



UNIVERSIDADE FEDERAL DE PELOTAS - UFPEL
PROGRAMA DE PÓS-GRADUAÇÃO EM
ORGANIZAÇÕES E MERCADOS
MESTRADO EM ECONOMIA APLICADA

CARINE MACHADO OLIVEIRA

Eficiência Técnica no Varejo Supermercadista Brasileiro:
Análise Envoltória de Dados e Análise de Redes

PELOTAS
2021

CARINE MACHADO OLIVEIRA

**Eficiência Técnica no Varejo Supermercadista Brasileiro:
Análise Envoltória de Dados e Análise de Redes**

Dissertação apresentada ao curso de Mestrado em Economia Aplicada da Universidade Federal de Pelotas – UFPEL, como requisito parcial para obtenção do título de Mestre em Economia.

Área de Concentração: Economia Aplicada

Orientador: Marcelo de Oliveira Passos

**PELOTAS
29/03/2021**

CARINE MACHADO OLIVEIRA

**Eficiência Técnica no Varejo Supermercadista Brasileiro:
Análise Envoltória de Dados e Análise de Redes**

Dissertação apresentada como
requisito parcial para obtenção do
grau de Mestre em Economia pelo
Programa de Pós-Graduação em
Organizações e Mercados da
Universidade Federal de Pelotas
(UFPEL)

Aprovada em ____/____/____

BANCA EXAMINADORA

Cesar Oviedo Tejada (membro interno)
Universidade Federal de Pelotas - UFPEL

Lívia Madeira Triaca (membro externo)
Universidade Federal do Rio Grande - FURG

Paulo de Andrade Jacinto (membro externo)
Universidade Federal do Paraná – UFPR

Marcelo de Oliveira Passos (orientador)
Universidade Federal de Pelotas – UFPEL

O11e Oliveira, Carine Machado

 Eficiência técnica no varejo supermercadista brasileiro : análise
envoltória de dados e análise de redes / Carine Machado Oliveira ;
Marcelo Oliveira Passos, orientador. — Pelotas, 2021.

 61 f.

 Dissertação (Mestrado) — Programa de Pós-Graduação em
Organizações e Mercados, Instituto de Ciências Humanas,
Universidade Federal de Pelotas, 2021.

 1. DEA. 2. Varejo. 3. Eficiência. I. Passos, Marcelo Oliveira, orient.
II. Título.

CDD : 658

Elaborada por Maria Inez Figueiredo Figas Machado CRB: 10/1612

SUMÁRIO

1. INTRODUÇÃO.....	9
2. REVISÃO DA LITERATURA TEÓRICA E EMPÍRICA.....	1
2.1 Conceitos e métricas de eficiência.....	3
2.2 .	1
2.3 Literatura empírica.....	3
2.4 .	2
3. METODOLOGIA, SOFTWARE E DADOS.....	7
3.1 Software e classificação utilizados.....	2
3.2 .	9
3.2 Metodologia.....	2
3.3 .	9
4 RESULTADOS.....	3
4.1 Resultados do modelo DEA.....	7
4.2 Rede de otimização de modularidade e <i>clusters</i> dos resultados do modelo DEA – VRS.....	3
5 CONSIDERAÇÕES FINAIS.....	4
REFERÊNCIAS.....	6
.	5

Lista de Gráficos

Gráfico 1 -	Utilização média dos insumos I.....	41
Gráfico 2 -	Utilização média dos insumos II.....	42
Gráfico 3 -	Distribuição percentual por formato entre as DMUs eficientes.....	44
Gráfico 4 -	Distribuição percentual por formato entre as DMUs ineficientes.....	44
Gráfico 5-	Rede dos Clusters das DMUs – otimização de modularidade.....	48
Gráfico 6 -	Rede dos graus ponderados das DMUs.....	51

Lista de Tabelas

Tabela 1 -	Formatos de loja por tamanho e número de checkouts.....	30
Tabela 2 -	Descrição das variáveis.....	36
Tabela 3 -	Estatísticas descritivas.....	37
Tabela 4 -	Resultados obtidos pelo modelo DEA-VRS.....	39
Tabela 5 -	DMUs classificadas segundo o parâmetro de Norman e Stoker (1991).....	40
Tabela 6 -	Classificação das DMUs eficientes segundo formato de loja e indicadores econômico-financeiros e de produtividade.....	43
Tabela 7 -	DMUs de eficiência mediana e ineficientes e seus conjuntos de referência	45
Tabela 8 -	Divisão dos clusters (agrupamentos) com as DMUs ineficientes e eficientes, conforme o critério da otimização de modularidade	47
Tabela 9 -	Indicadores econômico-financeiros, formato das DMUs e classes de modularidade.....	50
Tabela 10 -	Graus ponderados, modularidade e medidas de concentração de mercado.....	52

Lista de Siglas

BBC - Banker, Charnes e Cooper
CCR – Charnes, Cooper e Rhodes
CRS – *Constant returns to scale*
DMU – *Decision maker unit*
HHI - Herfindahl–Hirschman Index
PIB – Produto Interno Bruto
VRS - *Variable return to scale*

Eficiência Técnica no Varejo Supermercadista Brasileiro: Análise Envoltória de Dados e Análise de Redes

Carine Machado Oliveira

RESUMO - Neste trabalho, avaliamos os determinantes da eficiência técnica no setor de hipermercados. Com base na análise envoltória de dados com retornos em escala variável (DEA / VRS) e na estimativa de uma rede com clusters computados por otimização de modularidade, medimos a eficiência técnica de 47 empresas do setor, estimamos métricas para análise de benchmark e definimos uma rede complexa com 4 clusters compostos por empresas eficientes, ineficientes e de média eficiência. Também analisamos seus graus ponderados. Usando um modelo DEA-VRS, testamos a hipótese de que as empresas mais eficientes foram aquelas que centraram suas operações em formatos de grandes lojas ou naqueles que combinam atacado e varejo (popularmente chamado de "atacarejo"). Eles fizeram isso como uma forma de enfrentar a recessão de 2013-2015. Entre as empresas eficientes, predominou o formato hipermercado e "atacarejos" (73% do total dos pontos de venda). Verificamos também que, em média, as empresas ineficientes apresentaram maiores níveis de aproveitamento de insumos, número total de caixas, área de vendas, número de funcionários e área de loja por caixa em relação aos valores das empresas que estavam na fronteira da eficiência. A pesquisa também mostrou que se trata de um setor bastante heterogêneo e que, do ponto de vista das variáveis analisadas, as empresas mais eficientes não são necessariamente as que detêm maior participação de mercado.

Palavras-chaves: análise envoltória de dados, algoritmo de otimização de modularidade, análise de redes, comércio atacadista e varejista, organização industrial.

Códigos JEL: L81, C61, C44 e C45.

ABSTRACT - In this work, we evaluate the determinants of technical efficiency in the hypermarket sector. Based on the data envelopment analysis with variable scale returns (DEA / VRS) and the estimation of a network with clusters computed by modularity optimization, we measure the technical efficiency of 47 companies in the sector, estimate metrics for benchmark analysis and define a complex network with 4 clusters composed of efficient, inefficient and medium efficiency companies. We also analyzed their weighted degrees. Using a DEA-VRS model, we tested the hypothesis that the most efficient firms were those that selected business models that combine wholesale and retail (popularly called "cash and carry") or large stores. They did this as a way of facing the 2013-2015 recession. Among the efficient companies, the hypermarket and "atacarejos" format

predominated (73% of the total points of sale). We also found that, on average, inefficient companies had higher levels of input utilization, total number of boxes, sales area, number of employees and store area per box in relation to the values of companies that were on the frontier of efficiency. The survey also showed that it is a very heterogeneous sector and that the most efficient companies are not necessarily the ones with the greatest market share.

Keywords: data envelopment analysis (DEA), modularity optimization algorithm, network analysis, retail and wholesale trade, industrial organization.

JEL : L81, C61, C44 e C45.

1. INTRODUÇÃO

A partir das reformas econômicas iniciadas em 1995, o varejo de alimentos no Brasil passou por diversas mudanças estruturais que facilitaram a entrada de investimentos estrangeiros, uma maior consolidação através de fusões e aquisições e expansão do número de lojas. A estabilidade econômica e a baixa inflação experimentadas no início do Plano Real, impulsionaram o poder de compra já no início dos anos 2000. Com os efeitos do Plano e dado o aperfeiçoamento das políticas redistributivas, a classe média aumentou significativamente entre 2004 e 2008, tendo crescido 22,8%¹ no período. As elevadas taxas de crescimento da classe C e o aumento da classe B, que conjuntamente com a classe A cresceu 33,6% entre 2004 e 2008², fizeram com que o varejo tivesse que se adaptar para atender um perfil de consumidor mais preocupado com a qualidade dos produtos adquiridos, com maior acesso à informação e com um leque de consumo mais diversificado.

Este processo adaptativo impactou positivamente o setor. Fez com que o varejo hipermercadista crescesse, em termos reais, acima do PIB do período de 2011 a 2013³. Todavia, o cenário econômico, após esse período, registrou aumento do desemprego, alta dos juros reais e da inflação. O PIB recuou 7,5% no biênio 2015-2016, resultando em uma forte recessão. A taxa básica de juros Selic, que seguia em uma tendência de baixa até 2012 (7,5%) voltou a subir no ano seguinte (10%) chegando a 14,25% em 2015 (LACERDA, 2018; PONTES

¹ NERI, 2008.

² Ibidem.

³ 6,7% em 2011, 8,4% em 2012 e 5,36% em 2013 contra 2,1% anuais no mesmo período (ABRAS, 2016).

et al, 2017). Em 2015 o PIB apresentou queda real de 3,8% culminando na retração do poder de compra e da oferta de crédito. Tais mudanças impactaram no setor hipermercadista, que depende fortemente do nível de atividade econômica, traduzindo-se em queda real nas vendas de 3,25% (ABRAS, 2016).

A queda nas vendas e as incertezas oriundas do cenário político-econômico intensificaram a concorrência no varejo hipermercadista. Competição esta que foi acirrada pelo aumento da incerteza no setor, decorrente do aumento das taxas de desemprego e da diminuição do nível de confiança dos consumidores. Os consumidores tornaram-se ainda mais exigentes em relação aos preços, estimulando a competição por preços mais baixos entre os varejistas (PONTES *et al*, 2017; PONTES, 2016).

Ao longo dos últimos anos, o consumidor brasileiro tem se esforçado para manter o nível de consumo dos produtos de alta rotatividade. Com isto, a busca por ofertas promocionais, preocupações com saúde e bem-estar e a menor fidelidade a marcas e pontos comerciais passaram a caracterizar o comportamento dos compradores (NIELSEN, 2019).

O segmento denominado como “atacarejo”⁴ respondeu positivamente a este novo perfil de consumidores e ao novo cenário pós-recessivo, crescendo a taxas expressivas (11% em 2017 e 12,3% em 2018). Ele representou, em 2019, em média, a opção de compra de 60% das famílias brasileiras (NIELSEN, 2019; ABRAS, 2018).

No modelo de negócio de “atacarejo”, as lojas minimizam seus custos com base em uma estrutura bem simplificada e buscam atender as necessidades dos consumidores que enfrentam períodos de baixo crescimento econômico já que a demanda se torna mais elástica às flutuações de preços (PONTES *et al*, 2017).

Isto posto, testamos a hipótese de que as firmas mais eficientes, dado o cenário de retração da renda e baixa expectativa de crescimento econômico, são aquelas que priorizam os modelos de atacarejo ou de grandes lojas.

O varejo está em constante reinvenção. Já se passaram algumas décadas desde que a atividade deixou de resumir-se a simples compra e venda de produtos e as firmas travaram uma disputa por diferenciação, quer seja no preço,

⁴ Lojas grandes destinadas a venda ao atacado e varejo onde a maior parte do estoque fica no chão de loja, o *layout* é simplificado, sortimento restrito, número de funcionários reduzidos (a maioria dos repositores e demonstradores são funcionários dos fornecedores) e que oferecem precificação diferenciada para compra em maior quantidade de um mesmo item.

quer na agregação de serviços, na experiência de compra, nos canais de comunicação ou em outros fatores. A facilidade de acesso à internet trouxe para as lojas físicas mais um concorrente: as lojas virtuais de varejistas, que muitas vezes estão geograficamente distantes de seus mercados consumidores, mas que com a evolução da eficiência logística tornaram-se competidores locais (OLIVEIRA, 2020).

O comércio online deixou de ser, para muitos, uma “dor de cabeça concorrencial”, e tornou-se vantagem competitiva. E se mais recentemente a novidade era o esforço competitivo de estar presente nas mais diferentes formas de contato com o consumidor (loja virtual, loja física, televendas, redes sociais), rapidamente o desafio passou a ser conquistar singularidade de informações nos múltiplos canais e facilitar a interação entre eles. A estratégia competitiva, no sentido que Porter (1998) atribui ao termo, tornou-se cada vez mais responsável pelo aumento da lucratividade das empresas varejistas (PORTER, 1998).

A relevância do varejo na economia nacional é inegável: o setor é o que mais gera empregos formais e apresenta taxas consistentes de crescimento e modernização. Considerado pela literatura como um canal de distribuição (KOTLER, 2012; RICHTER, 1954), o varejo tem ganhado protagonismo nos últimos anos em função de seu posicionamento estratégico e da valorização de sua proximidade com o consumidor final.

A atividade varejista compreende supermercados, farmácias, concessionárias de veículos, lojas de vestuários, lojas de materiais de construção, lojas de móveis e decoração, postos de gasolina, lojas de eletroeletrônicos e livrarias que juntos correspondem por 63,4% do PIB nacional⁵. Já o varejo restrito, de onde excetuasse a venda de automóveis e materiais de construção, responde por 20,25%⁶ do produto nacional.

Em 2019 o varejo restrito cresceu 1,8%, contra 1,1% do PIB, movimentou cerca de 1,4 trilhão e, no que se refere ao mercado de trabalho, mantinha 26% dos trabalhadores com carteira assinada no País, o que equivale a mais de 8,5 milhões de pessoas⁷.

⁵ SBVC, dados referentes ao ano de 2007.

⁶ Idem.

⁷ SBVC, 2020.

O varejo é impactado pela tecnologia, pela moda, pela cultura, pelo câmbio, pela inflação, pela cadeia logística, pelo tamanho da loja, pelo *layout*, pelo cheiro da loja, pelo *market share*, pelo ponto de venda, pelo número de caixas (*checkouts*), enfim, no varejo múltiplas variáveis estão em jogo. No entanto, esses fatores têm relevância e impactos diferentes em cada ramo do varejo e variam também de acordo com o público alvo (OLIVEIRA, 2020).

Por isso, nós avaliamos os fatores determinantes da eficiência técnica em um ramo muito específico e importante do varejo brasileiro: o setor hipermercadista. Neste ramo, os principais *players* costumam adaptar *layouts*, tamanho de loja, localização, agregação de serviço, número de funcionários dentre outros fatores às oscilações econômicas e ao público a que se destinam (de acordo com a faixa de renda). Assim, nossa proposta é analisar a eficiência das 47 maiores firmas supermercadistas em operação no Brasil em termos de faturamento.

A literatura econômica, principalmente no ramo da economia industrial, dispõe de vários métodos para medir produtividade, sejam eles paramétricos ou não. Muitos desses métodos buscam comparar *inputs* e *outputs* com o objetivo de estimar uma fronteira que seja capaz de maximizar a produção usando o mínimo possível dos insumos.

Nesse sentido, escolhemos a metodologia da Análise Envoltória de Dados (a qual doravante nós chamaremos pela abreviatura mais conhecida de DEA – *Data Envelopment Analysis*). A DEA é um método não paramétrico muito utilizado para avaliar eficiência. Uma de suas vantagens é a possibilidade de examinar combinações distintas, porém igualmente eficientes, de *inputs* e *outputs*. Esse método também permite identificar uma fronteira de produção dentro do grupo analisado, bem como as unidades que dela fazem parte. Além disso, a DEA determina para cada unidade eficiente aquelas que fazem parte do seu conjunto de referência (ALMEIDA, 2017).

Para entender as características que permitem as firmas formarem determinados agrupamentos e o que as aproxima ou distancia, lançamos mão também da metodologia de redes complexas. Em essência, redes são abstrações que possibilitam codificar relações entre pares de objetos. Barabási (2003) define rede complexa como um grafo que apresenta estrutura topográfica

não usual formado por um conjunto de nodos (vértices) interligados por meios de arestas.

Adicionalmente são também apresentados o índice de Hirshman-Herfindahl (HHI) e o *Market share* com a finalidade de ilustrar mais claramente a organização do mercado e o posicionamento de cada firma. Estes índices permitem avaliar a repartição do mercado com base no faturamento.

Nós também pretendemos estimar dados úteis para a avaliação de *benchmarks* (ALMEIDA, 2017). Ao delimitarmos, por meio da DEA, uma fronteira formada pelas firmas mais eficientes, nós fornecemos uma análise que pode se constituir em uma ferramenta quantitativa contínua de avaliação de produtos, serviços e processos, a qual permite identificar as melhores práticas no mercado a fim de promover melhorias no âmbito das organizações aqui analisadas (STAPENHURST, 2009). Nesse sentido, dado a aplicação multidisciplinar do método, além de estarmos pesquisando economia aplicada - uma vez que a DEA é uma ferramenta amplamente utilizada por economistas industriais, como veremos na revisão da literatura da seção 2 – pretendemos também contribuir com a literatura de organizações e mercados (*business economics*), de pesquisa operacional e de marketing.

2. REVISÃO DA LITERATURA TEÓRICA E EMPÍRICA

2.1. Conceitos e métricas de eficiência

O modelo DEA foi proposto pela primeira vez por Charnes et al. (1978) como um instrumento para medir a eficiência relativa entre as unidades de tomada de decisão. Desde o seu desenvolvimento, o modelo tem sido amplamente utilizado como uma técnica popular de gestão de diagnóstico em diversas disciplinas relacionadas à *business economics*, tais como pesquisa operacional e produção (KEH et al., 2006; LOZANO et al., 2004), avaliação do setor de hospitais e de saúde (REYNOLDS & THOMSON, 2007; WOBER & FESENMEIER, 2004; HWANG e CHANG 2003; MOREY & DITTMAN, 1995), publicidade (LOU & DONTU, 2001) e lojas de varejo (DONTU & YOO, 1998) Medição de eficiência e benchmarking etc. A técnica também permite identificar uma fronteira de eficiência que representa os requisitos mínimos de recursos de

entrada para uma empresa alcançar um determinado nível de produção (ANDERSON et al., 2000).

A preocupação com métricas rigorosas para auferir a eficiência permeiam os estudos da Teoria da Produção há décadas e envolvem o estabelecimento de critérios que permitam identificar a melhor situação possível que pode ser alcançada pelos agentes econômicos na distribuição da produção. A perseguição do menor nível de utilização dos insumos, do nível de produção mais adequado e do menor custo associado a maior receita possível são elementos fundamentais na estratégia competitiva das firmas. Segundo Porter (1997), o objetivo de qualquer firma é maximizar a lucratividade e, para alcançar tal feito, os executivos prescrevem combinações de maximização dos recursos, eliminação de ineficiências, incrementos de produtividade e formas modernas de administração como qualidade total e *benchmarking*.

A Economia Industrial tem feito importantes contribuições para a compreensão das estratégias competitivas industriais alicerçadas na interação entre eficiência produtiva e rentabilidade do patrimônio líquido e a estrutura de mercado bem como com as estratégias de preço e inovação. De acordo com Jacquemin (1987), a análise de mercado, seja para firmas estabelecidas ou entrantes, necessita de adequada caracterização, assim, o principal objetivo da economia industrial tem sido fornecer essa caracterização com base em um esquema que relaciona a estrutura de mercado com os agentes e com os frutos que essa relação gera. Na visão da autora, essas questões são igualmente relevantes, sejam em firmas ou organizações públicas.

A melhor utilização dos recursos seja na prestação de serviços públicos, em atividades filantrópicas ou empresariais são cruciais para sobrevivência das firmas e instituições o que torna a análise envoltória de dados uma ferramenta de análise importante já que ela trata da otimização da utilização dos recursos enquanto condição relevante por si só (FERREIRA; GOMES, 2020).

A permanência de uma empresa no mercado depende, dentre outros fatores, de seu conjunto, de sua interação com o mercado e concorrentes. Assim, é preciso saber posicionar-se estrategicamente, o que, por vezes, significa perseguir as melhores práticas moldando seu comportamento de acordo com os movimentos dos líderes do setor (*benchmarking*) a fim de alcançar o chamado emparelhamento (*catch-up*), nivelar-se aos melhores padrões de eficiência

técnica de produção e de utilização da tecnologia (frontier-shift) (FERREIRA; GOMES, 2020).

De maneira geral, a ciência econômica se preocupa com a alocação eficiente de recursos escassos. Mais especificamente, a ciência econômica se preocupa em estudar a produção, a distribuição e o consumo de bens e serviços (ALMEIDA, 2017b). A empresa, enquanto unidade de produção, busca maximizar lucro adequando os fatores de produção (trabalho, capital e recursos naturais) para minimizar custos. Enquanto ferramenta que permite a análise da produção, a análise envoltória de dados orientada a insumos, por sua vez, indica a redução equiproporcional na utilização dos insumos para que se alcance a eficiência técnica. Essa redução no uso de insumos pode se traduzir em redução de custos e melhor posicionamento competitivo da firma já que a ineficiência, nesse caso, é oriunda da utilização inadequada dos recursos. Assim, a DEA se constituiu um instrumental relevante nos estudos econômicos como ferramenta para otimizar a alocação dos recursos.

As primeiras contribuições para o desenvolvimento da DEA são atribuídas a James Farrel que, em 1957, tentou desenvolver métodos mais eficientes para avaliar a produtividade. Para o autor, as medidas existentes até então eram limitadas uma vez que não permitiam combinar medidas de múltiplos insumos para gerar um indicador único de eficiência. Em seu estudo, Farrell limitou-se a utilização de apenas um produto já que o emprego de múltiplos produtos não atendia aos requisitos para a aplicação a grandes conjuntos de dados. O modelo permitia duas abordagens: com a proporção mínima dos inputs reais, dados os outputs e o mix de inputs, ou como a proporção de real máxima de outputs, dados os inputs (FARRELL, 1957; FARE & KNOX LOVELL, 1978; FERREIRA & GOMES, 2020).

De acordo com Fare e Lovell (1978) os trabalhos que se seguiram sobre mensuração da eficiência técnica são baseados em alguma extensão do artigo seminal de Farrell. Os autores pontuaram alguns problemas relevantes a cerca do modelo de Farrell: a abordagem possui duas interpretações que, exceto nas circunstâncias mais favoráveis, dão origem a diferentes inferências quanto à eficiência técnica; Farrell mede a eficiência técnica relativa a uma isoquanta, em vez de a um subconjunto eficiente, o que pode levar a avaliar uma firma como eficiente quando na verdade ela não o é; trata-se de uma medida radial, no

sentido de que assume um determinado conjunto de insumos, embora não exista motivo para medir a eficiência técnica radialmente, mesmo para tecnologias homotéticas.

Mais tarde, em 1978, Rhodes, em sua dissertação de PhD orientada por William W. Cooper, desenvolveu um método para avaliar a eficiência na tomada de decisão. Esse método foi especialmente pensando para avaliar programas públicos e era possível de ser trabalhado com múltiplos inputs e outputs mesmo que expressos em diferentes unidades de medida. A proposta de Cooper *et al* (1978) era baseada em retornos constantes de escala e não levava em consideração o tamanho das unidades tomadoras de decisão (doravante chamadas de DMUs ou *Decision Making Units*), já os *inputs* e *outputs* variavam em iguais proporções (CHARNES, COOPER, RHODES, 1978).

A grande contribuição de Rhodes veio ao fazer a conexão explícita entre um índice de produtividade mensurado pela soma ponderada dos outputs e dividido pela soma ponderal dos inputs com o método de mensuração da eficiência técnica de Farrell (1957). A inovação se deu em encontrar pesos que maximizassem a relação output/input unindo o conceito de produtividade da engenharia com o conceito de eficiência encontrado na economia (CHARNES, COOPER, RHODES, 1978; BARBOSA, 2019).

A teoria econômica (macro e micro) sempre trabalhou com problemas de otimização condicionada. No âmbito da teoria microeconômica da produção isso significar obter o maior benefício possível de uma dada quantidade de recursos (insumos). Eficiência econômica é aquela situação em que não há mais nada para melhorar, nenhum aperfeiçoamento a fazer, sem piorar alguma outra situação ou estado. Dependendo do contexto, o termo pode ser usado para duas concepções similares, porém não permutáveis: (i) eficiência alocativa ou eficiência paretiana; e (ii) eficiência produtiva.

Um arranjo entre duas partes é Pareto eficiente quando qualquer alteração para melhorar a situação de uma parte implica em piorar a situação da outra. Por outro lado, na eficiência produtiva, temos um cenário no qual não é possível nenhuma produção adicional sem incorrer em aumento de custos.

Nesse sentido, a economia do bem-estar pode ser definida como sendo o estudo do quanto são desejáveis, do ponto de vista social, os arranjos de atividades econômicas e alocações de recursos. É a análise da aceitabilidade

social de estados econômicos alternativos, sendo estes estados definidos por arranjos particulares das atividades econômicas e dos recursos da economia, como mencionamos. O teorema fundamental do Bem Estar postula que numa economia de trocas, desde que não existam externalidades, a distribuição de mercadorias promovida pelo equilíbrio competitivo do mercado é eficiente no sentido de Pareto. A recíproca do teorema também é válida: desde que não existam externalidades, toda distribuição eficiente no sentido de Pareto corresponde a um equilíbrio competitivo. Para que este teorema seja válido é preciso que as preferências e o conjunto de produção sejam convexos (ARROW, 1951, BATOR, 1957, MAS-COLLEL, WHINSTON e GREEN, 1995).

Ocorre que a eficiência pode ser abordada ainda sobre outras perspectivas. Os conceitos relacionados diretamente à atividade produtiva baseiam-se na concepção de que a relação entre insumo e produto não é determinística e que a quantidade produzida a partir de um montante fixo de insumos pode variar de acordo com aspectos motivacionais. Assim, mesmo que sejam atendidos os pressupostos do Teorema Fundamental do Bem Estar, podem existir ineficiências (BOUERI, 2015).

A Lei de Pareto foi adaptada pelo holandês Nobel de Economia Tjalling C. Koopmans que estabeleceu como princípio de eficiência produtiva que um gestor, que produz dois ou mais produtos, possui eficiência absoluta se, e somente se, for tecnologicamente impossível reduzir algum insumo sem incrementar algum outro para manter o nível de produção, ou ainda, se só for capaz de aumentar a produção de um produto diminuindo a produção de outro (Koopmans, 1951 apud RUGGIERO, 2000).

Cooper e Charnes utilizaram as contribuições de Pareto e Koopmans para sintetizar uma definição de eficiência (Definição Estendida de Pareto-Koopmans) segundo a qual a eficiência absoluta de uma unidade é alcançada se o nível de nenhum de seus insumos e ou produtos puder ser melhorado sem piorar os demais insumos ou produtos

Todavia, esse conceito pressupõe que o nível máximo de eficiência possível é conhecido previamente, assim, dado que na prática esses níveis são desconhecidos essa definição foi substituída pelo conceito de eficiência relativa segundo o qual a eficiência máxima de uma DMU é alcançada apenas se o desempenho das demais unidades em análise não mostrar que algum

insumo/produto possa ser melhorado sem afetar negativamente os demais (FERREIRA; GOMES, 2020).

Com os conceitos expostos, podemos agora definir o escopo do conceito de eficiência adotado neste trabalho. Resumimos as métricas de eficiência em uma classificação tripla: técnica, alocativa e econômica. A eficiência técnica diz respeito a capacidade de maximizar a produção a partir de um volume dado de recursos. Já a eficiência alocativa se refere a seleção de uma combinação de *inputs* que gerem uma dada quantidade de produto ao custo mínimo possível. A eficiência econômica, por sua vez, é o produto da eficiência técnica pela eficiência alocativa (COELLI *et al*, 1998; ALMEIDA, 2017).

Neste trabalho, nos debruçamos exclusivamente sobre a análise da eficiência técnica, aquela classe de eficiência que está intimamente associada à fronteira de produção, uma vez esta mensura o êxito da empresa na geração do maior nível de produtos possível a partir de um dado conjunto de insumos. Este tipo de análise é capaz de fornecer informações relevantes sobre como estão sendo utilizados os recursos por cada empresa. E isto é útil como um indicador estratégico (*benchmark*), uma vez que permite um critério objetivo de comparação da eficiência entre diferentes competidores (PERRIGOT; BARROS, 2008).

Os métodos de avaliação de eficiência não se restringem às indústrias e podem ser aplicados a qualquer empresa que transforme *inputs* em *outputs*. Estes métodos podem ser usados para avaliar firmas globalmente, suas subdivisões (as subdivisões de sua operação/aspectos específicos da operação) ou departamentos.

A Análise de Fronteira Estocástica e a Análise Envoltória de Dados são as principais técnicas utilizadas para avaliação de eficiência, são comumente aplicadas a uma amostra de firmas em um dado momento no tempo e mensuram a eficiência entre elas. Tais técnicas não assumem que todas as firmas da amostra são eficientes. Ao contrário da Análise de Fronteira Estocástica, a DEA é um método de programação matemática que não exige especificação da forma funcional dos dados. (COELLI *et al*, 1998).

As empresas produtoras são aqui chamadas de unidades de tomada de decisão (DMU), que possuem o objetivo de maximizar lucros. Esse processo, não determinístico, uma vez que é afetado por múltiplos fatores. Os conceitos de

eficiência e produtividade, nesse contexto, espelham a preocupação com a otimização da utilização dos recursos no intuito de obter o maior retorno econômico possível.

A eficiência relativa é dada pela comparação entre a produção de uma determinada DMU e a produção máxima possível utilizando o mesmo conjunto de *inputs*. A produção máxima, por sua vez, no caso da utilização de DEA, é determinada a partir da utilização de técnicas de programação matemática. Constroem-se referências a partir da amostra afim de estabelecer quais as DMUs eficientes e as ineficientes tendo por base o conjunto de *inputs* utilizados na produção e os *outputs* gerados por cada DMU. As DMUs identificadas como eficientes servem como referência para as demais DMUs possibilitando a análise de benchmark. (ALMEIDA, 2017).

Esta metodologia também permite a criação de uma DMU virtual usando uma combinação de dados das DMUs reais. Essa DMU virtual e eficiente também serviria como espelho para as demais firmas. Todavia, o problema de utilizar uma DMU virtual como benchmark reside no fato de que a DMU virtual pode ser construída a partir de uma combinação de *inputs* que não é factível. Essa possibilidade impõe a necessidade de determinar restrições ao modelo para que todas as combinações de inputs sejam de fato factível. Tais restrições serão detalhadas mais a frente (ALMEIDA, 2017).

Posto isso, uma vantagem da atualização da DEA é que, ao invés de utilizar uma única otimização como referência para todas as DMUs, como faria um modelo de MQO, utiliza-se uma série de otimizações, uma para cada DMU (THOMAS *et al*, 1998).

Zhu (2000) considera a DEA o modelo mais adequado para avaliar a eficiência quando comparada com as demais ferramentas uma vez que ela permite estabelecer uma medida de eficiência relativa entre diferentes unidades independentes estabelecendo uma combinação explícita de insumos e produtos. Já para Macedo (2004), a DEA representa uma alternativa aos modelos convencionais de análise uma que, para a construção da fronteira, realiza uma otimização para cada uma das DMUs, o que não acontece nos métodos tradicionais. Ademais, nesse modelo de análise, as DMUs eficientes não representam apenas desvios em relação à média, mas possíveis benchmarks que podem servir como modelo a ser perseguido por outras firmas.

Dentre os fatores que tornam a escolha da DEA uma alternativa interessante, pode-se acrescentar o fato de que ela a) fornece estimativas dos níveis de inputs e outputs que tornariam cada unidade eficiente; b) pode dispensar o sistema de preços; c) permite verificar os valores ótimos de insumos e produtos de acordo com restrições de factibilidade d) gera resultados ótimos no sentido de Pareto-koopmans e) não requer padronização de unidades de medida (GOMES; BAPTISTA, 2004; THANASSOULINS et al., 2008; ALMEIDA, 2017).

Muitos modelos de análise de eficiência são complexos e de difícil aplicação fora da esfera da pesquisa científica. Todavia, a DEA permite uma metodologia mais flexível sem perda de rigor científico. Tal flexibilidade permite uma adaptação mais compatível com os vários contextos que fazem parte do ambiente de negócios.

DEA está baseada na programação matemática, a qual é utilizada para mensurar a eficiência de unidades tomadoras de decisão (DMUs), as quais utilizam vários insumos em seus processos produtivos. Como o método é uma extensão da programação matemática fundamental do trabalho de Farrel (1957), o qual media a eficiência técnica em casos envolvendo um único insumo e único produto, a versão generalizada da técnica possui algumas vantagens sobre os métodos paramétricos tradicionais da análise de regressão. Enquanto a análise de regressão calcula a eficiência da unidade organizacional por meio de uma aproximação à média, DEA foca nas observações individuais e otimiza a eficiência de cada unidade (Gregoriou, 2006). Nesse sentido, DEA evita problemas em geral associados às aquelas regressões que requerem uma especificação, definida a priori, das relações entre as variáveis de insumo e produto. Trata-se, assim, de uma métrica de eficiência relativa pois mede o desempenho da unidade em análise em face das outras unidades. Apresentando, portanto, sensibilidade à inclusão ou exclusão de qualquer unidade da análise.

O modelo CRS (*constant returns to scale*) é adequado às unidades que apresentam retornos constantes à escala, ao passo em que o modelo VRS (*variable returns to scale*), desenvolvido por Banker, Charnes e Cooper em 1984 é apropriado para unidades que trabalham com retornos variáveis à escala.

Ademais, os modelos de análise envoltória de dados diferem dos métodos baseados em avaliação puramente econômica que requerem a conversão dos inputs e outputs em unidades monetárias, também apresentam a vantagem de serem baseados em dados reais ao invés de formulações teóricas, levando em consideração o fato de que *outliers* possam ser *benchmarks* e não apenas desvios e, ao contrário dos modelos paramétricos tradicionais, os modelos DEA fazem uma otimização para cada DMU no intuito de construir uma fronteira linear por partes compostas pelas unidades eficientes no sentido de Pareto.

A DEA permite ainda levar em consideração o conhecimento do pesquisador em relação ao objeto de análise. Através da agregação ao modelo de restrições que reflitam a realidade do setor avaliado é possível acomodar juízos de valor fazendo com que os resultados sejam factíveis (ALMEIDA, 2017).

No que se refere a amostra, não existe um consenso ou regra sobre a quantidade ou proporção entre o número *inputs* e DMUs, no entanto, recomenda-se que o número de observações seja o maior possível, ou, pelo menos, cinco vezes o número de variáveis. Para Gomes e Baptista (2004), o número de variáveis não deve superar a terça parte do número de observações.

Foram verificados, na literatura acerca da eficiência no varejo hipermercadista, estudos que trabalham com os mais diferentes tamanhos de amostra, algumas inclusive menores daquela aqui utilizada, são exemplos Keh e Chu (2003) que analisaram 13 lojas ao longo de 10 anos, Barros e Alves (2004) que estudaram 47 firmas ao curso de dois anos, Ferreira et al. (2009) que analisaram 300 empresas durante um ano e Gauri (2013) que analisou 50 hipermercados no período de um ano.

Ao trabalhar com DEA pode-se optar por um modelo com Retornos Constantes de Escala (CRS do inglês *constant returns to scale* ou CCR em referência aos seus percussores – Charnes, Cooper e Rhones, 1978) ou por um modelo com Retorno Variáveis de Escala (BCC - Banker, Charnes e Cooper, 1984). O modelo CCR constrói uma fronteira de produção de modo que as firmas mais eficientes fiquem situadas sobre ela e as menos eficientes abaixo dela, numa região chamada “envelope” (PÉRICO, 2009). Nesta fronteira não paramétrica e linearmente convexa, é impossível aumentar qualquer *output* sem aumentar *inputs* ou ainda reduzir *inputs* sem que isso também reduza *outputs*, assim que se alcance a eficiência no sentido de Pareto.

O conceito de rendimento em escala reflete a relação entre um acréscimo nas quantidades de *inputs* e o consequente aumento nos *outputs*. Assim, escala designa a taxa de crescimento do produto como resultado de um crescimento proporcional do produto. Desse modo, uma fronteira eficiente que represente as DMU's em cada nível de análise deve também mostrar aquelas DMU's que em seu nível de atuação alcançaram a escala máxima de produção possível (BENICIO e MELLO, 2014).

A existência de retornos variáveis de escala está associada ao comportamento não uniforme na fronteira. Dessa forma, a fronteira com retorno crescente deve apresentar segunda derivada maior que zero, assim tem-se uma curva convexa voltada para cima. Já para retornos decrescentes, está curva possui derivada menor que zero sendo convexa e voltada para baixo (BENICIO & MELLO, 2014).

Outro método muito frequentemente utilizado para análise de eficiência é a Fronteira Estocástica. Trata-se de uma abordagem que parte do princípio de que os dados são oriundos de uma distribuição de probabilidade (abordagem paramétrica) cuja fronteira é estimada através de métodos econométricos. Neste modelo o resíduo é dividido em dois componentes independentes entre si: o primeiro captura os aspectos referentes à ineficiência técnica e o segundo outros resíduos estatísticos que estão fora do controle da firma, como erro de medida, e choques exógenos, como greves e questões climáticas, por exemplo (ALMEIDA, 2017b).

A análise de fronteira estocástica pode ser aplicada tanto a dados individuais como a dados agregados, todavia, recomenda-se a aplicação a dados individuais pois existe o risco de que especificidades sejam perdidas com a agregação e a medida de eficiência obtida não seja uma representação autêntica da realidade.

Ademais essa abordagem requer especificação da forma funcional da tecnologia de produção e da distribuição de probabilidade do resíduo. É também necessário satisfazer algumas suposições, são elas (ALMEIDA, 2017b; SILVA, 2017):

- I. O erro simétrico tem média zero e variância homoscedástica;
- II. O erro possui distribuição meio-normal e é não negativo;

- III. Ambas as partes do termo de erro são distribuídas independentemente uma da outra e dos regressores.

Dentre as demais abordagens paramétricas disponíveis, a eficiência pode ser analisada por meio da análise de regressão simples (que não forma fronteira) e da fronteira determinista. Já na abordagem não paramétrica existe também o método de números índices.

2.2. Análise de redes complexas: graus ponderados, clusters e algoritmo de otimização de modularidade.

Para representar redes recorre-se à Matemática, mais especificamente ao ramo da Teoria dos Grafos. Porém, para analisar estas redes, é fundamental recorrer a outro ramo da Matemática, a Estatística. Utilizar medidas Estatísticas para analisar redes complexas (incluindo as redes sociais) permite compreender e avaliar a estrutura da rede sem a necessidade de conhecer a sua representação gráfica. O objetivo dessas medidas é quantificar a estrutura das redes, para possibilitar ao analista a compreensão do comportamento dos fenômenos sociais que geraram essas redes.

Na estatística de redes existem medidas ao nível dos agentes (ou nós) e ao nível da rede, como um todo. As mais utilizadas medidas ao nível do agente são as de centralidade do grau ou valência, as quais são definidas pelas seguintes expressões:

$$k_i = \sum_{j=1}^n a_{ij}, \quad 0 < k_i < n \quad (1) \quad \text{e} \quad k_v = |N_v| \quad 0 < k_v < n \quad (2)$$

Onde a_{ij} é a entrada da i -ésima linha e j -ésima coluna da matriz de adjacência A . E N_v é a vizinhança do agente (nó ou vértice) V .

Para redes direcionadas temos:

k_i^+ = grau de entrada (número de agentes de entrada, isto é, número das arestas ou relações que começam no agente v).

k_i^- = grau de saída (número de agentes de saída, isto é, número das arestas ou relações que terminam no agente v).

$$k_i^+ = \sum_{j=1}^n a_{ij} \quad (3) \quad , \quad k_i^- = \sum_{j=1}^n a_{ji} \quad (4)$$

A medida do grau em redes direcionadas é também conhecida como prestígio. É uma expressão muito usada em ARS (análise de redes sociais).

Existem dois tipos de prestígio: (i) o de suporte; e (ii) o de influência. O de suporte é o grau de entrada e o de influência é o grau de saída. Em redes pesadas (ou ponderadas) a força é equivalente ao grau. Ela é igual à soma dos pesos das arestas adjacentes a um dado agente (ou das relações vinculadas a este agente). Tal como em (5):

$$k_i^w = \sum_{j=1}^n a_{ij}^w \quad (5)$$

Uma das características únicas de muitas redes complexas, quer sejam econômicas, sociais, financeiras e biológicas, é que podem possuir estrutura de comunidades (agrupamentos ou *clusters*). Normalmente essa propriedade emerge como consequência da heterogeneidade global e local da distribuição das arestas num grafo, então nesse tipo de rede é possível encontrar concentrações elevadas de arestas em determinadas regiões, e baixa concentração de arestas entre essas regiões.

Comunidades ou *clusters* são grupos de vértices (ou agentes) densamente conectados com ligações esparsas entre eles. De acordo com Newman e Girvan (2004), existem duas linhas de investigação principais na descoberta de comunidades de redes. A primeira teve origem no Campo da Ciência da Computação e é conhecida como partição de grafos, enquanto a segunda foi essencialmente desenvolvida por sociólogos, sendo usualmente referida por *blockmodeling*, *agrupamento hierárquico* ou *deteção de estrutura de comunidades*. O processo básico subjacente aos algoritmos de deteção de comunidade baseia-se na divisão do grafo original num conjunto de subgrafos disjuntos, por via da otimização de uma dada função objetivo. O propósito de ambas as abordagens é descobrir grupos de vértices relacionados e, se possível, definir a respectiva organização hierárquica, tendo por base informação fornecida pela topologia da rede. Isto é geralmente realizado removendo

iterativamente as arestas pontes que ligam grupos de vértices, conforme sugerido por Girvan e Newman (2002).

Na vida real é possível encontrar uma variedade de exemplos de grupos coesos, ou comunidades. A sociedade é um ambiente rico em encontrar comunidades, uma vez que as pessoas têm uma tendência natural para formar grupos. Esses grupos podem ser famílias, círculos de amigos, grupos religiosos ou de trabalho, cidades, nações, etc. Se também consideramos grupos formados por empresas, ou consumidores de um dado produto, é possível identificar comunidades com relevância para a área da Economia e da Gestão.

A importância de estudar estas comunidades é intuitiva em domínios como a ARS (Análise de Redes Sociais). Para sublinhar esta importância, Fortunato (2010) afirmou que a análise da suposição estrutural dos nós, em cada comunidade da rede, pode ajudar a identificar *atores centrais*, associados a funções de estabilidade e controle do grupo, bem como *atores intermediários*, que são aqueles que se localizam nas fronteiras das comunidades e desempenham um papel fundamental na disseminação e troca de informações e novas ideias, criando pontes entre comunidades.

O agrupamento hierárquico é uma classe de métodos para detectar clusters, ou grupos. Algoritmos hierárquicos geram estruturas de grupos inseridos dentro de grupos maiores que, por sua vez, se encontram inseridos em grupos ainda maiores, que são representados por dendrogramas que mostram a estrutura multinível da rede. Esses métodos são eficazes na solução de problemas de análise de grupos e problemas semelhantes, como a partição de grafos e a identificação de comunidades.

O agrupamento hierárquico é bastante intuitivo, sendo baseado na definição de semelhança. Primeiro é necessário escolher uma medida de semelhança (ou dissemelhança) para avaliar quão semelhantes são dois nós, de acordo com uma dada propriedade global ou local⁸. Em seguida, deve-se calcular a matriz de semelhança entre todos os pares de nós, independente desses nós estarem conectados entre si. Depois, é preciso selecionar um método para agrupar os nós: os *métodos aglomerativos*, que focam nas regiões

⁸ Entre os exemplos dessa medida estão a semelhança do cosseno, o índice de Jaccard, a distância Euclidiana, a distância de Manhattan e a distância de Hamming entre pares de linhas numa matriz de adjacência.

mais densas da rede ao invés de focar nas ligações das fronteiras da rede⁹, ou os *métodos divisivos*, que focam na identificação e remoção das ligações que conectam regiões densamente conectadas a rede, sobretudo as *pontes* e as *pontes locais* (Easley e Kleinberg, 2010)¹⁰. Conforme a escolha, uma medida de distância é selecionada para calcular a semelhança entre grupos¹¹. O resultado final desse processo é um dendograma que ilustra a organização dos nós retornada pelo algoritmo hierárquico. Para selecionar os melhores métodos, uma estratégia é calcular o valor de modularidade (Newman e Girvan, 2004) para as comunidades e selecionar o número que maximiza essa função.

A *otimização (classe) de modularidade* é outro tipo de método utilizado para detectar comunidades em redes. A *modularidade Q* é uma função de qualidade que avalia e mede a importância de uma dada partição da rede em comunidades. Esta função é utilizada para comparar a qualidade das partições e também como uma função objetivo em problemas de otimização. Segundo Newman (2006), a modularidade é representada pela diferença normalizada entre o número de arestas observadas no interior de cada grupo de nós da rede e o número de arestas que seria provável observar no interior desse mesmo grupo numa rede equivalente onde as arestas são geradas aleatoriamente. A modularidade Q é calculada da seguinte forma:

$$Q = \frac{1}{2m} \sum_{ij} \left[A_{ij} \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \quad (5)$$

Onde m indica o número de arestas; k_i e k_j representam respectivamente o grau dos vértices i e j; A_{ij} é a entrada da matriz adjacência que indica que o número de ligações estabelecidas entre os vértices i e j; $\frac{k_i k_j}{2m}$ representa o número esperado de arestas que deveria existir entre o par de vértices (i, j); c_i e

⁹ O algoritmo *Walktrap* (elaborado por Pons e Latapy, 2005) é um exemplo deste tipo de método.

¹⁰ O conhecido algoritmo de Girvan e Newman (2002) é um exemplo bastante citado deste método.

¹¹ Alguns exemplos dessa medida: o *single linkage* (ou vizinho mais próximo), o *complete linkage* (ou vizinho mais afastado) e o *método de Ward*.

c_j denotam os grupos a que os vértices i e j pertencem; e $\delta(c_i, c_j)$ representa o delta de Kronecker.

A modularidade Q pode assumir valores positivos e negativos. Se $Q > 0$, então existe a possibilidade de encontrar estrutura de comunidade em rede. Se Q for um número positivo e elevado, então a respectiva partição tem maior probabilidade de refletir a estrutura de comunidade verdadeira. De acordo com Clauset et al. (2004), uma modularidade que assuma valores superiores ou iguais a 0,3 é um bom indicador da existência de comunidade com significado na rede.

2.3. Literatura empírica

Kahraman e colegas (2008) empregam o modelo de análise envoltória de dados com retornos variáveis orientado para insumos para avaliar a performance de uma cadeia de supermercados na Turquia. Os autores utilizam como inputs o tamanho da loja, número de empregados, número de entregas e o custo total. Por sua vez, o número de clientes, volume de vendas e avaliação dos clientes são utilizados como outputs.

Yu e Ramanathanm (2008), a partir dos dados de 41 firmas, avaliaram a eficiência operacional do setor varejista no Reino Unido no período de 2000 a 2005 utilizando DEA, o índice de Malquist e uma regressão Tobit. Os autores tomaram como outputs o volume de negócios e o lucro líquido e como inputs o número de empregados, o total de ativos e a participação acionária sobre os ativos. Os autores elucidam que a análise envoltória de dados foi importante na pesquisa para testar hipóteses relevantes sobre o impacto das variáveis ambientais (a localização da sede, tipo de propriedade, anos de constituição, forma jurídica e características de varejo) sobre o funcionamento do setor varejista. A análise mostrou que apenas 10 firmas eram eficientes usando o modelo de retornos constantes, já usando o modelo de retornos variáveis foram identificadas 16 firmas eficientes em 2005.

Ko e colegas (2017) avaliam a eficiência em uma cadeia de lojas na Coreia a partir da análise envoltória de dados. Na pesquisa, utilizam como inputs a área de vendas, o número de empregados, o mix de produtos e o custo do aluguel, como outputs foram usados o número de clientes e a receita de vendas.

Os autores concluíram que o número de itens por funcionário e um ambiente competitivo influenciam na eficiência das lojas. Um número excessivo de itens pode levar a uma redução da eficiência. Assim, de acordo com o estudo, seria necessário gerenciar o ciclo de vida do produto, considerando o trade-off entre variedade e eficiência. Os autores também concluíram que embora a competição ajude a melhorar eficiência, o excesso de competição pode a reduzir.

Baviera-Puig e colegas associam DEA e GIS (Geographical Information Systems) para avaliar eficiência nas 5 maiores cadeias de lojas do setor hipermercadista na Espanha (totalizando 61 pontos de vendas) e se propõem a identificar variáveis mais críticas para melhorar a eficiência das firmas. Como inputs foram utilizados a lealdade ao programa de fidelidade da rede de lojas, área da loja, o número de vagas de estacionamento da loja, perdas com produtos perecíveis em euros, número de *checkouts* e a idade da loja. O faturamento bruto anual e o número de vendas por metro quadrado foram usados como outputs. Os autores não utilizaram o número de empregados nem os salários na análise sob a justificativa de que os padrões das relações de trabalho são homogêneos em todas as lojas.

Em conclusão ao estudo os referidos autores ratificaram a relevância da DEA enquanto ferramenta para mensurar a eficiência em unidades de negócios e ressaltaram que a versatilidade do modelo permite gerir vários resultados (outputs) e classificar as DMUs com base nos dados disponíveis. Corroboram ainda com a literatura ao reafirmar a necessidade de associar os resultados obtidos com DEA a outros tipos de análise (como análises que considerem características demográficas e proximidade de concorrentes) dada às peculiaridades do mercado de alimentos. No que se refere às lojas analisadas, o estudo concluiu que os programas de fidelidade influenciam positivamente a eficiência das lojas.

Barros (2006) analisa supermercados e hipermercados portugueses em dois estágios, utilizando DEA e um modelo Tobit. O autor concluiu que, em média, as eficiências das firmas hipermercadista é superior a eficiência observada em outros setores; os grandes grupos varejistas são em média mais eficientes do que os grupos menores; a escala é relevante neste mercado e os insumos que impulsionam a eficiência são o *market share*, o número de pontos de venda e a localização.

Perrigot e Barros (2008) usam o mesmo método em dois estágios para avaliar a eficiência de 11 varejistas na França, os autores argumentam que o tamanho da amostra foi definido de acordo com a disponibilidade das informações necessárias a análise. O estudo identificou os conjuntos de referência para firmas ineficientes provendo assim subsídios para a análise de *benchmark* e concluiu que as os varejistas franceses são relativamente eficiente mas que nem todos aqueles com estratégia de expansão internacional são estão na fronteira de melhores práticas.

Keh e Chu (2003) utilizaram DEA para analisar uma cadeia de mercearias nos EUA composta por 13 lojas no período de 1988 a 1997. Os autores demonstraram que é possível obter resultados relevantes a partir de amostras pequenas mas reforçaram que a disponibilidade limitada de dados impede a obtenção de resultados mais específicos. Para eles, a principal contribuição do estudo foi permitir uma comparação multilateral da produtividade de vários pontos de venda ao longo do tempo sendo estas informações relevantes no processo gerencial de tomada de decisão. Foi inferido do estudo que a produtividade nas vendas em função do uso de capital e mão de obra está intimamente ligada à geração eficiente de serviços e que existem retornos crescentes de escalas no varejo de alimentos o que corrobora com a literatura.

3. Metodologia, software e dados

3.1. Software e classificação utilizados

O *software* utilizado para obter estimativas foi o RStudio Version 1.2.1335 com o auxílio do pacote *benchmarking*. Este pacote foi desenvolvido por Peter Bogetoft and Lars Otto. De acordo com os autores, ele é recomendado para implementação de métodos para análise de fronteira e *Data Envelopment Analysis* (DEA), sob diferentes suposições de tecnologia (fdh, vrs, drs, crs, irs, add / frh e fdh +) e também utilizando diferentes medidas de eficiência (com base no insumo, com base no produto, gráfico hiperbólico, aditiva, super e direcional). Os recursos de pares e folgas estão disponíveis, sendo que informações parciais de preços podem ser incluídas e o custo, a receita e o lucro ideais podem ser calculados. Avaliação de fusões também são suportadas. Métodos para

representar graficamente os conjuntos de tecnologia também estão incluídos, além de que há suporte para métodos comparativos baseados em análises de fronteira estocástica (SFA) e para análises não paramétricas convexas de mínimos quadrados para funções convexas (StoNED). Em geral, os métodos desse pacote podem ser usados para resolver não apenas modelos padrão, mas também muitas outras variantes do modelo. Complementa o livro, Bogetoft e Otto (2011), mas pode obviamente ser usado como um pacote independente.

O software utilizado para a classes de otimização de modularidade e também para a elaboração da rede complexa com os algoritmos de distribuição de Fruchterman Reingold proporcional, foi o Gephi 0.9.2¹².

Em relação à classificação, a Associação Brasileira de Supermercados (ABRAS) classifica os formatos de loja do varejo de alimentos segundo área vendas, número de *checkouts*, sortimento, participação percentual de produtos não alimentícios no sortimento e o número de sessões que compõem a loja. Tal classificação, publicada na revista SuperHiper em 1998, é amplamente utilizada na literatura a exemplo de FERREIRA e colegas (2009), MOURA (2005) e SESSO FILHO (2003). A tabela 1 apresenta uma versão resumida desta classificação considerando as variáveis utilizadas neste estudo.

Tabela 1 – Formatos de loja por tamanho e número de checkouts

Formato de loja	Área (m ²)	Número de <i>checkouts</i>
De conveniência	50 - 250	1 - 2
De sortimento limitado	200 - 400	2 - 4
Supermerc. compacto	300 - 700	2 - 6
Supermerc. convenc.	700 - 2.500	7 - 20
Superloja	3.000 - 5.000	25 - 36
Hipermercado	7.000 - 16.000	55 - 90
Loja de depósito	4.000 - 7.000	30 - 50
Clube atacadista	5.000 - 12.000	25 - 35

Fonte: adaptado de BRITO, 1998.

Embora tal classificação tenha sido publicada em 1998, e não tenham sido encontradas referências atualizadas, ela é amplamente utilizada na literatura e ainda consegue representar com fidedignidade os formatos de loja atuais.

¹² Disponível em <https://gephi.org/>

Para os fins desta pesquisa são considerados como “grandes lojas ou atacarejos” os formatos Clube atacadista, Loja de depósito, Hipermercado, Superloja e Supermercado convencional.

3.2. Metodologia

Os modelos de DEA supõem que as firmas busquem máxima eficiência, assim cada firma utiliza seus insumos, representados pelo vetor de *inputs* $x = \{x_1, x_2, \dots, x_n\}$, para gerar um vetor de *outputs* $y = \{y_1, y_2, \dots, y_n\}$. Todavia, admite-se que cada um desses insumos tenha grau de relevância distinto dentro do processo produtivo, assim, a participação relativa desses insumos pode representada pela seguinte expressão linear:

$$y = \{y_1, y_2, \dots, y_n\} = \varphi (v_1 x_1 + v_2 x_2 + \dots + v_n x_n) \quad (6)$$

Onde $v_1, v_2, \dots, v_n$ são os pesos relativos é φ uma constante. Da mesma maneira, a medida global da produção de uma DMU pode ser dada por uma combinação linear dos *outputs*:

$$y = u_1 x_1 + u_2 x_2 + \dots + u_n x_n \quad (7)$$

Onde $u_1, u_2, \dots, u_n$ são os pesos relativos dos *outputs*.

Igualando as equações (1) e (2) temos:

$$\frac{\sum_{i=1}^n u_i y_i}{\sum_{i=1}^n v_i x_i} = \varphi \quad (8)$$

Assim, quanto maior a variação de φ , maior a eficiência da DMU. O modelo CCR admite que o valor máximo para φ seja a unidade e o valor mínimo seja zero. Dessa forma, preocupa-se em encontrar os valões relativos de φ . A DMU que alcança o valor de $\varphi = 1$ é considerada eficiente e supõe-se que, portanto,

tenha otimizado seus insumos a fim de maximizar a produção. Essa DMU passa a servir de *benchmark* para as demais.

Uma vez que a função objetivo é o nível de eficiência da unidade i , o problema sugere que sejam selecionados os melhores pesos possíveis (u e v) para cada DMU e estabelece restrições. A primeira restrição impõe que os pesos escolhidos, quando aplicados aos dados de qualquer DMU, não podem gerar valores maiores que a unidade (restrição de consistência). As demais restrições (condições de positividade) garantem que os pesos sejam todos não negativos e que ao menos um *input* e um *output* sejam ponderados positivamente na construção do índice (BOUERI, 2015).

Uma interpretação do score de eficiência (φ) é quanto uma firma deveria reduzir em termos proporcionais a utilização de *inputs* para alcançar a fronteira de eficiência. Todavia, estar sobre a fronteira de eficiência é condição necessária, porém não suficiente, para que a DMU seja dita eficiente. Diz-se que a DMU é CCR-eficiente caso obtenha $\varphi = 1$ e caso nenhum de seus pesos (u e v) seja igual a zero (BOUERI, 2015).

Embora o avanço computacional já permita a resolução direta do modelo a partir do problema fractal, a época de seu desenvolvimento a facilidade computacional era muito relevante, dessa forma, a literatura de DEA se desenvolveu com base na resolução da versão linear do problema. Assim, sintetizando as restrições, deve-se resolver ao seguinte problema linear:

$$\begin{aligned} \text{Max: } \varphi_i &= u_{1i}y_{1i} + u_{2i}y_{2i} + \dots + u_{ni}y_{ni} & (9) \\ \text{s. a.: } v_{1i}x_{1i} + v_{2i}x_{2i} + \dots + v_{ni}x_{ni} &= 1 \\ u_{1i}, u_{2i}, \dots, u_{ni} &\geq 0 \text{ e } u_{ji} \neq \bar{0} \\ v_{1i}, v_{2i}, \dots, v_{ni} &\geq 0 \text{ e } v_{ji} \neq \bar{0} \end{aligned}$$

Na passagem do problema fractal para o problema linear agrega-se mais uma restrição: a soma ponderada dos insumos deve ser igual a unidade (restrição normalizadora).

Uma das limitações mais significativas do modelo CCR é a suposição de retornos constantes de escala. Banker, Charnes e Cooper (1984) superaram tal entrave no modelo que ficou conhecido como BCC. O modelo BBC difere do

CCR pela adição do axioma de convexidade (equação 4) entre as restrições do modelo, assim a fronteira só pode ser formada por combinações convexas. Essa restrição reduz o número de alternativas de produção viáveis e converte uma tecnologia de ganhos de escala constantes em uma de ganhos variáveis (BADIN, 1997; VILELA *et al*, 2007).

$$\sum \lambda_j = 1 \quad (10)$$

Ambos os modelos podem ser orientados a *inputs* ou *outputs*. No entanto o modelo CCR a mudança de orientação produz os mesmos resultados uma vez que a fronteira é linear, isso não acontece no modelo BCC que possui fronteira convexa e cujo valor de eficiência é sempre maior que o obtido no modelo CCR. Assim, uma vez que uma firma seja considerada eficiente no modelo BBC ela também será eficiente no modelo CCR.

O modelo BBC é o mais amplamente utilizado na literatura pois considera-se que a presunção de retornos variáveis é mais condizente com a realidade das firmas, uma de principais vantagens do seu uso é que esta formulação é capaz de isolar a influência da escala. Enquanto no modelo CCR encontra-se restrito pela condição de homogeneidade segundo a qual todas as DMU's devem estar sob as mesmas condições de contorno, o modelo BBC neutraliza a influência do tamanho e oferece uma medida de eficiência técnica pura (FERREIRA & GOMES, 2020).

Posto isso, dado que as DMU's aqui analisadas são ranqueadas segundo faturamento e por isso firmas de tamanhos diferentes compõem a amostra, o modelo BBC se adapta melhor as características dos dados estando condizente sua aplicação com a literatura e adequada à heterogeneidade das DMUs.

Na orientação a *inputs* busca-se a aproximação da fronteira por meio da redução de insumos mantendo-se os *outputs* constantes. Já na orientação a *outputs* busca-se alcançar a fronteira por meio de um acréscimo nos *outputs* mantendo constante os *inputs* (ALMEIDA, 2017). Via de regra, a escolha da orientação não exerce influência significativa sobre a magnitude da eficiência técnica. O critério de escolha entre as orientações se dá em função do objetivo que se busca alcançar através da análise. O pesquisador deve escolher a

orientação de acordo com sua prioridade: regularizar a utilização dos inputs ou viabilizar o aumento dos outputs (FERREIRA & GOMES, 2020).

Os dados utilizados nesta pesquisa foram retirados de publicações anuais da Associação Brasileira de Hipermercados (ABRAS). Foi considerado como *output* o faturamento de cada uma das 47 maiores DMUs do setor de hipermercados em operação no Brasil no ano de 2017, assim, temos como *output* y = faturamento. Por sua vez, foram considerados 8 *inputs*:

x_1 = Número total de *checkouts*

x_2 = Área de vendas

x_3 = Número de lojas

x_4 = Número de funcionários

x_5 = Número de *checkouts* por loja

x_6 = Área de loja por *checkouts*

x_7 = Número de funcionários por área de lojas

x_8 = Número de funcionários por *checkout*

Na literatura pouco se tem discutido a cerca de critérios para a escolha das variáveis do modelo. Na maioria dos estudos as variáveis são escolhidas de acordo com o julgamento de especialistas ou segundo a disponibilidade dos dados. Embora o conjunto de variáveis escolhido possa levar a resultados distintos, isto não é necessariamente uma falha do modelo e pode se dar em razão de o pesquisador estar buscando abordar aspectos diferentes do problema de pesquisa (SERNA et al, 2007).

Na maioria dos casos concretos as variáveis são escassas enquanto as DMUs encontram-se em proporção significativamente maior e, portanto, não faz sentido utilizar métodos específicos para selecionar as variáveis de modo que frequentemente todas as variáveis disponíveis são utilizadas sem que haja prejuízos ao modelo.

Todavia, nos casos em que o número de unidades observadas é pequeno a escolha do conjunto de variáveis pode trazer impactos significativos ao modelo sendo prudente considerar métodos formalizados de seleção. Norman e Stoker (1991) associaram análise de correlações simples à Análise Envoltória de Dados

para criar um procedimento iterativo que concomitantemente seleciona insumos e produtos relevantes e calcula medidas de ineficiência.

Golany & Roll (1989) utilizam um método de seleção em três estágios, são eles:

- I. *Judgmental screening*: diferenciação entre variáveis que determinam eficiência e variáveis que explicam ineficiência;
- II. *Non-DEA quantitative analysis*: uso de regressão para determinar qual variável deve ser *input* ou *output*;
- III. *DEA based analysis*: usado para identificar as variáveis que não agregam eficiência ao modelo e devem ser excluídas.

Existe ainda o I-O Stepwise que busca aumentar a eficiência média com um conjunto restrito de variáveis, segundo este método, a escolha deve ser pautada pela máxima relação causal entre insumos e produtos. Neste método, as variáveis que contribuem pouco para a eficiência média são excluídas (SENRA et al, 2007).

Visto que no ramo varejista o acesso a informações sobre produção e comercialização é escasso devido, dentre outros fatores, a questões ligadas à estratégia competitiva, a escolha das variáveis se deu de acordo com a disponibilidade de dados e com a literatura pré-existente. Ademais, dado o exposto anteriormente, como o número de DMUs é suficientemente grande, não há prejuízo para análise em incluir o número de variáveis escolhido.

Já o número de observações da amostra é considerado satisfatório uma vez que representa 28% do faturamento total do setor e é 5,8 vezes maior que o número de variáveis. Durante a seleção foram excluídos dados de 3 empresas que embora sejam líderes do setor em termos de vendas apresentam dados faltantes.

Aqui utiliza-se o modelo orientado a produto e são obtidas as eficiências de Farrell, deste modo algumas das eficiências calculadas poderão apresentar valores superiores a unidade. Assim, os resultados indicam a magnitude do acréscimo potencial no faturamento que poderia ser obtido utilizando-se as mesmas quantidades de *inputs*. Para simplificar a análise, as tabelas seguintes apresentam a medida recíproca da eficiência de Farrell.

A tabela 2 a apresenta a descrição das variáveis. S_{sales} denota a soma da metragem quadrada de todos os pontos de venda para uma determinada DMU. N_{chk} contabiliza o número total de *checkouts* (caixas) que uma DMU possui. N_{stores} refere-se ao número de lojas, enquanto Chk_{sales} é a razão entre número total de checkouts (N_{chk}) e o número de lojas (N_{stores}). Já S_{chk} representa a proporção entre a área de vendas (S_{sales}) e o número de caixas (N_{chk}). Wk_{byS} representa o número de funcionários para cada m^2 de loja, sendo que Wk_{chk} designa o número funcionários para cada *checkout*.

Tabela 2: Descrição das variáveis

Variável	Descrição
S_{sales}	Área total de vendas
N_{chk}	Número total de <i>checkouts</i>
N_{stores}	Número de lojas
$N_{workers}$	Número de funcionários
Chk_{sales}	Número de <i>checkouts</i> por loja
S_{chk}	Área de loja por <i>checkouts</i>
Wk_{byS}	Número de funcionários por m^2
Wk_{chk}	Número de funcionários por <i>checkout</i>

Fonte: Elaboração própria.

A Tabela 3 apresenta as estatísticas descritivas dos inputs e do output, discriminando as informações relativas a média, mediana, mínimo, máximo, 1ª e 3º quartis e desvio padrão de cada variável inserida no modelo. É possível notar a grande variabilidade entre as firmas, um exemplo está na variável $N_{workers}$ cujo valor mínimo observado é 1.795 enquanto o valor máximo alcança 27.182.

Tabela 3: Estatísticas descritivas

Variável	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
S_{sales}	20001	33562	47000	8281 1	100349	567074

N_{chk}	197.0	313.5	423.0	664.8	735.5	3814.0
N_{stores}	4.00	17.00	27.00	89.83	45.50	2493.0
$N_{workers}$	1795	2952	4079	6018	7080	27182
Chk_{sales}	2.00	13.50	17.00	18.53	21.00	66.00
S_{chk}	38.00	103.5	118.0	125.8	138.5	273.0
Wk_{byS}	3.000	6.000	8.000	8.702	10.09	11.00
Wk_{chk}	4.00	8.00	9.00	10.09	11.00	22.00

Fonte: Elaboração própria.

4. Resultados

Esta seção divide-se em duas partes: a primeira subseção trata dos resultados do modelo DEA - VRS; a segunda analisa os clusters definidos pelo processo de otimização de modularidade. Finalmente, ainda na segunda subseção, apresentamos a rede complexa gerada a partir do algoritmo de *layout* de Fruchterman-Reingold.

4.1. Resultados do modelo DEA-VRS

Pela tabela 4, observamos que, dentre as 47 DMUs analisadas, 26 atingiram 100% de eficiência de acordo com o modelo utilizado. A tabela 4 mostra cada uma das DMUs analisadas, seu respectivo conjunto de referência e seu índice de eficiência que é o índice recíproco ao índice de Farrell.

Tomando por exemplo o caso hipotético de duas DMUs (i e j) que geram o mesmo *output* mas uma delas utiliza quantidade maior de um dos seus insumos e nenhuma redução proporcional é possível, embora estas DMUs possam estar sobre a fronteira de eficiência, a DMU j que utiliza quantidade maior de um dos insumos gerando o mesmo resultado que a DMU i é ineficiente e possui um de seus pesos igual a zero. Neste caso, a atribuição de um peso positivo ao insumo utilizado em demasia iria violar a restrição de consistência quando os pesos escolhidos para j fossem utilizados em i , por isso o modelo atribui peso zero a tal insumo (BOUERI, 2015). No caso da amostra aqui analisada, nenhuma das DMUs sobre a fronteira apresentou qualquer folga.

Embora o modelo considere qualquer DMU com índice abaixo de 1 como ineficiente, na verdade, as DMUs apresentam graus de ineficiência distintos. Portanto, Norman e Stoker (1991) classificam como marginalmente ineficientes aquelas DMU's cujo índice de eficiência encontra-se entre 90 e 99%. Os autores as consideram capazes de tornarem-se eficientes mediante uma pequena alteração na distribuição de seus inputs. Já as DMUs com eficiência entre 70 e 90% são classificadas como e eficiência mediana e considera-se que utilizam inputs em demasia em relação ao seu produto (elas estão marcadas em negrito na tabela 4). Por fim, os autores taxam como ineficientes aquelas firmas com índice inferior a 70% (cujas linhas estão preenchidas com cor cinza na tabela 4).

Nesse sentido, as 26 empresas que alcançaram eficiência máxima concentram um *market share* total de 16,82% e o somatório dos seus HHI's (Índices de Herfindahl-Hirschmann) foi de 0,00241. As 5 empresas que atingiram eficiência mediana (que estão assinaladas em negrito) concentram um market share de 9,39% e somam HHI's de 0,00216. As 16 empresas consideradas ineficientes concentram um *market share* de 16,23% - próximo portanto ao das 26 que alcançaram eficiência máxima - e totalizam 0,00196 de HHI's (tabela 4).

Todavia, alcançar a fronteira de eficiência *é condição necessária, porém não suficiente para que uma DMU seja considerada eficiente*. Uma definição mais completa considera uma DMU eficiente *caso esta tenha score igual a unidade e nenhum de seus pesos (relativos a insumos ou produtos) seja zero* (BOUERI, 2015).

Tabela 4 - Resultados obtidos pelo modelo DEA-VRS

Cod.	DMU's	Unidade federativa	Market share	HHI	Índice
1	CENCOSUD	MG, SE, BA e RJ	3,86%	0,00149	1
6	DMA DISTRIBUIDOR	MG	1,53%	0,00023	1
10	MART MINAS	MG	1,19%	0,00014	1
17	AM/PM COMESTIVEIS LTDA.	RJ	0,96%	0,00009	1
25	CEMA	MG	0,62%	0,00004	1
27	NORDESTÃO	RN	0,56%	0,00003	1
28	JAD ZOGHEIB	SP	0,55%	0,00003	1
29	SILVA E BARBOSA	SP	0,51%	0,00003	1
30	KOCH	SC	0,51%	0,00003	1
32	TORRE E CIA.	RJ	0,46%	0,00002	1
34	SUPERMERCADOS JAU	SP	0,45%	0,00002	1

35	LUIZ TONIN	MG	0,44%	0,00002	1
36	D'AVÓ	SP	0,42%	0,00002	1
37	FORMOSA	PA	0,41%	0,00002	1
38	S. SUPERPAO	PA	0,41%	0,00002	1
39	REALMAR	ES	0,39%	0,00002	1
40	COVABRA	SP	0,37%	0,00001	1
41	MUNDIALMIX	SC	0,37%	0,00001	1
42	CIA. BEAL	PR	0,37%	0,00001	1
43	COMERCIAL ZAFFARI	RS	0,37%	0,00001	1
44	ORG. VERDEMAR	MG	0,31%	0,00001	1
45	IRMAOS BOA LTDA	SP	0,31%	0,00001	1
46	CENTRAL NEG. E ALIM.	SC	0,29%	0,00001	1
47	A.C.D.A	AC	0,29%	0,00001	1
23	ATAKAREJO	BA	0,67%	0,00005	1
24	HORTIGIL	ES	0,66%	0,00004	1
4	SUPERMERCADOS BH	MG	2,48%	0,00061	0.872
3	SDB	MS	2,61%	0,00068	0.825
2	MUFFATO	PR	2,72%	0,00074	0.759
21	UNIDASUL DIST.	RS	0,77%	0,00006	0.715
18	GMGB HOLDING	SP	0,81%	0,00006	0.708
7	SONDA	SP	1,52%	0,00023	0.671
5	ZAFFARI	RS	2,35%	0,00055	0.666
8	SAVEGNAGO	SP	1,32%	0,00018	0.633
12	ZARAGOZA	SP	1,00%	0,00010	0.6
9	LIDER	PA	1,24%	0,00015	0.574
14	MULTI FORMATO DIST.	MG	0,98%	0,00014	0.554
11	A. ANGELONI	SC	1,01%	0,00010	0.541
13	BAHAMAS	MG	0,99%	0,00010	0.508
16	COOP	SP	0,97%	0,00009	0.5
15	Co. SULAMERICANA	PR	0,97%	0,00009	0.482
19	PAGUE MENOS	SP	0,80%	0,00006	0.372
22	CARVALHO E FERNANDES	PI	0,77%	0,00006	0.347
20	GIASSI E CIA	SC	0,79%	0,00006	0.341
26	IRMÃOS LOPES	SP	0,59%	0,00006	0.243
31	S. CAVICCHIOLLI	SP	0,49%	0,00002	0.119
33	SÃO CRISTOVAO	SP	0,45%	0,00002	0.081

Fonte: Elaboração própria a partir dos resultados do modelo DEA - VRS.

Levando em consideração tal classificação, conforme exposto na tabela 5, temos que a amostra *não possui DMUs marginalmente eficientes* considerando retornos variáveis de escala; as empresas Muffato, SDB, Supermercados BH, Unidasul Dist. e GMGB Holding (grafadas em negrito na tabela 4 e descritas na tabela 5) possuem *eficiência mediana*, para essas empresas, correções nas quantidades de seus insumos podem levar a eficiência no médio prazo (JUNIOR; RODRIGUES, 2017).

Dentre as 26 firmas VRS-eficientes, as firmas 38 (Superpão) e 36 (D'avó) são as que mais participam dos conjuntos de referências das demais. Elas compõem, respectivamente, o conjunto de referência de 57,1% e 52,38% das firmas ineficientes. Na sequência, tem-se as firmas 43 (Zaffari) e 46 (Central Neg.

Alimentos) compondo o conjunto de referência de 33,33% das DMUs ineficientes.

Vale ressaltar que uma empresa possuir produtividade relativa igual a x é indicativo de sua eficiência apenas dentro da amostra na qual é avaliada, adicionar ou remover firmas da amostra altera as eficiências relativas para todas as DMUs. Tal fato faz com que esta metodologia de análise se caracterize como um modelo aberto e dinâmico compatível com os estudos de *benchmark* (BALDIN, 1997).

Tabela 5 – DMUs classificadas segundo o parâmetro de Norman e Stoker (1991)

Classificação	Códigos	Total
Eficientes	1, 6, 10, 17, 23, 24, 25, 27, 28, 29,30, 32, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47.	26
Marginalmente ineficientes	-	0
Eficiência mediana	2, 3, 4, 18, 21.	5
Ineficientes	7, 9, 11, 12, 13, 14, 15, 16, 19, 20, 22, 26, 31, 33.	16

Fonte: Elaboração própria a partir dos resultados do modelo DEA.

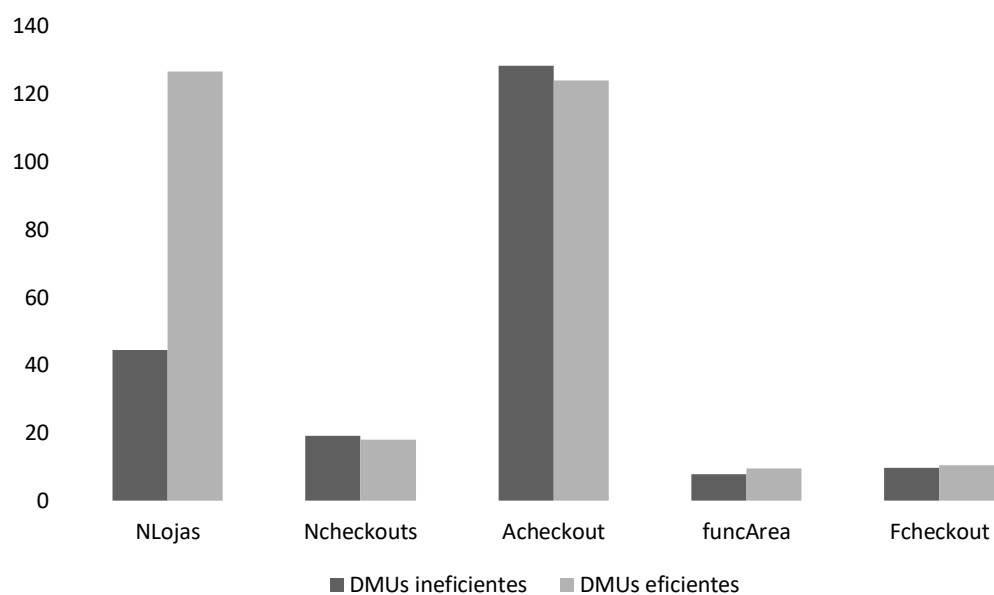
O Cencosud (cód. 1) é a firma que, dentre aquelas eficientes, possui o maior faturamento, comparando-o a firma de maior produto entre as ineficientes, a Muffato (cód. 2), o Cencosud possui uma área de venda por loja 18,99% menor, número de funcionários por m² 16,66% menor e um número de checkouts por loja 52% menor, no entanto seu faturamento é 23,06% maior.

De modo geral, as firmas eficientes possuem em média aproximadamente 1 checkout a cada 124m² de loja contra 1 checkout a cada 128m² nas firmas ineficientes, as firmas que compõe a fronteira têm também em média 71292,2m² de área total de vendas contra 97071,6m² das firmas abaixo da fronteira.

Como pode ser observado nos gráficos 1 e 2, em termos médios as firmas ineficientes possuem nível de utilização dos *inputs* N_{chk} , S_{sales} , $N_{workers}$ e S_{chk} superior aos níveis observados nas firmas que se posicionam sobre a fronteira.

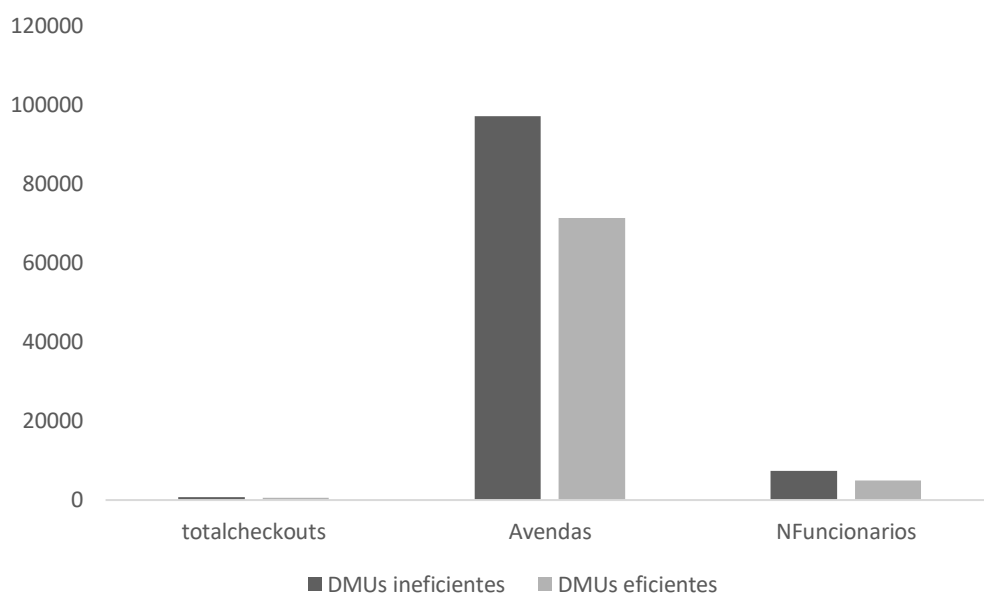
Já utilização dos insumos N_{stores} (-64,93%), Wk_{byS} (-17,12%) e Wk_{chk} (-5,65%) é em média menor nas DMUs ineficientes.

Gráfico 1 – Utilização média dos insumos I



Fonte: Elaboração própria.

Gráfico 2 – Utilização média dos insumos II



Fonte: Elaboração própria

A tabela 6 apresenta as DMUs eficientes classificadas de acordo com o formato de loja predominante. As firmas encontram-se dispostas de acordo com o faturamento em ordem decrescente. A classificação por formato foi feita com base em BRITO (1998).

Sabe-se que, em modo geral, as DMUs não possuem apenas um formato de loja, todavia, a tabela 6 mostra apenas o formato mais frequente para cada firma com base no tamanho médio de loja (razão entre a área de vendas e o número de lojas).

Tabela 6 – Classificação das DMUs eficientes segundo formato de loja e indicadores econômico-financeiros e de produtividade

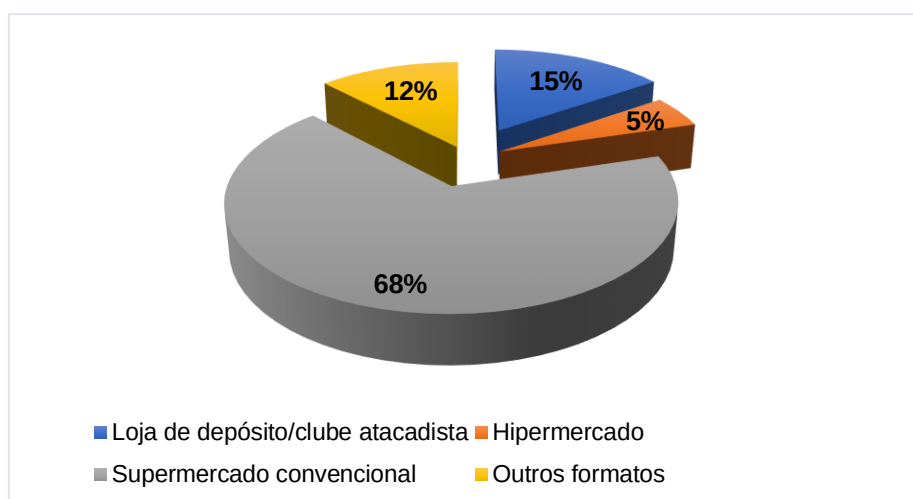
Cód.	DMUs	Área média por loja (m²)	Média checkouts por loja	Faturam. bruto	Nº de Lojas	Nº de funcionários	Área de vendas (m²)	Faturam./ loja	Faturam./ funcionário	Formato
1	CENCOSUD	2.821	12	8,535,697,000	204	27,452	572,664	41,841,652	310,932	Super. Convenc.

6	DMA DISTRIBUIDOR	1.493	12	3,356,067,550	129	12,572	184,491	26,016,028	266,948	Super. Convenc.
10	MART MINAS	5.249	19	2,226,910,505	25	3,968	130,081	89,076,420	561,217	Hiperm./ Clube atacad.
17	AM/PM COMESTIVEIS LTDA.	58	2	2,147,834,034	2415	13,887	135,240	889,372	154,665	Convêniencia
23	ATAKAREJO	3.125	27	1,209,807,345	9	2,245	25,955	134,423,038	538,890	Super loja
24	HORTIGIL	645	9	1,379,609,000	43	6,162	31,036	32,083,930	223,890	Super. Compacto
25	CEMA	4.847	22	1,309,000,000	17		82,399	77,000,000	#DIV/0!	Super loja
27	NORDESTÃO	2.923	22	1,294,544,170	9	3,559	26,310	143,838,241	363,738	Super loja
28	JAD ZOGHEIB	2.214	20	1,071,742,041	10	2,855	20,977	107,174,204	375,391	Super. Convenc.
29	SILVA E BARBOSA	1.184	15	1,119,469,805	30	3,429	34,615	37,315,660	326,471	Super. Convenc.
30	KOCH	1.975	17	858,707,261	19	1,950	31,804	45,195,119	440,363	Super. Convenc.
32	TORRE E CIA.	1.177	16	988,603,817	16	3,490	22,128	61,787,739	283,268	Super. Convenc.
34	SUPERMERCADOS JAU	1.215	9	982,119,400	35	3,500	41,620	28,060,554	280,606	Super. Convenc.
35	LUIZ TONIN	4.137	14	1,017,695,403	18	2,165	47,000	56,538,634	470,067	Super. Convenc.
36	D'AVÓ	9.307	31	931,849,706	11	2,480	45,510	84,713,610	375,746	Super loja
37	FORMOSA	1.637	66	908,498,841	4	3,446	42,000	227,124,710	263,639	Hiperm./ clube atacad.
38	S. SUPERPAO	1.359	14	810,590,819	25	2,968	40,775	32,423,633	273,110	Super. Convenc.
39	REALMAR	1.74	13	896,829,042	26	3,455	33,246	34,493,425	259,574	Super. Convenc.
40	COVABRA	1.413	17	819,291,469	15	2,420	25,905	54,619,431	338,550	Super. Convenc.
41	MUNDIALMIX	1.435	17	677,244,416	19	2,390	25,408	35,644,443	283,366	Super. Convenc.
42	CIA. BEAL	2.032	17	810,395,174	12	2,238	16,507	67,532,931	362,107	Super. Convenc.
43	COMERCIAL ZAFFARI	1.621	18	861,442,556	21.00		42,672	41,021,074	-	Super. Convenc.
44	ORG. VERDEMAR	1.699	15	710,807,678	12	4,200	19,578	59,233,973	169,240	Super. Convenc.
45	IRMAOS BOA LTDA	1.22	16	753,178,331	14.00		23,786	53,798,452	-	Super. Convenc.
46	CENTRAL NEG. E ALIM.	1.79	11	706,159,131	24.00		29,280	29,423,297	-	Super. Convenc.
47	A.C.D.A	4.137	18	635,608,018	12	2,334	22,576	52,967,335	272,326	Super. Convenc.

Fonte: elaboração própria a partir do Relatório da Abras 2017 e Revista Varejo S.A.

O gráfico 3 foi construído com base nos valores absolutos para os formatos clube atacadista, loja depósito, hipermercado e supermercados oferece uma melhor visualização da participação de cada formato no conjunto de lojas das DMUs eficientes.

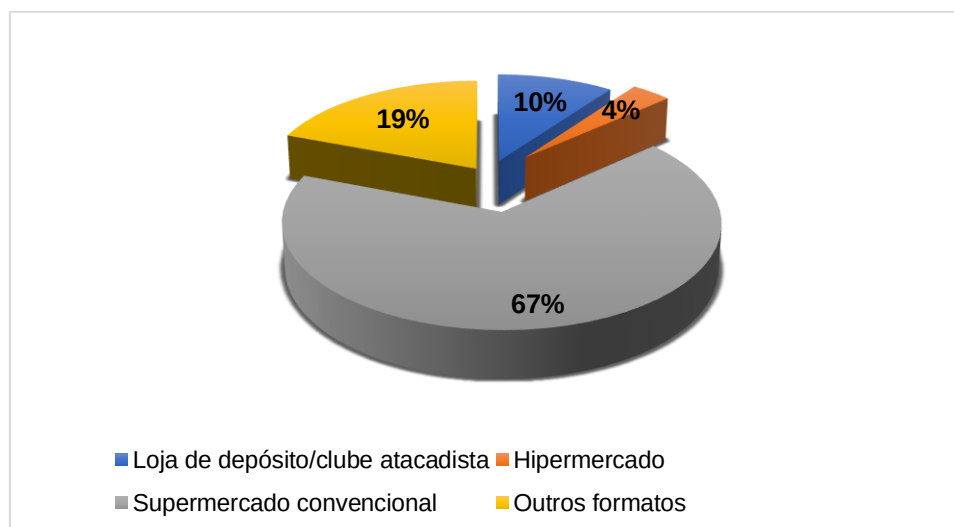
Gráfico 3 – Distribuição percentual por formato entre as DMUs eficientes



Fonte: dados da pesquisa.

O formato supermercado convencional (lojas com até 2.500m² e 20 *checkouts*) corresponde a 68% das firmas eficientes, já os formatos hipermercado e loja depósito/clube atacadista correspondem juntos a 20% das firmas.

Gráfico 4 – Distribuição percentual por formato entre as DMUs ineficientes



Fonte: dados da pesquisa.

Entre as DMUs ineficientes, como mostra o gráfico 4, o formato supermercado convencional predomina representando 67% das lojas enquanto os formatos de grandes lojas representam apenas 14%. As lojas de menor porte (conveniência, supermercado compacto etc) representam 19% do total estando em maior proporção do que entre as firmas eficientes (12%).

Passamos agora a investigar as empresas ineficientes e de eficiência mediana e suas interconexões. A tabela 7 apresenta as DMUs ineficientes e seus respectivos conjuntos de referência definidos pelo modelo DEA-VRS. Nas linhas de cor cinza claro estão as 5 empresas de eficiência mediana e seus conjuntos de referência. Nas de cor cinza escuro estão as 16 empresas ineficientes e novamente seus conjuntos de referência.

Tabela 7 - DMUs de eficiência mediana e ineficientes e seus conjuntos de referência.

Código	DMU	Índice	Referência
2	MUFFATO	0,759	1,36,38,43
3	SDB	0,825	43
4	SUPERMERCADOS BH	0,872	17,38,39,41,46
5	ZAFFARI	0,666	36,38
7	SONDA	0,671	36,41
8	SAVEGNAGO	0,633	36,39,41
9	LIDER	0,574	36
11	A. ANGELONI	0,541	36,38
12	ZARAGOZA	0,600	43
13	BAHAMAS	0,508	36,38
14	MULTI FORMATO DIST.	0,554	1,38,39,46
15	CIA. SULAMERICANA	0,482	36,38,39,41
16	COOP	0,500	43
18	GMGB HOLDING	0,709	10,35,43,46
19	PAGUE MENOS	0,373	36,38
20	GIASSI E CIA	0,341	36,38
21	UNIDASUL DIST.	0,715	17,24,34,46
22	CARVALHO E FERNANDES	0,347	1,38,43,46
26	IRMÃOS LOPES	0,243	17,38,41,43,46
31	S. CAVICCHIOLLI	0,119	36,38,40,43
33	SAO CRISTOVAO	0,081	24,39,42,46

Fonte: Dados da pesquisa.

Através desses resultados da análise envoltória, identificamos as unidades que serviram como “modelo” (*benchmark*) para as demais unidades, as quais devem perseguir suas práticas a fim de alcançar a fronteira de eficiência. *Essas unidades modelos constituem o chamado conjunto de referência da tabela 7.* Portanto, para cada DMU ineficiente da tabela 7 existe um conjunto de referência distinto formado por firmas eficientes que se assemelham a ela, em termos proporcionais, em termos de utilização dos

recursos. Porém, as ineficientes geram um resultado significativamente diverso. Quanto mais vezes uma firma eficiente é listada como “modelo” para as firmas ineficientes, maior é a relevância dessa firma dentre as firmas eficientes (FERREIRA, GOMES 2020). Na subseção seguinte, o processo de otimização de modularidade irá, com base nestes resultados da tabela 7, classificar em *clusters* tais empresas, o que acabará refinando ainda mais os resultados do modelo DEA-VRS.

Em suma, os resultados do modelo sinalizam que as empresas eficientes utilizam em média um menor número de funcionários e de caixas em relação à área de vendas. Tais empresas, em face das operações das ineficientes, apresentam a maior parcela de suas atividades operacionais no modelo de negócios baseado em grandes lojas. Esse resultado corrobora a hipótese inicial de que as firmas mais eficientes priorizam os modelos de atacarejo ou de grandes lojas, sendo que esse vínculo positivo entre tamanho e eficiência é confirmado por outras investigações da literatura nacional sobre varejo (PONTES, 2016; YU & ANGELO, 2001; SOUZA, MACEDO & FERREIRA, 2010).

4.2. Rede de otimização de modularidade e *clusters* dos resultados do modelo DEA - VRS

Os dados da tabela 7 servem como base para a construção da rede complexa desta subseção e para a análise dos clusters que foram estimados utilizando o processo de otimização de modularidade no software Gephi 0.9.2.

Identificam-se assim os *clusters* com as principais DMUs de eficiência mediana e ineficientes e os seus conjuntos de referência (nos os quais todas as DMUs são eficientes, conforme exposto na tabela 5).

A tabela 8 separa os quatro *clusters* que receberam classes de modularidade 0, 1, 2 e 3. Estes números possuem significado apenas classificatório, servindo para ordenar a distribuição e separar os *clusters* por meio de um critério de estatística de redes complexas, que é a otimização de modularidade (ou classe de modularidade).

Agora é possível identificar como se agrupam as DMUs e as regularidades empíricas das empresas que os compõem.

No cluster com o número zero, há apenas 7 empresas, sendo três delas ineficientes (em negrito), três eficientes e uma de eficiência mediana (em itálico). A eficiência média deste primeiro cluster é a segunda maior entre os quatro clusters e o desvio-padrão é o menor entre eles (tabela 9 e gráfico 5).

No cluster com número 1, há cinco empresas ineficientes, duas de eficiência mediana e quatro eficientes. A presença neste cluster do Zaffari e Comercial Zaffari (*market shares* somados de 2,72%, o segundo maior da amostra, junto com o Muffato) e do Supermercado BH (2,45%, o terceiro maior da amostra), elevam o *market share* médio do *cluster*. Porém, o número elevado de empresas ineficientes reduz a eficiência média desse agrupamento (tabela 8 e gráfico 5).

Tabela 8 – Divisão dos clusters (agrupamentos) com as DMUs ineficientes e eficientes, conforme o critério da otimização de modularidade

DMUs	Otimiz. de modularid.	Market Share médio	Eficiência média	Eficiência (desvio- padrão)
D'avó, Luiz Tonin, Jaú , <i>SDB</i> , Bahamas , Sonda e Giassi e Cia.	0	0,95%	0,764	0,266
Irmãos Lopes , S. Cavicchioli , Com. Zaffari, Zaffari , <i>Superm. BH</i> , Líder , A. Angeloni, Coop, AM/PM, Savegnago , <i>Unidasul Dist.</i>	1	1,14%	0,624	0,279
São Cristóvão , Hortigil, Cencosud, Zaragoza , Pague Menos , <i>Muffato</i> , Mart Minas, Multi Formato Dist.	2	1,60%	0,671	0,336
Superpão, Mundialmix, Cia. Beal, Realmar, Covabra, Central, <i>GMGB</i> <i>Holding</i> , Cia. Sulamericana de Dist. , Carvalho e Fernandes.	3	0,53%	0,837	0,260

Fonte: elaboração própria.

No cluster com número 2, há quatro empresas com eficiência baixa, uma com eficiência mediana e três eficientes. A predominância de empresas com baixa eficiência faz com que o agrupamento possua a segunda menor eficiência média entre os quatro clusters e o maior desvio-padrão de eficiência. Todavia, o *market share* médio desse agrupamento é o maior de todos. Esse patamar é garantido pela presença da Cencosud, que possui a maior fatia de mercado de toda amostra (3,86%), da eficiente Mart Minas (1,19%) e da ineficiente Zaragoza (1,0%) (tabela 8 e gráfico 5).

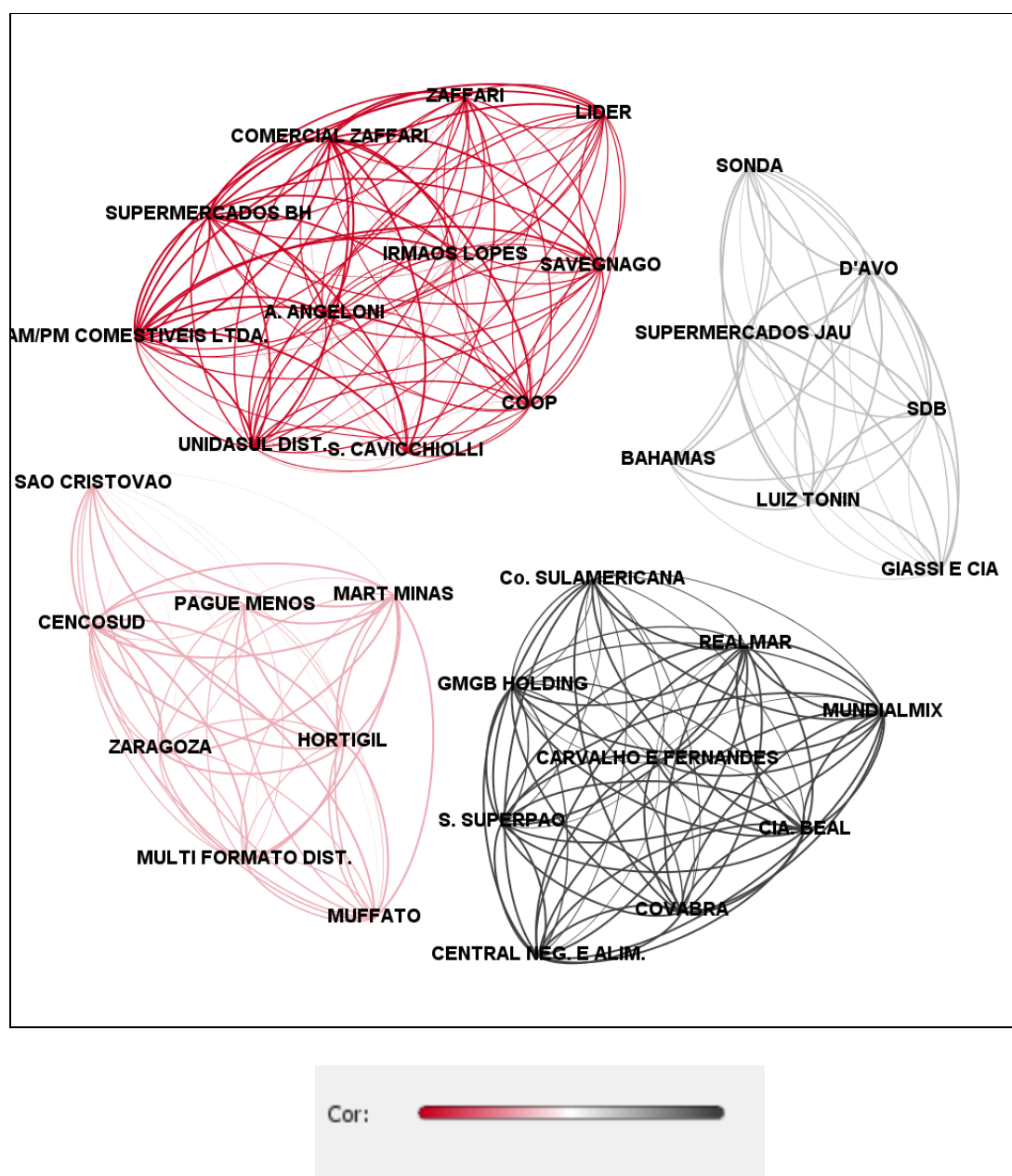
No último cluster de número 3, há a frequência mais elevada de empresas eficientes de todos os quatro agrupamentos (6 empresas) e também o número mais baixo de ineficientes (apenas 2 empresas), além de uma empresa de eficiência mediana. Isso faz com que o agrupamento apresente o menor desvio-padrão de eficiência e o maior grau de eficiência média entre os quatro clusters (tabela 8 e gráfico 5).

A rede complexa do gráfico 5 descreve os clusters descritos anteriormente. A rede é “pesada”, isto é, suas ligações ou arestas são ponderadas pela eficiência transmitida de uma DMU para as outras, conforme as conexões da tabela 9 e os pesos dos índices de eficiência da tabela 4.

Cruzando as informações das empresas eficientes que compõem os *clusters* (descritas nas tabelas 8, 9, e gráfico 5) com os formatos e os indicadores econômico-financeiros e de produtividade da tabela 6, pode-se afirmar que as características de cada cluster são as seguintes:

- Cluster zero: predominância de supermercados convencionais, com números de funcionários entre 2 mil e 3,5 mil; área de vendas de 41,6 mil a 47 mil m²; faturamento por m² e market shares muito semelhantes. Predominância do formato de supermercados convencionais.
- Cluster 1: há apenas duas empresas eficientes e com formatos, *market shares*, áreas de vendas e demais indicadores bem distintos. Trata-se de um cluster de duas empresas eficientes que destoam das demais em várias métricas.
- Cluster 2 : as empresas Cerconsud, Mart Minas e Hortigil, apesar de terem formatos diferentes de negócio, possuem indicadores altos de market share (os três mais altos desta amostra). Esta é a razão de estarem no mesmo cluster. No mais, a Hortigil difere das outras duas, as quais possuem semelhanças entre si, tais como: faturamento por m² similar, amplas áreas de vendas e altos faturamentos por funcionário.
- Cluster 3: este é agrupamento mais homogêneo de todos, pois todas as empresas são supermercados convencionais, apresentam faturamentos por loja e por funcionário semelhantes, market shares próximos e números de funcionários também parecidos.

Gráfico 5 – Rede dos Clusters das DMUs – otimização de modularidade



Fonte: elaboração própria utilizando o Gephi 0.9.2.

Estes resultados podem ser observados a partir do exame da tabela 9 que é similar aos descritos na tabela 6. Porém, agora dividimos e selecionamos apenas as empresas que compõem os clusters mencionados.

Tabela 9 – Indicadores econômico-financeiros, formato das DMUs e classes de modularidade

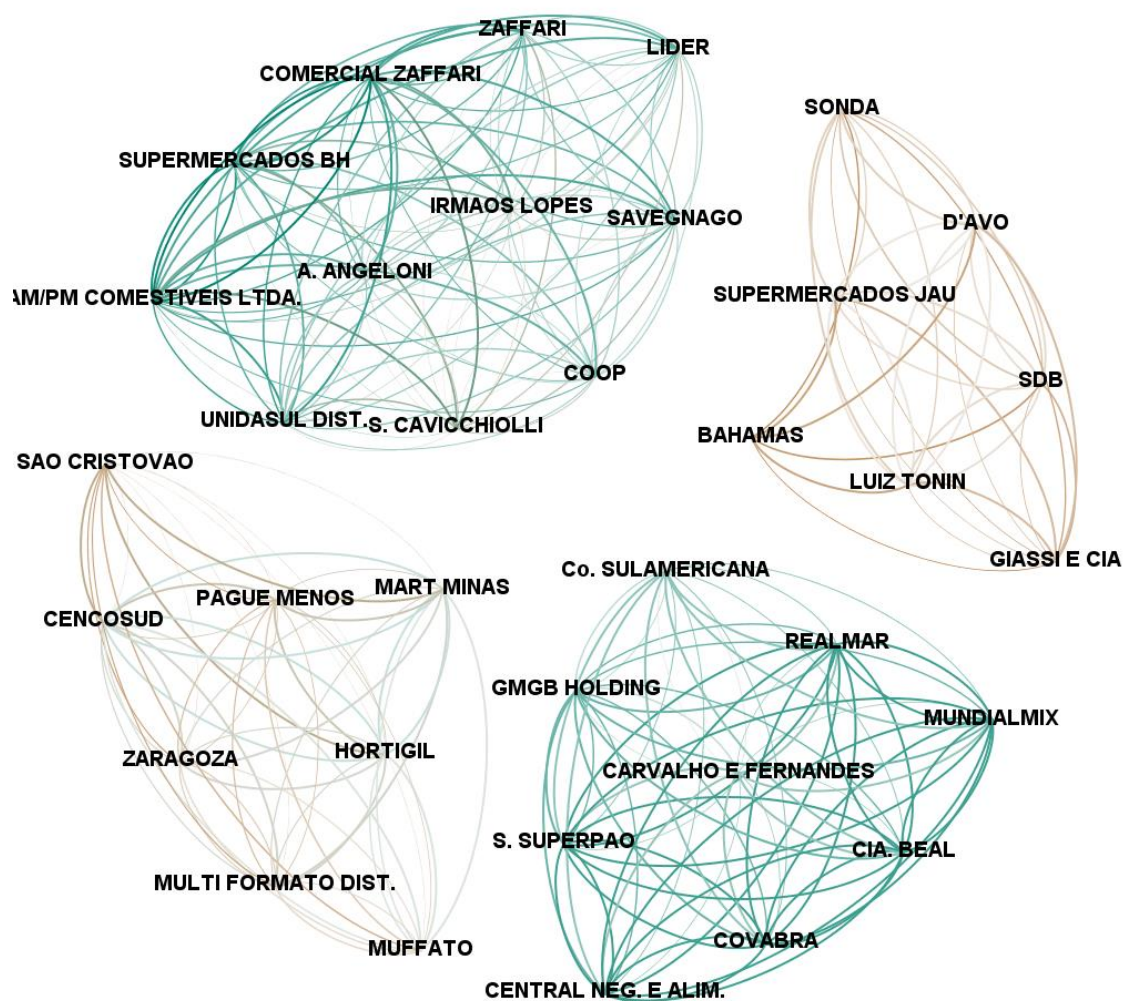
Cód.	DMUs	Nº de funcionários	Área de vendas M²	Faturam./ loja	Faturam./ funcionário	Faturam./m²	Market shares	Formato	Modularidade
------	------	--------------------	-------------------	----------------	-----------------------	-------------	---------------	---------	--------------

34	SUPERMERCADOS JAU	3,500	41,620	28,060,554	280,606	23,597	0.45%	Super. Convenc.	0
35	LUIZ TONIN	2,165	47,000	56,538,634	470,067	21,653	0.44%	Super. Convenc.	0
36	D'AVÓ	2,480	45,510	84,713,610	375,746	20,476	0.42%	Super loja	0
17	AM/PM COMESTIVEIS LTDA.	13,887	135,240	889,372	154,665	15,882	0.96%	Convêniencia	1
43	COMERCIAL ZAFFARI	1,463	42,672	41,021,074	588,819	20,188	0.37%	Super. Convenc.	1
1	CENCOSUD	27,452	572,664	41,841,652	310,932	14,905	3.86%	Super. Convenc.	2
10	MART MINAS	3,968	130,081	89,076,420	561,217	17,119	1.19%	Hiperm/ Clube atacad.	2
24	HORTIGIL	6,162	31,036	32,083,930	223,890	44,452	0.66%	Super. Compacto	2
38	S. SUPERPAO	2,968	40,775	32,423,633	273,110	19,880	0.41%	Super. Convenc.	3
39	REALMAR	3,455	33,246	34,493,425	259,574	26,976	0.39%	Super. Convenc.	3
40	COVABRA	2,420	25,905	54,619,431	338,550	31,627	0.37%	Super. Convenc.	3
41	MUNDIALMIX	2,390	25,408	35,644,443	283,366	26,655	0.37%	Super. Convenc.	3
42	CIA. BEAL	2,238	16,507	67,532,931	362,107	49,094	0.37%	Super. Convenc.	3

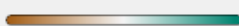
Fonte: elaboração própria a partir dos cálculos no Gephi 0.9.2, do Relatório da Abras 2017 e da Revista Varejo S.A.

Finalmente, rodamos a rede dos graus ponderados, a partir da qual corroboramos um ponto importante da análise que fizemos no modelo DEA-VRS (gráfico 6).

Gráfico 6 – Rede dos graus ponderados das DMUs



Cor:



Fonte: elaboração própria utilizando o Gephi 0.9.2.

Tabela 10 – Graus ponderados, modularidade e medidas de concentração de mercado

Id	Label	Grau pond. de entrada	Grau pond. de saída	Grau ponderado	Market Share (%)	HHI	Modularidade
17	AM/PM COMESTIVEIS LTDA.	5.863	10	15.863	0.96%	0,00009	0
43	COMERCIAL ZAFFARI	5.863	10	15.863	0.37%	0,00001	0
4	SUPERMERCADOS BH	5.991	8.72	14.711	2.48%	0,00061	0
38	S. SUPERPAO	6.537	8	14.537	0.41%	0,00002	3
39	REALMAR	6.537	8	14.537	0.39%	0,00002	3
40	COVABRA	6.537	8	14.537	0.37%	0,00001	3
41	MUNDIALMIX	6.537	8	14.537	0.37%	0,00001	3
42	CIA. BEAL	6.537	8	14.537	0.37%	0,00001	3
46	CENTRAL NEG. E ALIM.	6.537	8	14.537	0.29%	0,00001	3
21	UNIDASUL DIST.	6.148	7.15	13.298	0.77%	0,00006	0
5	ZAFFARI	6.197	6.66	12.857	2.35%	0,00055	0
8	SAVEGNAGO	6.23	6.33	12.56	1.32%	0,00018	0
18	GMGB HOLDING	6.829	5.664	12.493	0.81%	0,00006	3
9	LIDER	6.289	5.74	12.029	1.24%	0,00015	0
11	A. ANGELONI	6.322	5.41	11.732	1.01%	0,00010	0
1	CENCOSUD	4.366	7	11.366	3.86%	0,00149	1
10	MART MINAS	4.366	7	11.366	1.19%	0,00014	1
24	HORTIGIL	4.366	7	11.366	0.66%	0,00004	1
16	COOP	6.363	5	11.363	0.97%	0,00009	0
15	Co. SULAMERICANA CARVALHO E FERNANDES	7.055	3.856	10.911	0.97%	0,00009	3
22	MUFFATO	7.19	2.776	9.966	0.77%	0,00006	3
2	MUFFATO	4.607	5.313	9.92	2.72%	0,00074	1
34	SUPERMERCADOS JAU	3.837	6	9.837	0.45%	0,00002	2
35	LUIZ TONIN	3.837	6	9.837	0.44%	0,00002	2
36	D'AVÓ	3.837	6	9.837	0.42%	0,00002	2
26	IRMAOS LOPES	6.62	2.43	9.05	0.59%	0,00006	0
12	ZARAGOZA	4.766	4.2	8.966	1.00%	0,00010	1
3	SDB	4.012	4.95	8.962	2.61%	0,00068	2
14	MULTI FORMATO DIST.	4.812	3.878	8.69	0.98%	0,00014	1
7	SONDA	4.166	4.026	8.192	1.52%	0,00023	2
31	S. CAVICCHIOLLI	6.744	1.19	7.934	0.49%	0,00002	0
19	PAGUE MENOS	4.994	2.604	7.598	0.80%	0,00006	1
20	GIASSI E CIA	4.496	2.046	6.542	0.79%	0,00006	2
33	SAO CRISTOVAO	5.285	0.567	5.852	0.45%	0,00002	1
13	BAHAMAS	4.837	0	4.837	0.99%	0,00010	2

Fonte: elaboração própria a partir dos resultados do modelo DEA-VRS e dos cálculos no Gephi 0.9.2

A análise dos graus ponderados de entrada, de saída e total (que corresponde à soma dos outros dois) nos ajuda a verificar a influência do market share e do índice HHI sobre a influência que a eficiência de cada empresa gera

sobre as demais. Observando os dados da tabela 10, verifica-se que é desprezível o grau de correlação (-6,71%) entre os graus ponderados das empresas e os seus níveis de participação no mercado. O mesmo ocorre com a correlação entre seus graus ponderados e seus índices HHI, que é ainda menor (- 1,71%). Portanto, estes dois resultados corroboram os resultados do modelo DEA-VRS, o qual mostrou que as empresas mais eficientes não são necessariamente as que detêm maior participação de mercado. Isto foi confirmado pela análise da rede de graus ponderados do gráfico 6 e pelas estatísticas de rede da tabela 10.

Examinando a tabela 10, percebemos também que as classes de modularidade foram estimadas conforme a semelhança entre os graus ponderados de saída e de entrada de cada agente/DMU da rede. Isto é uma característica presente no algoritmo de otimização de modularidade. Obviamente, uma vez que o grau de saída revela a força da influência gerada pelo índice de eficiência de uma DMU para as demais e o grau de entrada mostra o quanto uma DMU recebeu de eficiência transmitida por outra, obviamente as classes de modularidade agregam empresas com perfis semelhantes de indicadores utilizados no modelo DEA-VRS, conforme vimos anteriormente.

5. Considerações finais

O varejo, no mundo todo, é fortemente marcado pelo dinamismo e pela inovação. A globalização trouxe para o setor acirramento da competitividade e estreitamento das margens de lucro. Com isso, a estratégia competitiva se tornou ferramenta chave para assegurar a manutenção ou o aumento de margens líquidas.

Nesse cenário, acompanhar os movimentos de outros competidores e, quando não for possível superá-los, ao menos, igualar-se a eles nas boas práticas é tão importante quanto manter uma boa gestão de sua própria operação. Nesse intuito as ferramentas de avaliação de eficiência são de fundamental importância para a análise de *benchmark*.

A pesquisa demonstrou que o varejo hipermercado no Brasil é bastante heterogêneo. Uma evidência disto está na discrepância entre os valores médios e mínimos das variáveis apresentadas nas estatísticas descritivas. Dentre as

firmas julgadas eficientes, 68% delas concentram suas atividades em lojas de até 2500m² e 20% em lojas com até 12mil m².

O modelo DEA-VRS apontou que as empresas eficientes são aquelas que usam em média um menor número de funcionários e de caixas em relação a área de vendas. Estas empresas, em comparação com as ineficientes, possuem um percentual maior de sua operação concentrado em grandes lojas. Isso corrobora com a hipótese inicial de que as firmas mais eficientes são aquelas que priorizam os modelos de atacarejo ou de grandes lojas. Além disso, essa relação positiva entre tamanho e eficiência é ratificada por outros estudos na literatura nacional (PONTES, 2016; YU & ANGELO, 2001; SOUZA, MACEDO & FERREIRA, 2010).

É importante observar que as firmas da amostra julgadas ineficientes também concentram uma parcela significativa de suas operações no formato de “supermercados convencionais” (67% contra 68% das eficientes), assim se faz necessário, em pesquisas futuras, investigar outros aspectos da operação das firmas para que se possa confirmar o protagonismo do formato como determinante ou não da eficiência. Todavia, se forem considerados apenas os formatos de hipermercado, clube atacadista e loja depósito, que são os maiores formatos, as firmas eficientes concentram neles 20% da sua operação contra 14% nas firmas ineficientes.

Além de colaborar para ampliação da literatura acerca do tema, esta pesquisa contribui com o mercado varejista, uma vez que evidencia a necessidade correção na proporção de utilização dos insumos como forma de aumentar a eficiência. Também traz um conjunto de informações relevantes para elaboração de estratégias de *benchmark*, uma vez que os conjuntos de referência estabelecidos podem ser utilizados tanto como parâmetros referenciais (benchmarks) quanto para a construção de outros indicadores não necessariamente baseados em índices de eficiência.

No que se refere a metodologia empregada, uma das principais limitações da Análise Envoltória de Dados é a atribuição de todas as falhas à ineficiência operacional. Assim, características específicas das empresas não relacionadas aos fatores de produção ficaram de fora da análise. Isso vale também para os fatores exógenos (relacionados ao ambiente de negócios ou aos stakeholders das empresas) que podem influenciar o resultado. Por isso, embora seja uma

ferramenta de grande utilidade, os resultados obtidos através da DEA devem ser usados com cautela e quando possível associados a outras ferramentas.

Nesse sentido, recorreremos à duas métricas da análise de redes para reforçar a análise, DEA-VRS e Análise de Redes. Estas duas métricas são o cálculo de graus ponderados de saída, de entrada e totais e também o algoritmo de otimização de modularidade, que separou as empresas eficientes, ineficientes e de eficiência mediana em clusters de acordo com os índices de eficiência obtidos pelo modelo DEA-VRS.

A aplicação deste algoritmo confirmou o resultado do modelo DEA-VRS de que existe uma relação positiva entre tamanho, eficiência e modelo de negócio adotado pelo hipermercado. Além disso, o algoritmo permitiu definir clusters que agruparam tais empresas de acordo com suas semelhanças, o que facilitou muito a visualização dos resultados e a identificação de tais empresas.

Já o cômputo dos graus de entrada, saída e totais nos permitiram confirmar outro resultado do modelo DEA-VRS: o de que são desprezíveis, em termos de estatística de redes, as correlações entre a transmissão de índices de eficiência entre os hipermercados que compõem cada cluster, seus *market shares* e seus índices HHI (Herfindahl-Hirschman Index).

Finalmente, no que diz respeito às limitações encontradas por este estudo, a escassez de dados impediu que fossem incorporados mais insumos à função de produção, bem como limitou o número de firmas participantes da amostra. O acesso a informações financeiras mais detalhadas, como o por exemplo, o lucro que poderia ser utilizado como *output* ao invés do faturamento bruto, tornaria os resultados da pesquisa mais robustos.

Outra limitação refere-se à não estimação da eficiência para as firmas analisadas em anos anteriores. Isso permitiria avaliar melhor a influência do formato das lojas no desempenho das unidades. Embora os dados referentes ao produto e à utilização dos insumos estivessem disponíveis para períodos anteriores, os dados referentes a distribuição da operação de cada firma segundo formato de loja não foram publicados para os mesmos períodos. Dessa forma, seria possível estimar apenas a eficiência da unidade sem estabelecer qualquer relação com o formato de loja, o que não atenderia aos propósitos da pesquisa.

Por fim, vale ressaltar que os critérios de eficiência técnica embora sejam extremamente relevantes para dar suporte as decisões de alocação eficiente dos recursos, não podem ser considerados isoladamente para indicar com precisão o nível mais eficiente de produção. Assim, sugere-se que em estudos futuros medidas de eficiência econômica sejam adicionadas.

Referências

ALMEIDA, Paulo Nazareno Alves. Análise envoltória de dados. In: **Econometria na Prática**. ANDRADE, C. M.; TIRYAKI, G. F. (eds.). São Paulo: Alta Books, 2017, p. 419-451.

ALMEIDA, Paulo Nazareno Alves. Análise de Fronteira estocástica. In: **Econometria na Prática**. ANDRADE, C. M.; TIRYAKI, G. F. (eds.). São Paulo: Alta Books, 2017b, p. 379-417.

ANDERSON, R. I.; FOK, R.; SCOTT, J. **American Business Review**; West Haven Vol. 18, Ed. 1, Jan 2000: 40-48.

ARROW, K. J. **Social Choice and Individual Values**. Mansfield Centre (Conn.): Martino Fine Books, 1951

BADIN, Neiva Teresinha. **Avaliação da Produtividade de Supermercados e Seu Benchmarking**. Florianópolis: UFSC, 1997, 133p.

BARBOSA, F. C. **Análise envoltória de dados: teoria e Aplicações Práticas**. Itumbiara: Conhecimento Livre, 2019.

BARROS, C. P. ALVES, C. An empirical analysis of productivity in a Portuguese retail chain using the Malmquist index. **Journal of Retailing and Consumer Services**: 2004, vol.1, n.5, p. 269-278.

BARROS, C. P. Efficiency measurement among hypermarkets and supermarkets and the identification of the efficiency drivers: A case study. **International Journal of Retail & Distribution Management**, v. 34, n. 2, p. 135–154, 2006.

BATOR, F. M. On Capital Productivity, Input Allocation and Growth*. **The Quarterly Journal of Economics**, v. 71, n. 1, p. 86–106, 1957.

BAVIERA-PUIG, A.; BAVIERA, T; BUITRAGO-VERA, J. et al. “Internal benchmarking in retailing with DEA and GIS: the case of a loyalty-oriented supermarket chain”. **Journal of Business Economics and Management**, v. 21, n. 4, p. 1035–1057, 2020.

BENICIO, J.; MELLO, J. C. et al. “Retornos de escala em DEA: críticas ao BCC e novo modelo”. **XLVI Simpósio Brasileiro De Pesquisa Operacional**. Pesquisa Operacional na Gestão da Segurança Pública. Salvador, 2014.

BOGETOFT, P.; OTTO, L. Data Envelopment Analysis DEA. *In*: BOGETOFT, P.; OTTO, L. (Orgs.). **Benchmarking with DEA, SFA, and R**. New York, NY: Springer, 2011, p. 81–113. (International Series in Operations Research & Management Science). Disponível em: <https://doi.org/10.1007/978-1-4419-7961-2_4>. Acesso em: 24 mar. 2021.

BOUERI, R. “Modelos não paramétricos: Análise Envoltória de Dados (DEA)” *In*: BRASIL. Ministério da Fazenda. **Avaliação da Qualidade do Gasto Público e Mensuração da Eficiência**. Brasília, 2015.

BOUERI, R.; ROCHA, F. F.; RODOPOULOS, F. M. A. Avaliação da qualidade do gasto público e mensuração da eficiência. 2015. Disponível em: <<https://repositorio.usp.br/item/002732373>>. Acesso em: 24 mar. 2021.

BRITO, D. Qual o formato correto de sua loja? **SuperHiper**, ed. 24, p. 74–77, 1998.

CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision making units. **European Journal of Operational Research**, v. 2, n. 6, p. 429–444, 1978.

CLAUSET, A., NEWMAN, M. E. J. e MOORE, C. Finding community structure in very large networks. **Physical Review E**, 70(6): 066111, 2004. <https://arxiv.org/pdf/cond-mat/0408187.pdf>

COELLI, T. J. RAO, D. S. P. BATTESE, G. E. **An introduction to efficiency and productivity analysis**. Boston: KluwerAcademic, 1998.

DONTHU, N.; YOO, B.. Cultural Influences on Service Quality Expectations. **Journal of Services Research**, v. 2, p. 178–186, 1998.

EASLEY, D. e KLEINBERG, J. **Networks, crowds and markets: reasoning about a highly connected world**. Cambridge, Massachussets: Cambridge University Press, 2010.

FÄRE, R.; KNOX, C. A. Measuring the technical efficiency of production. **Journal of Economic Theory**, v. 19, n. 1, p. 150–162, 1978.

FARRELL, M. J. The Measurement of Productive Efficiency. **Journal of the Royal Statistical Society: Series A (General)**, v. 120, n. 3, p. 253–281, 1957.

FERREIRA, C. M. C. ; GOMES, A. P. **Introdução à Análise Envoltória de Dados** – Teoria, modelos e aplicações. Editora UFV, 2020.

FERREIRA, Carlos Maurício de Carvalho; GOMES, Adriano Provezano. **Introdução à Análise Envoltória de Dados**. 2. ed. [s.l.]: Editora UFV, 2020.

FERREIRA, M. A. M.; VENÂNCIO, M. M. e ABRANTES, L. A. Análise da Eficiência do Setor de Supermercados no Brasil. **Economia Aplicada**, v. 13, n. 2, 2009, p. 333-347.

FILHO, S.; ANTONIO, U. **O setor supermercadista no Brasil nos anos 1990**. text, Universidade de São Paulo, 2003. Disponível em: <http://www.teses.usp.br/teses/disponiveis/11/11132/tde-11072003-140924/>. Acesso em: 24 mar. 2021.

FORTUNATO, S. Community detection in graphs. **Physics Report**, 486(3-5): 75-174, 2010. <https://arxiv.org/pdf/0906.0612.pdf>

GAURI, D. K., 2013. Benchmarking Retail Productivity Considering Retail Pricing and Format Strategy. **Journal of Retailing**: 2013, vol. 89, n. 1, p. 1-14.

GIRVAN, M.; NEWMAN, M.E. J. Community structure in social and biological networks. *Proc. Natl Acad. Sci. USA* 99, 7821-7826. **Proceedings of the National Academy of Sciences of the United States of America**, v. 99, p. 7821–6, 2002.

GOMES, Adriano Provezano. BAPTISTA Antonio José Medina dos Santos. Análise envoltória de dados: conceitos e aplicações. In: SANTOS, Maurinho Luiz dos; VIEIRA, Wilson da Cruz. **Métodos Quantitativos em Economia** 1ed.Viçosa: UFV, 2004, v. 1, p. 121-160.

HWANG, Shiuh-Nan; CHANG, Te-Yi. Using data envelopment analysis to measure hotel managerial efficiency change in Taiwan. **Tourism Management**, v. 24, n. 4, p. 357–369, 2003.

JACQUEMIN, Alexis. **The New Industrial Organization: Market Forces and Strategic Behavior**. Cambridge, Mass: MIT Press, 1987.

JUNIOR, S. P. C.; RODRIGUES, L. F. **Análise de Eficiência de Distribuidoras de Energia Elétrica: Um Estudo Envolvendo o Modelo DEA Baseado em Folgas**. XXXVII Encontro Nacional de Engenharia de Produção: Joinville, SC, Brasil, 10 a 13 de outubro de 2017.

KAHRAMAN, C.; ULUĞ, İ.; OTHAN, C. B.; et al. Efficiency Analysis in Retail Sector: Implementation of Data Envelopment Analysis in a Local Supermarket Chain. In: DURAKBASA, Numan M.; GENÇYILMAZ, M. Günes (Orgs.). **Proceedings of the International Symposium for Production Research** 2018. Cham: Springer International Publishing, p. 884–897, 2019.

KEH, H. T. e CHU, S. Retail productivity and scale economies at the firm level: a DEA approach. In: **OMEGA International Journal of Management Science**: 2003, Vol. 31, n. 2. p. 75-82.

KEH, H. T.; CHU, S. Retail productivity and scale economies at the firm level: a DEA approach. **Omega**, v. 31, n. 2, p. 75–82, 2003.

KO, K.; CHANG, M.; BAE, Eun-Song; et al. Efficiency Analysis of Retail Chain Stores in Korea. **Sustainability**, v. 9, p. 1629, 2017.

KOTLER, Philip; KELLER, Kevin Lane; YAMAMOTO, Sonia Midori; *et al.* **Administração de Marketing**. 15ª edição. [s.l.]: Pearson Universidades, 2019.

MACEDO, M. A. S. A Utilização da Análise Envoltória de Dados (DEA) na Consolidação de Medidas de Desempenho Organizacional. In: **CONGRESSO BRASILEIRO DE CUSTOS**, 11, 2004, Porto Seguro. Anais do XI Congresso Brasileiro de Custos. Porto Seguro: ABC; 2004.

MAS-COLELL, A.; WHINSTON, M. D.; GREEN, J. **Microeconomic Theory**. Illustrated edição. New York: Oxford University Press, USA, 1995.

MOURA, T. L. Formatos de varejo de alimentos: um estudo sobre as preferências do consumidor. UFSCar, 2005, 213pág. Disponível em: <<https://repositorio.ufscar.br/handle/ufscar/3796>>. Acesso em: 24 mar. 2021.

MOREY, R. C.; DITTMAN, D. A. Evalatn a Hotel GM's Performance: A Case Study in Benchmarking. **Cornell Hotel and Restaurant Administration Quarterly**, 1995.

NERI, M. A nova classe média. **Revista Conjuntura Econômica**, v. 62, n. 9, p. 48–51, 2008.

NEWMAN, M. E. J. e GIRVAN, M. Finding and evaluating community structure in networks. **Physical Review E**, 69(2): 026113, 2004.

NORMAN, M. e STOKER B. **Data envelopment analysis: the assessment of performance**. John Wiley & Sons Publishers, New York, 1991.

OLIVEIRA, C. M. Concentração de mercado: um diagnóstico do varejo hipermercadista no brasil entre 2009-2017. **Braz. J. of Develop.**, Curitiba, v. 6, n.6, p.38418-38430, jun.2020.

PERRIGOT, R.; BARROS, C. P. “Technical efficiency of French retailers”. **Journal of Retailing and Consumer Services**, v. 15, n. 4, p. 296–305, 2008.

PONS, P. e LATAPY, M. Computing commuties in large networks using random networks. **Journal of Graph Algorithms and Applications**, v. 10, n.2, p. 191-218, 2006.

PONTES, T. T. S. **Determinantes da eficiência técnica no setor de supermercados brasileiro: uma análise utilizando métodos de fronteira de produção**. Tese (doutorado). Universidade de São Paulo, 213p., 2017.

PONTES, T. T. S. “Eficiência Técnica no Setor Supermercadista Brasileiro”. **XX SEMEAD Seminários em Administração**: novembro, 2017.

PORTER, M. “Os caminhos da lucratividade: como implementar uma verdadeira vantagem competitiva”. **HSM Management**, nº 5, nov.-dez. 1997.

RICHTER, H. G. **Retailing: principles and practices**. New York: Mc Graw Hill, 1954.

RUGGIERO, J. “Measuring technical efficiency”. **European Journal of Operational Research**, v. 121, n. 1, p. 138–150, 2000.

SBVC. Sociedade Brasileira de Varejo e Consumo. **O papel do varejo na economia brasileira**. São Paulo, 2020. Disponível em: <<http://sbvc.com.br/o-papel-do-varejo-na-economia-brasileira-atualizacao-2020/>> Acesso em 27 jan. 2020.

SENRA, L. F. A. C.; NANJI, L. C. MELLO, J. C. C. B. S. *et al.* “Estudo sobre métodos de seleção de variáveis em DEA”. **Pesquisa Operacional**, v. 27, n. 2, p. 191–207, 2007.

SOUZA, M. A. F. ; MACEDO, M. V. S. e FERREIRA, M. S. “Desempenho organizacional no setor supermercadista brasileiro: uma análise apoiada em DEA”. **REGE**, São Paulo, v. 17, n. 2, p.151-167, abr./jun. 2010.

STAIR, R. M.; REYNOLDS, G. W. **Fundamentals of Information Systems**. [s.l.]: Thomson Course Technology, 2007.

TAPENHURST, T. “A Background to BM”. In: STAPENHURST, T. **The Benchmarking Book**. Butterworth-Heinemann, 2009, p. 1-2.

THANASSOULIS, E. “A Comparison of Regression Analysis and Data Envelopment Analysis as Alternative Methods for Performance Assessments”. **The Journal of the Operational Research Society**, v. 44, n. 11, p. 1129–1144, 1993.

THANASSOULIS, E. “Assessing the Effectiveness of Schools with Pupils of Different Ability Using Data Envelopment Analysis”. **The Journal of the Operational Research Society**, v. 47, n. 1, p. 84–97, 1996.

THOMAS, R. R. *et al.* “A process for evaluating retail store efficiency: a restricted DEA approach”. **Intern. J. of Research in Marketing**, p. 487–503, 1998.

WANG, J. “Corporate performance efficiency investigated by data envelopment analysis and balanced scorecard”. **Journal of American Academy of Business**, Cambridge, v. 9, n. 2, p. 312-8, sept. 2006.

WÖBER, K.; FESENMAIER, D. “A multi-criteria approach to destination benchmarking: A case study of state tourism advertising programs in the United States”. **Journal of Travel and Tourism Marketing**, v. 16, p. 1–18, 2014.

YU, A. S. O.; ANGELO, C. F. “Performance of Brazilian supermarkets: A comparative analysis between large and small store chains”. **Journal of Small Business and Enterprise Development**, 8 ed., v. 4, p. 3339-348, 2001.

YU, W.; RAMANATHAN, R. "An assessment of operational efficiencies in the UK retail sector". **International Journal of Retail & Distribution Management**, v. 36, n. 11, p. 861–882, 2008.

ZHU, J.; Zhu, J. "Multi-factor performance measure model with an application to Fortune 500 companies". **European Journal of Operational Research**, v. 123, p. 105–124, 2000.