GIRO7^2	-0.005095	0.004378	-1.163785	0.2564
GIRO7*MARVEN7	0.000622	0.002054	0.302691	0.7648
GIRO7*ENDLP7	-0.001882	0.001514	-1.243159	0.2263
GIRO7*DINMAR	-0.000975	0.001144	-0.851819	0.4031
GIRO7*DSEGIR	-0.049267	0.018531	-2.658602	0.0140
GIRO7*DACLQ	-0.001461	0.003171	-0.460581	0.6494
GIRO7*ENDGE7	0.000857	0.000910	0.941369	0.3563
GIRO7*RENTAJU7	0.000537	0.000917	0.585244	0.5641
MARVEN7	0.009242	0.006213	1.487456	0.1505
MARVEN7^2	-0.000450	0.000176	-2.558092	0.0176
MARVEN7*ENDLP7	-0.000654	0.000298	-2.192079	0.0388
MARVEN7*DINMAR	0.000300	0.000107	2.806968	0.0100
MARVEN7*DSEGIR	-0.005443	0.003751	-1.450979	0.1603
MARVEN7*DACLQ	-2.35E-05	0.000470	-0.049892	0.9606
MARVEN7*ENDGE7	0.000210	0.000185	1.133786	0.2686
MARVEN7*RENTAJU7	0.000760	0.000320	2.374600	0.0263
ENDLP7	0.001155	0.002344	0.492467	0.6271
ENDLP7^2	0.000267	0.000123	2.172564	0.0404
ENDLP7*DINMAR	-8.04E-05	0.000181	-0.443438	0.6616
ENDLP7*DSEGIR	-0.008883	0.005002	-1.775924	0.0890
ENDLP7*DACLQ	0.001432	0.000446	3.213373	0.0039
ENDLP7*ENDGE7	-0.000468	0.000186	-2.511040	0.0195
ENDLP7*RENTAJU7	0.000622	0.000250	2.483461	0.0207
DINMAR	0.010313	0.003768	2.737122	0.0117
DINMAR*DACLQ	-0.000454	0.000215	-2.115698	0.0454
DINMAR*ENDGE7	0.000176	0.000107	1.649902	0.1126
DINMAR*RENTAJU7	-0.000741	0.000246	-3.014528	0.0062

Tabela 72: Teste de heterocedasticidade de White para os resíduos do *cluster* 3 com todas as variáveis

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DSEGIR	0.386459	0.115858	3.335625	0.0029
DSEGIR*DACLIQ	-0.004501	0.007196	-0.625443	0.5378
DSEGIR*ENDGE7	0.005877	0.002641	2.225362	0.0361
DSEGIR*RENTAJU7	-0.007048	0.003657	-1.927189	0.0664
DACLIQ	-0.033527	0.010798	-3.105100	0.0050
DACLIQ^2	0.000764	0.001100	0.694733	0.4942
DACLIQ*ENDGE7	-0.000801	0.000260	-3.087418	0.0052
DACLIQ*RENTAJU7	0.000537	0.000361	1.489354	0.1500
ENDGE7^2	0.000171	6.70E-05	2.560321	0.0175
ENDGE7*RENTAJU7	-0.000328	0.000155	-2.117803	0.0452
RENTAJU7	-0.015762	0.004798	-3.284953	0.0032
RENTAJU7^2	3.88E-05	5.59E-05	0.693104	0.4952
R-squared	0.893271	Mean dependent var	0.003035	
Adjusted R-squared	0.661250	S.D. dependent var	0.003713	
S.E. of regression	0.002161	Akaike info criterion	-9.226613	
Sum squared resid	0.000107	Schwarz criterion	-7.638677	
Log likelihood	392.3847	F-statistic	3.849964	
Durbin-Watson stat	2.056593	Prob(F-statistic)	0.000417	

Devido ao problema detectado quanto à homocedasticidade dos resíduos, fez-se a correção conforme processo de White. Comparando os resultados do procedimento da tabela 73 com o da regressão da tabela 70, percebe-se que os erros padrões dos coeficientes angulares aumentam e os valores dos testes t diminuem, mas, mesmo assim, os valores de Prob relativos a sete das oito variáveis do modelo são inferiores a 0,01, evidenciando que essas variáveis são significativas nesse nível de significância.

Tabela 73: Processo de White para heterocedasticidade referente ao *cluster* 3

Dependent Variable: SCORE

Method: Least Squares

Date: 12/16/08 Time: 07:45

Sample: 1 74

Included observations: 74

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.003966	0.001046	-3.790219	0.0003
GIRO7	0.219863	0.024095	9.124661	0.0000
MARVEN7	0.027046	0.004087	6.617157	0.0000
ENDLP7	0.015426	0.002927	5.270939	0.0000
DINMAR	-0.006594	0.001490	-4.425204	0.0000
DSEGIR	-0.075749	0.018065	-4.193176	0.0001
DACLIQ	-0.029433	0.006314	-4.661886	0.0000
ENDGE7	-0.005650	0.001850	-3.054169	0.0033
RENTAJU7	-0.007360	0.002207	-3.334629	0.0014
C	0.247728	0.082568	3.000291	0.0038
R-squared	0.926760	Mean dependent var	0.484919	
Adjusted R-squared	0.916461	S.D. dependent var	0.204940	
S.E. of regression	0.059234	Akaike info criterion	-2.689553	
Sum squared resid	0.224555	Schwarz criterion	-2.378193	
Log likelihood	109.5135	F-statistic	89.98238	
Durbin-Watson stat	2.244051	Prob(F-statistic)	0.000000	

Avaliando os dados obtidos, é possível afirmar as seguintes evidências:

Rejeita-se H_0 , de que os indicadores obtidos da através da técnica Delphi não têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA. Portanto, os indicadores composição do endividamento, giro do ativo, margem líquida, rentabilidade sobre o patrimônio líquido são significativos em um nível de significância inferior a 0,01.

Dos demais indicadores e variáveis testados, são significativos em um nível de significância de 0,01, os seguintes: endividamento em longo prazo, *dummy* margem líquida do ramo indústria, *dummy* giro do ativo do ramo serviços, *dummy* de liquidez corrente de empresas com ações e endividamento geral.

3.25.3.1 Cluster 3: Análise detalhada das outras variáveis significativas

Algumas variáveis – endividamento em longo prazo, *dummy* margem líquida do ramo indústria, *dummy* giro do ativo do ramo serviços, *dummy* de liquidez corrente de empresas com ações e endividamento geral – mostram-se significativas na regressão e devem ser analisadas.

3.25.3.1.1 Endividamento em longo prazo e endividamento geral

De acordo com a tabela 74, há significativa correlação em um nível de 0,01 entre o endividamento em longo prazo e o endividamento geral. Além disso, ambas variáveis apresentam-se significativas na regressão.

A variável endividamento em longo prazo contribui com 1,9% do poder explicativo da regressão enquanto o endividamento geral contribui apenas com -0,2%.

Tabela 74: Correlação entre Endividamento a longo prazo e endividamento geral

Correlations		ENDGE7	ENDLP7
ENDGE7	Pearson Correlation	1,000	,896**
	Sig. (2-tailed)		,000
	N	74,000	74
ENDLP7	Pearson Correlation	,896**	1,000
	Sig. (2-tailed)	,000	
	N	74	74,000

** . Correlation is significant at the 0.01 level (2-tailed).

Através da tabela 74, percebe-se que o coeficiente de correlação de Pearson é de 0,896 e, portanto, significativo em um nível de 0,01. Isso ocasionaria colinearidade na regressão. O método *stepwise* utilizado na extração da regressão teria rejeitado uma dessas variáveis em função da colinearidade. Entretanto, isso não ocorreu provavelmente devido ao reduzido acréscimo no poder explicativo de uma dessas variáveis, o endividamento geral, com -0,2%.

3.25.3.1.2 *Dummy* margem líquida do ramo indústria

A variável *dummy* margem líquida do ramo indústria (DINMAR) é significativa na regressão. Esse fato indica que a margem líquida do ramo industrial constitui-se significativamente diferente dos ramos comercial e de serviços. A hipótese a ser testada é:

H_0 : não há diferença significativa na margem líquida entre os ramos comercial (1), industrial (2) e de serviços.

Para testar essa hipótese, é empregado a ANOVA – Análise de variância. No entanto, para a aplicação desse teste, é necessário que sejam atendidos dois pressupostos: o da normalidade das variáveis e homogeneidade das variâncias. Então, inicialmente serão atendidos tais pressupostos.

Conforme resultados apresentados na tabela 75, não há evidências para se rejeitar a hipótese da normalidade da variável para os ramos de atividade industrial e de serviços, pois os valores

Sig apresentados pelo teste Kolmogorov-Smirnov são maiores que 0,05. Já o ramo comercial apresenta um número de dados muito pequeno.

Tabela 75: Teste de normalidade

Tests of Normality							
CODRAMO		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
MARVEN7	1	,260	2	.			
	2	,097	55	,200*	,972	55	,220
	3	,188	17	,110	,914	17	,117

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

De acordo com resultados apresentados na tabela 76 não há evidências para se rejeitar a hipótese de homogeneidade das variâncias pois o valor Sig = 0,089, confirmado pelo teste de Levene, é maior que 0,05.

Tabela 76: Teste de Homogeneidade das variâncias

Test of Homogeneity of Variances			
MARVEN7			
Levene Statistic	df1	df2	Sig.
2,509	2	71	,089

A tabela 77 mostra o resultado do teste ANOVA. O valor Sig = 0,385, maior que 0,05, indica não haver evidências para rejeitar a hipótese nula de que não há diferença significativa no giro do ativo entre os três ramos de atividade.

Tabela 77: Teste ANOVA

ANOVA					
MARVEN7					
	Sum of Squares	DF	Mean Square	F	Sig.
Between Groups	33,989	2	16,994	,969	,385
Within Groups	1245,687	71	17,545		
Total	1279,676	73			

Com base nesses resultados, é possível constatar as seguintes evidências:

No *cluster* 3 não há evidências para se rejeitar a hipótese nula de que as margens dos três ramos de atividade não são significativamente diferentes em um nível de 0,05. Cabe salientar, entretanto, que nesse *cluster* o ramo comércio apresenta apenas 2 empresas, impossibilitando conclusões mais precisas.

3.25.3.1.3 *Dummy* giro do ativo do ramo serviços

A variável *dummy* giro do ativo do ramo serviços (DSEGIR) mostra-se significativa na regressão. Esse fato comprova que o giro do ativo do ramo serviços é significativamente diferente dos ramos comercial e industrial. A hipótese a ser testada é:

H_0 : não há diferença significativa no giro do ativo entre os ramos comercial (1), industrial (2) e de serviços (3).

Conforme resultados expostos na tabela 78, não há evidências para se rejeitar a hipótese da normalidade da variável para o ramo de atividades de serviços, pois o valor Sig é maior que 0,05, de acordo com o teste Kolmogorov-Smirnov. Para o ramo industrial há evidências para a rejeição da hipótese nula, pois o valor Sig é menor que 0,05. Já o ramo comercial apresenta um número de dados muito pequeno, não permitindo qualquer conclusão.

Tabela 78: Teste de Normalidade

CODRAMO		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	DF	Sig.	Statistic	DF	Sig.
GIRO7	1	,260	2	.			
	2	,139	55	,010	,953	55	,032
	3	,146	17	,200*	,936	17	,269

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

De acordo com resultados da tabela 79, não se rejeita a hipótese de homogeneidade das variâncias, pois o valor $Sig = 0,227$, confirmado pelo teste de Levene, é maior que 0,05.

Tabela 79: Teste de Homogeneidade das Variâncias
Test of Homogeneity of Variances

GIRO7

Levene Statistic	df1	DF2	Sig.
1,515	2	71	,227

O pressuposto da normalidade não foi atendido e, a partir disso, aplicou-se o teste de Kruskal-Wallis. A tabela 80 mostra que ambos resultaram num valor $Asymp Sig = 0,057$, confirmando não existirem evidências para rejeitar a hipótese nula.

Tabela 80: Teste de Kruskal-Wallis
Test Statistics^{a,b}

	GIRO7
Chi-Square	5,722
DF	2
Asymp. Sig.	,057

a. Kruskal Wallis Test

b. Grouping Variable: CODRAMO

Avaliando os resultados, é possível afirmar que:

Não há evidências para a rejeição da hipótese nula de que não há diferença significativa no giro entre os ramos de atividade comercial, industrial e de serviços. Entretanto, o setor de atividade comercial apresenta uma quantidade de dados muito pequena para conclusões mais precisas.

3.25.3.1.4 *Dummy* de liquidez corrente de empresas com ações

A variável *dummy* de liquidez corrente de empresas com ações (DACLIQ) mostra-se significativa na regressão. Esse fato comprova que a liquidez corrente de empresas com ações na bolsa é significativamente diferente das empresas sem ações. A hipótese a ser testada é:

H_0 : não há diferença significativa na liquidez corrente de empresas com ações na bolsa e sem ações na bolsa.

Conforme resultados da tabela 81, não há evidências para se rejeitar a hipótese da normalidade da variável para empresas com ações (1) e sem ações (0), pois o valor Sig para ambos os casos é de 0,20, maior que 0,05, de acordo com o teste Kolmogorov-Smirnov.

Tabela 81: Teste de normalidade

Tests of Normality						
AC	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	DF	Sig.	Statistic	Df	Sig.
LIQCOR7 0	,078	55	,200*	,968	55	,156
1	,122	19	,200*	,924	19	,133

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Atendeu-se o pressuposto da normalidade e, pela tabela 82, não se pode rejeitar a hipótese de igualdade de variâncias, pois o valor Sig = 0,06 do teste de Levene é maior que 0,05. Dessa maneira, aplica-se então o teste t para diferença de médias para amostras independentes.

A tabela 82 mostra que o valor Sig (2-tailed) para o teste t é de 0,019, menor que 0,05. Portanto, rejeita-se a hipótese nula de que o índice de liquidez corrente médio das empresas com ações não é significativamente diferente das empresas sem ações.

Tabela 82: Teste t para diferença de médias para amostras independentes

Independent Samples Test				
			LIQCOR7	
			Equal variances assumed	Equal variances not assumed
Levene's Test for Equality of Variances	F		3,661	
	Sig.		0,060	
t-test for Equality of Means	T		2,409	2,753
	DF		72,000	41,134
	Sig. (2-tailed)		0,019	0,009
	Mean Difference		0,314	0,314
	Std. Error Difference		0,131	0,114
	95% Confidence Interval of the Difference	Lower	0,054	0,084
		Upper	0,575	0,545

Tabela 83: Índice de liquidez corrente de empresas com ações (1) e sem ações (0)

Group Statistics				
	AC	N	Mean	Std. Deviation
LIQCOR7	1	19	2,549505	,3936404
	0	55	2,863925	,5186740
				Std. Error Mean
				,0903073
				,0699380

Considerando os resultados, é possível obter as seguintes evidências:

Segundo a tabela 83, o índice de liquidez corrente de empresas com ações na bolsa é significativamente menor (2,54) do que o das empresas sem ações (2,86), em um nível de significância de 0,05. O valor do coeficiente angular da regressão aparece com valor negativo, devido ao índice de liquidez corrente significativamente menor para as empresas com ações. Isso reflete negativamente nos escores de eficiência.

3.25.4 Regressão e análise dos resultados do *cluster* 4: incluindo todas as variáveis

Como na análise dos resultados dos outros *clusters*, empregaram-se o programa SPSS e o método *stepwise* para fazer uma pré-seleção das variáveis do modelo. Estabeleceu-se que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

Com o uso do método *stepwise* e inclusão de todas as variáveis – mencionadas na lista de variáveis, no início deste trabalho –, são selecionadas aquelas significativas em um nível de significância de 5% conforme quadro 3.

Quadro 3: Variáveis selecionadas pelo método *stepwise*

Variável	Sigla
Composição do endividamento	COMPEND7
Giro do ativo	GIRO7
Liquidez corrente	LIQCOR7
Endividamento em longo prazo	ENDLP7
Crescimento das vendas	CRESVEN7
Capital circulante líquido	CCL7

A seguir, foi utilizado o programa *Eviews* para continuar com a análise de regressão. Esse programa tem opções que possibilitam efetuar automaticamente os testes dos pressupostos básicos da regressão.

A tabela 84 demonstra os resultados da regressão em que a variável dependente é constituída pelos escores e as independentes são as variáveis do quadro 3 .

Analisando os resultados, verifica-se que as cinco primeiras variáveis independentes incluídas na tabela 84 são significativas em um nível de significância menor que 1%. Pode-se constatar que os valores Prob para as cinco variáveis são todos menores que 0,01. Já a sexta variável da Tabela Capital Circulante Líquido tem o valor Prob = 0,0156, comprovando que ela é significativa apenas em um nível de 0,05 e não no de 0,01.

Os sinais dos coeficientes angulares das cinco primeiras variáveis são coerentes, como a variável composição do endividamento, a qual tem sinal negativo, pois o aumento da composição do endividamento, isto é, o aumento das dívidas de curto prazo, acarreta redução no nível de eficiência. Cabe salientar que a sexta variável da Tabela Capital Circulante Líquido tem o sinal negativo, demonstrando que, quanto menor o valor desse tipo de conta, melhor é o escore de desempenho, fato ainda discutível.

Mais adiante neste trabalho serão feitos testes para avaliação da coerência dos sinais das variáveis *dummies* introduzidas.

O modelo de regressão como um todo é significativo (*Prob(F-statistic)*) e menor que 0,01. O poder explicativo da regressão (*Adjusted R-squared*) é de 0,9055, considerado muito bom.

Tabela 84: Resultados da regressão para o cluster 4 com todas as variáveis

Dependent Variable: SCORE

Method: Least Squares

Date: 12/16/08 Time: 08:02

Sample: 1 69

Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.005098	0.001059	-4.814857	0.0000
GIRO7	0.090763	0.009220	9.844508	0.0000
LIQCOR7	0.100200	0.015517	6.457413	0.0000
ENDLP7	0.004985	0.001383	3.605615	0.0006
CRESVEN7	0.002298	0.000654	3.513851	0.0008
CCL7	-1.38E-07	5.53E-08	-2.486509	0.0156
C	0.356323	0.106894	3.333429	0.0015
R-squared	0.913892	Mean dependent var		0.551826
Adjusted R-squared	0.905559	S.D. dependent var		0.163044
S.E. of regression	0.050106	Akaike info criterion		-3.053442
Sum squared resid	0.155655	Schwarz criterion		-2.826794
Log likelihood	112.3438	F-statistic		109.6709
Durbin-Watson stat	2.304701	Prob(F-statistic)		0.000000

A figura 55 mostra os resultados do teste de normalidade dos resíduos e a utilização do teste Jarque-Bera. O valor para a *Probability* = 0,00736, menor que 0,05 indica que há evidências para a rejeição da hipótese de normalidade dos resíduos.

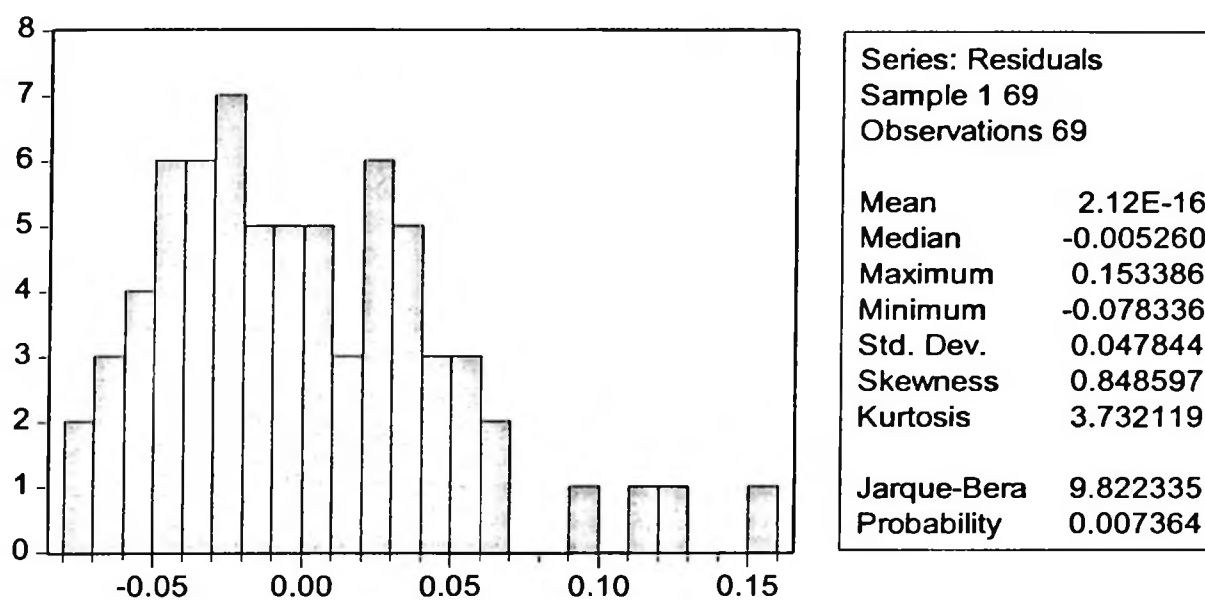


Figura 55: Teste da normalidade dos resíduos da regressão para o *cluster 4* com todas as variáveis

O teste de correlação serial de Breusch_Godfrey da Tabela 85 indica *Probability* = 0,0137, significando que há evidências para se rejeitar a hipótese de ausência de correlação serial, pois esse valor é menor que o nível de significância de 0,05.

Tabela 85: Teste de correlação serial dos resíduos da regressão para o *cluster* 4 com todas as variáveis**Breusch-Godfrey Serial Correlation LM Test:**

F-statistic	4.604348	Probability	0.013795
Obs*R-squared	9.180928	Probability	0.010148

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/16/08 Time: 08:06

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	0.000483	0.001015	0.475240	0.6363
GIRO7	0.001678	0.008774	0.191231	0.8490
LIQCOR7	0.003841	0.014768	0.260108	0.7957
ENDLP7	0.000912	0.001346	0.677444	0.5007
CRESVEN7	-0.000146	0.000621	-0.234517	0.8154
CCL7	2.04E-09	5.25E-08	0.038951	0.9691
C	-0.053583	0.102868	-0.520890	0.6044
RESID(-1)	-0.214327	0.122704	-1.746697	0.0858
RESID(-2)	-0.350518	0.128056	-2.737228	0.0081
R-squared	0.133057	Mean dependent var	2.12E-16	
Adjusted R-squared	0.017465	S.D. dependent var	0.047844	
S.E. of regression	0.047424	Akaike info criterion	-3.138253	
Sum squared resid	0.134944	Schwarz criterion	-2.846848	
Log likelihood	117.2697	F-statistic	1.151087	
Durbin-Watson stat	2.091871	Prob(F-statistic)	0.343605	

A tabela 86 expõe os resultados do teste de heterocedasticidade de White. O valor *Probability* = 0,000001 demonstra a existência de evidências para se rejeitar a hipótese de homocedasticidade dos resíduos, em um nível de significância menor que 0,01.

Tabela 86: Teste de heterocedasticidade de White para os resíduos do *cluster 4* com todas as variáveis

White Heteroskedasticity Test:

F-statistic	5.367678	Probability	0.000001
Obs*R-squared	53.78438	Probability	0.001617

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 12/16/08 Time: 08:07

Sample: 1 69

Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.015108	0.082463	0.183213	0.8555
COMPEND7	-0.000327	0.001734	-0.188446	0.8515
COMPEND7^2	4.03E-06	9.84E-06	0.409900	0.6840
COMPEND7*GIRO7	-0.000103	0.000107	-0.960501	0.3424
COMPEND7*LIQCOR7	3.24E-05	0.000181	0.179496	0.8584
COMPEND7*ENDLP7	3.53E-06	2.33E-05	0.151854	0.8800
COMPEND7*CRESVEN7	-1.53E-06	8.56E-06	-0.178874	0.8589
COMPEND7*CCL7	-2.45E-10	8.28E-10	-0.295768	0.7689
GIRO7	0.004674	0.010582	0.441655	0.6611
GIRO7^2	0.000641	0.000774	0.828153	0.4124
GIRO7*LIQCOR7	0.001826	0.002257	0.809173	0.4231
GIRO7*ENDLP7	-2.89E-05	0.000145	-0.199716	0.8427
GIRO7*CRESVEN7	-5.86E-05	6.55E-05	-0.894388	0.3763
GIRO7*CCL7	6.69E-09	9.03E-09	0.741090	0.4629
LIQCOR7	-0.014304	0.019278	-0.742023	0.4623
LIQCOR7^2	0.000613	0.001641	0.373845	0.7104
LIQCOR7*ENDLP7	0.000257	0.000253	1.013159	0.3169
LIQCOR7*CRESVEN7	1.84E-05	0.000146	0.126163	0.9002
LIQCOR7*CCL7	1.54E-08	1.16E-08	1.329204	0.1911
ENDLP7	-0.000104	0.002087	-0.049737	0.9606
ENDLP7^2	2.96E-06	1.29E-05	0.228747	0.8202
ENDLP7*CRESVEN7	-9.96E-06	1.31E-05	-0.758411	0.4525
ENDLP7*CCL7	-1.20E-10	1.04E-09	-0.115468	0.9086
CRESVEN7	1.95E-05	0.000850	0.022910	0.9818
CRESVEN7^2	4.02E-06	3.30E-06	1.216864	0.2306
CRESVEN7*CCL7	-1.78E-09	9.42E-10	-1.893114	0.0654
CCL7	7.68E-08	1.01E-07	0.758956	0.4522
CCL7^2	-3.83E-14	2.57E-14	-1.490138	0.1438
R-squared	0.779484	Mean dependent var	0.002256	
Adjusted R-squared	0.634266	S.D. dependent var	0.003756	
S.E. of regression	0.002272	Akaike info criterion	-9.045671	
Sum squared resid	0.000212	Schwarz criterion	-8.139077	
Log likelihood	340.0756	F-statistic	5.367678	
Durbin-Watson stat	1.846841	Prob(F-statistic)	0.000001	

Quanto ao problema detectado da correlação serial e também quanto à homocedasticidade dos resíduos, foi efetuada a correção, conforme o processo de Newey-West. Comparando os resultados obtidos a partir desse procedimento da tabela 87 com aqueles da regressão da tabela 84, verifica-se que os erros padrões dos coeficientes angulares aumentam e os valores dos testes t reduzem, no entanto, mesmo assim, os valores de Prob relativos às cinco primeiras

variáveis são inferiores a 0,05, mostrando que essas variáveis são significativas. Já a sexta variável Capital Circulante Líquido demonstra o valor Prob = 0,0566, superior a 0,05, e, consequentemente, não significativa após o processo de Newey-West.

Tabela 87: Processo de Newey-West para heterocedasticidade referente ao *cluster* 4

Dependent Variable: SCORE

Method: Least Squares

Date: 12/16/08 Time: 08:09

Sample: 1 69

Included observations: 69

Newey-West HAC Standard Errors & Covariance (lag truncation=3)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.005098	0.001454	-3.505701	0.0009
GIRO7	0.090763	0.011068	8.200740	0.0000
LIQCOR7	0.100200	0.021423	4.677159	0.0000
ENDLP7	0.004985	0.002189	2.276936	0.0263
CRESVEN7	0.002298	0.000629	3.651371	0.0005
CCL7	-1.38E-07	7.08E-08	-1.942490	0.0566
C	0.356323	0.150333	2.370219	0.0209
R-squared	0.913892	Mean dependent var		0.551826
Adjusted R-squared	0.905559	S.D. dependent var		0.163044
S.E. of regression	0.050106	Akaike info criterion		-3.053442
Sum squared resid	0.155655	Schwarz criterion		-2.826794
Log likelihood	112.3438	F-statistic		109.6709
Durbin-Watson stat	2.304701	Prob(F-statistic)		0.000000

A partir dessas análises, é possível afirmar as seguintes evidências:

Rejeita-se H_0 , de que os indicadores obtidos da através da técnica Delphi não têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA. Portanto, os indicadores a seguir relacionados são significativos em um nível de significância inferior a 0,01: composição do endividamento, giro do ativo, liquidez corrente, crescimento das vendas.

Dos demais indicadores e variáveis testados e relacionados na lista de variáveis, do início deste trabalho, mostrou-se significativo em um nível de significância de 0,05 o indicador endividamento em longo prazo com valor Prob = 0,0263.

Entretanto, cabe salientar que foram encontradas evidências para rejeitar a hipótese de da normalidade dos resíduos em um nível de significância menor que 5%.

3.25.4.1 *Cluster 4: Análise mais detalhada de outras variáveis significativas*

Nesse item, serão feitos comentários sobre as variáveis endividamento em longo prazo (ENDLP7) e capital circulante líquido 1 (CCL7), as quais demonstraram ser significativas na regressão.

3.25.4.1.1 Endividamento em longo prazo

A variável endividamento em longo prazo (ENDLP7) apresentou-se significativa na regressão quando foram incluídas todas as variáveis explicativas. Essa variável contribui com 2,5% do poder explicativo da regressão naquela ocasião.

A tabela 88 expõe uma regressão linear simples na qual a variável dependente é representada pelos valores dos escores do *cluster 4* e a variável independente é o endividamento em longo prazo.

Avaliando os resultados, nota-se que a variável endividamento em longo prazo, se considerada como única variável explicativa do comportamento dos escores, teria um poder explicativo de 0,40 ou 40% .

Tabela 88: Regressão: escores em função do endividamento a longo prazo

Dependent Variable: SCORE

Method: Least Squares

Date: 01/12/09 Time: 16:29

Sample: 1 69

Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.382041	0.029626	12.89557	0.0000
ENDLP7	0.007665	0.001145	6.693761	0.0000
R-squared	0.400750	Mean dependent var		0.551826
Adjusted R-squared	0.391806	S.D. dependent var		0.163044
S.E. of regression	0.127153	Akaike info criterion		-1.258294
Sum squared resid	1.083248	Schwarz criterion		-1.193537
Log likelihood	45.41115	F-statistic		44.80644
Durbin-Watson stat	1.813246	Prob(F-statistic)		0.000000

A figura 56 mostra o teste da normalidade dos resíduos da regressão. Como o valor Probability = 0,0239 é menor que 0,05, fica rejeitada a hipótese de normalidade dos resíduos.

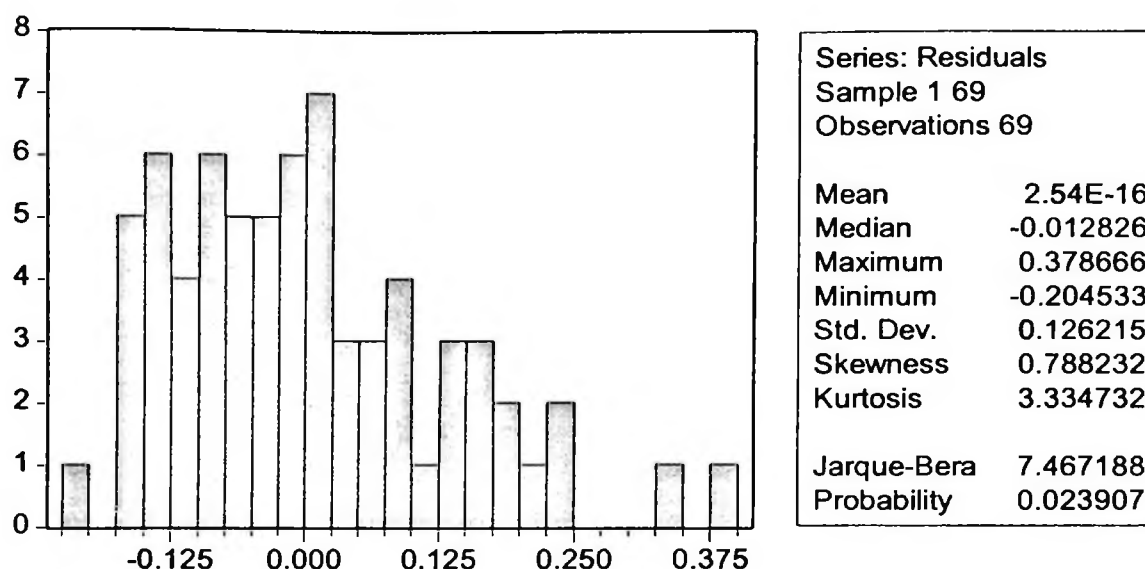


Figura 56: Teste de normalidade dos resíduos da regressão

Para corrigir o problema da normalidade dos resíduos, as variáveis foram transformadas. Utilizou-se $\log(\text{escores} + 1)$ e $\log(\text{endividamento a longo prazo} + 1)$. Os resultados da nova regressão são visualizados na tabela 89.

A variável independente continua significativa em um nível de 0,01. Mas o poder explicativo da regressão cai para 0,289.

Tabela 89: Regressão com as variáveis transformadas

Dependent Variable: LOG(SCORE+1)

Method: Least Squares

Date: 01/12/09 Time: 16:43

Sample: 1 69

Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.238639	0.038826	6.146402	0.0000
LOG(ENDLP7+1)	0.067257	0.012863	5.228792	0.0000
R-squared	0.289805	Mean dependent var		0.434227
Adjusted R-squared	0.279205	S.D. dependent var		0.101786
S.E. of regression	0.086416	Akaike info criterion		-2.030732
Sum squared resid	0.500337	Schwarz criterion		-1.965976
Log likelihood	72.06026	F-statistic		27.34027
Durbin-Watson stat	1.935341	Prob(F-statistic)		0.000002

A figura 57 demonstra que o problema da normalidade dos resíduos foi corrigida. Então, dessa forma, não se rejeita a hipótese de normalidade dos resíduos, pois o valor *Probability* = 0,189 é maior que 0,05.

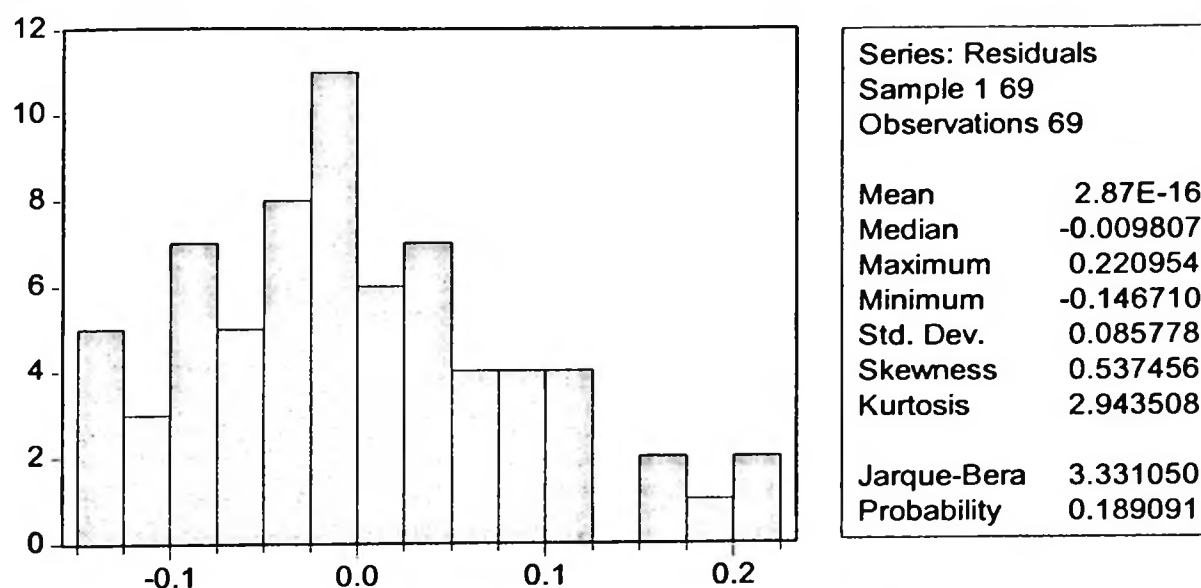


Figura 57: Teste de normalidade dos resíduos da regressão

A tabela 90 explana os resultados do teste de correlação serial. Não se rejeita a hipótese de ausência de correlação serial, pois o valor *Probability* = 0,857 é maior que 0,05.

Tabela 90: Teste e correlação serial dos resíduos da regressão

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.153633	Probability	0.857897
Obs*R-squared	0.324640	Probability	0.850169

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 01/12/09 Time: 16:45

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000504	0.039387	-0.012800	0.9898
LOG(ENDLP7+1)	0.000138	0.013051	0.010567	0.9916
RESID(-1)	0.026094	0.124037	0.210369	0.8340
RESID(-2)	0.063712	0.125944	0.505875	0.6147
R-squared	0.004705	Mean dependent var	2.87E-16	
Adjusted R-squared	-0.041232	S.D. dependent var	0.085778	
S.E. of regression	0.087529	Akaike info criterion	-1.977477	
Sum squared resid	0.497983	Schwarz criterion	-1.847964	
Log likelihood	72.22297	F-statistic	0.102422	
Durbin-Watson stat	2.005710	Prob(F-statistic)	0.958349	

A tabela 91 sugere a não rejeição da hipótese de ausência de heterocedasticidade dos resíduos da regressão, pois o valor *Probability* = 0,353 é maior que 0,05.

Tabela 91: Teste de heterocedasticidade dos resíduos da regressão

White Heteroskedasticity Test:

F-statistic	1.055964	Probability	0.353659
Obs*R-squared	2.139464	Probability	0.343101

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 01/12/09 Time: 16:46

Sample: 1 69

Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.013570	0.006969	1.947179	0.0558
LOG(ENDLP7+1)	-0.007684	0.005868	-1.309592	0.1949
(LOG(ENDLP7+1))^2	0.001759	0.001232	1.427823	0.1581
R-squared	0.031007	Mean dependent var		0.007251
Adjusted R-squared	0.001643	S.D. dependent var		0.010183
S.E. of regression	0.010175	Akaike info criterion		-6.295330
Sum squared resid	0.006833	Schwarz criterion		-6.198195
Log likelihood	220.1889	F-statistic		1.055964
Durbin-Watson stat	2.101128	Prob(F-statistic)		0.353659

Ao avaliar os resultados obtidos, é pertinente evidenciar que:

A variável endividamento em longo prazo tem influência significativa sobre os valores dos escores de eficiência do *cluster* 4, em um nível de significância de 1%.

3.25.4.1.2 Capital Circulante Líquido

A variável capital circulante líquido (CCL7) mostrou ser significativa na regressão. Essa variável contribui com 1,% do poder explicativo da regressão considerando todas as variáveis explicativas.

A tabela 92 expõe a regressão linear simples em que a variável dependente é representada pelos valores dos escores do *cluster* 4 e a variável independente é o capital circulante líquido.

Ao avaliar os resultados, verifica-se que a variável capital circulante líquido, se considerada como única variável explicativa do comportamento dos escores, teria um poder explicativo de apenas 0,0003 ou 0,03%. O valor do *Prob (F-statistic)* = 0,88 maior que 0,05 sugere que o modelo não é significativo. O valor *Prob* = 0,88 do coeficiente angular maior que 0,05 indica que a variável capital circulante líquido não é significativa.

Tabela 92: Regressão: escores em função do capital circulante líquido

Dependent Variable: SCORE
 Method: Least Squares
 Date: 01/12/09 Time: 16:47
 Sample: 1 69
 Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.550586	0.021514	25.59232	0.0000
CCL7	2.32E-08	1.59E-07	0.146144	0.8842
R-squared	0.000319	Mean dependent var		0.551826
Adjusted R-squared	-0.014602	S.D. dependent var		0.163044
S.E. of regression	0.164230	Akaike info criterion		-0.746536
Sum squared resid	1.807098	Schwarz criterion		-0.681780
Log likelihood	27.75550	F-statistic		0.021358
Durbin-Watson stat	1.898462	Prob(F-statistic)		0.884246

Sendo assim, a partir dos dados obtidos, é possível obter as seguintes evidências:

A variável capital circulante líquido, quando considerada como única variável explicativa dos escores, não é significativa no nível de 0,05.

3.25.4.2 Cluster 4: Mais testes para a normalidade dos resíduos da regressão com variáveis adicionais

Na análise da regressão referente ao *cluster* 4, foi feita uma ressalva quanto à normalidade dos resíduos da regressão. A hipótese nula de normalidade dos resíduos foi rejeitada.

Assim, o assunto foi melhor examinado. Foi obtida nova regressão com a inclusão de todas as variáveis que haviam sido selecionadas pela técnica Delphi e também as variáveis adicionais selecionadas pelo método *stepwise*. Os resultados demonstraram que a hipótese nula da normalidade dos resíduos continuou a ser rejeitada, considerando o nível de significância de 5%.

Mais uma tentativa de resolver o problema da normalidade foi realizada. Dessa vez, cada uma das variáveis mencionadas no parágrafo anterior foi transformada de acordo com a seguinte expressão geral: $\log(x + 1)$. Os resultados estão na tabela 93.

Percebe-se que as seguintes variáveis não são significativas: log (capital circulante líquido + 1), log (endividamento a longo prazo + 1), log (margem líquida + 1), pois os valores Prob de cada uma delas é maior que 0,05. Todas as outras cinco variáveis são significativas. O poder explicativo da regressão é de *Adjusted R-squared* = 0,964.

Tabela 93: Cluster 4: Regressão com log de todas as variáveis da técnica Delphi e das variáveis adicionais

Dependent Variable: LOG(SCORE+1)

Method: Least Squares

Date: 01/19/09 Time: 09:42

Sample: 1 69

Included observations: 51

Excluded observations: 18

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.970754	0.154566	6.280529	0.0000
LOG(CCL7+1)	-0.000772	0.003074	-0.251167	0.8029
LOG(COMPEND7+1)	-0.347265	0.017598	-19.73348	0.0000
LOG(CRESVEN7+1)	0.117650	0.016537	7.114160	0.0000
LOG(ENDLP7+1)	-0.004485	0.006210	-0.722196	0.4742
LOG(GIRO7+1)	0.173439	0.019607	8.845665	0.0000
LOG(LIQCOR7+1)	0.130763	0.023926	5.465255	0.0000
LOG(MARVEN7+1)	-0.054166	0.032960	-1.643399	0.1078
LOG(RENTAJU7+1)	0.084278	0.027745	3.037642	0.0041
R-squared	0.969931	Mean dependent var		0.449787
Adjusted R-squared	0.964204	S.D. dependent var		0.104981
S.E. of regression	0.019862	Akaike info criterion		-4.841197
Sum squared resid	0.016570	Schwarz criterion		-4.500286
Log likelihood	132.4505	F-statistic		169.3491
Durbin-Watson stat	2.293693	Prob(F-statistic)		0.000000

A figura 58 apresenta teste Jarque-Bera para normalidade dos resíduos. Como o valor *Probability* = 0,415 é maior que 0,05, não há evidências para a rejeição da hipótese da normalidade dos resíduos.

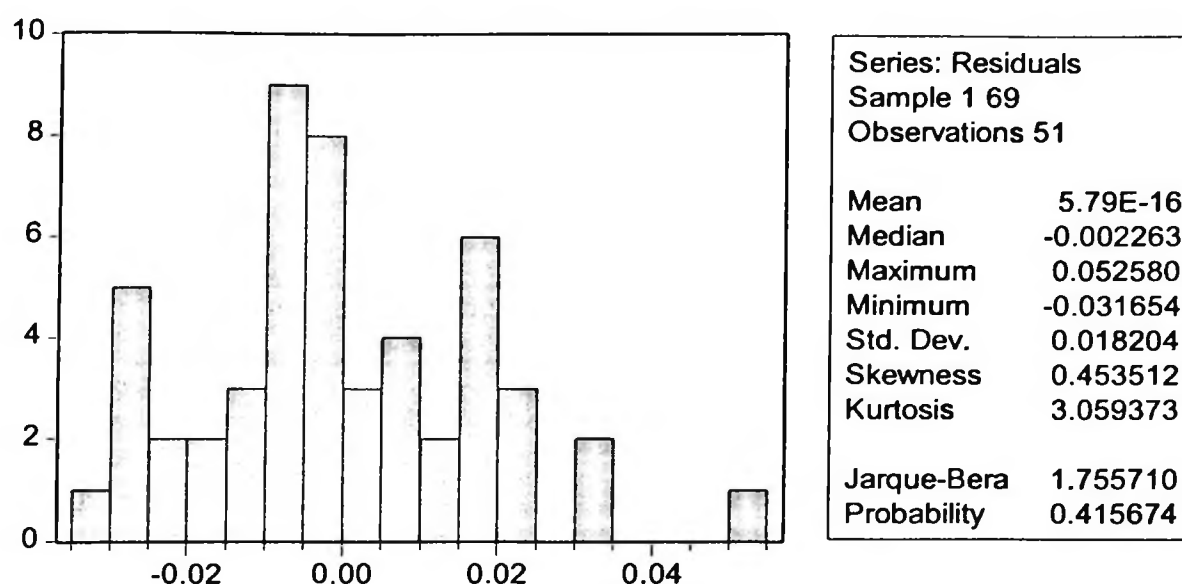


Figura 58: Normalidade dos resíduos da regressão

A tabela 94 expõe o teste de correlação serial de Breusch-Godfrey . Como o valor *Probability* = 0,50 é maior que 0,05, não se pode rejeitar a hipótese nula de ausência de correlação serial.

Tabela 94: Teste de correlação serial

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.399617	Probability	0.673224	
Obs*R-squared	0.999061	Probability	0.606815	
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Date: 01/19/09 Time: 09:47				
Presample and interior missing value lagged residuals set to zero.				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.014704	0.157779	0.093196	0.9262
LOG(CCL7+1)	0.000501	0.003159	0.158637	0.8748
LOG(COMPEN7+1)	-0.000285	0.018346	-0.015528	0.9877
LOG(CRESVEN7+1)	-0.002279	0.017105	-0.133267	0.8947
LOG(ENDLP7+1)	-0.001571	0.006598	-0.238042	0.8131
LOG(GIRO7+1)	0.000432	0.019991	0.021603	0.9829
LOG(LIQCOR7+1)	0.001944	0.025430	0.076452	0.9394
LOG(MARVEN7+1)	-0.002149	0.033598	-0.063947	0.9493
LOG(RENTAJU7+1)	-0.000361	0.028152	-0.012809	0.9898
RESID(-1)	-0.206025	0.213420	-0.965350	0.3402
RESID(-2)	-0.047317	0.213393	-0.221737	0.8256
R-squared	0.019589	Mean dependent var	5.79E-16	
Adjusted R-squared	-0.225513	S.D. dependent var	0.018204	
S.E. of regression	0.020153	Akaike info criterion	-4.782549	
Sum squared resid	0.016245	Schwarz criterion	-4.365881	
Log likelihood	132.9550	F-statistic	0.079923	
Durbin-Watson stat	2.038194	Prob(F-statistic)	0.999910	

A tabela 95 demonstra o teste de White para a heterocedasticidade dos resíduos. Como o valor *Probability* = 0,19 é maior que 0,05, não se rejeita a hipótese nula de homocedasticidade dos resíduos.

Tabela 95: Teste de heterocedasticidade dos resíduos

White Heteroskedasticity Test:

F-statistic	2.013182	Probability	0.192688
Obs*R-squared	47.76464	Probability	0.322378

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 01/19/09 Time: 09:48

Sample: 1 69

Included observations: 51

Excluded observations: 18

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.036904	0.512251	0.072043	0.9449
LOG(CCL7+1)	0.000128	0.013151	0.009720	0.9926
(LOG(CCL7+1))^2	0.000161	0.000147	1.095537	0.3153
(LOG(CCL7+1))*(LOG(COMPEN7+1))	-0.000670	0.002477	-0.270496	0.7958
(LOG(CCL7+1))*(LOG(CRESVEN7+1))	0.000612	0.001765	0.346916	0.7405
(LOG(CCL7+1))*(LOG(ENDLP7+1))	-0.000193	0.001362	-0.142043	0.8917
(LOG(CCL7+1))*(LOG(GIRO7+1))	0.000393	0.001214	0.324070	0.7569
(LOG(CCL7+1))*(LOG(LIQCOR7+1))	-0.001254	0.002197	-0.570802	0.5889
(LOG(CCL7+1))*(LOG(MARVEN7+1))	-0.002177	0.002734	-0.796275	0.4562
(LOG(CCL7+1))*(LOG(RENTAJU7+1))	0.001203	0.002245	0.535865	0.6113
LOG(COMPEN7+1)	-0.032963	0.118416	-0.278364	0.7901
(LOG(COMPEN7+1))^2	0.002746	0.010415	0.263647	0.8009
(LOG(COMPEN7+1))*(LOG(CRESVEN7+1))	0.002806	0.007370	0.380658	0.7166
(LOG(COMPEN7+1))*(LOG(ENDLP7+1))	0.000670	0.009745	0.068749	0.9474
(LOG(COMPEN7+1))*(LOG(GIRO7+1))	-0.008832	0.009674	-0.912986	0.3965
(LOG(COMPEN7+1))*(LOG(LIQCOR7+1))	0.011354	0.023358	0.486075	0.6442
(LOG(COMPEN7+1))*(LOG(MARVEN7+1))	-0.009853	0.016281	-0.605217	0.5672
(LOG(COMPEN7+1))*(LOG(RENTAJU7+1))	0.008992	0.009260	0.971061	0.3690
LOG(CRESVEN7+1)	0.032544	0.044799	0.726452	0.4949
(LOG(CRESVEN7+1))^2	-0.003105	0.002529	-1.227741	0.2655
(LOG(CRESVEN7+1))*(LOG(ENDLP7+1))	0.003403	0.004081	0.833762	0.4363
(LOG(CRESVEN7+1))*(LOG(GIRO7+1))	0.009740	0.004632	2.102686	0.0802
(LOG(CRESVEN7+1))*(LOG(LIQCOR7+1))	-0.012875	0.012331	-1.044116	0.3367
(LOG(CRESVEN7+1))*(LOG(MARVEN7+1))	0.002279	0.012985	0.175492	0.8665
(LOG(CRESVEN7+1))*(LOG(RENTAJU7+1))	-0.010733	0.008836	-1.214736	0.2701
LOG(ENDLP7+1)	0.005767	0.057462	0.100358	0.9233
(LOG(ENDLP7+1))^2	-0.000200	0.001327	-0.150952	0.8850
(LOG(ENDLP7+1))*(LOG(GIRO7+1))	-0.006748	0.007258	-0.929700	0.3884
(LOG(ENDLP7+1))*(LOG(LIQCOR7+1))	0.004873	0.007553	0.645193	0.5427
(LOG(ENDLP7+1))*(LOG(MARVEN7+1))	-0.006660	0.009799	-0.679709	0.5220
(LOG(ENDLP7+1))*(LOG(RENTAJU7+1))	0.001452	0.005558	0.261227	0.8027
LOG(GIRO7+1)	0.096686	0.084800	1.140167	0.2977
(LOG(GIRO7+1))^2	-0.007848	0.006274	-1.250798	0.2576
(LOG(GIRO7+1))*(LOG(LIQCOR7+1))	0.000132	0.008212	0.016032	0.9877
(LOG(GIRO7+1))*(LOG(MARVEN7+1))	-0.014933	0.019676	-0.758915	0.4766
(LOG(GIRO7+1))*(LOG(RENTAJU7+1))	-0.004732	0.015657	-0.302234	0.7727
LOG(LIQCOR7+1)	-0.040933	0.096371	-0.424749	0.6858
(LOG(LIQCOR7+1))^2	0.002802	0.006046	0.463363	0.6594

Tabela 95: Teste de heterocedasticidade dos resíduos

Variable	Coefficient	Std. Error	t-Statistic	Prob.
(LOG(LIQCOR7+1))*(LOG(MARVEN7+1))	-0.001324	0.015841	-0.083549	0.9361
(LOG(LIQCOR7+1))*(LOG(RENTAJU7+1))	0.010595	0.015147	0.699486	0.5104
LOG(MARVEN7+1)	0.151232	0.172292	0.877766	0.4138
(LOG(MARVEN7+1))^2	0.002693	0.020568	0.130925	0.9001
(LOG(MARVEN7+1))*(LOG(RENTAJU7+1))	-0.018723	0.028789	-0.650352	0.5396
LOG(RENTAJU7+1)	-0.155054	0.116607	-1.329709	0.2319
(LOG(RENTAJU7+1))^2	0.024222	0.015483	1.564462	0.1687
R-squared	0.936562	Mean dependent var	0.000325	
Adjusted R-squared	0.471347	S.D. dependent var	0.000471	
S.E. of regression	0.000342	Akaike info criterion	-13.49673	
Sum squared resid	7.03E-07	Schwarz criterion	-11.79218	
Log likelihood	389.1667	F-statistic	2.013182	
Durbin-Watson stat	2.029700	Prob(F-statistic)	0.192688	

A partir desses dados, é possível ressaltar as seguintes evidências:

Usando as variáveis transformadas de acordo com a expressão $\log(X + 1)$ para as variáveis obtidas pela técnica Delphi e também para as demais variáveis selecionadas através da técnica *stepwise* do programa SPSS, houve a possibilidade de se obterem resultados que permitissem atender os pressupostos da regressão quanto aos resíduos: presença de normalidade, ausência de autocorrelação serial e ausência de heterocedasticidade. Entretanto, algumas variáveis – endividamento em longo prazo e capital circulante líquido – tornaram-se não significativas em relação ao modelo apresentado anteriormente. Testes adicionais feitos anteriormente apontaram que o capital circulante líquido não é significativo. Por outro lado, a variável rentabilidade sobre o patrimônio líquido não se mostrava significativa no outro modelo, mas, de acordo com os novos resultados, é significativa.

3.25.5 Regressão e análise dos resultados do *cluster 5*: incluindo todas as variáveis

Como na análise dos resultados dos outros *clusters*, empregaram-se o programa SPSS e o método *stepwise* para fazer uma pré-seleção das variáveis do modelo. Foi instituído que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

Utilizando o método *stepwise* e incluindo todas as variáveis – mencionadas na lista de variáveis –, são selecionadas aquelas significativas em um nível de significância de 5% conforme quadro 4.

Quadro 4: Variáveis selecionadas pelo método *stepwise*

Variável	Sigla
Composição do endividamento	COMPEND7
Giro do ativo	GIRO7
Liquidez corrente	LIQCOR7
Rentabilidade sobre o patrimônio líquido	RENTAJU7
<i>Dummy</i> composição do endividamento de empresas nacionais	DNACCOM

Em seguida, a fim de prosseguir com a análise de regressão, foi usado o programa *Eviews*, pois este oferece alternativas que possibilitam efetuar automaticamente os testes dos pressupostos básicos da regressão.

A tabela 96 traz os resultados da regressão em que a variável dependente é constituída pelos escores e as independentes são as variáveis do quadro 4 .

Avaliando os resultados, constata-se que as quatro primeiras variáveis independentes incluídas na tabela são significativas em um nível de significância menor que 1%. Pode-se verificar que os valores Prob para as quatro variáveis são todos menores que 0,01. Já a quinta variável da tabela, *Dummy* composição do endividamento de empresas nacionais, tem o valor Prob = 0,0262, evidenciando que ela é significativa apenas em um nível de 0,05 e não no de 0,01.

Os sinais dos coeficientes angulares das cinco primeiras variáveis são coerentes, como, por exemplo, da variável composição do endividamento, que tem sinal negativo, pois o aumento da composição do endividamento, isto é, o aumento das dívidas de curto prazo, implica redução no nível de eficiência.

Mais adiante, serão realizados testes para avaliação da coerência dos sinais das variáveis *dummies* introduzidas

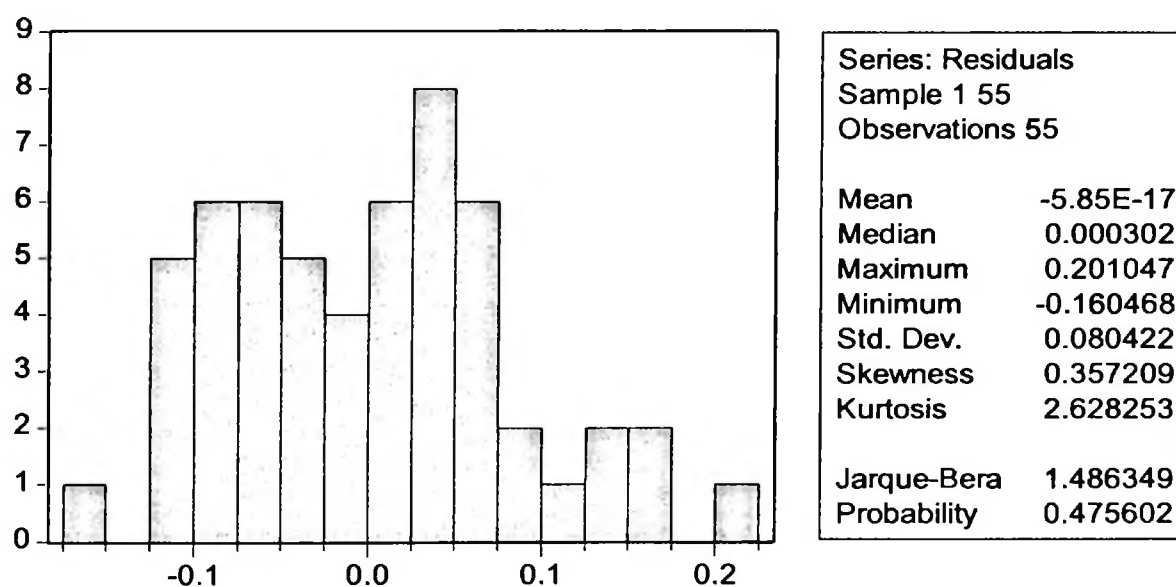
O modelo de regressão como um todo é significativo (*Prob(F-statistic)*) e menor que 0,01. O poder explicativo da regressão (*Adjusted R-squared*) é de 0,86, considerado bastante razoável.

Tabela 96: Resultados da regressão para o *cluster 5* com todas as variáveis

Dependent Variable: SCORE
 Method: Least Squares
 Date: 12/16/08 Time: 08:21
 Sample: 1 55
 Included observations: 55

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.009228	0.000563	-16.39961	0.0000
GIRO7	0.231947	0.054133	4.284748	0.0001
LIQCOR7	0.072292	0.023140	3.124062	0.0030
RENTAJU7	0.005797	0.002130	2.721925	0.0090
DNACCOM	-0.001003	0.000438	-2.292610	0.0262
C	0.646790	0.080371	8.047536	0.0000
R-squared	0.873307	Mean dependent var	0.632000	
Adjusted R-squared	0.860379	S.D. dependent var	0.225943	
S.E. of regression	0.084425	Akaike info criterion	-2.003227	
Sum squared resid	0.349255	Schwarz criterion	-1.784245	
Log likelihood	61.08874	F-statistic	67.55233	
Durbin-Watson stat	1.604974	Prob(F-statistic)	0.000000	

A figura 59 mostra os resultados do teste de normalidade dos resíduos, tendo sido utilizado o teste Jarque-Bera. O valor para a *Probability* = 0,475, maior que 0,05, indica que não há evidências para a rejeição da hipótese de normalidade dos resíduos.

**Figura 59: Teste da normalidade dos resíduos da regressão para o *cluster 5* com todas as variáveis**

A tabela 97 expõe o teste de correlação serial de Breusch_Godfrey, com *Probability* = 0,133, significando que não se pode rejeitar a hipótese de ausência de correlação serial, pois esse valor é superior a 0,05..

Tabela 97: Teste de correlação serial dos resíduos da regressão para o *cluster* 5 com todas as variáveis

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	2.102525	Probability	0.133490
Obs*R-squared	4.516699	Probability	0.104523

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/16/08 Time: 08:24

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-1.45E-05	0.000551	-0.026259	0.9792
GIRO7	-0.014406	0.053509	-0.269231	0.7889
LIQCOR7	0.006566	0.022866	0.287147	0.7753
RENTAJU7	0.000216	0.002090	0.103472	0.9180
DNACCOM	-1.21E-05	0.000428	-0.028148	0.9777
C	-0.005852	0.078798	-0.074264	0.9411
RESID(-1)	0.226898	0.149200	1.520762	0.1350
RESID(-2)	-0.251172	0.151212	-1.661054	0.1034
R-squared	0.082122	Mean dependent var	-5.85E-17	
Adjusted R-squared	-0.054583	S.D. dependent var	0.080422	
S.E. of regression	0.082588	Akaike info criterion	-2.016190	
Sum squared resid	0.320574	Schwarz criterion	-1.724214	
Log likelihood	63.44523	F-statistic	0.600722	
Durbin-Watson stat	1.831731	Prob(F-statistic)	0.752122	

A tabela 98 demonstra os resultados do teste de heterocedasticidade de White. O valor *Probability* = 0,259 sugere não haver evidências para se rejeitar a hipótese de homocedasticidade dos resíduos, em um nível de significância menor que 0,01.

Tabela 98: Teste de heterocedasticidade de White para os resíduos do *cluster 5* para todas as variáveis

White Heteroskedasticity Test:

F-statistic	1.276357	Probability	0.259047
Obs*R-squared	22.51098	Probability	0.259568

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 12/16/08 Time: 08:24

Sample: 1 55

Included observations: 55

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.011729	0.063430	-0.184911	0.8544
COMPEND7	0.000467	0.000591	0.789659	0.4350
COMPEND7^2	4.36E-07	3.25E-06	0.134256	0.8940
COMPEND7*GIRO7	-5.99E-05	0.000291	-0.205742	0.8382
COMPEND7*LIQCOR7	-0.000143	0.000138	-1.036216	0.3072
COMPEND7*RENTAJU7	-7.97E-06	1.11E-05	-0.719029	0.4769
COMPEND7*DNACCOM	5.11E-06	2.24E-06	2.277491	0.0290
GIRO7	-0.068599	0.060772	-1.128802	0.2667
GIRO7^2	-0.019570	0.026027	-0.751901	0.4571
GIRO7*LIQCOR7	0.007051	0.018712	0.376805	0.7086
GIRO7*RENTAJU7	0.001867	0.001336	1.397080	0.1712
GIRO7*DNACCOM	2.05E-05	0.000250	0.081964	0.9351
LIQCOR7	0.034553	0.026599	1.298998	0.2024
LIQCOR7^2	-0.006063	0.005133	-1.181171	0.2455
LIQCOR7*RENTAJU7	-0.000361	0.000611	-0.590270	0.5588
LIQCOR7*DNACCOM	-2.05E-05	0.000105	-0.194543	0.8469
RENTAJU7	0.000638	0.002987	0.213452	0.8322
RENTAJU7^2	-1.09E-05	3.32E-05	-0.329369	0.7438
RENTAJU7*DNACCOM	-7.26E-06	1.36E-05	-0.532916	0.5975
DNACCOM	-6.81E-05	0.000451	-0.151171	0.8807
R-squared	0.409290	Mean dependent var	0.006350	
Adjusted R-squared	0.088620	S.D. dependent var	0.008178	
S.E. of regression	0.007807	Akaike info criterion	-6.592345	
Sum squared resid	0.002133	Schwarz criterion	-5.862406	
Log likelihood	201.2895	F-statistic	1.276357	
Durbin-Watson stat	1.956357	Prob(F-statistic)	0.259047	

Avaliando os resultados, são plausíveis as seguintes evidências:

Rejeita-se H_0 , de que os indicadores obtidos da através da técnica Delphi não têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA. Portanto, os indicadores a seguir relacionados são significativos em um nível de significância inferior a 0,01: composição do endividamento, giro do ativo, liquidez corrente e rentabilidade sobre o patrimônio líquido.

Dos demais indicadores e variáveis testados e relacionados na lista de variáveis, comprovou-se significativo, em um nível de significância de 0,05, o indicador *dummy* composição do endividamento de empresas nacionais com valor Prob = 0,0262.

3.25.5.1 Cluster 5: Análise detalhada das outras variáveis significativas

A variável *dummy* composição do endividamento de empresas nacionais (DNACCOM) com valor Prob = 0,0262 confirmou ser significativa em um nível de significância de 0,05. Esse fato comprova que o indicador composição do endividamento de empresas nacionais é significativamente diferente quando o mesmo se refere às empresas estrangeiras. A hipótese a ser testada é:

H_0 : não há diferença significativa na composição do endividamento entre empresas nacionais e empresas estrangeiras.

Será utilizado o teste t para diferença de médias para amostras independentes. Mas, ao aplicar o teste, é necessário que sejam atendidos dois pressupostos: da normalidade das variáveis e da homogeneidade das variâncias.

Conforme resultados expostos na tabela 99, não há evidências para se rejeitar a hipótese da normalidade da variável composição do endividamento tanto para as empresas estrangeiras (0) com Sig = 0,200, quanto para as nacionais (1) com Sig = 0,148. Esses valores foram obtidos através do teste Kolmogorov-Smirnov e são maiores que 0,05.

Tabela 99: Teste de normalidade

Tests of Normality							
NAC		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
COMPEND7	0	,144	18	,200*	,908	18	,080
	1	,126	37	,148	,936	37	,035

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Na tabela 100, através do teste de Levene, nota-se que a hipótese nula de igualdade de variâncias não pode ser rejeitada, pois o valor Sig = 0,1 é maior que 0,05. A mesma tabela expõe os resultados do teste t de diferença de médias para amostras independentes. Como o valor *Sig (2-tailed)* = 0,022 é menor que 0,05, rejeita-se a hipótese nula de que não há diferença significativa entre as médias da composição do endividamento das empresas nacionais e estrangeiras.

Tabela 100: Teste de diferença de médias para amostras independentes

Independent Samples Test				
			COMPEND7	
			Equal variances assumed	Equal variances not assumed
Levene's Test for Equality of Variances	F		2,807	
	Sig.		0,100	
t-test for Equality of Means	T		2,363	2,165
	DF		53	27,315
	Sig. (2-tailed)		0,022	0,039
	Mean Difference		15,886	15,886
	Std. Error Difference		6,722	7,339
	95% Confidence Interval of the Difference	Lower	2,402	0,836
		Upper	29,370	30,936

Conforme tabela 101, pode-se constatar que o indicador composição do endividamento das empresas nacionais (1) apresenta média de 45,09, portanto, bem inferior ao mesmo indicador das estrangeiras (0), com média de 60,98.

Tabela 101: Estatística descritiva

Group Statistics					
	NAC	N	Mean	Std. Deviation	Std. Error Mean
COMPEND7	0	18	60,98	27,38	6,45
	1	37	45,09	21,25	3,49

Ao avaliar os resultados, é possível destacar as seguintes afirmações:

As empresas nacionais apresentam valor do indicador composição do endividamento médio significativamente menor que o das empresas estrangeiras. Na regressão, esse indicador é significativo em um nível de 0,05 e aparece com o sinal negativo, demonstrando que seu comportamento segue no sentido inverso ao do desempenho, de forma coerente do ponto de vista econômico.

3.25.6 Regressão e análise dos resultados do *cluster* 6: incluindo todas as variáveis

Como foi feito com os resultados dos outros *clusters*, foram utilizados o programa SPSS e o método *stepwise* para fazer uma pré-seleção das variáveis do modelo. Foi estabelecido que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

Usando o método *stepwise* e compreendendo todas as variáveis – mencionadas na lista de variáveis –, são selecionadas aquelas significativas em um nível de significância de 5% conforme quadro 5.

Quadro 5: Variáveis selecionadas pelo método *stepwise*

Variável	Sigla
Composição do endividamento	COMPEN7
Giro do ativo	GIRO7
Crescimento das vendas	CRESVEN7
Investimento no imobilizado	INVIMOB7
Rentabilidade sobre o patrimônio líquido	RENTAJU7
Endividamento geral	ENDGE7

A fim de efetuar a análise de regressão, empregou-se o programa *Eviews*, o qual possibilita a execução automática dos testes dos pressupostos básicos da regressão.

A tabela 102 traz os resultados da regressão em que a variável dependente é constituída pelos escores e as independentes são as variáveis que constam na tabela .

Analisando os resultados, verifica-se que todas as variáveis independentes incluídas na tabela, exceto a variável investimento no imobilizado, são significativas em um nível de significância menor que 1%. Com exceção desta, pode-se perceber que os valores Prob para todas as variáveis são todos menores que 0,01. Dessa forma, a variável investimento no imobilizado tem o valor Prob = 0,013, mostrando-se significativa apenas em um nível de 0,05 e não no de 0,01.

Os sinais dos coeficientes angulares das cinco primeiras variáveis são coerentes, como a variável composição do endividamento, a qual tem sinal negativo, pois o aumento da composição do endividamento, ou o aumento das dívidas de curto prazo, ocasiona redução no nível de eficiência.

Mais adiante serão feitos testes para avaliação da coerência dos sinais das variáveis *dummies* introduzidas

O modelo de regressão como um todo é significativo (*Prob(F-statistic)*) e menor que 0,01. O poder explicativo da regressão (*Adjusted R-squared*) é de 0,797, sendo considerado razoável.

Tabela 102: Resultados da regressão para o cluster 6 com todas as variáveis

Dependent Variable: SCORE

Method: Least Squares

Date: 12/16/08 Time: 08:57

Sample: 1 70

Included observations: 69

Excluded observations: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.008275	0.000576	-14.37243	0.0000
GIRO7	0.141074	0.016215	8.699960	0.0000
CRESVEN7	0.003300	0.000864	3.818622	0.0003
INVIMOB7	0.000490	0.000192	2.557601	0.0130
RENTAJU7	0.003669	0.001040	3.528748	0.0008
ENDGE7	-0.001597	0.000576	-2.773821	0.0073
C	0.718361	0.066257	10.84196	0.0000
R-squared	0.815231	Mean dependent var		0.608464
Adjusted R-squared	0.797350	S.D. dependent var		0.151681
S.E. of regression	0.068282	Akaike info criterion		-2.434420
Sum squared resid	0.289069	Schwarz criterion		-2.207772
Log likelihood	90.98749	F-statistic		45.59234
Durbin-Watson stat	1.913060	Prob(F-statistic)		0.000000

A figura 60 mostra os resultados do teste de normalidade dos resíduos. Empregou-se o teste Jarque-Bera, com o valor para a *Probability* = 0,0002, menor que 0,05, indicando a existência de evidências para a rejeição da hipótese de normalidade dos resíduos.

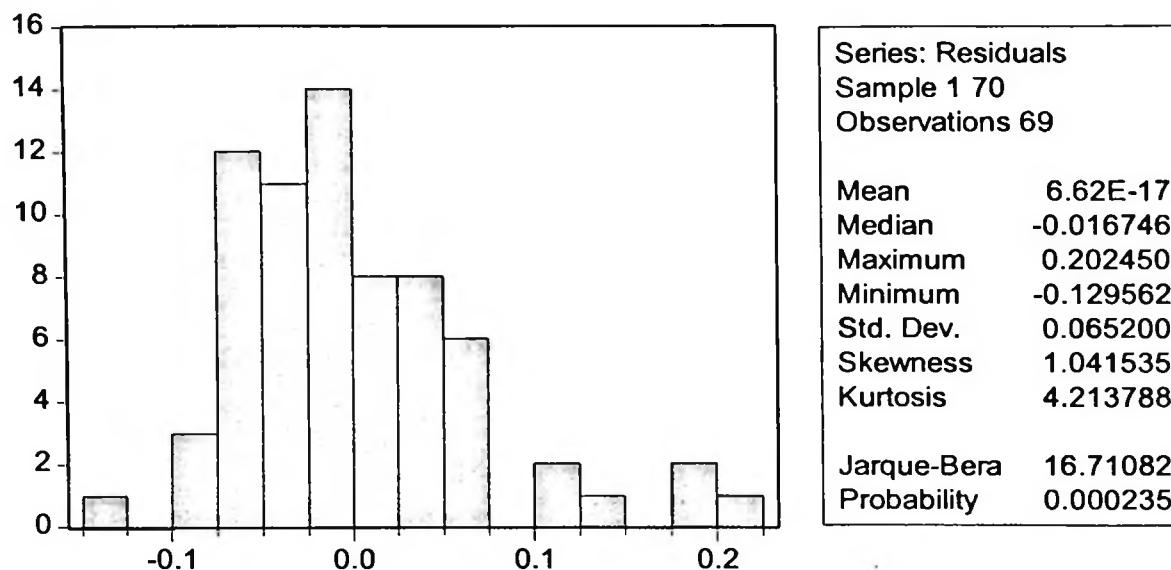


Figura 60: Teste da normalidade dos resíduos da regressão para o *cluster 6* com todas as variáveis

A tabela 103 expõe o teste de correlação serial dos resíduos da regressão de Breusch_Godfrey, com *Probability* = 0,88, sugerindo que não se pode rejeitar a hipótese de ausência de correlação serial em um nível de significância menor que 0,01..

Tabela 103: Teste de correlação serial dos resíduos da regressão para o *cluster 6* com todas as variáveis

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.126709	Probability	0.881225
Obs*R-squared	0.290204	Probability	0.864934

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/16/08 Time: 09:00

Presample and interior missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	3.74E-05	0.000595	0.062856	0.9501
GIRO7	-0.000135	0.016603	-0.008130	0.9935
CRESVEN7	5.96E-05	0.000886	0.067194	0.9467
INVIMOB7	-1.13E-05	0.000196	-0.057674	0.9542
RENTAJU7	-0.000119	0.001088	-0.109616	0.9131
ENDGE7	2.39E-05	0.000597	0.039988	0.9682
C	-0.003086	0.068471	-0.045070	0.9642
RESID(-1)	0.043666	0.140324	0.311178	0.7567
RESID(-2)	-0.047449	0.135385	-0.350474	0.7272
R-squared	0.004206	Mean dependent var	6.62E-17	
Adjusted R-squared	-0.128567	S.D. dependent var	0.065200	
S.E. of regression	0.069264	Akaike info criterion	-2.380664	
Sum squared resid	0.287853	Schwarz criterion	-2.089259	
Log likelihood	91.13290	F-statistic	0.031677	
Durbin-Watson stat	1.976613	Prob(F-statistic)	0.999988	

A tabela 104 expõe os resultados do teste de heterocedasticidade de White. O valor *Probability* = 0,00012, indica que há evidências para se rejeitar a hipótese de homocedasticidade dos resíduos, em um nível de significância menor que 0,01.

Tabela 104: Teste de heterocedasticidade de White para os resíduos do *cluster* 6 com todas as variáveis

White Heteroskedasticity Test:

F-statistic	3.561285	Probability	0.000126
Obs*R-squared	48.37366	Probability	0.006973

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 12/16/08 Time: 09:01

Sample: 1 70

Included observations: 69

Excluded observations: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.032270	0.055036	0.586338	0.5609
COMPEND7	-0.001299	0.000526	-2.470463	0.0177
COMPEND7^2	7.98E-06	3.34E-06	2.391271	0.0215
COMPEND7*GIRO7	-2.90E-05	0.000144	-0.201206	0.8415
COMPEND7*CRESVEN7	3.72E-06	8.06E-06	0.462155	0.6464
COMPEND7*INVIMOB7	1.86E-06	1.77E-06	1.052603	0.2987
COMPEND7*RENTAJU7	-1.74E-06	8.21E-06	-0.212354	0.8329
COMPEND7*ENDGE7	-7.69E-07	3.84E-06	-0.199949	0.8425
GIRO7	-0.006894	0.014311	-0.481737	0.6326
GIRO7^2	0.002319	0.002896	0.801017	0.4277
GIRO7*CRESVEN7	0.000187	0.000213	0.874296	0.3870
GIRO7*INVIMOB7	-8.62E-05	7.10E-05	-1.213539	0.2319
GIRO7*RENTAJU7	-0.000563	0.000244	-2.305955	0.0262
GIRO7*ENDGE7	0.000153	0.000112	1.368036	0.1788
CRESVEN7	0.001768	0.001033	1.712223	0.0944
CRESVEN7^2	-1.63E-05	8.59E-06	-1.897153	0.0649
CRESVEN7*INVIMOB7	-1.29E-07	3.46E-06	-0.037185	0.9705
CRESVEN7*RENTAJU7	4.96E-07	1.41E-05	0.035126	0.9722
CRESVEN7*ENDGE7	-2.74E-06	5.84E-06	-0.469684	0.6411
INVIMOB7	-0.000386	0.000376	-1.024947	0.3114
INVIMOB7^2	5.26E-08	3.13E-07	0.168079	0.8673
INVIMOB7*RENTAJU7	6.75E-06	6.66E-06	1.013555	0.3167
INVIMOB7*ENDGE7	2.67E-06	3.04E-06	0.880272	0.3838
RENTAJU7	-6.98E-05	0.000863	-0.080889	0.9359
RENTAJU7^2	9.67E-06	1.18E-05	0.819902	0.4170
RENTAJU7*ENDGE7	3.60E-06	8.11E-06	0.444330	0.6591
ENDGE7	-0.001050	0.000537	-1.954444	0.0575
ENDGE7^2	7.19E-06	3.41E-06	2.106656	0.0413
R-squared	0.701067	Mean dependent var	0.004189	
Adjusted R-squared	0.504210	S.D. dependent var	0.007565	
S.E. of regression	0.005327	Akaike info criterion	-7.341007	
Sum squared resid	0.001163	Schwarz criterion	-6.434414	
Log likelihood	281.2648	F-statistic	3.561285	
Durbin-Watson stat	1.960986	Prob(F-statistic)	0.000126	

Em razão do problema detectado quanto à homocedasticidade dos resíduos foi executada a correção conforme o processo de White. Na comparação dos resultados obtidos no procedimento da tabela 105 com o da regressão da tabela 102, constata-se que os erros padrões dos coeficientes angulares aumentam e os valores dos testes t diminuem, mas, mesmo assim, os valores de Prob relativos a todas as variáveis, exceto a variável endividamento geral, são inferiores a 0,01, demonstrando que essas variáveis são significativas. A variável endividamento geral também se revela significativa, pois seu valor Prob = 0,044 é inferior aos 0,05.

Tabela 105: Processo de White para heterocedasticidade referente ao *cluster* 6

Dependent Variable: SCORE

Method: Least Squares

Date: 12/16/08 Time: 09:02

Sample: 1 70

Included observations: 69

Excluded observations: 1

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.008275	0.000845	-9.789727	0.0000
GIRO7	0.141074	0.021047	6.702695	0.0000
CRESVEN7	0.003300	0.000656	5.029393	0.0000
INVIMOB7	0.000490	0.000147	3.332008	0.0015
RENTAJU7	0.003669	0.001281	2.863104	0.0057
ENDGE7	-0.001597	0.000777	-2.055749	0.0440
C	0.718361	0.081233	8.843253	0.0000
R-squared	0.815231	Mean dependent var		0.608464
Adjusted R-squared	0.797350	S.D. dependent var		0.151681
S.E. of regression	0.068282	Akaike info criterion		-2.434420
Sum squared resid	0.289069	Schwarz criterion		-2.207772
Log likelihood	90.98749	F-statistic		45.59234
Durbin-Watson stat	1.913060	Prob(F-statistic)		0.000000

Avaliando os resultados, são consideradas as seguintes evidências:

Rejeita-se H_0 , de que os indicadores obtidos da através da técnica Delphi não têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA. Portanto, os indicadores composição do endividamento, giro do ativo, crescimento das vendas e rentabilidade sobre o patrimônio líquido são significativos em um nível de significância inferior a 0,01.

Dos demais indicadores e variáveis testados e relacionados na lista de variáveis, presente no início do trabalho, constituem-se significativas, em um nível de significância de 0,05, as variáveis investimento no imobilizado e endividamento geral. Entretanto, cabe salientar que

foram encontradas evidências para rejeitar a hipótese da normalidade dos resíduos em um nível de significância menor que 5%.

3.25.6.1 *Cluster 6: Análise detalhada das outras variáveis significativas*

Nesse item, serão feitos comentários sobre as variáveis investimento no imobilizado (INVIMOB7) e endividamento geral (ENDGE7).

3.25.6.1.1 *Variável Investimento no Imobilizado*

Essa variável demonstra ser significativa na regressão e contribui com 0,034 do poder explicativo da regressão considerando todas as variáveis explicativas.

A tabela 106 apresenta a regressão linear simples na qual a variável dependente é representada pelos valores dos escores do *cluster 6* e a variável independente é investimento no imobilizado.

Segundo os resultados, a variável investimento no imobilizado, se considerada como única variável explicativa do comportamento dos escores, teria um poder explicativo de apenas 0,089 ou 8,9%, considerado muito reduzido. O valor do *Prob (F-statistic)* = 0,0125 é menor que 0,05 e indica que o modelo é significativo. O valor *Prob* = 0,0125 do coeficiente angular menor que 0,05 mostra que a variável investimento no imobilizado é significativa.

Tabela 106: Regressão: escores em função da variável investimento no imobilizado

Dependent Variable: SCORE

Method: Least Squares

Date: 01/16/09 Time: 17:42

Sample: 1 70

Included observations: 69

Excluded observations: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.575348	0.021785	26.41016	0.0000
INVIMOB7	0.000956	0.000372	2.566649	0.0125
R-squared	0.089522	Mean dependent var		0.608464
Adjusted R-squared	0.075932	S.D. dependent var		0.151681
S.E. of regression	0.145809	Akaike info criterion		-0.984484
Sum squared resid	1.424434	Schwarz criterion		-0.919727
Log likelihood	35.96470	F-statistic		6.587689
Durbin-Watson stat	2.054182	Prob(F-statistic)		0.012509

A figura 61 explana o teste da normalidade dos resíduos da regressão. Como o valor *Probability* = 0,008 é menor que 0,05, fica rejeitada a hipótese de normalidade dos resíduos.

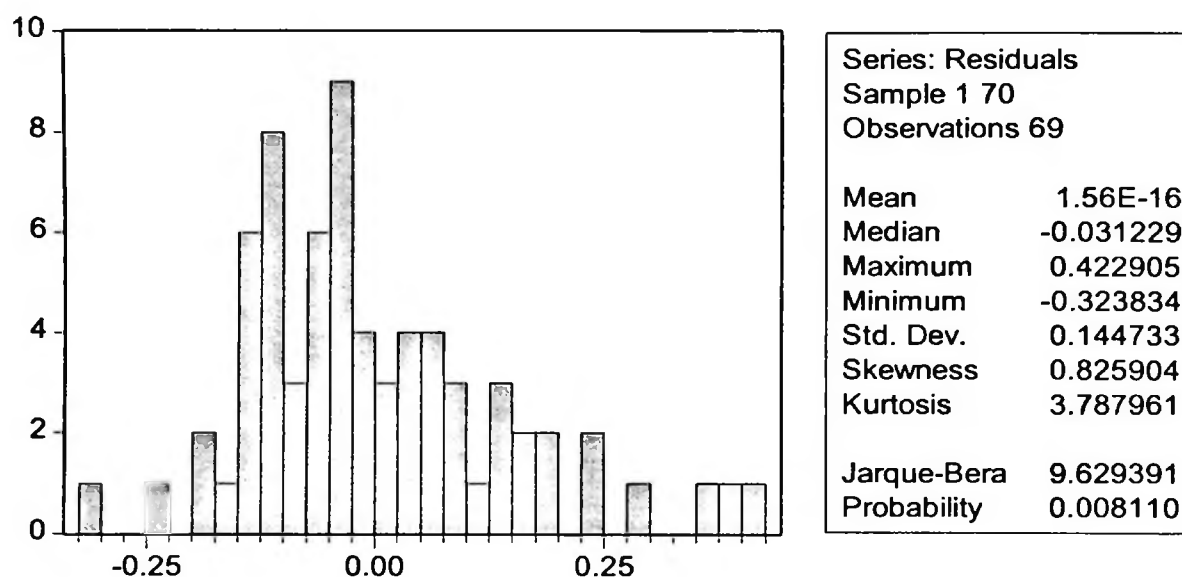


Figura 61: Teste de normalidade dos resíduos da regressão

A Tabela 107 exhibe os resultados do teste de correlação serial. Não se rejeita a hipótese de ausência de correlação serial, pois o valor *Probability* = 0,639 é maior que 0,05.

Tabela 107: Teste e correlação serial dos resíduos da regressão

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.450885	Probability	0.639041
Obs*R-squared	0.944166	Probability	0.623702

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 01/16/09 Time: 17:45

Presample and interior missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000438	0.022002	0.019900	0.9842
INVIMOB7	-1.30E-05	0.000378	-0.034352	0.9727
RESID(-1)	-0.024780	0.124998	-0.198241	0.8435
RESID(-2)	0.130377	0.123821	1.052952	0.2963
R-squared	0.013684	Mean dependent var	1.56E-16	
Adjusted R-squared	-0.031839	S.D. dependent var	0.144733	
S.E. of regression	0.147019	Akaike info criterion	-0.940291	
Sum squared resid	1.404942	Schwarz criterion	-0.810778	
Log likelihood	36.44005	F-statistic	0.300590	
Durbin-Watson stat	1.973987	Prob(F-statistic)	0.824846	

Segundo a tabela 108, não se rejeita a hipótese de ausência de heterocedasticidade dos resíduos da regressão, pois o valor *Probability* = 0,734 é maior que 0,05.

Tabela 108: Teste de heterocedasticidade dos resíduos da regressão

White Heteroskedasticity Test:

F-statistic	0.309705	Probability	0.734724
Obs*R-squared	0.641544	Probability	0.725589

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 01/16/09 Time: 17:46

Sample: 1 70

Included observations: 69

Excluded observations: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.016230	0.007069	2.295972	0.0249
INVIMOB7	0.000185	0.000236	0.786312	0.4345
INVIMOB7^2	-5.86E-07	8.20E-07	-0.714489	0.4774
R-squared	0.009298	Mean dependent var		0.020644
Adjusted R-squared	-0.020724	S.D. dependent var		0.034722
S.E. of regression	0.035080	Akaike info criterion		-3.819862
Sum squared resid	0.081220	Schwarz criterion		-3.722727
Log likelihood	134.7852	F-statistic		0.309705
Durbin-Watson stat	2.334352	Prob(F-statistic)		0.734724

A partir dos dados obtidos, são plausíveis as seguintes evidências:

A variável investimento no imobilizado tem influência significativa sobre os valores dos escores de eficiência do cluster 6, em um nível de significância de 1%. O poder explicativo dessa variável, entretanto, é bastante reduzido. Cabe ressaltar, assim, que o pressuposto da normalidade não foi atendido no nível de significância de 5%.

3.25.6.1.2 Endividamento geral

Essa variável mostra-se significativa na regressão e contribui com 0,023 do poder explicativo da regressão, considerando todas as variáveis explicativas.

A tabela 109 expõe a regressão linear simples na qual a variável dependente é representada pelos valores dos escores do *cluster* 6 e a variável independente é endividamento geral.

De acordo com os resultados, a variável endividamento geral, se compreendida como única variável explicativa do comportamento dos escores, teria um poder explicativo de apenas 0,002 ou 0,2% . O valor do *Prob (F-statistic)* = 0,706 maior do que 0,05 indica que o modelo é não é significativo. O valor *Prob* = 0,706 do coeficiente angular maior que 0,05 ressalta que a variável endividamento geral não é significativa.

Tabela 109: Regressão: escores em função da variável endividamento geral

Dependent Variable: SCORE

Method: Least Squares

Date: 01/16/09 Time: 17:51

Sample: 1 70

Included observations: 70

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.584919	0.066120	8.846299	0.0000
ENDGE7	0.000426	0.001125	0.378467	0.7063
R-squared	0.002102	Mean dependent var		0.608986
Adjusted R-squared	-0.012573	S.D. dependent var		0.150641
S.E. of regression	0.151585	Akaike info criterion		-0.907179
Sum squared resid	1.562514	Schwarz criterion		-0.842937
Log likelihood	33.75128	F-statistic		0.143238
Durbin-Watson stat	2.000964	Prob(F-statistic)		0.706262

Dessa forma, é possível constatar as seguintes evidências:

O endividamento geral, considerado como única variável explicativa da regressão, não se mostrou significativa ao nível de significância de 5%. Em outro modelo também testado, no qual foram compreendidas todas as variáveis obtidas na técnica Delphi, a variável endividamento geral apresentou-se significativo, embora isso só tenha ocorrido quando se introduziram todas as variáveis. Ainda assim, o poder explicativo dessa variável, mesmo naquelas circunstâncias, era baixo.

3.25.6.2 Cluster 6: Mais testes para a normalidade dos resíduos da regressão com variáveis adicionais

Na análise da regressão referente ao *cluster* 6 foi feita uma ressalva quanto à normalidade dos resíduos da regressão. A hipótese nula de normalidade dos resíduos foi rejeitada.

Desse modo, pesquisou-se melhor o assunto e, a partir disso, foi obtida nova regressão com a inclusão de todas as variáveis selecionadas pela técnica Delphi e também das variáveis

adicionais escolhidas pelo método *stepwise*. Os resultados confirmam que a hipótese nula da normalidade dos resíduos continuou a ser rejeitada, considerando o nível de significância de 5%, através da utilização do teste de normalidade de Jarque-Bera.

Foi realizada mais uma tentativa de resolver o problema da normalidade. A partir disso, cada uma das variáveis mencionadas no parágrafo anterior foi transformada de acordo com a seguinte expressão geral: $\log(x + 1)$. O teste de normalidade utilizado foi o de Jarque-Bera.

Testou-se mais uma vez a normalidade dos resíduos e então se empregou o teste de Kolmogorov-Smirnov. Os resultados comprovam que a hipótese nula da normalidade dos resíduos continuou a ser rejeitada, considerando o nível de significância de 5%.

Conforme Gujarati (2006, p.88), de acordo com o teorema central do limite, se houver um grande número de variáveis aleatórias independentes e com distribuição idêntica, a distribuição de sua soma tende à distribuição normal à medida que o número de variáveis aumenta indefinidamente.

Ainda segundo Gujarati (2006, p.89), mesmo que não haja um grande número de variáveis ou que essas variáveis não sejam estritamente independentes, sua soma ainda pode ser normalmente distribuída.

Gujarati (2006, p. 89) salienta que, no caso de uma amostra de tamanho pequeno, ou finito, ou seja, de menos de cem observações, a premissa da normalidade assume um papel fundamental porque permite recorrer aos testes, t , F . Como em amostras pequenas esses testes exigem a normalidade, é necessário que essa premissa seja testada principalmente se o objetivo for não só a estimação, mas também a inferência (Gujarati, 2006, p.2001).

Para Gujarati (2006, p.273), se os resíduos não forem distribuídos normalmente, é possível contar com o apoio do teorema central do limite e usar os testes t e F em grandes amostras, mas não nas amostras pequenas.

Como citado nos parágrafos anteriores, esse autor considera amostra pequena aquela com menos de cem observações. Entretanto, não há unanimidade entre outros autores quanto à definição do que seja uma amostra grande. Assim, por exemplo, McClave *et al* (1998, p. 254)

afirma que, para a maior parte das populações testadas, uma amostra de tamanho igual ou maior que 30 será suficiente para uma aproximação da normal ser considerada razoável. De acordo com Hildebrand *et al* (1998, p. 241), um bom guia para o uso do teorema central do limite se constitui quando o tamanho da amostra excede 30 dados. Esses autores alegam que uma regra melhor seria considerar os efeitos da assimetria: se um histograma indicasse uma curva com pequeno grau de assimetria, poderia ser aproximada de uma normal mesmo com amostras de tamanho 15 ou 20 (Hildebrand, 1998, p. 242).

De acordo com Evans *et al* (1998, p.77), coeficientes de assimetria entre 0,5 e 1 e -0,5 e -1 podem ser considerado moderadamente assimétricos.

Os testes dos resíduos da regressão na utilização de todas as variáveis no formato $\log(X + 1)$ apresentarm um grau de assimetria de 0,69, conforme figura 62, enquadrando-se na regra de Evans. Além disso, deve-se destacar que o tamanho da amostra é de 69 observações.

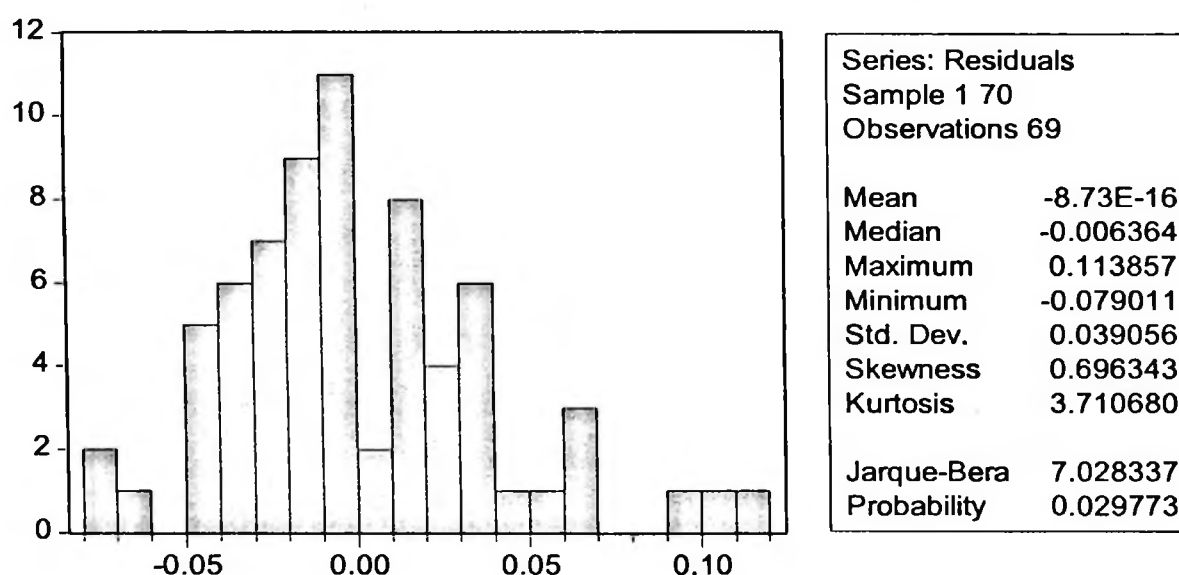


Figura 62: Teste da normalidade dos resíduos

Com base nesses dados, é possível obter as seguintes conclusões:

No modelo do *cluster 6*, todas as variáveis foram testadas e o único problema não solucionado foi o atendimento do pressuposto da normalidade. Entretanto, dadas todas as considerações e justificativas efetuadas, em função das demais correções e testes realizados nos itens anteriores, talvez seja plausível considerar as conclusões obtidas ao longo dos comentários sobre o *cluster 6* como uma aproximação razoável da realidade.

3.25.7 Regressão e análise dos resultados do *cluster 7*: incluindo todas as variáveis

Conforme a avaliação dos dados dos outros *clusters*, foram usados o programa SPSS e o método *stepwise* para fazer uma pré-seleção das variáveis do modelo. Foi estabelecido que variáveis com nível de significância superior a 5% fossem excluídas do modelo de regressão.

Com o emprego do método *stepwise* e inclusão de todas as variáveis – mencionadas na lista de variáveis –, são selecionadas aquelas significativas em um nível de significância de 5% conforme quadro 6.

Quadro 6: Variáveis selecionadas pelo método *stepwise*

Variável	Sigla
Composição do endividamento	COMPEND7
Giro do ativo	GIRO7
<i>Dummy</i> composição do endividamento de empresas com ações	DACCOM
<i>Dummy</i> giro do ativo de empresas com ações	DACGIR
Liquidez corrente	LIQCOR7
Crescimento das vendas	CRESVEN7
Capital circulante líquido	CCL7
Ativo	ATIVO7
Vendas	VENDAS7

A seguir, foi utilizado o programa *Eviews* para continuar com a análise de regressão, pois esse programa tem opções que possibilitam efetuar automaticamente os testes dos pressupostos básicos da regressão.

A Tabela 110 expõe os resultados da regressão em que a variável dependente é constituída pelos escores e as independentes são as variáveis do quadro 6.

Ao avaliar os resultados, nota-se que sete variáveis independentes inclusas no modelo são significativas em um nível de significância menor que 1%. Pode-se constatar que os valores Prob para sete variáveis – composição do endividamento, giro do ativo, *dummy* composição do endividamento de empresas com ações, *dummy* giro do ativo de empresas com ações, liquidez corrente, ativo e vendas – são todos menores que 0,01.

Na aplicação do método *stepwise*, quando estavam todas as variáveis consideradas em conjunto, havia uma variável significativa ao nível de significância de 5%. Porém, na atual circunstância, ela não mais se apresenta significativa no nível de 0,05. É o caso da variável crescimento das vendas, cujo valor Prob = 0,0587 não a torna mais significativa.

A variável composição do endividamento tem sinal negativo, o que é coerente, pois o aumento da composição do endividamento, isto é, o aumento das dívidas de curto prazo, acarreta redução no nível de eficiência. A variável vendas demonstra sinal negativo para o coeficiente angular, o que não é coerente. O mesmo ocorre com a variável Capital circulante líquido.

Serão realizados, ainda neste estudo, testes para avaliação da coerência dos sinais das variáveis introduzidas.

O modelo de regressão como um todo é significativo (*Prob(F-statistic)*) e menor que 0,01. O poder explicativo da regressão (*Adjusted R-squared*) é de 0,9234, considerado muito bom.

Tabela 110: Resultados da regressão para o *cluster 7* com todas as variáveis

Dependent Variable: SCORE

Method: Least Squares

Date: 12/16/08 Time: 09:19

Sample: 1 87

Included observations: 87

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-0.014519	0.000601	-24.14373	0.0000
GIRO7	0.301622	0.020138	14.97806	0.0000
DACCOM	-0.005054	0.000930	-5.435118	0.0000
DACGIR	0.187379	0.039280	4.770362	0.0000
LIQCOR7	0.077733	0.015808	4.917418	0.0000
CRESVEN7	0.001102	0.000574	1.918553	0.0587
CCL7	-7.10E-08	2.81E-08	-2.524587	0.0136
ATIVO7	1.96E-08	4.45E-09	4.405712	0.0000
VENDAS7	-2.14E-08	5.76E-09	-3.718082	0.0004
C	0.842994	0.040730	20.69694	0.0000
R-squared	0.931502	Mean dependent var		0.652540
Adjusted R-squared	0.923495	S.D. dependent var		0.184833
S.E. of regression	0.051124	Akaike info criterion		-3.001349
Sum squared resid	0.201251	Schwarz criterion		-2.717911
Log likelihood	140.5587	F-statistic		116.3461
Durbin-Watson stat	1.773388	Prob(F-statistic)		0.000000

A figura 63 exhibe os resultados do teste de normalidade dos resíduos. Foi utilizado o teste Jarque-Bera, sendo o valor para a *Probability* = 0,819, maior que 0,05, sugerindo a existência de evidências para a rejeição da hipótese de normalidade dos resíduos.

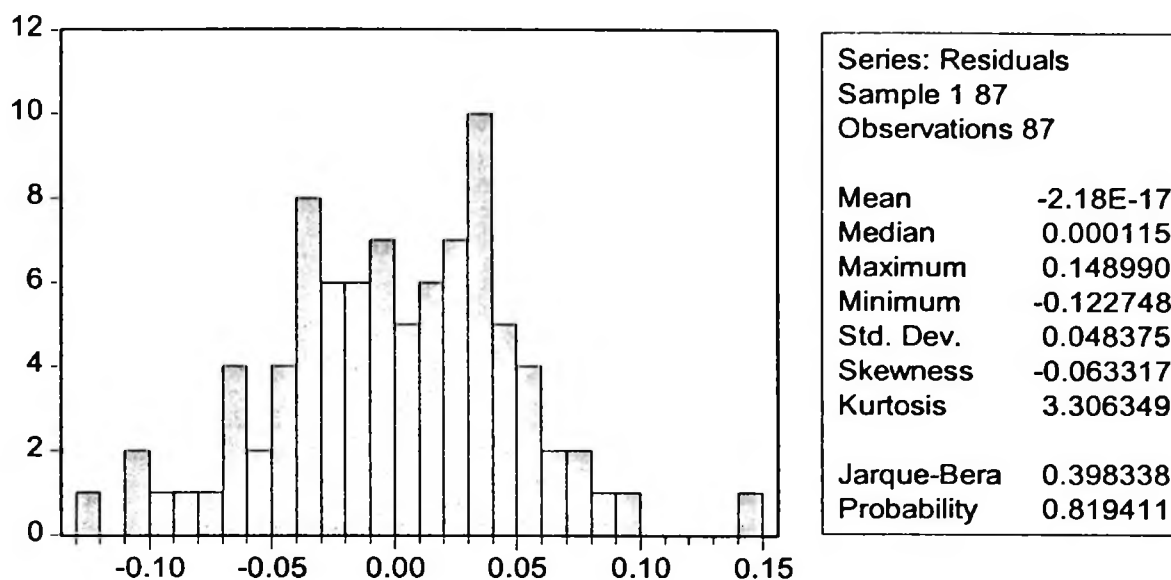


Figura 63: Teste da normalidade dos resíduos da regressão para o *cluster* 3 com todas as variáveis

A tabela 111 mostra os resultados da aplicação do teste de correlação serial dos resíduos da regressão de Breusch_Godfrey, que indica *Probability* = 0,665, evidenciando que não se pode rejeitar a hipótese de ausência de correlação serial.

Tabela 111: Teste de correlação serial dos resíduos da regressão para o *cluster* 7 com todas as variáveis
Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.410067	Probability	0.665084
Obs*R-squared	0.941065	Probability	0.624669

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/16/08 Time: 09:22

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COMPEND7	-2.20E-05	0.000607	-0.036181	0.9712
GIRO7	0.003402	0.021051	0.161597	0.8721
DACCOM	4.74E-05	0.000944	0.050176	0.9601
DACGIR	-0.000747	0.039919	-0.018714	0.9851
LIQCOR7	0.000655	0.015956	0.041073	0.9673
CRESVEN7	2.86E-05	0.000585	0.048901	0.9611
CCL7	4.13E-09	2.93E-08	0.140847	0.8884
ATIVO7	2.54E-11	4.53E-09	0.005600	0.9955
VENDAS7	-1.34E-11	5.92E-09	-0.002269	0.9982
C	-0.004215	0.041926	-0.100533	0.9202
RESID(-1)	0.105106	0.122888	0.855297	0.3951
RESID(-2)	0.028268	0.136309	0.207380	0.8363
R-squared	0.010817	Mean dependent var	-2.18E-17	
Adjusted R-squared	-0.134263	S.D. dependent var	0.048375	
S.E. of regression	0.051520	Akaike info criterion	-2.966248	
Sum squared resid	0.199074	Schwarz criterion	-2.626123	
Log likelihood	141.0318	F-statistic	0.074558	
Durbin-Watson stat	1.957938	Prob(F-statistic)	0.999976	

A tabela 112 expõe os resultados do teste de heterocedasticidade de White. O valor *Probability* = 0,127 comprova que não há evidências para se rejeitar a hipótese de homocedasticidade dos resíduos, pois esse valor é superior a 0,05.

Tabela 112: Teste de heterocedasticidade de White para os resíduos do *cluster* 7 com todas as variáveis

White Heteroskedasticity Test:

F-statistic	1.438383	Probability	0.127585
Obs*R-squared	57.97828	Probability	0.204731

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 12/16/08 Time: 09:23

Sample: 1 87

Included observations: 87

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.061024	0.024572	-2.483455	0.0178
COMPEND7	0.001244	0.000461	2.696130	0.0106
COMPEND7^2	-1.07E-05	5.04E-06	-2.115330	0.0414
COMPEND7*GIRO7	-0.000149	0.000212	-0.699934	0.4885
COMPEND7*DACCOM	3.53E-06	1.39E-05	0.253310	0.8015
COMPEND7*DACGIR	-0.000205	0.001086	-0.188672	0.8514
COMPEND7*LIQCOR7	-0.000181	0.000228	-0.794586	0.4321
COMPEND7*CRESVEN7	1.36E-06	4.88E-06	0.278718	0.7821
COMPEND7*CCL7	-1.57E-09	2.39E-09	-0.658930	0.5141
COMPEND7*ATIVO7	-1.69E-10	2.50E-10	-0.676429	0.5031
COMPEND7*VENDAS7	3.28E-10	4.15E-10	0.790992	0.4341
GIRO7	-0.010791	0.023583	-0.457583	0.6500
GIRO7^2	0.008021	0.005171	1.551079	0.1296
GIRO7*DACGIR	-0.008197	0.027078	-0.302717	0.7638
GIRO7*LIQCOR7	0.004161	0.009186	0.453034	0.6532
GIRO7*CRESVEN7	-0.000192	0.000341	-0.563358	0.5767
GIRO7*CCL7	-1.31E-07	1.03E-07	-1.269695	0.2123
GIRO7*ATIVO7	-4.56E-08	7.96E-08	-0.572573	0.5705
GIRO7*VENDAS7	-1.16E-09	1.19E-08	-0.097102	0.9232
DACCOM	0.000611	0.000756	0.808335	0.4242
DACCOM*LIQCOR7	-0.000548	0.000443	-1.235182	0.2248
DACCOM*CRESVEN7	-1.15E-05	1.65E-05	-0.698738	0.4892
DACCOM*CCL7	1.43E-09	2.55E-09	0.558962	0.5796
DACCOM*ATIVO7	1.78E-10	3.82E-10	0.464991	0.6447
DACCOM*VENDAS7	-2.65E-10	4.72E-10	-0.560115	0.5789
DACGIR	0.002733	0.035014	0.078058	0.9382
DACGIR*LIQCOR7	0.012316	0.020129	0.611867	0.5445
DACGIR*CRESVEN7	0.000272	0.000773	0.352491	0.7265
DACGIR*CCL7	-3.87E-09	1.38E-07	-0.028063	0.9778
DACGIR*ATIVO7	-2.00E-08	2.31E-08	-0.864618	0.3930
DACGIR*VENDAS7	2.23E-08	2.52E-08	0.886316	0.3813
LIQCOR7	0.033529	0.022284	1.504606	0.1411
LIQCOR7^2	0.000978	0.005136	0.190437	0.8500
LIQCOR7*CRESVEN7	-0.000760	0.000263	-2.892383	0.0064
LIQCOR7*CCL7	-3.85E-08	4.64E-08	-0.829588	0.4122
LIQCOR7*ATIVO7	8.65E-11	6.29E-09	0.013747	0.9891
LIQCOR7*VENDAS7	-2.27E-10	1.04E-08	-0.021808	0.9827
CRESVEN7	0.001185	0.000406	2.922144	0.0060
CRESVEN7^2	-2.65E-06	4.72E-06	-0.561638	0.5778
CRESVEN7*CCL7	5.48E-09	2.06E-09	2.657255	0.0117

Tabela 112: Teste de heterocedasticidade de White para os resíduos do *cluster 7* com todas as variáveis

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CRESVEN7*ATIVO7	-3.06E-11	2.59E-10	-0.118211	0.9066
CRESVEN7*VENDAS7	-3.79E-11	3.80E-10	-0.099693	0.9211
CCL7	1.09E-08	1.61E-07	0.067572	0.9465
CCL7^2	5.34E-15	5.30E-14	0.100764	0.9203
CCL7*ATIVO7	-2.95E-14	2.39E-14	-1.234816	0.2249
CCL7*VENDAS7	3.21E-14	2.83E-14	1.133353	0.2646
ATIVO7	7.08E-09	1.27E-08	0.557671	0.5805
ATIVO7^2	-1.28E-15	1.05E-15	-1.224687	0.2286
ATIVO7*VENDAS7	6.23E-15	3.83E-15	1.624915	0.1129
VENDAS7	3.74E-08	7.76E-08	0.481933	0.6328
VENDAS7^2	-5.16E-15	3.06E-15	-1.683706	0.1009
R-squared	0.666417	Mean dependent var	0.002313	
Adjusted R-squared	0.203107	S.D. dependent var	0.003533	
S.E. of regression	0.003154	Akaike info criterion	-8.390129	
Sum squared resid	0.000358	Schwarz criterion	-6.944597	
Log likelihood	415.9706	F-statistic	1.438383	
Durbin-Watson stat	1.795469	Prob(F-statistic)	0.127585	

Ao averiguar os dados obtidos, são plausíveis as seguintes evidências:

Rejeita-se H_0 , de que os indicadores obtidos através da técnica Delphi não têm influência significativa sobre os níveis de eficiência obtidos por empresa e por *cluster* segundo a técnica DEA. Portanto, os indicadores a seguir relacionados são significativos a um nível de significância inferior a 0,01: composição do endividamento, giro do ativo e liquidez corrente.

Não há evidências para rejeitar a hipótese nula no caso do indicador crescimento das vendas, em um nível de significância de 0,05. Seu valor Prob é de 0,587.

Quanto aos demais indicadores relacionados na lista de variáveis, no início do trabalho, revelaram-se significativos em um nível de significância de 0,01, os seguintes: *dummy* composição do endividamento de empresas com ações, *dummy* giro do ativo de empresas com ações e ativo.

Cabe salientar que as variáveis capital circulante líquida e vendas tiveram coeficientes angulares negativos, o que parece ser incoerente com relação à teoria.

3.25.7.1 Cluster 7: Análise detalhada das outras variáveis significativas

A seguir, serão feitas algumas explanações sobre as seguintes variáveis que se mostraram significativas na regressão e que constam no quadro 7:

Quadro 7: Variáveis significativas da regressão

Variável	Sigla
Dummy composição do endividamento de empresas com ações na bolsa	DACCOM
Dummy giro do ativo de empresas com ações na bolsa	DACGIR
Capital circulante líquido	CCL7
Ativo	ATIVO7
Vendas	VENDAS7

3.25.7.1.1 Variável *Dummy* Composição do Endividamento de Empresas com Ações

A variável *dummy* composição do endividamento de empresas com ações (DACCOM) mostrou-se significativa em um nível de significância de 0,01 na regressão. Esse fato confirma que a composição do endividamento de empresas com ações na bolsa é significativamente diferente quando o mesmo se refere às empresas sem ações. A hipótese a ser testada é:

H_0 : não há diferença significativa na composição do endividamento entre empresas com ações e empresas sem ações na bolsa.

Será utilizado o teste t para diferença de médias para amostras independentes. Porém, ao aplicar esse teste, é necessário que sejam atendidos dois pressupostos: da normalidade das variáveis e da homogeneidade das variâncias.

Conforme os resultados da tabela 113, não há evidências para se rejeitar a hipótese da normalidade da variável composição do endividamento tanto para as empresas sem ações (0) com Sig = 0,200, quanto para as com ações (1) com Sig = 0,200. Esses valores foram obtidos através do teste Kolmogorov-Smirnov e são maiores que 0,05.

Tabela 113: Teste de normalidade

Tests of Normality						
AC	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	Df	Sig.
COMPEND7 0	,084	52	,200*	,972	52	,262
1	,102	35	,200*	,965	35	,327

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Conforme a tabela 114, o teste de Levene confirma a rejeição da hipótese nula de igualdade de variâncias, pois o valor Sig = 0,013 é menor que 0,05. A mesma tabela apresenta os resultados do teste t de diferença de médias para amostras independentes. Como o valor Sig (2-tailed) = 0,01 é menor que 0,05, rejeita-se a hipótese nula de que não há diferença significativa entre as médias da composição do endividamento das empresas com ações e sem ações.

Tabela 114: Teste de diferença de médias para amostras independentes

Independent Samples Test				
			COMPEND7	
			Equal variances assumed	Equal variances not assumed
Levene's Test for Equality of Variances	F		6,365	
	Sig.		0,013	
t-test for Equality of Means	T		2,443	2,637
	DF		85,000	84,998
	Sig. (2-tailed)		0,017	0,010
	Mean Difference		6,095	6,095
	Std. Error Difference		2,495	2,312
	95% Confidence Interval of the Difference	Lower	1,135	1,499
		Upper	11,055	10,692

Como não se confirmou a hipótese nula quanto à homogeneidade das variâncias, procedeu-se ao teste não paramétrico correspondente cujos resultados encontram-se na tabela 115. Sendo o valor *Asymp. Sig. (2-tailed)* = 0,018 menor que 0,05, rejeita-se a hipótese de igualdade das médias da composição do endividamento das empresas com ações e sem ações.

Tabela 115: Teste não paramétrico de Mann-Witney

Mann-Whitney Test				
Ranks				
	AC	N	Mean Rank	Sum of Ranks
COMPEND7	0	52	49,27	2562,00
	1	35	36,17	1266,00
	Total	87		

Test Statistics ^a	
	COMPEND7
Mann-Whitney U	636,000
Wilcoxon W	1266,000
Z	-2,372
Asymp. Sig. (2-tailed)	,018

a. Grouping Variable: AC

A partir dessas análises, é possível constatar as seguintes evidências:

O valor médio da composição do endividamento das empresas com ações na bolsa é significativamente menor (34,78%) do que a das empresas sem ações (40,87%) em um nível de significância de 0,05. Por isso justifica-se o fato de o coeficiente angular dessa variável ser negativo na regressão.

3.25.7.1.2 Variável *Dummy* Giro do Ativo de Empresas com Ações

A variável *dummy* giro do ativo de empresas com ações (DACGIR) demonstrou ser significativa em um nível de significância de 0,01 na regressão. Isso comprova que o giro do ativo de empresas com ações na bolsa é significativamente diferente quando o mesmo se refere às empresas sem ações. A hipótese a ser testada é:

H_0 : não há diferença significativa no giro do ativo entre empresas com ações e empresas sem ações na bolsa.

Será utilizado o teste t para diferença de médias para amostras independentes. Apesar disso, na aplicação desse teste, é preciso que sejam atendidos dois pressupostos: da normalidade das variáveis e da homogeneidade das variâncias.

Conforme resultados expostos na tabela 116, não há evidências para se rejeitar a hipótese da normalidade da variável giro do ativo para as empresas sem ações (0) com Sig = 0,200. Quanto ao giro do ativo para as empresas com ações (1), o Sig = 0,029 menor que 0,05 indica rejeição da hipótese nula de normalidade. Esses valores foram obtidos através do teste Kolmogorov-Smirnov e são maiores que 0,05.

Tabela 116: Teste de normalidade

Tests of Normality						
AC	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	DF	Sig.	Statistic	Df	Sig.
GIRO7 0	,089	52	,200 [*]	,973	52	,291
1	,157	35	,029	,952	35	,126

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Através da tabela 117, nota-se que o teste de Levene comprova a não rejeição da hipótese nula de igualdade de variâncias, pois o valor Sig = 0,084 é maior que 0,05.

Tabela 117: Teste de Levene para a igualdade das variâncias

Levene's Test for Equality of Variances	F	3,053
	Sig.	0,084

Não havendo a confirmação da hipótese nula quanto à normalidade da variável giro do ativo para empresas com ações, aplicou-se o teste não paramétrico correspondente, cujos resultados encontram-se na tabela 118. Como o valor *Asymp. Sig. (2-tailed)* = 0,604 é maior que 0,05, não se rejeita a hipótese de igualdade das médias do giro do ativo das empresas com ações e sem ações.

Tabela 118: Teste não paramétrico de Mann-Witney**Mann-Whitney Test**

Ranks			
AC	N	Mean Rank	Sum of Ranks
GIRO7 0	52	45,15	2348,00
1	35	42,29	1480,00
Total	87		

Test Statistics ^a	
	GIRO7
Mann-Whitney U	850,000
Wilcoxon W	1480,000
Z	-,519
Asymp. Sig. (2-tailed)	,604

a. Grouping Variable: AC

Ao se avaliarem os resultados, são constatadas as seguintes evidências:

Não se rejeita a hipótese de igualdade das médias do giro do ativo para empresas sem ações e com ações na bolsa em um nível de significância de 0,05.

3.25.7.1.3 Variável Vendas

A variável Vendas (VENDAS7) apresentou-se significativa na regressão e contribui com 0,012 do poder explicativo da regressão, considerando todas as variáveis explicativas.

A tabela 119 expõe a regressão linear simples na qual a variável dependente é representada pelos valores dos escores do *cluster 7* e a variável independente é a de vendas.

Analisando os resultados, verifica-se que a variável vendas, se considerada como única variável explicativa do comportamento dos escores, teria um poder explicativo de apenas 0,0007 ou 0,07% . O valor do *Prob (F-statistic)* = 0,79 é maior que 0,05 e indica que o modelo não é significativo. O valor *Prob* = 0,79 do coeficiente angular maior que 0,05 sugere que a variável capital circulante líquido não é significativa.

Tabela 119: Regressão: escores em função das vendas

Dependent Variable: SCORE

Method: Least Squares

Date: 01/16/09 Time: 15:51

Sample: 1 87

Included observations: 87

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.655299	0.022609	28.98392	0.0000
VENDAS7	-1.97E-09	7.62E-09	-0.258152	0.7969
R-squared	0.000783	Mean dependent var		0.652540
Adjusted R-squared	-0.010972	S.D. dependent var		0.184833
S.E. of regression	0.185844	Akaike info criterion		-0.505095
Sum squared resid	2.935738	Schwarz criterion		-0.448408
Log likelihood	23.97164	F-statistic		0.066643
Durbin-Watson stat	1.674185	Prob(F-statistic)		0.796913

Avaliando os dados obtidos, é possível obter as seguintes evidências:

A variável vendas, considerada como única explicativa da regressão, não se expôs significativa no nível de significância de 5%. Em outro modelo testado, em que foram incluídas, além de vendas, todas as variáveis obtidas na técnica Delphi, as vendas também se apresentaram não significativas. Ela só se mostrou significativa quando foram introduzidas todas as variáveis, porém, seu poder explicativo, mesmo naquela situação, era muito baixo.

3.25.7.1.4 Capital circulante líquido

A variável Capital circulante líquido (CCL7) demonstrou ser significativa na regressão e contribui com 0,6% do poder explicativo da regressão, considerando todas as variáveis explicativas.

A tabela 120 exibe a regressão linear simples na qual a variável dependente é representada pelos valores dos escores do *cluster 7* e a variável independente é o capital circulante líquido.

Conforme os resultados obtidos, a variável capital circulante líquido, se considerada como única variável explicativa do comportamento dos escores, teria um poder explicativo de apenas 0,0188 ou 1,88% . O valor do *Prob (F-statistic)* = 0,205 é maior que 0,05, indicando que o modelo não é significativo. O valor *Prob* = 0,205 do coeficiente angular maior que 0,05 confirma que a variável capital circulante líquido não é significativa.

Tabela 120: Regressão: escores em função do capital circulante líquido

Dependent Variable: SCORE

Method: Least Squares

Date: 01/16/09 Time: 15:55

Sample: 1 87

Included observations: 87

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.647378	0.020154	32.12156	0.0000
CCL7	1.13E-07	8.83E-08	1.276495	0.2053
R-squared	0.018809	Mean dependent var		0.652540
Adjusted R-squared	0.007266	S.D. dependent var		0.184833
S.E. of regression	0.184160	Akaike info criterion		-0.523300
Sum squared resid	2.882777	Schwarz criterion		-0.466612
Log likelihood	24.76355	F-statistic		1.629439
Durbin-Watson stat	1.664155	Prob(F-statistic)		0.205257

Dessa maneira, é possível afirmar as seguintes evidências:

O capital circulante líquido, considerado como única variável explicativa da regressão, também não se mostrou significativo no nível de significância de 5%. Em outro modelo testado, em que se incluíram, além do capital circulante líquido, todas as variáveis obtidas na técnica Delphi, o capital circulante líquido também se apresentou não significativo. Ele só se expôs significativo quando foram introduzidas todas as variáveis, mas seu poder explicativo, mesmo naquelas circunstâncias, era muito baixo.

3.25.7.1.5 Ativo

A variável ativo (ATIV07) confirmou ser significativa na regressão e contribui com 0,5% do poder explicativo da regressão considerando todas as variáveis explicativas.

A tabela 121 demonstra a regressão linear simples na qual a variável dependente é representada pelos valores dos escores do *cluster 7* e a variável independente pelo ativo.

Conforme os resultados, a variável ativo, se considerada como explicativa do comportamento dos escores, teria um poder explicativo de apenas 0,00126 ou 0,126% . O valor do *Prob (F-statistic)* = 0,74 é maior que 0,05, comprovando que o modelo não é significativo. O valor

Prob = 0,74 do coeficiente angular maior que 0,05, indica que a variável capital circulante líquido não é significativa,

Tabela 121: Regressão: escores em função do ativo

Dependent Variable: SCORE

Method: Least Squares

Date: 01/16/09 Time: 15:56

Sample: 1 87

Included observations: 87

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.648653	0.023173	27.99124	0.0000
ATIVO7	1.96E-09	5.98E-09	0.328319	0.7435
R-squared	0.001267	Mean dependent var		0.652540
Adjusted R-squared	-0.010483	S.D. dependent var		0.184833
S.E. of regression	0.185799	Akaike info criterion		-0.505579
Sum squared resid	2.934318	Schwarz criterion		-0.448891
Log likelihood	23.99268	F-statistic		0.107793
Durbin-Watson stat	1.676874	Prob(F-statistic)		0.743478

Ao avaliar os resultados obtidos, é possível constatar as seguintes evidências:

O ativo, considerado como única variável explicativa da regressão, não se mostrou significativo no nível de significância de 5%. Em outro modelo examinado, no qual se incluíram, além do ativo, todas as variáveis obtidas na técnica Delphi, o valor do ativo também se apresentou não significativo. O ativo só se expôs significativo quando foram introduzidas todas as variáveis, embora, mesmo assim, seu poder explicativo, naquelas circunstâncias, fosse muito baixo.

4 Conclusão

Este estudo teve como principal objetivo verificar a eficiência de empresas através do uso de indicadores contábeis e da utilização de diversos instrumentos quantitativos como: Análise por Envoltória de Dados, Análise de Agrupamentos, Análise de Regressão e Testes de Hipóteses Paramétricos e Não Paramétricos.

A preocupação inicial deste trabalho relacionava-se aos indicadores contábeis a serem usados na avaliação das empresas. Em função da experiência adquirida ao longo dos anos na avaliação de empresas através de indicadores contábeis, sempre houve o cuidado em mostrar que não havia a necessidade da utilização de um grande número de indicadores para discriminar empresas eficientes das não eficientes.

Surgiu então a oportunidade de divulgar que, com poucos indicadores contábeis devidamente selecionados, seria possível avaliar empresas de forma eficiente. O trabalho de Lyra (2008) constituiu-se em grande avanço nesse sentido quando foi empregada a técnica Delphi para a seleção dos indicadores contábeis.

Através da utilização da técnica Delphi e da consulta a professores especializados em análise de demonstrações financeiras, Lyra (2008) chegou à seleção de sete indicadores contábeis significativos. A pesquisa atual partiu da seleção de indicadores efetuada por Lyra (2007), com a utilização da técnica Delphi.

Utilizaram-se dados contábeis publicados anualmente pela edição de Melhores & e Maiores da Revista Exame. Foi empregada a publicação mais recente referente ao ano de 2008, com dados relativos às demonstrações financeiras de 2007, referentes a mil empresas.

Dos sete indicadores selecionados por Lyra (2008), foi excluído o indicador rentabilidade sobre o ativo, sendo usados, neste estudo, seis indicadores: rentabilidade sobre o patrimônio líquido, crescimento das vendas, liquidez corrente, composição do endividamento, margem líquida e giro do ativo.

Para a avaliação da eficiência, foi escolhida a técnica *Data Envelopment Analysis* (DEA) criada por Charnes, Cooper e Rhodes (1978). Essa técnica permite estabelecer a eficiência relativa de cada empresa e escalonar os respectivos desempenhos através de escores de

eficiência, podendo definir-se a fronteira eficiente como aquela em que se situam as empresas 100% eficientes.

Um dos pressupostos para a utilização do DEA é que as empresas apresentassem comportamento homogêneo. Para que esse pressuposto fosse atendido, utilizou-se a técnica *Cluster Analysis*, resultando na separação das empresas em sete grupos homogêneos (*clusters*).

Cada grupo (*cluster*) apresentou como principal característica o fato de atingir o valor extremo em apenas um dos indicadores. Assim, cada um dos seis indicadores teve seu valor máximo em apenas um dos grupos. Um sétimo grupo se caracterizou por apresentar os piores desempenhos; dos seis indicadores utilizados, obteve os piores desempenhos em três: crescimento das vendas, rentabilidade sobre o patrimônio líquido e margem líquida.

Pelo menos três indicadores sempre se apresentaram como significativos para a formação de cada um dos sete grupos, mas não foram sempre os mesmos indicadores. Dependendo de cada grupo específico houve alternância de indicadores. A separação entre os grupos se deu em função da diferença entre os graus de importância dos indicadores na formação dos grupos.

Verificou-se que algumas empresas, embora pertencentes a setores diferentes, apresentaram comportamento semelhante em relação aos indicadores de desempenho, ou seja, mostraram conduta semelhante e assim acabaram se situando num mesmo *cluster*. Da mesma forma, empresas do mesmo setor acabaram sendo classificadas em *clusters* diferentes porque suas condutas foram distintas em função dos indicadores considerados.

Dos seis indicadores utilizados para a aplicação do DEA, o de composição do endividamento foi considerado como um insumo (*input*) e os demais como produtos (*outputs*). As empresas da fronteira eficiente foram as que obtiveram a maior razão entre a soma ponderada dos *outputs* (retorno sobre o patrimônio líquido, crescimento das vendas, margem líquida, liquidez corrente, giro do ativo) em relação ao *input* (composição do endividamento). Então, dado o endividamento assumido (*input*), constituíram-se as empresas mais eficientes aquelas que conseguiram a maior cesta de produtos (*outputs*) relativamente às demais.

Como o DEA não possibilita a realização de testes de hipóteses, efetuou-se a regressão linear múltipla para tal finalidade. Foram aplicadas regressões para cada um dos sete agrupamentos (*clusters*). Ao analisar os escores de eficiência e as seis variáveis obtidas pela técnica Delphi, concluiu-se que, das seis variáveis, pelo menos quatro e no máximo cinco se mostraram significativas no nível de significância de 5% em todos os agrupamentos. Assim, de acordo com as evidências encontradas, os indicadores obtidos através da técnica Delphi têm influência significativa sobre os escores de eficiência obtidos por meio do modelo DEA.

Verificou-se que as regressões relativas a todos os agrupamentos (*clusters*) demonstraram bom poder explicativo medido pelo R-quadrado ajustado, com valor médio situado em torno de 84,7%. Esse resultado sugere que se possa usar o procedimento abordado neste estudo para projetar os escores de eficiência de uma determinada empresa em função dos seis indicadores obtidos das suas demonstrações contábeis. Cabe salientar que os pressupostos exigidos pelos modelos de regressão foram, de maneira geral, obedecidos, possibilitando, dessa forma, a utilização dos resultados não só para projeção como também para os testes de hipóteses.

Na última etapa do trabalho foram testados mais uma vez os modelos de regressão linear múltipla para cada agrupamento. Dessa vez foram incluídas como variáveis independentes não só os seis indicadores definidos pela técnica Delphi, mas todas as outras variáveis publicadas pela Edição de Melhores & Maiores. Além desses indicadores, foram também compreendidas variáveis *dummies* para testar diferença de comportamento devido a diversas situações específicas, para cada um dos seis indicadores e para cada um dos sete agrupamentos. Dependendo do tipo de empresa poderia haver diferença de comportamento em um ou mais indicadores devido a, por exemplo, algumas empresas de um dado agrupamento terem ações na bolsa de valores ou não, ou então por serem auditadas pelas grandes empresas de auditoria consideradas como “*big four*” ou não, ou ainda por diversas outras razões. Desse modo, foram criadas variáveis *dummies* para diferentes situações e para cada um dos seis indicadores. A seguir relacionam-se as questões suscitadas, considerando cada um dos seis indicadores definidos pela técnica Delphi e cada um dos sete agrupamentos (*clusters*):

- empresas que possuem ações negociadas na bolsa de valores têm diversidade de comportamento em relação às que não possuem?
- empresas estrangeiras têm diversidade de comportamento em relação às empresas nacionais?

- empresas estatais têm diversidade de comportamento em relação às demais ?
- empresas auditadas pelas grandes empresas de auditoria conhecidas como “*big four*” têm diversidade de comportamento em relação às que não são auditadas?
- há diferença de comportamento entre empresas dos setores comércio, indústria e serviços?

O objetivo dessas regressões, nas quais se incluíram todas as variáveis da edição da Melhores & Maiores, foi testar as hipóteses de que outras variáveis, além dos seis indicadores definidos pela técnica Delphi, poderiam ter influência significativa nos níveis de eficiência estabelecidos através do DEA e medidos pelos escores de eficiência. Além disso, com a inclusão das *dummies*, poderiam ser analisados outros aspectos interessantes.

Assim, uma vez compreendidas todas as variáveis na regressão, observou-se que:

- os indicadores obtidos pela técnica Delphi continuaram significativos;
- após a realização de diversos testes e análises confirmatórias, apenas duas variáveis da lista da Melhores & Maiores se mostraram significativas, excetuando-se os indicadores da técnica Delphi: endividamento a longo prazo (apenas no *cluster* 3) e investimento no imobilizado (apenas no *cluster* 6);
- algumas variáveis *dummies* apresentaram-se significativas nos *clusters* 1, 3, 5 e 7.

Quanto à significância das variáveis *dummies*, é importante ressaltar as seguintes constatações:

Cluster 1:

- giro do ativo do ramo industrial é significativamente menor que o dos ramos comercial e de serviços;
- a margem líquida do ramo industrial é significativamente maior do que a dos ramos comercial e de serviços;

Cluster 3:

- o índice de liquidez corrente de empresas com ações na bolsa é significativamente menor do que o das empresas sem ações;

Cluster 5:

- as empresas nacionais apresentam o valor médio do indicador composição do endividamento significativamente menor que o das empresas estrangeiras;

Cluster 7:

- o valor médio da composição do endividamento das empresas com ações na bolsa é significativamente menor do que o das empresas sem ações em um nível de significância de 0,05.;

Após essas considerações, é possível concluir que:

- a técnica Delphi, parte do estudo de Lyra (2007), foi avaliada pelo presente estudo e gerou indicadores muito significativos. Essa técnica, portanto, pode ser indicada para os futuros pesquisadores de técnicas de análise;
- através dos resultados apresentados neste estudo, não é necessário utilizar grande número de indicadores para avaliar eficiência de empresas desde que os indicadores utilizados sejam selecionados de forma eficiente;
- é possível separar as empresas da economia em grupos homogêneos embora possa haver controvérsia no processo de seleção. Os artigos publicados até o momento não chegaram a um consenso quanto ao método de agrupamento ideal para a utilização do DEA. O método empregado neste estudo serviu muito bem aos objetivos propostos. A separação tradicional em setores pode apresentar problemas ainda maiores do que os observados com a técnica de *cluster analysis*;
- tem aumentado muito nos últimos anos a quantidade de publicações com a utilização do modelo DEA para a avaliação de eficiência. Esse modelo vem sendo recomendado pelos pesquisadores do mundo inteiro e o número de estudiosos especializados tem

crescido bastante. Verificou-se neste estudo que o modelo DEA pode ser complementado por outras técnicas, inclusive estatísticas paramétricas para testar hipóteses, o que não é possível através do DEA;

- é possível criar modelos de previsão de níveis de eficiência através da utilização de indicadores contábeis conjugados com o modelo DEA e a análise de regressão.

Neste estudo foi feita uma análise parcial de eficiência através da técnica DEA. Sugere-se que os pesquisadores interessados efetuem uma análise DEA mais ampla, empregando a técnica Delphi para definir outros *inputs* e *outputs*. Recomenda-se também que seja considerada a utilização do modelo BCC, pois os problemas de heterocedasticidade encontrados em alguns casos podem ser causados por diferenças de escala.

Por fim, é importante destacar que esta pesquisa, através de sua fundamentação teórica e análise dos resultados, respondeu às questões formuladas, atingindo os objetivos propostos de maneira coerente, constituindo-se, dessa forma, em referência para futuros estudos sobre análise da eficiência das empresas por meio de indicadores contábeis.

REFERÊNCIAS

AHN, T., A.; CHARNES, ; COOPER, W.W. *Efficiency Characterization in Different DEA Models*, **Socio-Economic Planning Sciences** 22, pp. 253-257, 1988.

AIGNER, D.J.; LOVELL, C.A.K.; SCHMIDT, P. *Formulation and estimations of stochastic frontier production function models*. **Journal of Econometrics**, v. 6, p. 21-37, 1977.

ALI, I ; SEIFORD, L.M.. *Computational Accuracy and Infinitesimals in Data Envelopment Analysis*. **INFORMS** 31, pp. 290-297, 1990a.

ALI, I; SEIFORD. L.M. *Translational Invariance in Data Envelopment Analysis*, **Operation Research Letter** 9, pp. 403-405, 1990b.

ALLEN, R.; ATHANASSOPOULOS, A.; DYSON, R.G.; THANASSOULIS, E. *Weights restrictions and value judgments in data envelopment analysis: evolution, development and future directions*. **Annals of Operations Research**, n. 73, p. 13-34, 1997.

ALMEIDA, H. R. **Análise de envoltória de dados na tomada de decisão para concessão de crédito**. Tese (Mestrado) Instituto Tecnológico da Aeronáutica. ITA. São José dos Campos, SP, 2000.

ANGELO, Cláudio Felisoni de; SANVICENTE, Antonio Zoratto. Agrupamento de empresas por semelhança: uma crítica à análise setorial convencional. **Revista de Administração**, São Paulo, v. 25, n. 2, p. 20-27, abr./jun.1990.

ANGELO, Cláudio Felisoni de; SANVICENTE, Antonio Zoratto. Conduta e desempenho de empresas: uma aplicação de “cluster analysis” à segmentação da indústria do cimento. **Estudos Econômicos**, São Paulo, v. 22, n. 2, p. 20-27, abr./jun.1990.

ANGULO MEZA, L. *Data envelopment analysis na determinação da eficiência dos programas de pós-graduação da COPPE/UFRJ*. Tese (Mestrado em Engenharia de Produção) – COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1998.

ANGULO MEZA, L. **Um enfoque multiobjetivo para os modelos de determinação de alvos na Análise Envoltória de Dados**. Tese (Doutorado em Engenharia de Produção) – COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2002a.

ANGULO MEZA, L.; LINS, M. P. E. *Review of methods for increasing discrimination in data envelopment analysis*. **Annals of Operations Research**, Netherlands, v.116, p.225-242, 2002b.

ANGULO MEZA, L.; BIONDI NETO, L.; SOARES DE MELLO, J.C.C.B.; GOMES, E.G.; COELHO, P.H.G. *FSDA – Free Software for Decision Analysis (SLAD – Software Livre de Apoio à Decisão): A Software Package for Data Envelopment Analysis Models*. In: **XII CONGRESO LATINO-IBEROAMERICANO DE INVESTIGACIÓN OPERATIVA**, 12., 2004, La Habana. Memorias..., La Habana, 2004.

ASSAF Neto, Alexandre. **Estrutura e análise de balanços: um enfoque econômico-financeiro**. 7. ed. São Paulo: Atlas, 2002.

ATHANASSOPOULOS, A. D.; THANASSOULIS, E. *Separating market efficiency from profitability and its implications for planning*, **Journal of Operational Research Society** 46 (1), p. 20–34, 1995.

ATKINSON, Anthony A.; BANKER, Rajiv D.; KAPLAN, Robert S.; YOUNG, S. Mark. **Management accounting**. New Jersey: Prentice-Hall, 1995.

BANKER, R.D.; NATARAJAN, R. *Statistical Tests Based on DEA Efficiency Scores in W.W. Cooper, L.M. Seiford and J. Zhu, eds., Handbook on Data Envelopment Analysis* (Norwell, Mass.: Kluwer Academic Publishers), 2004.

BANKER, R.D.; CHARNES, A.; COOPER, W.W. *Some Models for Estimating Technical and Scale Efficiencies in DEA*. **Management Science** 30, pp. 1078-1092, 1984.

BANKER, R.D.; COOPER, W W.; SEIFORD, L.M.; ZHU, J.. *Returns to Scale in DEA, Chapter 2 in Cooper, Seiford and Zhu, eds., Handbook on Data Envelopment Analysis* (Norwell, Mass.: Kluwer Academic Publishers), 2004.

BANKER, R.D.; POTTER, G.; SRINIVASAN, D. *An empirical investigation of an incentive plan that includes nonfinancial performance measures*. **The Accounting Review**, Vol. 75 No. 1, pp. 65-92, 2000.

BANKER, R.D.; CHARNES, A.; COOPER, W.W. *Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis*. **Management Science**, Vol. 30, No 9, pp. 1078-1092, September 1984.

BARBA-ROMERO, S.; POMEROL, J.C. **Decisiones multicriterio: fundamentos teóricos e utilización práctica**. Madrid: Universidad de Alcalá, (Colección de Economía), 1997.

BARDHAN, I.; BOWLIN, W.F.; COOPER, W.W.; SUEYOSHI, T. *Models and Measures for Efficiency Dominance in DEA, Part II: Free Disposal Hull (FDH) and Russell Measure (RM) Approaches*. **Journal of the Operational Research Society of Japan** 39, pp. 333-345, 1996.

BARDHAN, I.; COOPER, W.W.; KUMBHAKAR, SC. *A Simulation Study of Joint Uses of DEA and Statistical Regression for Production Function Estimation and Efficiency Evaluation*. **Journal of Production Analysis** 9, pp. 249-278, 1998.

BARROW, M Wagstaff A. *Efficiency measurement in the public sector: an appraisal*. **Fiscal Studies**, 72-97, 1989.

BELLONI, José Ângelo. **Uma metodologia de avaliação da eficiência produtiva de Universidade Federais Brasileiras**. Tese (Doutorado) – Programa de Pós-Graduação em Engenharia de Produção, do Departamento de Engenharia de Produção e Sistemas, da Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2000.

BESSENT, A.; BESSENT, W.; ELAM, J. *Efficiency frontier determination by constraint facet analysis*. **Operation Research**, v. 36, p. 785-796, 1988.

BLATT, Adriano. **Análise de Balanços: estruturação e avaliação das demonstrações financeiras e contábeis**. São Paulo: Makron Books, 2001.

BOUSSOFIANE A.; DYSON, RG; THANASSOULIS E. *Applied Data Envelopment Analysis*. **Eur J. Ops Res.** 52(1), 1-15, 1991.

BOWLIN W.F.; CHARNES A.; COOPER, W.W.; SHERMAN, H.D. *Data envelopment analysis and regression approaches to efficiency estimation and evaluation*. **Ann. Ops Res.** 2, 113-138, 1985.

BULLA, S.; COOPER, W.W.; WILSON, D. *Evaluating Efficiencies of Turbofan Jet Engines: A Data Envelopment Analysis Approach*. **Journal of Propulsion and Power** Vol. 16, pp. 431-439, 2000.

BUSSAB, Wilton de Oliveira; MIAZAKI, Édina Shizue; ANDRADE, Dalton Francisco. **Introdução à análise de agrupamentos**. São Paulo: IME-USP, 105 p., 1990.

CALLEN, J.L.. *Data envelopment analysis: partial survey and applications for management accounting*. **Journal of Management Accounting Research** 3 (Fall), 35–57, 1991.

CASANOVA, S. C. *Utilização da Análise por Envolvória de Dados (DEA) na Análise de Demonstrações Contábeis*. Tese (Doutorado) Universidade de São Paulo. São Paulo: FEA/USP, 2002.

CATELLI, Armando (Coord.). *Controladoria: uma abordagem da gestão econômica – GECON*. São Paulo: Atlas, 1999.

CHARNES, A.; COOPER, W.W. *Programming with Linear Fractional Functionals*. **Naval Research Logistics Quarterly**, 1962.

CHARNES, A.; COOPER, W.W.; RHODES, E.. *Short communication: Measuring the efficiency of decision making units*. **European Journal of Operational Research** 3, 339, 1979.

CHARNES, A.; COOPER, W.W.; SUEYOSHI T.. *Least Squares/Ridge Regression and Goal Programming/Constrained Regression Alternatives*. **European Journal of Operational Research** 27, pp. 144-157, 1985.

CHARNES, A.; COOPER, W.W.; GOLANY, B.; SEIFORD, L.; STUTZ, J.. *Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions*. **Journal of Econometric**, 30, pp. 91-107, 1985.

CHARNES, A.; COOPER, W.W. SEIFORD, L.M.; STATZ, J.. *Invariant Multiplicative Efficiency and Piecewise Cobb-Douglas Envelopments*. **Operations Research Letters** 2, pp. 101-103, 1983.

CHARNES, A.; COOPER, W.W.; HUANG, Z.M.; SUN, D.B. *Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks*. **Journal of Econometrics**, v. 46, p. 73- 91, 1990.

CHARNES, A.; COOPER, W.W.; LEWIN, A.Y.; SEIFORD, L.M. **Data envelopment analysis: Theory, methodology and applications**. Boston: Kluwer Academic Publishers, 1994.

CHARNES, A.; COOPER, W.W.; RHODES, E. *Measuring the efficiency of decision-making units*. **European Journal of Operational Research**, v. 2, p. 429-444, 1978.

CHRISTENSEN, L. R.; JORGENSEN, D. W.; LAU, L. J. *Transcendental logarithmic production frontiers*. **Review of Economics and Statistics**. Vol. 55 No. 1, pp. 28-45, 1973.

COELLI, T.J. *Recent developments in frontier modeling and efficiency measurement*. **Australian Journal of Agricultural Economics**, v. 39, n. 3, p. 219-245, 1995.

COELLI, T.J.; RAO, D.S.P.; BATTESE, G.E. **An Introduction to Efficiency and Productivity Analysis**, Boston: Kluwer Academic Publishers, 1998.

COOPER, W. W.; HUANG, Z.; LI, S. *Satisficing DEA models under chance constraints*. **The Annals of Operational Research**, forthcoming, 1995.

COOPER, W. W.; SEIFORD, L. M.; TONE, K.; ZHU, J. *DEA: Past Accomplishments and Future Prospects*. **McCombs Research Paper Series No. IROM – 01-05**. June 2005.

COOPER, W.; SEIFORD, L.M.; ZHU, J. *Handbook on Data Envelopment Analysis (International Series in Operations Research & Management Science)*. Springer, 2004.

COOPER, W.W.; PARK, K. S.; PASTOR, J.T.. *RAM: A Range Adjusted Measure of Inefficiency for Use with Additive Models and Relations to Other Models and Measures in DEA*. **Journal of Productivity Analysis**. 11, pp. 5-42, 1999.

COOPER, W.W.; PARK, K. S.; PASTOR, J.T. *The Range Adjusted Measure (RAM) in DEA: A Response to the Comments by Steinmann and Zweifl*. **Journal of Productivity Analysis** 15, pp. 145-152, 2000.

COOPER, W.W., SEIFORD, L.M.; ZHU, J.. **Handbook on Data Envelopment Analysis** (Norwell, Mass: Kluwer Academic Publishers), 2004.

COOPER, W.W.; SEIFORD, L. M.; TONE, K. **Data Envelopment Analysis: a comprehensive text with models, applications, references and DEA-solver software**. Second Edition, New York: Springer, 2007.

COOPER, W.W.; SEIFORD, L.M.; TONE, K. **Data Envelopment Analysis: a comprehensive text with models, applications, references and DEA-solver software**. Boston: Kluwer Academic Publishers, 2000.

CORRAR, L. J.; PAULO, E.; DIAS FILHO, J. M. (Coordenadores). **Análise Multivariada para os Cursos de Administração, Ciências Contábeis e Economia**. São Paulo: Atlas, 2007.

- DANTZIG, G. **Linear Programming and Extensions** (Princeton, N.J., Princeton, 1963.
- DEBREU, G. *The Coefficient of Resource Utilization*. **Econometrica**, 9, pp. 273-292, 1951.
- DESPOTIS, D.K. *Measuring Human Development via Data Envelopment Analysis. the Case of Asia and the Pacific*. **Omega** 33, pp. 285-390, 2005.
- DYSON, R. G.; ALLEN, R.; CAMANHO, V.V.; PODINOVSKI, A. S. V. V.; SARRICO, C. S.; SHALE, E. A. *Pitfalls and protocols in DEA*, **European Journal of Operational Research** 132 (2), pp. 245–259, 2001.
- DYSON, R.G.; THANASSOULIS, E. *Reducing weight flexibility in data envelopment analysis*. **Journal of the Operational Research Society**, v. 39, n. 6, p. 563-576, 1988.
- ENTANI, T.; MAEDA, Y.; TANAKA, H. *Dual models of interval DEA and its extensions to interval data*. **European Journal of Operational Research**, v. 136, p. 32-45, 2002.
- EPSTEIN, M.K.; HENDERSON, J.C. *Data envelopment analysis for managerial control and diagnosis*. **Decision Sciences** 20, 90–119, 1989.
- EVANS, J. R.; OLSON, D. L. **Introduction to simulation and risk analysis**. New Jersey: Prentice-Hall, 1998.
- EVERITT, B.S. **Cluster Analysis** (third edition). Edward Arnold, 1993.
- EXAME MELHORES E MAIORES. São Paulo: Editora Abril, 2008.
- FARE, R.; GROSSKOPF, S.. **Intertemporal production functions with dynamic DEA** (Norwell, Mass. Kluwer Academic Publishers), 1996.
- FARREL, M. J. *The measurement of productive efficiency*. **Journal of the Royal Statistical Society, Series A, (General)**, Vol. 120, No. 3, 253-290, 1957.
- FIZEL, J.L.; NUNNIKHOVEN, T.S. *Technical efficiency of for-profit and non-profit nursing homes*. **Managerial and Decision Economics** 13, 429–439, 1992.

FIZEL, J.L.; NUNNIKHOVEN, T.S. *The efficiency of nursing home chains*. **Applied Economics** 25, 49–55, 1993.

FLEURIET, M.; KEHDY, R.; BLANC, G. **A dinâmica financeira das empresas brasileiras**. Belo Horizonte: Fundação Don Cabral, 1978.

GATTOUFI, S.; ORAL, M.; REISMAN, A. *Data Envelopment Analysis Literature, A Bibliography Update (1951-2001)*, **Socio-Economic Planning Sciences** 38, pp.159-229, 2004.

GUERREIRO, Reinaldo. **A meta da empresa: seu alcance sem mistérios**. São Paulo: Atlas, 1996.

GUJARATI, D. **Econometria básica**. Rio de Janeiro: Elsevier, 2006.

HAAS, D. A. MURPHY, F. H. *Compensating for non-homogeneity in decision-making units in data envelopment analysis*, **European Journal of Operational Research** 144 (3) 530–544, 2003.

HAAS, D.A. **Evaluating the efficiency of municipal reverse logistics channels: an application of data envelopment analysis**. Ph.D. Diss., Temple University, 1998.

HAIR JUNIOR, Joseph F.; ANDERSON, Rolph E.; TATHAM, Ronald L. et al. **Multivariate data analysis: with readings**. 5.ed. Upper Saddle River, New Jersey: Prentice-Hall, 730 p., 1998.

HENDRIKSEN, E. S.; VAN BREDA, M. F. **Teoria da Contabilidade**. São Paulo: Atlas, 1999.

HILDEBRAND, D. K.; OTT, R. L. **Statistical thinking for managers**. USA: Brooks/Cole Publishing Company, 4th ed., 1998.

HUANG, Y.-G.L.; McLAUGHLIN, C.P., *Relative efficiency in rural primary health care: An application of data envelopment analysis*. **Health Services Research** 24, 144–158, 1989.

IUDÍCIBUS, Sérgio de. **Análise de Balanços**. 7 ed. São Paulo: Atlas, 1998.

IUDÍCIBUS, Sérgio de. **Teoria da Contabilidade**. 6. ed. São Paulo: Atlas, 2000.

IUDÍCIBUS, Sérgio de; MARTINS, Eliseu; GELBCKE, Ernesto Rubens. **Manual de contabilidade das sociedades por ações**. 7. ed. São Paulo: Atlas, 2007.

JOHNSON, Richard A.; WICHERN, Dean W. **Applied multivariate statistical analysis**. 4^a ed. Upper Saddle River, New Jersey: Prentice-Hall, 1998.

KANITZ, Stephen Charles. **Indicadores contábeis e financeiros de previsão de insolvência: a experiência da pequena e média empresa brasileira**. São Paulo. Tese (Livre-Docente) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, 1976.

KERLINGER, Fred Nichols. **Metodologia da pesquisa em ciências sociais: um tratamento conceitual**. Tradução de Helena Mendes Rotundo. Revisão Técnica de José Roberto Malufe. 6. Reimpressão. São Paulo: EPU, ©1979.

KOOPMANS, T. **Activity analysis of production and allocation** (New York: John Wiley & Sons), 1951.

LAND, K. ; LOWEL, C.A.K.; THORE, S. Chance constrained data envelopment analysis. **Managerial and Decisions Economics**, Vol. 14, pp. 541-54, 1990.

LETA, F.R.; SOARES DE MELLO, J.C.C.B.; GOMES, E.G.; ANGULO MEZA, L. Métodos de melhora de ordenação em DEA aplicados à avaliação estática de tornos mecânicos. **Investigação Operacional**, v. 25, 2005.

LEVIN, Jack. **Estatística aplicada a ciências humanas**. 2.ed. Tradução e adaptação de Sérgio Francisco Costa. São Paulo: Harper & Row do Brasil, 392 p., 1985

LINS, M.P.E.; ANGULO MEZA, L. **Análise envoltória de dados e perspectivas de integração no ambiente de apoio à decisão**. Rio de Janeiro: Editora da COPPE/UFRJ, 2000.

LINS, M.P.E.; ANGULO MEZA, L.; SILVA, A.C.M. *A multi-objective approach to determine alternative targets in data envelopment analysis*. **Journal of the Operational Research Society**, v. 55, p. 1090-1101, 2004.

LINS, M.P.E.; MOREIRA, M.C.B. **Método I-O stepwise para seleção de variáveis em modelos de análise envoltória de dados**. **Pesquisa Operacional**, v. 19, n. 1, p. 39-50, 1999.

LINS, M.P.E.; NOVAES, L.F.L.; LEGEY, L.F.L. *Real estate value assessment: a double perspective data envelopment analysis*. **Annals of Operations Research**, (forthcoming), 2005.

LYRA, R. L. W. C. **Análise hierárquica dos indicadores contábeis sob a óptica do desempenho empresarial**. Tese (Doutorado) Universidade de São Paulo. São Paulo, 171 p., 2008.

LOVELL, C.A.K. ; PASTOR, J.T.. *Units Invariant and Translation Invariant DEA Models*. **Operational Research Letters**, 18, pp. 147-151, 1995.

LOVELL, C.A.K; SCHMIDT, S.S (eds.), **The measurement of productive efficiency**. New York, Oxford University, 1993.

MACEDO, M. A. S.; Utilização da análise envoltória de dados (DEA) na consolidação de medidas de desempenho organizacional. In Congresso Brasileiro de Custos. Porto Seguro. **Anais do XI Congresso Brasileiro de Custos**, Porto Seguro: ABC, 2004

MACEDO, M. A.S.; SILVA, F. F.; SANTOS, R. M. Análise do mercado de seguros no Brasil: uma visão desempenho organizacional das seguradoras no ano de 2003. São Paulo: **Revista de Contabilidade & Finanças**. Ano XVII, Vol. 5, Especial Atuária, dezembro de 2006.

MACEDO, M. A. S.; CIPOLA, F. C. FERREIRA, A. F. R. Análise do desempenho organizacional sob as perspectivas sócio-ambiental e de imagem corporativa: um estudo apoiado em DEA sobre os seis maiores bancos no Brasil. **REGE. Revista de Gestão USP**, v.15, p. 1-20, 2008.

MARION, José Carlos. **Análise das demonstrações contábeis**. 3. ed. São Paulo: Atlas, 2005a.

MARION, José Carlos. **Contabilidade empresarial**. 11. ed. São Paulo: Atlas, 2005b.

MARTINS, Eliseu. **Análise da correção monetária das demonstrações financeiras: implicações no lucro e na alavancagem financeira**. São Paulo: Atlas, 1979.

MARTINS, Eliseu; ASSAF Neto, Alexandre. **Administração financeira: as finanças das empresas sob condições inflacionárias**. São Paulo: Atlas, 1996.

MARTINS, Eliseu. Avaliando a empresa. **Boletim IOB. Temática Contábil e Balanços**. Bol. 10/1998.

MARTINS, Eliseu (Coord.). **Avaliação de empresas: da mensuração contábil à econômica**. São Paulo: Atlas, 2001.

MARTINS, Eliseu. Análise crítica de balanços. Parte 1. **Boletim IOB. Temática Contábil e Balanços**. Bol. 26. 2005b.

MARTINS, Eliseu. Análise crítica de balanços. Problemas na análise da estrutura dos balanços no Brasil. **Boletim IOB. Temática Contábil – Manual de Procedimentos**. Semana 49/2005. 2005c.

MARTINS, Gilberto de Andrade. **Manual para elaboração de monografias e dissertações**. 2. ed. São Paulo: Atlas, 1994.

MARTINS, G. A. **Estatística Geral e Aplicada**. São Paulo: Atlas, 2001.

MARTINS, Gilberto de Andrade; THEÓPHILO, Carlos Renato. **Metodologia da investigação científica para ciências sociais aplicadas**. São Paulo: Atlas, 2007.

MATARAZZO, Dante C. **Análise financeira de balanços: abordagem básica e gerencial**. São Paulo: Atlas, 1985.

MATARAZZO, Dante C. **Análise financeira de balanços: abordagem básica e gerencial**. 5. ed. São Paulo: Atlas, 1998.

MATARAZZO, Dante C. **Análise financeira de balanços. abordagem básica e gerencial**. 6. ed. São Paulo: Atlas, 2003.

MATIAS, Alberto Borges. **Indicadores contábeis e financeiros de previsão de insolvência: a experiência da pequena e média empresa**. Tese (Livre-Docência) - Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo. São Paulo: FEA-USP, 1996.

McCLAVE, J. T.; BENSON, P. G.; SINCICH, T. **Statistics for business and economics**. USA: Prentice-Hall, 7th ed, 1998.

MELLO, J. C. C. B. S.; MEZA, L. A.; GOMES, E. G.; BIONDI NETO, L. **Curso de análise de envoltória de dados**. Pesquisa Operacional e o Desenvolvimento Sustentável. Gramado, RS, 27 a 30/09/05.

MESTER, L. *Measuring heterogeneity at US banks: accounting for heterogeneity is Important*, **European Journal of Operational Research** 98 (2) 230–242, 1997.

MIGON, M. N. **Eficiência da indústria de transporte aéreo no Brasil: uma aplicação de Análise de Envoltória de Dados (DEA)**. Dissertação (Mestrado) – Programa de Pós-Graduação em Engenharia, da Universidade Federal do Rio de Janeiro. Rio de Janeiro: COPPE/UFRJ, 2000.

MOITA, M. H. V. Um modelo para avaliação da eficiência técnica de professores universitários utilizando análise de envoltória de dados: o caso dos professores da área de engenharias. Tese (Doutorado) Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2002.

MURPHY, D.J.; PEARSON, J.N.. *Evaluating performance of the purchasing department using data envelopment analysis*. **Journal of Business Logistics** 17, 77–91, 1996.

NETER, J.; WASSERMAN, W., WHITMORE, G.A., **Applied Statistics**, third ed. Allyn & Bacon, Boston, 1988.

NIEDERAUER, Carlos Alberto Pittaluga. **Avaliação dos bolsistas de Produtividade em Pesquisa da Engenharia da Produção utilizando Data Envelopment Analysis**. Dissertação (Mestrado) – Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina. Florianópolis: UFSC, 1998.

NIEDERAUER, Carlos Alberto Pittaluga. **ETHOS: um modelo para medir produtividade relativa de pesquisadores baseado na análise por envoltória de dados**. Tese (Doutorado) – Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2002.

NOLAN, J.F. *Determinants of productive efficiency in urban transit*. **Logistics and Transportation Review** 32, 319–342, 1996.

NORMAN, Michel; STOKER, Barry. **Data Envelopment Analysis: the assessment of performance**. Chichester: John Wiley, 1991.

NORUSIS, M. J. **SPSS 13.0 Statistical procedures companion**. New Jersey: Prentice-Hall.

NOVAES, L.F.L. **Função fronteira de produção aplicada para avaliação de eficiência entre plataformas de petróleo – Data Envelopment Analysis – (D.E.A.)**. Dissertação (Mestrado) – Programa de Pós-Graduação em Engenharia, da Universidade Federal do Rio de Janeiro. Rio de Janeiro: COPPE/UFRJ, 1998.

OLINQUEVITCH, José Leonidas; SANTI FILHO, Armando de. **Análise de balanços para controle gerencial**. 4. ed. São Paulo: Atlas, 2004.

OUM, T.H.; YU, C., *Economic efficiency of railways and implications for public policy*. **Journal of Transport Economics and Policy** 28, 121–138, 1994.

PAIVA, F.C. **Eficiência produtiva de programas de ensino de pós-graduação em engenharias: uma aplicação do método Análise Envolvente de Dados – DEA**. Dissertação (Mestrado). Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2000.

PARADI, J. C. et al. *Applications of DEA to measure the efficiency of software production at two large Canadian banks*. **Annals of Operations Research**, v. 73, p. 91-115, 1997.

PARETO, V. **Manuel d'économie politique**. (Paris: Giars & Briere), 1909.

PARK, K.S. *Simplification of the Transformations and Redundancy of Assurance Regions in IDEA (Imprecise DEA)*, **Journal of the Operational Research Society** 55, pp. 1333-1336, 2004.

PASTOR, J. T. *Translation invariance in data envelopment analysis: a generalization*. **Annals of Operations Research**, v. 66, p. 93-112, 1996

PEREIRA DA SILVA, José. **Análise financeira de empresas**. 6. ed. São Paulo: Atlas, 2004.

PILLE, Peter; PARADI, Joseph C. *Financial performance analysis of Ontario (Canada) Credit Unions: an application of DEA in the regulatory environment*. **European Journal of Operational Research**, 139, p. 339-350, 2002.

ANDYCK, R.; RUBINFELD, D. L. **Econometria: modelos & previsões**. Rio de Janeiro: Elsevier, 2004.

HOHLMANN, M. C. Análise de Conglomerados apud CORRAR, L. J.; PAULO, E.; DIAS FILHO, J. M. (Coordenadores). **Análise multivariada para os Cursos de Administração, Ciências Contábeis e Economia**. São Paulo: Atlas, p. 326, 2007.

HOHLMANN, M. C. ; CORRAR, L. J. Análise *cluster* de setores da economia brasileira utilizando indicadores contábeis. XXVIII Encontro da ANPAD, Curitiba, Paraná, setembro, 2004.

PORTER, Michael E. **Estratégia competitiva: técnicas para análise de indústrias e da concorrência**. 7.ed. Rio de Janeiro: Campus, 362 p., 1986

LAGSDALE, C. T. **Spreadsheet modeling and decision analysis**. Cincinnati, Ohio: South-Western. Thomson Learning, 3rd ed , 2001.

RAY, S. C. **Data Envelopment Analysis: theory and techniques for economics and operations research**. Cambridge: Cambridge University Press, 2004.

RAY, S.C., *Resource-use efficiency in public schools: A study of Connecticut data*. **Management Science** 37, 1620–1628, 1991.

RODES, E.L. **Data Envelopment Analysis and approaches for measuring the efficiency of decision making units with an application to program follow-through in U.S. education**. Ph.D. Dissertation, Carnegie Mellon university, 1978.

ROBEIRO, Osni Moura. **Estrutura e análise de balanços: fácil**. 5.ed. São Paulo: Saraiva, 1997.

RODRIGUES, Júlio Cesar R. **Análise de Dados Qualitativos: estratégias metodológicas para as ciências da saúde, humanas e sociais**. 3.ed. São Paulo: Edusp, 156 p., 2001

ROLL, Y.; COOK, W.D.; GOLANY, B. *Controlling factor weights in DEA*. **IEEE Transactions**, v. 23, n. 1, p. 2-9, 1991.

ROSEN, D.; SCHAFFNIT, C.; PARADI, J.C. *Marginal rates and two-dimensional level curves in DEA*. **Journal of Productivity Analysis**, v. 9, n. 3, p. 205-232, 1998.

SAATY, T. L. **Decision Making for Leaders**, RWS Publications, Pittsburgh, 1990.

SAATY, T. L. **Método de análise hierárquica**. São Paulo: McGraw-Hill, 1991.

SAEN, R. F.; MEMARIANI, A. *Determining relative efficiency of slightly non-homogeneous decision making units by data envelopment analysis: a case study in IROST*. **Applied Mathematics and Computation** 165 , 313-328, 2005.

SANT'ANNA, L. A. F. P, **Análise Envoltória de Dados aplicada à avaliação de performance no Sistema Elétrico Brasileiro**. Dissertação (Mestrado) – Programa de Pós-Graduação em Engenharia, da Universidade Federal do Rio de Janeiro. Rio de Janeiro: COPPE/UFRJ, 1999.

SARKIS, J., *Evaluating flexible manufacturing systems alternatives using data envelopment analysis*. **Engineering Economist** 43 (Fall), 25–47, 1997.

SARRICO, C. S.; DYSON, R. G. *Using DEA for planning in UK university—an institutional perspective*, **Journal of Operational Research Society** 51 (7), 789–800, 2000.

SAVYTZKY, Taras. **Análise de Balanços: método prático**. 4. ed. Curitiba: Juruá, 2007.

SENGUPTA, J.K. **Dynamics of Data Envelopment Analysis: Theory of System Efficiency** (Norwell, Mass.: Kluwer Academic Publishers), 1995.

SENRA, L.F.A.C. **Métodos de seleção de variáveis em DEA: estudo de caso no setor elétrico brasileiro**. Dissertação (Mestrado em Engenharia de Produção) - Universidade Federal Fluminense, 88 p., 2004.

SEXTON, T.R.; SLEEPER, S.; TAGGART JR., R.E., *Improving pupil transportation in North Carolina*. **Interfaces** 24 (January–February), 87–103, 1994.

SHAFFER, Scott M.; BYRD, Terry A. *A framework for measuring the efficiency of organizational investments in information technology using data envelopment analysis*. **Omega – The International Journal of Management Science**, 28, p. 125-141, 1999.

SHERMAN HD. *Data envelopment analysis as a new managerial audit methodology--test and evaluation*. **Auditing: a Journal of Practice and Theory**, 4(1), 1994,

SHIMIZU, Tamio. **Decisão nas organizações: introdução aos problemas de decisão** centrados nas organizações e nos sistemas de apoio à decisão. São Paulo: Atlas, 317 p., 2001.

SILVA, A.C.M. **Análise da eficiência das instituições financeiras brasileiras, segundo a metodologia do Data Envelopment Analysis (DEA)**. Dissertação (Mestrado) – Instituto de Pós-Graduação em Administração, da Universidade Federal do Rio de Janeiro. Rio de Janeiro: COPPEAD/UFRJ, 2000.

SIMAK, Paul C. **DEA based analysis of corporate failure**. Thesis (Masters of Applied Science) – Graduate Department of Mechanical and Industrial Engineering, University of Toronto. Toronto (Canadá): University of Toronto, 1997.

SIMAK, Paul C. **Inverse and Negative DEA and their application to credit risk evaluation**. Thesis (Doctor of Philosophy) – Graduate Department of Mechanical and Industrial Engineering, University of Toronto. Toronto (Canadá): University of Toronto, 2000.

SINGH, H.; MOTWANI, J.; KUMAR A. *A review and analysis of state-of-the-art research on productivity measurement*. **Industrial Management & Data Systems** 100/5, 234-241, 2000.

SMITH, P. *Data Envelopment Analysis applied to financial statements*. **Omega – International Journal of Management Science**, 18, 2, p. 131-138, 1990.

SOARES DE MELLO, J.C.C.B.; ANGULO MEZA, L.; GOMES, E.G.; SERAPIÃO, B.P.; LINS, M.P.E. Análise de Envoltória de Dados no estudo da eficiência e dos benchmarks para companhias aéreas brasileiras. **Pesquisa Operacional**, v. 23, n. 2, p. 325-345, 2003b.

SOARES DE MELLO, J.C.C.B.; GOMES, E.G.; ANGULO MEZA, L.; SOARES DE MELLO, M.H.C.. Uma análise da qualidade e da produtividade de Programas de Pós-Graduação em Engenharia. Ensaio - **Avaliação e Políticas Públicas em Educação**, v. 11, n. 39, p. 167-179, 2003a.

SOARES DE MELLO, J.C.C.B.; GOMES, E.G.; SOARES DE MELLO, M.H.C., LINS, M.P.E. Método multicritério para seleção de variáveis em modelos DEA. **Revista Pesquisa Naval**, v. 15, p. 55-66, 2002.

SOARES DE MELLO, J.C.C.B.; LINS, M.P.E.; GOMES, E.G. Estimativa de planos tangentes à fronteira DEA em DMUs extremo-eficientes. In: XXXIII SBPO – **Simpósio Brasileiro de Pesquisa Operacional**, Campos do Jordão. Anais..., São Paulo, 2001.

SOTERIOU, A.; ZENIOS, S.A.. *Operations, quality, and profitability in the provision of banking services*. **Management Science** 45 (September), 1221–1238, 1999.

SOTERIOU, Andreas C.; ZENIOS, Stravos A. *Using data envelopment analysis for costing bank products*. **European Journal of Operational Research**, 114, p. 234-248, 1999.

SUDIT, E. *Productivity measurement in industrial operations*. **European Journal of Operational Research**, Vol. 85, 435-453, 1995.

THANASSOULIS E *A comparison of regression analysis and data envelopment analysis as alternative methods for performance assessments*. **J. Opl Res. Soc.** 44(11), 1129-1144, 1993.

THANASSOULIS E ; DUNSTAN, P .*Guiding schools to improved performance using data envelopment analysis: An illustration with data from a local education authority*. **J. Opl Res. Soc.** 45(11), 1247-1262, 1994.

THANASSOULIS E ;DYSON, R.G. *Estimating preferred input output levels using data envelopment analysis*. **Eur J. Ops Res.** 56, 80-97, 1992.

THANASSOULIS, E. *A data Envelopment Analysis Approach to Clustering Operating Units for Resource Allocation Purposes*. **Omega, Int. J. Mgmt Sci.** Vol. 24, No 4, pp. 463-476, 1996

THANASSOULIS, E. ; ALLEN, R. *Simulating Weight restrictions in Data Envelopment Analysis by means of unobserved DMUs*. **Management Science**, v.44, 1998.

THORE, Stern; KOZMETSKY, George; PHILLIPS, Fred. *DEA of financial statements data: the U. S. computer industry*. **The Journal of Productivity Analysis**, 5, p. 229-248, 1994.

THRALL, R.M.. *Duality, Classification and Slacks in DEA*. **Annals of Operations Research** 66, pp. 109-128, 1996.

TONE, K. *A Slacks Based Measure of Efficiency in Data Envelopment Analysis*. **European Journal of Operational Research** 130, pp. 498-509, 2001.

WONG, Y.H.B.; BEASLEY, J.E. *Restricting weight flexibility in DEA*. **Journal of the Operational Research Society**, v. 41, p. 829-835, 1990.

WONNACOTT, T.H.; WONNACOTT, R. J. **Introductory statistics for business and economics**. USA: John Wiley & Sons, fourth edition, 1990

WOOLDRIDGE, J. M. **Introdução à econometria: uma abordagem moderna**. São Paulo: Pioneira Thomson Learning, 2006.

YAMADA, Y.; MATUI, T.; SUGIYAMA, M. *New analysis of efficiency based on DEA*. **Journal of the Operations Research Society of Japan**, v. 37, n. 2, p. 158-167, 1994.

ZDANOWICZ, José Eduardo. **Estrutura e análise das demonstrações contábeis**. Porto Alegre: Sagra Luzzatto, 1998.

ZHU, J. **Quantitative models for performance evaluation and benchmarking: Data Envelopment Analysis with spreadsheets and DEA excel solver**. Springer, 2002.

ZHU, Joe. *Multi-factor performance measure model with an application to Fortune 500 companies*. **European Journal of Operational Research**, 123, p. 105-124, 2000.