

UNIVERSIDADE DO VALE DO RIO DOS SINOS — UNISINOS
UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS CONTÁBEIS
NÍVEL MESTRADO

EDUARDO ROSA KRÁS BORGES

INDICADORES DE COMPETITIVIDADE DO MERCADO DE ENSINO SUPERIOR
PARA A AVALIAÇÃO DA TRANSFERÊNCIA INTERINSTITUCIONAL DISCENTE:
UM ESTUDO PARA OS CURSOS DE GRADUAÇÃO EM CIÊNCIAS CONTÁBEIS

PORTO ALEGRE-RS
2022

Eduardo Rosa Krás Borges

INDICADORES DE COMPETITIVIDADE DO MERCADO DE ENSINO SUPERIOR
PARA A AVALIAÇÃO DA TRANSFERÊNCIA INTERINSTITUCIONAL DISCENTE:
UM ESTUDO PARA OS CURSOS DE GRADUAÇÃO EM CIÊNCIAS CONTÁBEIS

Dissertação apresentada como requisito
parcial para a obtenção de título de Mestre
em Ciências Contábeis: Controle de
Gestão, pelo Programa de Pós-Graduação
em Ciências Contábeis da Universidade
do Vale do Rio dos Sinos – UNISINOS

Orientador:
Prof. Dr. Alexsandro M. Carvalho

Porto Alegre-RS
2022

B732i

Borges, Eduardo Rosa Krás.

Indicadores de competitividade do mercado de ensino superior para a avaliação da transferência interinstitucional discente : um estudo para os cursos de graduação em ciências contábeis / por Eduardo Rosa Krás Borges. – 2022.
74 f. : il. ; 30 cm.

Dissertação (mestrado) — Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Ciências Contábeis, Porto Alegre, RS, 2022.

“Orientador: Dr. Alessandro M. Carvalho”.

1. Mercado de educação superior. 2. Redes complexas. 3. Teoria da decisão. 4. Migração de alunos. 5. Controle de gestão. 6. Curso de ciências contábeis. I. Título.

CDU: 657:378

EDUARDO ROSA KRÁS BORGES

**INDICADORES DE COMPETITIVIDADE DO MERCADO DE ENSINO SUPERIOR
PARA A AVALIAÇÃO DA TRANSFERÊNCIA INTERINSTITUCIONAL DISCENTE**

Dissertação apresentada como requisito parcial para a obtenção do título de Mestre em Ciências Contábeis: Controle de Gestão, pelo Programa de Pós-Graduação em Ciências Contábeis da Universidade do Vale do Rio dos Sinos - UNISINOS

Orientador: Prof. Dr. Alexsandro M. Carvalho

Aprovado em 12 de dezembro de 2022.

BANCA EXAMINADORA

Prof^a. Dr^a. WENDY BEATRIZ WITT HADDAD CARRARO (UFRGS)

Prof. Dr. CARLOS ALBERTO DIEHL (UNISINOS)

Prof. Dr. TIAGO WICKSTROM ALVES (UNISINOS)

Prof. Dr. ALEXSANDRO MARIAN CARVALHO (Orientador)

À memória de meus pais, Enio e Esther Teresinha, com amor eterno.

AGRADECIMENTOS

Ao meu orientador, professor Doutor Alexsandro Carvalho, pelos ensinamentos, dedicação, paciência e amizade durante esse processo.

Ao professor Doutor Carlos Diehl, pelo incansável incentivo e por ser um dos grandes responsáveis por eu estar concluindo o mestrado.

À minha colega de trabalho e amiga Fernanda Robin, por ter me incentivado a ingressar no mestrado e pela parceria durante este período.

Ao Dr. Nilo Frantz, pelo importante apoio nesta jornada, e por ser um exemplo de busca constante pelo conhecimento.

Aos meus colegas de mestrado, pelo convívio e aprendizado mútuos, em especial aos colegas Camila Gandini, Cassandra Martinelli, Jéssica Maske e Luís Guilherme Raposo.

Às minhas filhas, por serem fonte de inspiração para que eu busque constantemente ser uma pessoa melhor nesta existência.

RESUMO

Nas últimas décadas, o mercado de educação superior no Brasil apresentou um cenário de forte expansão na oferta de cursos. Esta evidência, deve-se principalmente a aspectos socioeconômicos, mudanças na legislação e programas governamentais voltados para a popularização do acesso ao ensino superior. Em meio a este mercado de concentração e concorrência em plena evolução, a sustentabilidade econômica das instituições, em parte, passa pelo desafio de implementar a adoção de práticas de controle e gestão que minimizam os efeitos do desequilíbrio entre oferta e demanda de vagas (oferta » demanda). No contexto da gestão de indicadores, o monitoramento da demanda por vagas dos cursos representa um indicador chave na tomada de decisão dos gestores. Diante disso, a presente pesquisa teve como objetivo avaliar as relações entre demanda por vagas (número de migrações) e a competição entre os cursos no que tange as transferências interinstitucionais. A concorrência entre as instituições foi monitorada tanto pela preferência dos discentes (probabilidade de retenção) como pelo número de rotas de imigração | emigração entre as instituições (número de conexões entre as instituições). Para tanto, analisou-se dados longitudinais das bases do INEP e do IBGE no período de 2009 a 2017. A população da pesquisa foi representada pela totalidade das IES no Brasil, e a amostra constituída pelos alunos do curso de Ciências Contábeis. A análise diagnóstica foi baseada no desenvolvimento de modelos econômicos nos níveis discente e cursos. Para os discentes, desenvolveu-se um modelo Logit de Dados Empilhados que trata da relação entre a probabilidade de retenção e aspectos socioeconômicos. Já no nível dos cursos, determinou-se um modelo de Dados em Painel de Efeitos Aleatórios para explicação das migrações em função do número de conexões entre as instituições. Em termos de achados, em nível da decisão do aluno, a probabilidade de retenção discente no curso depende principalmente, em ordem de maior probabilidade, de aspectos econômicos, sociais e qualidade do curso. No que se refere à concorrência entre os cursos, constatou-se que o fluxo de migração discente responde a sua centralidade na competição interinstitucional. Em outras palavras, há uma relação positiva entre a taxa de migração de alunos e a diferença entre número/dispersão do número de instituições de origem e destino. Por fim, foi apresentada sugestão para agregar variável de receita média por aluno aos modelos, dado este que não é de domínio público atualmente.

Palavras-chave: Mercado de educação superior. Redes complexas. Teoria da decisão. Migração de alunos. Controle de gestão.

ABSTRACT

In recent decades, the higher education market in Brazil has seen a strong expansion in the offer of courses. This evidence is mainly due to socioeconomic aspects, changes in legislation and government programs aimed at popularizing access to higher education. The economic sustainability of the institutions is in the midst of this market of concentration and competition. Competition happens to be in full evolution; and the economic sustainability, in part, involves the challenge of implementing the adoption of control. As well, management practices will minimize the effects between supply and demand for vacancies. Thus, creating an imbalance. In the context of managing indicators, monitoring the demand for vacancies in courses represents a key indicator in decision-making by managers. Therefore, this research aimed to evaluate the relationship between demand for places (number of migrations) and competition between courses in terms of interinstitutional transfers. Competition between institutions was monitored both by students' preference (probability of retention) and by the number of immigration routes | emigration between institutions (number of connections between institutions). For this purpose, longitudinal data from the INEP and IBGE databases from 2009 to 2017 were analyzed. The research population was represented by all HEIs in Brazil, and the sample consisted of Accounting students. The diagnostic analysis was based on the development of econometric models at student and course levels. For the students, a Logit model of Stacked Data was developed that deals with the relationship between the probability of retention and socioeconomic aspects. At the level of courses, a Random Effects Panel Data model was determined to explain migrations as a function of the number of connections between institutions. In terms of findings, at the level of the student's decision, the probability of student retention in the course mainly depends, in order of greater probability, on economic, social and quality aspects of the course. With regard to competition between courses, it was found that the student migration flow responds to its centrality in inter-institutional competition. In other words, there is a positive relationship between the student migration rate and the difference between the number/dispersion of the number of source and target institutions. Finally, a suggestion was presented to add a variable of average revenue per student to the models, given that it is not currently in the public domain.

Keywords: Higher education market. Complex networks. Decision theory. Student migration. Management control.

LISTA DE FIGURAS

1	Função de valor na teoria da prospecção.	23
2	Nº de Instituições de Ensino Superior que possuem curso de graduação em Ciências Contábeis.	29
3	Representação das interações sociais de pessoas de um mesmo clube.	31
4	Grau de competição do mercado	41
5	Diagrama do Processamento de Dados - Dimensão Aluno	50
6	Diagrama do Processamento de Dados - Dimensão IES	53
7	Percentual de Migração dos Alunos - 2011 a 2017	55
8	Número de Alunos de Ciências Contábeis - 2011 a 2017	55
9	Matriz de Correlação de Variáveis Independentes	58
10	Número de alunos de ciências contábeis por modalidade de 2010 a 2017	61
11	Crescimento do número de alunos de ciências contábeis de 2010 a 2017	62
12	Número de alunos por curso de ciências contábeis - 2010 a 2017 . .	63
13	Número de cursos de ciências contábeis por modalidade - 2010 a 2017	63

LISTA DE TABELAS

1	Nº de IES, matriculados, concluintes, vagas e inscritos - 1995 e 2018	28
2	Análise Descritiva das Variáveis Independentes - Dimensão Aluno .	56
3	Teste VIF das variáveis independentes	57
4	Resultado da Regressão Logística - Brasil	58
5	Tabela de <i>odd/ratio</i>	59
6	Análise Descritiva das Variáveis Independentes - Dimensão IES . . .	61
7	Dados em painel da variação de entrantes nos cursos de ciências contábeis	65
8	Teste VIF para multicolinearidade	66
9	Teste Hausmann	66
10	Dados em painel com efeitos aleatórios - Dimensão IES	67

LISTA DE QUADROS

1	Resumo dos fatores na teoria de Herzberg	21
2	Comparação entre os dois fatores da Teoria de Herzberg	21
3	Classificação de Marchal: as estruturas de mercado diferenciadas também por fatores associados ao comportamento dos agentes a natureza dos produtos.	25
4	Variáveis de pesquisa	44

LISTA DE ABREVIATURAS

CES	Censo da Educação Superior
EaD	Ensino à Distância
ENADE	Exame Nacional de Desempenho dos Estudantes
ENEM	Exame Nacional do Ensino Médio
Fies	Fundo de Financiamento Estudantil
HEI	High Education Institution
HME	Hipótese do Mercado Eficiente
IBGE	Instituto Brasileiro de Geografia e Estatística
IES	Instituições de Ensino Superior
LDB	Lei de Diretrizes Básicas da Educação
LGPD	Lei Geral de Proteção aos Dados
PROUNI	Programa Universidade para Todos
SiSU	Sistema de Seleção Unificada
VIF	Fator de Inflação da Variância

SUMÁRIO

1 INTRODUÇÃO	14
1.1 Tema e contextualização	14
1.2 Problema de Pesquisa	15
1.3 Objetivos	15
1.3.1 Objetivo geral	15
1.3.2 Objetivos específicos	16
1.4 Justificativa	16
1.5 Delimitações	17
2 FUNDAMENTAÇÃO TEÓRICA E EMPÍRICA	19
2.1 Teoria da Decisão	19
2.1.1 Teoria Utilitarista	19
2.1.2 Teoria dos Dois Fatores	20
2.1.3 Teoria da Perspectiva	21
2.1.4 Escolha Intertemporal	23
2.2 Mercado	24
2.2.1 Eficiência e Ineficiência	24
2.2.2 Estrutura	25
2.2.3 Ensino Superior	26
2.3 Mercado como Redes	29
2.3.1 Redes Complexas	30
2.4 Estudos Relacionados	33
2.5 Desenvolvimento das Hipóteses	35
3 METODOLOGIA	36
3.1 Delineamento	36
3.2 População e Amostra	37
3.3 Coleta e Tratamento dos Dados	37
3.3.1 Coleta dos Dados	37
3.4 Variáveis da Pesquisa	38
3.4.1 Dimensão Aluno	38
3.4.2 Dimensão IES	41
3.4.3 Variáveis de controle	42
3.5 Síntese das Variáveis	43
3.6 Modelos	43
3.6.1 Dimensão Aluno	43
3.6.2 Dimensão IES	45
3.7 Processamento dos Dados	46
3.7.1 Dimensão Aluno	46
3.7.2 Dimensão IES	49
3.7.3 Limitações do Método	52
4 ANÁLISE DOS RESULTADOS	54
4.1 Dimensão Aluno	54
4.1.1 Estatística Descritiva das Variáveis	54
4.1.2 Análise do Modelo Econométrico	56
4.2 Dimensão IES	60

4.2.1 Estatística Descritiva das Variáveis	60
4.2.2 Análise do Modelo Econométrico	63
4.2.3 Síntese da Seção de Análise de Resultados	68
5 CONSIDERAÇÕES FINAIS	69
REFERÊNCIAS	72

1 INTRODUÇÃO

1.1 Tema e contextualização

A história do ensino superior privado no Brasil teve início oficialmente na década de 1940 durante o governo de Getúlio Vargas. Este marco inicial da educação superior se deu a partir da publicação do Decreto 6.409/1940, o qual autorizou a abertura da primeira Instituição de Ensino Superior (IES) privada do Brasil: a Faculdade de Direito da Pontifícia Universidade Católica do Rio de Janeiro (antiga Faculdade Católica do Rio de Janeiro) (PEREIRA; BRITO, 2014).

Posteriormente na década de 1960, ocorreu outro fato relevante na história evolutiva das IES, durante o regime de governo militar. Conforme Pereira e Brito (2014), por intermédio do então Conselho Federal de Educação, foi promovida uma visão pró privatização das universidades e interrompida a federalização de estabelecimentos de ensino superior. Todavia, o maior impacto no segmento de educação superior se deu a partir do início dos anos 1990, sob os efeitos da globalização do capital e expansionismo financeiro. No mercado brasileiro, as mudanças foram ainda mais fortemente alavancadas com o advento da promulgação da Lei de Diretrizes Básicas da educação (LDB), de 1996 (DINIZ; OLIVEIRA; LIMA, 2021). Este fato, aliado aos aspectos conjunturais de globalização e abertura, proporcionou o interesse e a viabilidade para a entrada de novas organizações a partir das mudanças promovidas pela nova legislação, entre elas, a permissão da atuação com fins lucrativos neste mercado (SAMPAIO, 2013).

Neste contexto, eclodiu um grande número de IES, conforme dados do setor publicamente divulgados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Aliado a estes fatores, de acordo com Sampaio (2013), houve uma transformação do mercado para um ambiente propício ao interesse de grandes grupos na aquisição ou fusão de instituições, o que aqueceu o mercado de educação no Brasil. Esta dinâmica exigiu que as empresas do segmento passassem a ser mais competitivas e eficientes na condução dos seus negócios, fazendo com que buscassem aprimorar a sua estrutura nos níveis estratégico e gerencial. Para tanto, lançaram mão de ferramentas de gestão, tais como planejamento estratégico, controle de indicadores de desempenho, gestão orçamentária e planejamento financeiro, entre outras. Conforme Breitenbach, Alves e Diehl (2010), a busca pela sofisticação da gestão das IES também teve o efeito de promover o aumento no nível de competição entre elas. Este movimento é bidirecional, ou seja, à medida em que as práticas de gestão das empresas evoluem, o nível de competição do mercado aumenta, e isto acarreta em um maior grau de exigência no nível de gestão das empresas.

O principal produto de uma IES é representado pelos serviços educacionais

prestados para os alunos. Estas tarefas se caracterizam pela estreita ligação entre o propósito da organização e seu objeto (alunos) (BREITENBACH; ALVES; DIEHL, 2010). Por exemplo, a mensuração do desempenho requer que a gestão tenha uma visão abrangente e detalhada tanto interna quanto externa. Isto significa, portanto, que se deve proceder a mensuração dos aspectos tanto endógenos (fatores internos) quanto exógenos (fatores externos).

Há um aspecto que se destaca em relação aos fatores externos, o qual tem impacto direto na demanda: a migração de alunos entre as IES. Este movimento é especialmente relevante no contexto de mercado de educação superior no Brasil, onde houve uma escalada na ampliação da oferta de cursos e instituições. Também há que se considerar os impactos na demanda decorrentes das alterações no formato de seleção utilizados por diversas universidades. Isto se deu a partir de 2009, pelos critérios centralizados de seleção com base no Exame Nacional do Ensino Médio (ENEM) e pela implementação da plataforma do Sistema de Seleção Unificada (SiSU) (LI; CHAGAS, 2017). O mapeamento e medição dos determinantes da migração de alunos entre as IES é útil tanto para as organizações quanto para academia. Para as organizações, o domínio desta informação é de grande valia no seu planejamento, principalmente no que tange ao processo orçamentário (BERGMANN et al., 2020). Já para a academia, a adição de fatores como a concorrência entre as IES e a preferência dos estudantes no estudo da demanda por vagas, pode auxiliar no desenvolvimento de modelos mais realísticos.

1.2 Problema de Pesquisa

Diante do contexto exposto, foi gerada a seguinte questão de pesquisa: qual a relação entre as decisões dos alunos e a concorrência entre as instituições na demanda por vagas via transferência interinstitucional discente?

1.3 Objetivos

1.3.1 Objetivo geral

Analisar a relação entre as decisões dos alunos e a concorrência entre as instituições na demanda por vagas nos cursos de Ciências Contábeis no Brasil, via transferência interinstitucional discente.

1.3.2 Objetivos específicos

- a) Determinar as variáveis "decisão de migração dos alunos", "concorrência entre as instituições" e "demanda por vagas";
- b) Desenvolver um modelo que relacione as variáveis decisão de migração dos alunos e a concorrência entre as instituições, com a demanda por vagas;
- c) Avaliar a relação existente entre as variáveis associadas à decisão dos alunos e a competição do mercado.

1.4 Justificativa

Conhecer o comportamento da demanda em qualquer mercado é uma vantagem que impacta diretamente o processo orçamentário empresarial, tanto no âmbito do planejamento das receitas, quanto na otimização das operações para se atingir os resultados pretendidos.

Neste sentido, faz-se necessário contar com boas práticas de controles de gestão e de desempenho que assegurem às empresas a obtenção e tratamento adequado dos dados de mercado disponíveis. O emprego de infraestruturas sofisticadas para o tratamento de dados é positivo para o processo de análise de negócios na gestão orçamentária (BERGMANN et al., 2020; PRATOLO; SOFYANI; ANWAR, 2020).

Isto se aplica também ao mercado de ensino superior. A migração de alunos entre as IES é um problema de foro privado das instituições públicas e particulares. E, adicionalmente, também pode se constituir em um problema de interesse das políticas públicas educacionais. A evasão se constitui em um desperdício de ordem social, econômica e acadêmica, que onera o erário público pelos investimentos em instituições sem retorno (SILVA FILHO et al., 2007). Como contraponto interessante, Coimbra, Silva e Costa (2021) postulam que grande parte da literatura especializada na evasão restringe seus achados com base em um fator isolado, seja no vínculo com um curso, instituição ou sistema. Na opinião destes autores, o ponto-de-partida para a análise migratória de alunos deveria se dar a partir da análise conjuntural dos alunos, seus anseios e aspirações, os quais extrapolam os limites das instituições. A recente pesquisa dos autores revelou que os periódicos apresentam um baixo número de pesquisas sobre evasão, ao passo que o número de teses e dissertações sobre o tema é bastante elevado. Todavia, conforme Coimbra, Silva e Costa (2021, p.7), "[...] é constatável que a maioria das definições usa expressões diferentes, mas alinha-se na descrição da evasão como sendo simplesmente a perda de vínculo, a saída da instituição, o abandono do curso,[...] instituição ou sistema, sejam eles atos voluntários ou não."

Sendo assim, a presente pesquisa trata do tema da migração de alunos empreendendo uma abordagem ampla das relações, dando enfoque tanto para os fatores de decisão dos alunos, quanto para os fatores de competição entre as IES. A pesquisa se justifica pelos seguintes elementos:

- **Relevância:** a migração de alunos tem impacto direto sobre a avaliação da demanda no mercado de educação superior, tendo reflexos também no processo orçamentário;
- **Contribuição:** a aplicação dos conceitos da teoria de decisão no embasamento da análise dos agentes, bem como da teoria de redes complexas no mapeamento e abordagem da competição entre as IES;
- **Oportunidade:** há poucos trabalhos que abordam o tema de forma semelhante ao tipo de análise proposto. Os trabalhos encontrados que tratam do tema abordam com mais frequência a evasão dos discentes e/ou outros métodos de análise;
- **Acessibilidade:** o estudo é viável, tendo em vista a disponibilidade das bases de dados públicas contendo informações anuais referentes aos discentes e instituições em cursos superiores.

1.5 Delimitações

A presente pesquisa foi construída à luz das teorias da decisão e redes complexas, a fim de buscar embasamento teórico para a explicação dos fenômenos empíricos observados na transferência interinstitucional discente.

No entanto, a pesquisa se restringirá à observação de um determinado número de parâmetros que se entende por serem valorizados e levados em conta ou não na visão dos discentes, no seu processo de tomada de decisão. Apesar de se propor a avaliar este processo decisório, é importante destacar que a pesquisa não tem a pretensão de estabelecer um modelo que esgote sua capacidade preditiva, mas sim, buscar, com base nos elementos disponíveis, o entendimento dos fatores que motivam a tomada de decisão por parte dos alunos, e sua relação com a concorrência interinstitucional.

Do ponto-de-vista da concorrência entre as IES, a abordagem realizada abrange apenas os desdobramentos decorrentes do fluxo de alunos do curso de Ciências Contábeis, entre outras características das instituições, como tamanho, natureza jurídica e modalidade de ensino. Não são tratados outros fatores, tais como estrutura de capital, estrutura docente e ambiente competitivo.

Outro aspecto a ser destacado é o fato da pesquisa aplicar os preceitos da teoria de redes complexas apenas no âmbito da aplicação do conceito de topologia de

rede. Esta conduta foi necessária para viabilizar a construção dos modelos propostos, ou seja, a malha de conexões existentes entre as instituições, tendo em vista a característica intrínseca da pesquisa, a qual necessita mapear os fluxos interinstitucionais para promover os achados.

2 FUNDAMENTAÇÃO TEÓRICA E EMPÍRICA

2.1 Teoria da Decisão

A tomada de decisão é uma característica inerente à humanidade desde os primórdios de sua existência, simbolizada por escolhas feitas tendo por base determinada interpretação de sinais, astros e símbolos, e avançando, a partir do surgimento da escrita, sob a forma de decisões tomadas com base em consultas a oráculos, profetas e videntes (BUCHANAN; O CONNELL, 2006).

Segundo Steele e Stefánsson (2015), o conceito central da tomada de decisão gira em torno de dois aspectos principais: preferências e perspectivas. Segundo estes autores, de forma elementar, a preferência é uma atitude comparativa à medida em que uma opção é escolhida mediante a maior desejabilidade do agente pela alternativa escolhida em detrimento de outra, ou de outras.

Nas próximas seções, serão abordadas quatro vertentes teóricas que tratam da teoria da decisão: teoria utilitarista, teoria dos dois fatores, teoria da perspectiva e escolha intertemporal.

2.1.1 Teoria Utilitarista

A teoria utilitarista (TU), ou teoria da utilidade esperada, consiste na teoria mais ortodoxa de decisão, e reside no conceito de que, em condições de incerteza, um agente deve preferir a opção com maior desejabilidade (valor esperado), ou seja, deve maximizar sua utilidade esperada mediante sua tomada de decisão (THOMA, 2019). Neste sentido, poderia se afirmar que o valor das coisas baseia-se na utilidade gerada e não nos seus preços.

As origens da TU remontam ao ano de 1738, quando o matemático Daniel Bernoulli, ao observar as chances em jogos de azar, verificou que as pessoas tipicamente não maximizam a expectativa monetária de valor, quando, por exemplo, abrem mão de apostar para uma chance de ganhar 10 unidades monetárias incertas a uma certeza de ganhar 5 unidades monetárias (OKASHA, 2016).

Um aspecto relevante da TU diz respeito à ordenação numérica das preferências, às quais de acordo com (STEELE; STEFÁNSSON, 2015) são representadas como "funções da utilidade", e se dividem em ordinais e cardinais. A função ordinal parte do princípio de que, em um conjunto finito de opções, há uma relação de força/fraqueza entre elas neste conjunto que possibilita a sua ordenação, de acordo com percepção do agente. Por sua vez, a função cardinal considera a distância de desejabilidade entre as opções, o que requer uma melhor estruturação em comparação à utilidade ordinal.

Existem vários atributos inerentes às alternativas disponíveis em um processo de tomada de decisão, e nesse contexto as funções de utilidade podem ser usadas para indicar a alternativa mais aderente com as preferências do agente decisor (DIEHL, 1997). Ainda segundo o referido autor, as funções de utilidade atribuem um valor entre 0 e 1 para mensurar a importância de um atributo, tendo por base a perspectiva do agente decisor, sendo que atributos com importância mais próximos de 0 corresponderão ao pior resultado, ao passo que os atributos de valor mais elevado corresponderão ao melhor resultado. Desta forma, complementa Diehl (1997) que é possível indicar a alternativa mais compatível com os anseios do agente decisor, por meio da ponderação das alternativas aplicando estas funções. O modelo geral para as funções de utilidade é expresso por:

$$U(X) = U(X_1, X_2, \dots, X_n) \quad (2.1)$$

em que $U(X)$ é a função utilidade para multiatributos, de acordo com os modelos aditivo ou multiplicativo e X_1, X_2, \dots, X_n são os resultados das funções utilidade para os atributos $X_{(1-n)}$.

A partir da ideia central representada pelas funções de utilidade, são derivadas equações e modelos para buscar tangibilizar as opções disponíveis ao tomador de decisão.

Não obstante este ser um modelo importante para análise de alternativas sob o ponto-de-vista do agente tomador de decisão, Diehl (1997, p. 65) conclui que "os modelos de utilidade para multiatributos não podem ser vistos somente como métodos de decisão, mas principalmente de apoio à decisão", dadas as limitações deste modelo, principalmente no que diz respeito à provável interdependência dos resultados dos atributos e as funções de utilidade.

2.1.2 Teoria dos Dois Fatores

Também conhecida como teoria de motivação-higiene, a teoria dos dois fatores foi concebida a partir da investigação de Herzberg sobre os fatores que causavam satisfação e insatisfação no trabalho.

Em suas observações, Herzberg concluiu que os fatores extrínsecos ao indivíduo, tais como: conteúdo e natureza das tarefas, descrição do cargo, ambiente e condições de trabalho, salário, constituem-se como fatores que apenas previnem a insatisfação, no que foi denominado como fatores higiênicos, ou de manutenção (PILATTI, 2004).

A resultante em relação ao atendimento ou não destes fatores é a ausência de insatisfação, ou a insatisfação, respectivamente. Por outro lado, Herzberg atribuiu aos fatores motivacionais os aspectos intrínsecos do indivíduo ligados ao trabalho e a

necessidade de crescimento profissional e aprimoramento pessoal.

Estes fatores estão relacionados, portanto, às condições de possibilidade de crescimento no trabalho, satisfação do indivíduo com seu trabalho, reconhecimento e realização. No caso dos fatores motivacionais, a resultante de acordo com o atendimento ou não destes fatores irá gerar satisfação ou ausência de satisfação, respectivamente.

O Quadro 1 resume os fatores de motivação e os fatores de higiene, e o Quadro 2 demonstra os efeitos na satisfação dos indivíduos de acordo com cada fator de motivação (ALSHMEMRI; SHAHWAN-AKL; MAUDE, 2017).

Quadro 1: Resumo dos fatores na teoria de Herzberg

Fatores de Motivação	Fatores de Higiene
Status / posição	Relacionamento interpessoal
Conteúdo do trabalho	Salário
Possibilidade de crescimento	Políticas e administração
Responsabilidade	Supervisão
Reconhecimento	Condições de trabalho
Realizações	

Fonte: adaptado de Alshmemri, Shahwan-Akl e Maude (2017)

Quadro 2: Comparação entre os dois fatores da Teoria de Herzberg

	Fatores de Motivação	Fatores de Higiene
Ausência	Não satisfação	Insatisfação
Presença	Satisfação	Não insatisfação
Descrição de Herzberg	Intrínseco ao trabalho	Extrínseco ao trabalho
Importância na satisfação	Alta	Baixa

Fonte: adaptado de Alshmemri, Shahwan-Akl e Maude (2017)

2.1.3 Teoria da Perspectiva

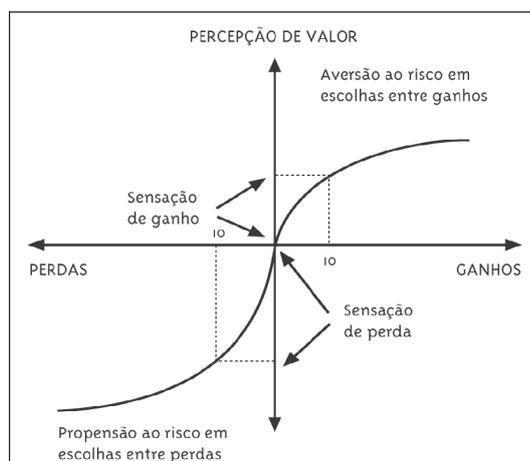
A Teoria da Perspectiva foi formulada por Kahneman e Tversky (1979), a partir de um artigo emblemático que teve grande influência na ciência comportamental. O experimento conduzido pelos autores testou escolhas financeiras sob situações de risco e incerteza, e concluiu que tais julgamentos se desviaram significativamente dos pressupostos da teoria da utilidade esperada, o que causou um grande impacto científico.

Pode se afirmar que este estudo trouxe uma clara crítica alusiva à teoria da utilidade esperada, tendo como uma de suas premissas a incerteza e o risco na tomada de decisão (KROHLING; SOUZA, 2012). Recentemente em 2019, Ruggeri et al. (2020) replicaram os métodos da pesquisa original de Kahneman e Tversky em um estudo multinacional, apenas ajustando as bases monetárias e valores presentes, e obtiveram resultados praticamente idênticos ao estudo original, concluindo que "as bases empíricas para a teoria da perspectiva se replicam além de quaisquer limites imagináveis"(RUGGERI et al., 2020, p. 1).

A aversão ao risco faz com que os indivíduos em geral tenham maior sensibilidade às perdas do que aos possíveis ganhos em um processo de tomada de decisão, e este é o ponto chave da crítica de Kahneman e Tversky ao modelo de utilidade esperada (ARNOTT; GAO, 2019). Portanto, pode se afirmar que a teoria da perspectiva é um modelo descritivo de uma tomada de decisão feita sob condições de risco (KROHLING; SOUZA, 2012), não obstante haver uma dificuldade na mensuração da aversão ao risco, eis que faz-se necessária uma medição simultânea da utilidade tanto para os ganhos quanto para as perdas (ARNOTT; GAO, 2019).

A teoria da perspectiva busca, por meio de referenciais comparativos, demonstrar os resultados expressos sob a forma de ganhos e perdas com base no parâmetro de referência, sendo que Kahneman e Tversky desenvolveram um modelo matemático para representar tais resultados. Os referidos autores postularam que a diferenciação da utilidade frente aos ganhos e perdas pode ser representada graficamente em um formato de "S"(Figura 1), onde os ganhos são expressos em uma forma côncava e as perdas em uma forma convexa, sendo que a curva de utilidade para perdas é mais íngreme em relação aos ganhos, evidenciando uma aversão ao risco. De acordo com Arnott e Gao (2019), há diversas evidências empíricas que vão ao encontro desta perspectiva em relação à curva côncava dos ganhos, porém, o mesmo não se aplica em relação às perdas, onde evidências empíricas retratam um cenário mais misto, com estudos apontando para um formato linear ou até côncavo, ainda que predominem as evidências de convexidade da curva.

Figura 1: Função de valor na teoria da prospecção.



Fonte: Krohling e Souza (2012)

2.1.4 Escolha Intertemporal

As escolhas intertemporais se caracterizam por decisões tomadas cujos custos e benefícios poderão ter seus efeitos em diferentes momentos no tempo (READ; MCDONALD; HE, 2004). Esta definição é bastante abrangente, e considera que o agente pode escolher contrair um custo de forma antecipada no tempo ao benefício, ou vice-versa, tal como optar por se vacinar contra alguma doença, contraindo um "custo" inicial para posterior benefício, ou aderir a um plano de previdência ao invés de gastar o dinheiro cujo benefício seja imediato.

Tomando por base o conceito exposto, pode-se elencar dois tipos de escolhas intertemporais:

- Posição credora: é a situação em que o agente aceita um custo inicial, para auferir um benefício futuro. Exemplo: plano de previdência privada, contratação de seguro etc.
- Posição devedora: é a situação em que o agente usufrui de um benefício no momento inicial, empenhando, para tanto, um custo futuro. Exemplo: compra de um bem de consumo a prazo.

Segundo Read, McDonald e He (2004), a análise convencional da escolha intertemporal considera que o efeito do atraso no valor subjetivo dos resultados futuros pode ser capturado por uma função de desconto, que desempenha o mesmo papel na escolha intertemporal que a função de ponderação de probabilidade na escolha de risco. Este efeito, caracterizado por um desconto na utilidade esperada, foi alvo de proposição de vários modelos com base no foco presente para descrever como as escolhas intertemporais são feitas pelos humanos (ERICSON; LAIBSON, 2019).

A abordagem das escolhas intertemporais é relevante no *framework* da presente

pesquisa, uma vez que o objeto de estudo envolve custos e benefícios intertemporais inerentes ao serviço contratado e às escolhas das IES por parte dos alunos de graduação em Ciências Contábeis.

2.2 Mercado

Esta seção abordará os aspectos teóricos referente ao mercado, os conceitos de eficiência e ineficiência, estrutura, e especial destaque ao mercado de ensino superior no Brasil.

2.2.1 Eficiência e Ineficiência

A eficiência e ineficiência de mercado é um conceito originado a partir da construção da Hipótese do Mercado Eficiente (HME) por Eugene F. Fama, em artigo publicado em 1970 sob o título: *Efficient capital markets: A review of theory and empirical work*, no qual definiu que um mercado é denominado "eficiente" se os preços sempre refletirem fielmente as informações disponíveis. Em outras palavras, A eficiência diz respeito à capacidade de precificação dos ativos no mercado face ao surgimento de novos fatos, e que, em um mercado ativo, onde há muitos investidores bem informados, os preços dos ativos serão apropriados e refletirão todas as informações disponíveis.

De acordo com Malkiel (2003), a HME está associada ao conceito de passeio aleatório (*random walk*), que em finanças caracteriza-se pela definição de que os preços sofrem mudanças subsequentes por desvios aleatórios em comparação com os preços anteriores. Esta lógica assume que, se o fluxo de informações for desimpedido, tendo o seu reflexo imediatamente precificado, então as mudanças nos dias posteriores refletirão tão somente as notícias destes dias, independente das mudanças do preço de hoje.

A HME possui três formas de caracterização, conforme o grau de disponibilidade das informações no mercado: eficiência fraca, semi-forte e forte (CLARKE; JANDIK; MANDELKER, 2001). Na hipótese fraca de eficiência de mercado, os preços dos ativos refletem as informações do passado (demonstrações financeiras), sendo que os preços variam sempre que alguma informação nova é divulgada. Na hipótese semi-forte, os preços dos ativos incorporam todas as informações disponíveis publicamente, sendo alterado logo após a divulgação de novas notícias. Por fim, na hipótese forte, o preço passa a refletir todas as informações do ativo, sejam elas de caráter público ou não.

2.2.2 Estrutura

Os mercados caracterizam-se pelos ambientes em que atuam as forças de oferta e demanda de bens e serviços, por meio dos agentes compradores e vendedores. Neste sentido, o entendimento da estrutura de mercado da educação superior no Brasil é relevante para a presente pesquisa, eis que o movimento migratório dos alunos impacta de forma relevante e direta sobre a demanda, o que afeta o equilíbrio das forças neste mercado.

Os elementos principais para a definição do tipo de estrutura dos mercados passam pelo entendimento das suas características quanto à quantidade de vendedores/ofertantes, a quantidade de demandantes, o tipo de produto ou serviço comercializado, e o acesso à informação. O número de agentes envolvidos nos lados de oferta e demanda é a base das classificações mais simples de estruturas de mercado (ROSSETTI, 2016), e é assim que Heinrich Stackelberg propôs uma matriz clássica de estrutura de mercado, baseada apenas em um elemento de diferenciação: o número dos agentes de mercado, tanto na oferta quanto na demanda, conforme ilustrado no Quadro 3.

De acordo com ROSSETTI (2016), os oligopólios podem ser divididos em quatro tipos, conforme a sua característica: oligopólio concentrado, oligopólio diferenciado, oligopólio diferenciado-concentrado e oligopólio competitivo. Em um oligopólio concentrado, há a ausência de diferenciação dos produtos e altas barreiras de entrada para novos competidores. Em um oligopólio diferenciado, há espaço para a diferenciação do produto/serviço, que normalmente é associada à segmentação de clientes, seja por nível de renda, padrão de consumo, hábitos e aspirações sociais. Por sua vez, um oligopólio diferenciado-concentrado é resultante da combinação dos dois primeiros tipos de oligopólio. Por fim, o oligopólio competitivo caracteriza-se pela alta concentração da produção e baixas barreiras de entrada para novos competidores,

Quadro 3: Classificação de Marchal: as estruturas de mercado diferenciadas também por fatores associados ao comportamento dos agentes a natureza dos produtos.

Oferta Procura	Um só vendedor	Pequeno número de vendedores	Grande número de vendedores
Um só comprador	Monopólio bilateral	Quase monopsônio	Monopsônio
Pequeno número de compradores	Quase monopólio	Oligopólio bilateral	Oligopsônio
Grande número de compradores	Monopólio	Oligopólio	Concorr. perfeita

Fonte: adaptado pelo autor (ROSSETTI, 2016, p. 417)

que competem normalmente por preços, com objetivo de ampliar a participação de mercado das empresas mais bem posicionadas.

O forte cenário de expansão das IES privadas, que caracterizou o mercado de ensino superior nas últimas décadas no Brasil, favoreceu a formação de uma estrutura econômica típica de um oligopólio diferenciado. Neste sentido, corroboram Marques, Ximenes e Ugino (2018), ao afirmarem que no início do século XXI, o governo federal fomentou o setor privado por meio de condições inéditas para ampliação da oferta e lucratividade, o que acabou resultando na formação de um oligopólio.

Nesta mesma linha, Chaves, Reis e Guimarães (2018) postulam que a valorização global do capital no contexto da expansão do mercado de educação superior favoreceu a entrada de grandes grupos de capital aberto e fechado neste mercado, e que "estudos realizados sobre esse processo de mercantilização da educação superior no Brasil [...] indicam, como resultado da compra/fusão de IES, que a tendência é de formação de oligopólios"(CHAVES; REIS; GUIMARÃES, 2018, p.7). É farto o número de pesquisas e artigos que classificam o setor de ensino como um oligopólio, em função da dinâmica do cenário da educação brasileira a partir do final da década de 90, e mais acentuadamente, a partir de 2007. Conforme mencionam em seu artigo Diniz, Oliveira e Lima (2021), houve, especialmente a partir de 2007, um processo de consolidação de pequenas IES privadas que foram compradas ou se viram forçadas a sair do mercado, tendo em vista o movimento de grandes conglomerados empresariais que passaram a explorar mercantilmente a educação, favorecendo o surgimento de um oligopólio nesse mercado. A financeirização de empresas via abertura de capital na bolsa de valores também é um elemento catalisador desse processo. Com isso, houve espaço para a redução de custos pelo aproveitamento de sinergias, e por conseguinte, uma redução de preços e aumento da oferta de cursos, o que facilitou o acesso de uma maior parcela da população.

2.2.3 Ensino Superior

O ensino superior no Brasil sofreu profundas transformações ao longo das últimas três décadas. O final da década de 1980 e início da década de 1990 foi marcado por mudanças na política nacional, as quais foram ao encontro de uma maior abertura para a iniciativa privada em todos os setores da economia. Estas mudanças acabaram por impactar, de forma relevante, tanto na demanda por serviços de educação, quanto na oferta por parte de instituições privadas de ensino (SAMPAIO, 2013).

Seguindo uma tendência mundial, simbolizada pela abertura e intensificação das relações econômicas internacionais em um cenário de globalização, tais mudanças culminaram em um contexto que propiciou a ampliação da participação de instituições privadas de ensino superior no cenário educacional no Brasil (SAMPAIO, 2013).

Especialmente a partir da implementação da LDB, promulgada em 1996, a qual se estabeleceu como um novo marco constitucional para a educação brasileira em todos os níveis, o cenário da educação no Brasil passou por importantes modificações que propiciaram a ampliação da oferta de IES. Isto se deve ao fato de que na antiga legislação as instituições obrigatoriamente deveriam ser constituídas sem fins lucrativos, e a partir da promulgação da nova LDB, houve a permissão para que as IES passassem a ser constituídas sob qualquer forma de sociedade civil, inclusive com finalidade lucrativa.

Na esteira das mudanças na legislação, houve um aumento expressivo na oferta de educação superior, inicialmente pela atribuição de autonomia e competência conferida às universidades, e posteriormente pelos movimentos de fusões e aquisições. Após o advento da LDB, conforme postula Sampaio (2013), as IES passaram a ter autonomia para criar e extinguir cursos na própria sede, bem como remanejar o número de vagas nos cursos ofertados. Essa prerrogativa proporcionou às instituições privadas a se adequarem com mais agilidade às demandas de massa por ensino superior. Também segundo Corbucci, Kubota e Meira (2016), na década de 1990 foi observada uma expansão horizontal do setor privado no mercado de ensino superior, caracterizada pelo aumento do número de instituições de pequeno e médio porte. Na primeira década deste século, iniciou-se um movimento de concentração das instituições privadas por meio de operações de fusões e aquisições. Segundo PEREIRA e BRITO (2014), estes movimentos se intensificaram a partir de 2005, com a compra do controle acionário da rede Anhembi-Morumbi pelo grupo americano Laureate, no ano de 2005. Desde então houve um incremento significativo no processo de mercantilização no setor, oportunizando a entrada de fundos de investimentos que passaram a investir no mercado de educação superior, amparados na abertura dada pela legislação brasileira ao capital estrangeiro.

Em outra frente, as transformações oriundas de novas tecnologias, bem como a ampliação da conectividade a partir de um processo de abertura da internet para fins comerciais (1995) e formatos de trabalho em constante evolução, também contribuíram para o surgimento de novas IES, cursos e modalidades de ensino, principalmente após a regulamentação dos cursos na modalidade de Ensino à Distância (EaD) promovida pela Decreto nº 5.622/2005, posteriormente substituído pelo Decreto nº 9.057/2017.

Esta nova realidade culminou na ampliação da oferta de cursos e de IES, se tornando um elemento importante a ser levado em consideração em qualquer pesquisa que aborde a evasão e retenção de alunos de graduação no Brasil, visto que a maior oferta de cursos e instituições, entre outros fatores, potencializa o poder de escolha dos estudantes de graduação, ampliando a gama de alternativas de acordo com os interesses e conveniências individuais (preço, modalidade, forma de ingresso, curso etc.).

Os números analisados e demonstrados na Tabela 1, referentes à oferta de va-

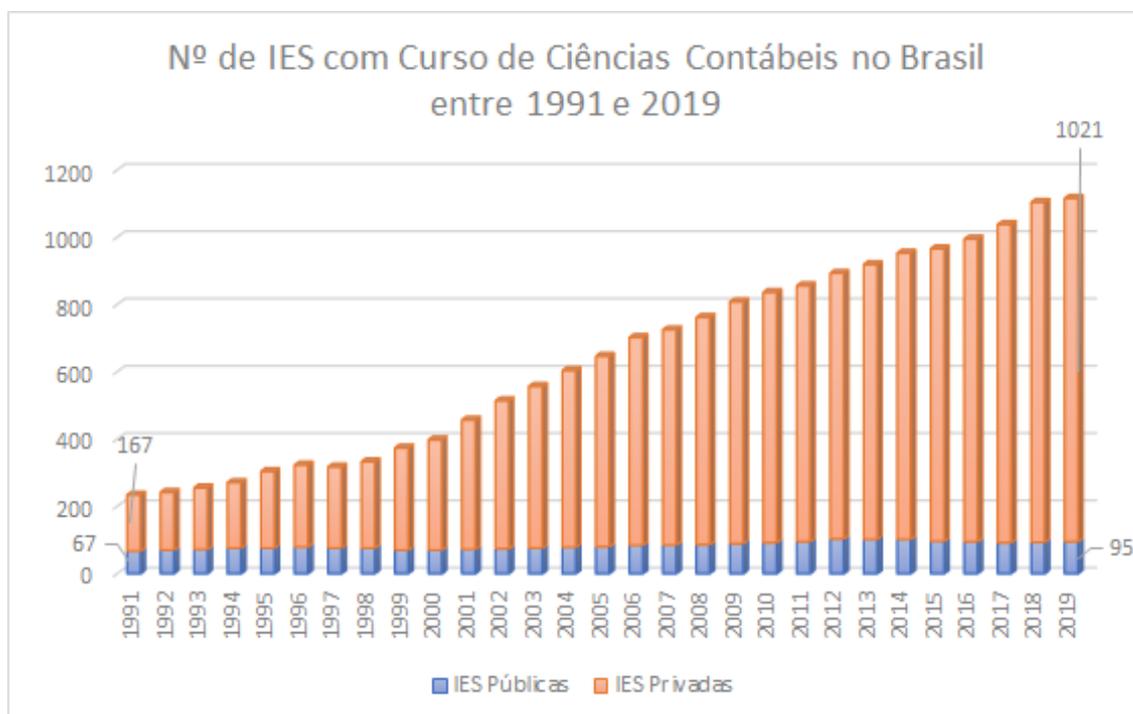
gas, número de instituições e alunos matriculados, bem como a população brasileira ao confrontar os cenários de 1995 e 2018, corroboram a descrição do atual contexto. Segundo dados do INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira) (2019) e IBGE (IBGE - Instituto Brasileiro de Geografia e Estatística), entre 1995 e 2018 houve um crescimento de 184% no número de instituições, sendo 42% de crescimento nas IES públicas e 227% nas IES privadas; já em relação ao número de matriculados, o crescimento total foi de 380%, e o número de vagas houve um salto de 2.117%, sendo 2.837% na rede privada, e 369% na rede pública. Em contrapartida, a estimativa de crescimento populacional no mesmo período foi de 29,6%, segundo dados estimados dos referidos anos, com base nos censos populacionais do Brasil. Na graduação em Ciências Contábeis, o comportamento observado é refletido também nas IES que oferecem este curso de graduação. A Figura 2 demonstra que entre 1991 e 2019 houve um crescimento de 377% no número de instituições ofertantes do curso de graduação em Ciências Contábeis (em 1991 eram 234, e em 2019 era de 1.116), sendo que quando ao ser estratificado em instituições públicas ou particulares, fica evidente o salto no número de instituições privadas no período analisado: 511% de crescimento, contra 42% na rede pública. Estes números corroboram o esforço do país na busca da ampliação do acesso e inclusão dos seus habitantes na educação superior, bem como denotam um efetivo aumento nesta inserção da população brasileira em cursos de nível superior, ainda que o país esteja aquém de um nível ideal de inclusão.

Tabela 1: Nº de IES, matriculados, concluintes, vagas e inscritos - 1995 e 2018

Dados	1995	2018	Var. %	Nº p/ 100 mil hab.	
				1995	2018
Instituições	894	2.537	184%	0,55	1,21
Públicas	210	299	42%	0,13	0,14
Particulares	684	2.238	227%	0,42	1,07
Matriculados	1.759.703	8.450.755	380%	1.086	4.024
Públicas	631.314	2.077.481	229%	390	989
Particulares	1.128.389	6.373.274	465%	697	3.035
Concluintes	245.887	1.264.288	414%	152	602
Públicas	87.862	259.302	195%	54	123
Particulares	158.025	1.004.986	536%	98	479
Vagas oferecidas	610.355	13.529.101	2.117%	377	6.442
Públicas	178.145	835.569	369%	110	398
Particulares	432.210	12.693.532	2.837%	267	6.045
Inscritos 1ª opção	2.653.853	17.213.064	549%	1.638	8.197
Públicas	1.399.092	6.596.808	372%	864	3.141
Particulares	1.254.761	10.616.256	746%	775	5.055

Fonte: base de dados do INEP (2021)

Figura 2: Nº de Instituições de Ensino Superior que possuem curso de graduação em Ciências Contábeis.



Fonte: dados do INEP.

O importante crescimento do mercado de ensino superior no Brasil observado nas últimas décadas, tanto na demanda de alunos de graduação quanto na oferta e disponibilidade de cursos e instituições, aqueceu a dinâmica competitiva deste setor, ampliando a necessidade dos gestores disporem de ferramentas gerenciais que lhes proporcionem informações, não só a respeito do seu desempenho econômico-financeiro propriamente dito (representado nas demonstrações financeiras), como também informações gerenciais referentes à dinâmica do mercado de atuação.

2.3 Mercado como Redes

O mercado como redes pode ser explicado a partir da relação tanto de cooperação quanto de competição como forças motrizes do mercado com as empresas, envolvendo a troca de informações entre as partes e pela importância do contexto da transação (SALMAN, 2020).

O avanço tecnológico das últimas décadas transformou as relações sócio-econômicas, facilitando as transações nos mercados por meio da convergência de plataformas tecnológicas (CORDEIRO; VOLDNES, 2021).

2.3.1 Redes Complexas

Nas últimas décadas, as redes complexas surgiram como um novo movimento na ciência no campo multidisciplinar, possibilitando a modelagem por meio de uma estrutura de rede para os mais variados fenômenos de qualquer campo de conhecimento (JALILI; PERC, 2017). Esta também é a visão de Varela et al. (2015), ao contextualizar que o surgimento das redes complexas, com suas características estruturais específicas, foram introduzidas a partir do final da década de 1990.

Conforme definem Cimini et al. (2019), uma rede representa a modelagem de uma vasta gama de sistemas, como sociais, econômicos, tecnológicos e biológicos, em uma representação gráfica de um conjunto de entidades (denominadas vértices ou nós, que interagem por meio de conexões (*links* ou arestas). Segundo estes autores, as interações representam os graus de liberdade da rede e podem ser de diferentes tipos: direcionadas ou não-direcionadas e binárias ou com graus de peso, dependendo da natureza do sistema e a resolução usada para descrevê-la.

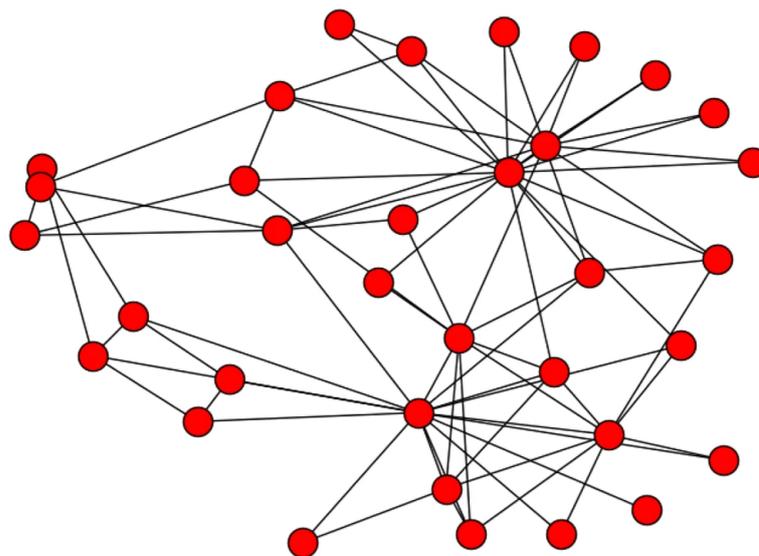
De acordo com Mata (2020), qualquer mecanismo funcional, seja de ordem social, biológica ou tecnológica, é passível de ser representado sob a forma de uma topologia de rede. Por exemplo, o corpo humano e seus diversos sistemas, a sociedade e os relacionamentos interpessoais, os conglomerados de empresas etc., são sistemas que podem ser representados sob a forma de rede. Estes sistemas são considerados complexos, uma vez que não é possível prever o comportamento coletivo a partir dos componentes individuais. Desta forma, segundo o referido autor, a análise de redes complexas permite o entendimento do funcionamento de diversos sistemas reais, e, apesar de não ser possível a previsão do comportamento coletivo a partir do individual, o entendimento da descrição matemática destes sistemas possibilita uma maior capacidade de previsão e de controle dos fenômenos associados. Para Metz et al. (2007), as redes complexas receberam esta denominação a partir do avanço da tecnologia computacional e da disponibilidade de poder de processamento. Estes avanços possibilitaram a análise de dados em grandes quantidades, e as pesquisas, que outrora tinham seu foco na análise de pequenas redes de vértices individuais, passaram a considerar propriedades estatísticas em larga escala.

No âmbito da análise de redes complexas em modelos sócio-econômicos, Varela et al. (2015) defendem que a aplicação dos métodos da física de sistemas complexos permitiu a obtenção de um novo olhar sobre os domínios econômicos, a partir do consenso crescente na comunidade científica de que uma rede pode ser representada pela visão da economia como um sistema complexo em evolução, baseado em agentes.

Ainda que haja diferenças entre os mais diversos sistemas complexos existentes na sociedade, estes acabam se tornando bastante similares, uma vez que utilizam

as mesmas regras e princípios ao terem suas estruturas representadas sob a forma de rede. Isto então permite a utilização do mesmo conjunto de ferramentas de computação e matemática na sua análise (MATA, 2020). A Figura 3 ilustra a estrutura de uma rede clássica conhecida como o "Zachary's Karate club". Os nós na Figura representam as pessoas, enquanto que as arestas representam a interação entre elas.

Figura 3: Representação das interações sociais de pessoas de um mesmo clube.



Fonte: Mata (2020, p. 2)

2.3.1.1 Propriedades das redes complexas

De acordo com Metz et al. (2007), as redes complexas possuem algumas propriedades que as caracterizam e as distinguem de redes simples. Dentre as principais propriedades, os autores destacam:

- Coeficiente de aglomeração: também denominado fenômeno de transitividade, ocorre quando um nó A está ligado a um nó B, que por sua vez está ligado a um nó C, o que aumenta a probabilidade do nó A estar conectado ao nó C. A transitividade irá apontar para uma elevada presença de triângulos em uma rede;
- Distribuição de graus: o número de arestas possíveis em um vértice é representado pelo seu grau, e a distribuição de graus é uma função de distribuição probabilística que indica a probabilidade de um determinado vértice ter grau fixo;
- Resistência: é a capacidade de resistência da rede em função da remoção de alguns vértices, sem que afete sua funcionalidade;

- Misturas de padrões: os vértices em uma rede complexa guardam características diferentes entre si, o que afeta a probabilidade de conexão entre os mesmos;
- Correlação de graus: também conhecido como assortatividade, indica a relação entre os nós em uma rede;
- Intermediação: conhecida também como *betweenness*, é uma medida de centralidade que considera que um nó é mais importante à medida em que é mais acessado pelo caminho mais curto na rede (AMANCIO, 2015).

2.3.1.2 Tipos de redes complexas

As redes complexas possuem uma classificação de acordo com o seu coeficiente de agregação. Conforme pontuam Jalili e Perc (2017) e Mata (2020), as redes extraídas do mundo real evidenciam o compartilhamento de algumas propriedades estruturais que permitem sua classificação em diversos tipos de modelos, tais como: redes aleatórias regulares, redes pequeno-mundo, redes livres de escala, grades regulares, redes de crescimento e conexão preferencial, redes aleatórias não correlacionadas, redes temporais, entre outras. As características dos principais tipos de redes complexas, conforme Metz et al. (2007) e Mata (2020), são:

- Redes aleatórias: modelo proposto pelos matemáticos Paul Erdős e Alfréd Rényi em 1958, em que são adicionadas arestas não direcionadas de forma aleatória entre um número fixo de vértices N . Cada aresta é representada de forma independente com base em uma probabilidade de conexão p . O modelo gera grafos aleatórios com n vértices e k arestas, onde inicialmente os vértices desconectados são conectados de forma aleatória até que o número de arestas do grafo seja igual a k .
- Redes pequeno-mundo: modelo proposto em 1998 por Watts e Strogatz, onde postulam que, em uma rede muito grande e altamente conectada, a distância média entre dois vértices não ultrapassa um número muito elevado de vértices ao se adicionarem algumas conexões aleatórias. Exemplos de redes com essas características estão presentes em conexões aeroportuárias, plantas de distribuição de energia, redes sociais, entre outras.
- Redes livres de escala: modelo proposto em 1999 por Albert-László Barabási, onde a principal característica consiste na tendência de que um novo vértice na rede se conecte a um vértice que possua um alto grau de conexões, formando assim conglomerados de vértices denominados de *hubs*.

2.4 Estudos Relacionados

Nesta seção serão apresentados alguns estudos anteriores que abordaram a migração e evasão de alunos no ensino superior e que estão, de alguma forma, relacionados ao tema da presente pesquisa.

Hoed (2016) pesquisou sobre a evasão em cursos de nível superior na área de computação, obtendo evidências de que a evasão é maior nestes cursos quanto menor a relação de candidatos por vagas, bem como a evasão é maior quanto mais elevado for o grau de complexidade em conhecimentos matemáticos e de abstração algorítmica. Também evidenciou que o sexo, forma de ingresso na instituição e condição de cotista também foram fatores relacionados à evasão nestes cursos.

Por sua vez, Tontini e Walter (2014) propuseram um método para identificação do risco de evasão de alunos em cursos de nível superior. Para tanto, basearam sua proposta em uma determinada IES, sendo que aplicaram um questionário estruturado e definiram seu modelo por meio de redes neurais artificiais e análise de *cluster*. A aplicação do modelo foi feita na IES pesquisada, e os resultados das taxas de evasão após aplicação prática do modelo proposto, foram comparados a outras IES, observando uma queda de 18% no índice de evasão da IES pesquisada. Concluíram que a dimensão com maior influência na retenção ou evasão foi a colocação profissional e vocação do aluno, seguida pela disponibilidade de tempo para estudos e fatores da vida pessoal.

Outro artigo que pesquisou o fenômeno migratório de alunos na educação superior foi elaborado por Furtado e Alves (2012). Neste artigo, os autores analisaram os fatores determinantes da evasão universitária na Universidade do Vale do Rio dos Sinos (UNISINOS), com base nos alunos ingressantes entre 2005 e 2008. Avaliaram a relação entre variáveis socioeconômicas e a evasão dos alunos por meio da aplicação de regressão logística, e concluíram que as variáveis mais relevantes na explicação da evasão dos alunos foram as disciplinas canceladas, média de graus, porcentual de disciplinas concluída, entre outras. Destacaram também que o estudo poderia ser melhorado se conseguissem levar em consideração o recebimento de auxílio ou não por parte dos alunos, uma vez que os dados analisados não continham estas informações, bem como sugeriram a replicação do estudo em outras IES. De forma similar, Silva (2013) pesquisou sobre a evasão no ensino superior de uma IES entre os anos de 2006 a 2009, baseando sua pesquisa em um modelo de *duration* e concluindo seus achados amplamente sobre fatores que recaem sobre as escolhas ou condições dos indivíduos, tais como: idade, sexo, participação no PROUNI, renda, entre outros.

Anteriormente, Silva Filho et al. (2007) estudaram a evasão no ensino superior por meio de um estudo com base nos dados do Censo da Educação Superior, entretanto com uma clara limitação das bases de dados públicas da época, os quais ainda

não continham o detalhamento individualizado dos alunos¹. A limitação dos dados impôs uma simplificação na modelagem proposta pelos autores, sendo apresentados resultados referentes à taxas agregadas de evasão nas IES, concluindo também que quanto maior o número de candidatos por curso, menor a evasão.

O tema também é abordado em publicações internacionais, tal como a pesquisa de Martins et al. (2020) que estudou a evasão dos alunos de licenciatura do Instituto Politécnico de Bragança (IPB) entre os anos de 2007 e 2008, e 2015 e 2016, abordando o conjunto de alunos que iniciaram e finalizaram suas jornadas neste período. Utilizando a mineração dos dados como ferramenta, propuseram a criação de um modelo de previsão da evasão, baseado em fatores estritamente ligados às condições dos indivíduos, tais como fatores socioeconômicos, incentivos, acesso e demográficos. Concluíram que o desempenho curricular do aluno nos primeiros semestres é um fator relevante para determinar a evasão. Também neste estudo não foi destacado nenhum fator que envolvesse a concorrência entre as IES. De maneira similar, a pesquisa de Lourenço e Sá (2019) aplicou um modelo gravitacional para estudar o fluxo de estudantes universitários em Portugal, com uma abordagem acerca dos fatores geográficos envolvendo os alunos e as instituições. Concluíram que os candidatos ao ensino superior não se distribuem de forma homogênea no espaço, pois tendem a se concentrar nas maiores cidades situadas na costa, e, conseqüentemente, este efeito provoca uma baixa na demanda nas IES localizadas fora das grandes cidades.

Por sua vez, Sa, Florax e Rietveld (2004) pesquisaram o assunto no mercado de ensino superior da Holanda, e propuseram um modelo gravitacional da mobilidade estudantil em perspectiva, onde focaram no aspecto geográfico das IES na determinação da sua relevância em relação ao processo de tomada de decisão dos alunos. Um aspecto importante na pesquisa foi o fato de os autores proporem um modelo para atribuir um índice de competitividade entre as IES, baseado na razão entre os aspectos de atratividade e a distância entre as instituições de origem e destino. Entre outros resultados, destacaram principalmente que as escolhas dos futuros estudantes universitários são guiadas principalmente por motivos de consumo, enquanto as razões de investimento não são predominantes, porque a qualidade dos programas educacionais naquele país não desempenha um papel significativo nos comportamentos de escolha.

Diversos outros artigos e pesquisas envolvendo a migração de alunos no ensino superior foram encontrados após a revisão de literatura nesta temática, entretanto constatou-se que a ampla maioria dos achados baseou suas análises sob uma ótica mais isolada e restrita a apenas aspectos relacionados aos cursos ou determinados fatores individuais dos alunos. Não foram encontradas análises mais aprofundadas que abordassem a demanda de alunos também sob a perspectiva de fatores competi-

¹O detalhamento dos dados ao nível individual dos alunos somente foi adotado pelo INEP a partir do CES do ano de 2009

tivos entre as IES. Neste sentido, a presente pesquisa busca explorar estas prováveis lacunas e apresentar uma resposta científica englobando tanto os fatores envolvendo as escolhas individuais dos alunos, quanto a concorrência entre as IES.

2.5 Desenvolvimento das Hipóteses

Conforme abordado anteriormente, as pesquisas existentes sobre o tema exploram largamente os aspectos ligados aos indivíduos enquanto agentes na relação aluno-IES (HOED, 2016; TONTINI; WALTER, 2014; FURTADO; ALVES, 2012; SILVA FILHO et al., 2007; SILVA, 2013; MARTINS et al., 2020), sendo que apenas uma das pesquisas encontradas propõe, de forma simplificada, uma medida para a concorrência entre as IES.

Pesquisadores têm buscado criar modelos de predição para a evasão no ensino superior, aqui também entendida como a migração dos alunos entre as IES, onde relacionam, invariavelmente, fatores relacionados às escolhas dos alunos, em que se baseiam em variáveis explicativas inerentes às suas condições socioeconômicas, curriculares, renda, fatores geográficos, entre outras. Não obstante, segundo eles, existe uma clara relação entre as escolhas dos alunos e a demanda por vagas nas IES. A busca de uma resposta para os movimentos migratórios dos alunos passa, inexoravelmente pela via dos ofertantes, representada pelos fatores de concorrência entre as IES. No sentido de obter o conjunto de elementos necessários para explorar o papel da decisão dos alunos na transferência interinstitucional, formulou-se as seguintes hipóteses:

H1: A retenção dos discentes no curso depende de benefícios geo-socioeconômicos

Espera-se haver uma relação positiva dos benefícios geo-socioeconômicos (auxílio econômico, fatores sociais e distância entre instituição de ensino e cidade de origem do aluno) e a retenção dos discentes.

H2: A migração interinstitucional dos estudantes é potencializada por uma concorrência oligopolista diferenciada

Espera-se haver uma relação positiva entre o número de concorrentes e o fluxo migratório de alunos, potencializando sua decisão de transferência.

O conjunto formulado de hipóteses remete à expectativa de que haja uma relação que mutuamente potencialize a utilidade como fator de tomada de decisão dos alunos, no que tange à migração interinstitucional. Nesse sentido, a utilidade esperada consiste na combinação dos fatores geo-socioeconômicos dos alunos, com a concorrência entre as IES.

3 METODOLOGIA

Este capítulo aborda os procedimentos metodológicos utilizados neste estudo, com a finalidade de responder ao questionamento da pesquisa, bem como satisfazer os objetivos geral e específicos. Primeiramente é feito o detalhamento sobre o delineamento da pesquisa, e, em seguida, são abordadas as técnicas de coleta, tratamento e análise dos dados.

3.1 Delineamento

Conforme Gil et al. (2017), as pesquisas podem seguir múltiplos sistemas de classificação, sendo definidas conforme a área de conhecimento, a finalidade, o nível de explicação e os métodos adotados.

A presente pesquisa classifica-se, quanto à finalidade, como aplicada, uma vez que irá propor a solução de um problema específico, gerando conhecimentos para aplicação prática (GIL et al., 2017).

Quanto à forma de abordagem, classifica-se como quantitativa e qualitativa, uma vez que a investigação aborcará tanto a questão factual, quanto fenomênica. Conforme postula Kauark, Manhães e Medeiros (2010, p. 27), uma pesquisa fenomênica "prevê a coleta de dados a partir de interações sociais e sua análise a partir da hermenêutica do pesquisador".

Quanto aos objetivos, a pesquisa é explicativa, tendo em vista que a investigação irá buscar a identificação de fatores que expliquem a movimentação de alunos de IES em termos de evasão e migração de cursos de graduação em Ciências Contábeis no Brasil.

Quanto ao método, classifica-se como uma pesquisa documental, uma vez que foi elaborada a partir de informações brutas as quais não receberam nenhum tratamento analítico. A pesquisa foi realizada diretamente com base nos dados compilados da educação superior pelo INEP, sob o formato eletrônico de texto CSV (valores separados por vírgula) para importação. Normalmente a pesquisa documental é realizada com base em textos escritos em papel, porém, documentos eletrônicos estão se tornando mais comuns para amparar este tipo de pesquisa (GIL et al., 2017). Ainda de acordo com os referidos autores, os documentos eletrônicos constituem-se como fonte de informações no âmbito da pesquisa documental, sendo que os registros estatísticos e documentos institucionais de entidades privadas e governamentais figuram como exemplo dentre os tipos mais utilizados nas pesquisas.

3.2 População e Amostra

A população da pesquisa é representada pela totalidade das IES no Brasil, e a amostra constituída pelos alunos de graduação em Ciências Contábeis. A pesquisa foi realizada a partir de modelos econométricos criados para contrastar as hipóteses de pesquisa, bem como da construção de diagramas de rede, com base no fluxo de entrada e saída de alunos ao longo do período pesquisado.

3.3 Coleta e Tratamento dos Dados

3.3.1 Coleta dos Dados

Os dados referentes à movimentação dos alunos ao longo do tempo foram coletados a partir dos microdados do Censo da Educação Superior (CES), no segundo semestre do ano de 2021¹. Os microdados, segundo a instituição, se constituem no menor nível de segregação de dados recolhidos por suas pesquisas estatísticas, avaliações e exames. Os arquivos contendo os microdados do censo da educação superior, utilizados como base de dados para esta pesquisa, foram acessados a partir da referência Borges (2021)². As informações pertinentes aos PIBs *per capita* foram obtidas do site do IBGE, na seção de "Downloads de Estatísticas", subseção "PIB Municípios" (IBGE, 2021). Em relação aos dados dos conceitos das IES, os mesmos foram extraídos a partir dos arquivos dos microdados do ENADE, disponibilizados pelo INEP em seu sítio na internet (INEP, 2021).

A base histórica disponível compreendeu o período entre os anos de 2009 a 2019, contendo informações detalhadas referentes aos alunos, instituições, cursos, entre outros dados que foram cruzados e explorados nesta pesquisa. Estas informações foram essenciais para viabilizar o modelo empírico proposto. Entretanto, a pesquisa contemplou os dados de 2009 a 2017, tendo em vista que o INEP promoveu modificações nos anos de 2018 e 2019 que impossibilitaram a inclusão deste período na pesquisa. A mudança promovida pelo INEP alterou a codificação dos alunos, impossibilitando o rastreamento do movimento de migração entre os anos a partir da modificação realizada, sob a justificativa de atender aos requisitos da LGPD.

Além dos dados mencionados, também foram utilizadas informações públicas disponíveis relacionadas ao PIB das cidades envolvidas no mesmo período, bem como os conceitos das IES a partir das notas obtidas nas avaliações do ENADE (Exame

¹Os microdados do CES ficaram indisponíveis no portal do INEP entre o final de 2021 e início de 2022. Após o retorno, as informações disponíveis não mais detalhavam a movimentação dos alunos, devido à alegação de potencial risco à Lei Geral de Proteção de Dados (LGPD)

²Os arquivos disponíveis neste endereço eletrônico são os arquivos que eram disponibilizados pelo INEP até o segundo semestre de 2021. A partir de 2022, o INEP voltou a disponibilizar os microdados, porém sem a base de alunos.

Nacional de Desempenho dos Estudantes), aplicadas pelo INEP.

3.4 Variáveis da Pesquisa

Para atingir os objetivos propostos nesta pesquisa, as informações referentes à migração de alunos intra e extra instituição foram compostas sob a ótica de duas dimensões: (1) alunos e (2) IES.

A pesquisa contou com duas variáveis dependentes, relativamente às dimensões propostas. Pela ordem, primeiramente foi abordada a dimensão dos alunos. Na sequência, foi abordada a dimensão das instituições.

3.4.1 Dimensão Aluno

Esta dimensão aborda os aspectos relacionados ao comportamento dos estudantes, baseado nas variáveis descritas a seguir.

3.4.1.1 Variáveis Independentes

Para explicar a variável dependente, foram utilizadas quatro variáveis independentes, às quais estão relacionadas às condições geo-socioeconômicas. Nesta pesquisa, elas se constituem de auxílios econômico-financeiros, aspectos sociais, aspecto geográfico e diferenciação do curso entre as IES, a seguir especificadas:

- a) Fator econômico-financeiro: Esta variável reúne, de forma agregada, as informações pertinentes aos alunos no tocante aos auxílios e benefícios que recebem ou não, especificamente de: financiamento, bolsa e auxílios (moradia, alimentação etc.). Para cada um destes elementos foi atribuído um valor binário, que assumiu o valor 1 no caso de o indivíduo possuir o benefício, ou 0, se não possuir benefícios. O cálculo se deu pela variação entre a condição econômico-financeira da instituição de origem i e destino j no tempo t .

É expresso pela seguinte equação:

$$\Delta E_{kij}^t = E_{kj}^t - E_{ki}^{t-1} \quad (3.1)$$

onde:

$$E_{ki}^t = \frac{F_{kj}^t + B_{kj}^t + A_{kj}^t}{3} \quad (3.2)$$

em que:

$$E_{kj}^t = \text{Condição econômico-financeira do aluno } k \text{ no tempo } t \text{ na instituição } j$$

F_{kj}^t = Financiamento do aluno k na instituição j no tempo t . Assumirá o número 1 se possuir, ou 0 se não possuir.

B_{kj}^t = Bolsa de estudo do aluno k na instituição j no tempo t . Assumiu o valor de 1 caso o aluno possuísse qualquer tipo de bolsa de estudo, ou 0 caso não possuísse qualquer tipo de bolsa;

A_{kj}^t = Auxílio ao aluno k na instituição j no tempo t . Assumiu o número 1 se possuísse auxílio, ou 0 se não possuísse;

b) Fator Social:

A finalidade desta variável é compor a função de utilidade, por meio da comparação entre os atributos de sexo, idade e deficiência do aluno na instituição de origem e os mesmos atributos na instituição de destino. Foi calculada conforme a seguinte equação:

$$\Delta S_{kij}^t = S_{kj}^t - S_{ki}^{t-1} \quad (3.3)$$

Sendo que os componentes S_{kj}^t e S_{ki}^{t-1} referem-se respectivamente aos fatores sociais do aluno k nas instituições de origem (i) e destino (j), entre os tempos t e $t - 1$.

O fator social S foi calculado conforme a equação abaixo:

$$S_{kj}^t = \frac{Se_{kj}^t + Id_{kj}^t + De_{kj}^t}{3} \quad (3.4)$$

em que:

Se_{kj}^t = Percentual de alunos do mesmo sexo do aluno k sobre o total de alunos da instituição j , no tempo t ;

Id_{kj}^t = Percentual de alunos na mesma faixa etária do aluno k na instituição j , no tempo t ;

De_{kj}^t = Percentual de alunos com a mesma deficiência do aluno k sobre o total de alunos da instituição j no tempo t .

c) Fator Distância:

Esta variável tem por finalidade atribuir um fator de peso com base na distância entre a cidade do aluno e a cidade das instituições de origem e destino. É representada pela equação abaixo:

$$\Delta D_{kij}^t = D_{kj}^t - D_{ki}^{t-1} \quad (3.5)$$

em que:

ΔD_{kij}^t = Variação da distância entre cidade do aluno k e as cidades das instituições de origem i e destino j no tempo t ;

D_{kj}^t = Distância entre a cidade de nascimento do aluno k e a instituição de destino j no tempo t ;

D_{ki}^{t-1} = Distância entre a cidade de nascimento do aluno k e a instituição de origem i no tempo $t - 1$.

d) Fator Nota do Curso:

A última variável explicativa desta dimensão de análise diz respeito à pontuação do curso de Ciências Contábeis nas instituições de origem e destino. Partiu-se da premissa que a qualidade do curso oferecido, com base na sua pontuação divulgada pelo MEC, é um fator que aumenta a atratividade e a probabilidade de um aluno migrar de uma instituição para outra.

O cálculo desta variável é representado pela Equação 3.6.

$$\Delta N_{ij}^t = N_j^t - N_i^{t-1} \quad (3.6)$$

em que:

ΔN_{ij}^t = Variação da nota entre as instituições de origem i e destino j no tempo t ;

N_j^t = Nota do curso na instituição de destino j no tempo t ;

N_i^{t-1} = Nota do curso na instituição de origem i no tempo $t - 1$.

3.4.1.2 Variável Dependente

A variável dependente utilizada nesta dimensão foi definida como a probabilidade do aluno migrar de IES ao longo do tempo. É representada por P_{kij}^t , cujo resultado está entre 0 e 1, sendo que quanto mais próximo de 1, maior a probabilidade do aluno migrar de IES ou curso, ao passo que quanto mais próximo de 0, maior a probabilidade de retenção do aluno na IES.

em que:

P_{kij}^t = probabilidade do aluno k migrar da IES i para j no tempo t ;

ΔU_{kij}^t = variação da função utilidade (modelo "dimensão aluno").

3.4.2 Dimensão IES

Esta dimensão investiga a relação entre a competição e o fluxo migratório de alunos entre os cursos.

3.4.2.1 Variáveis Independentes

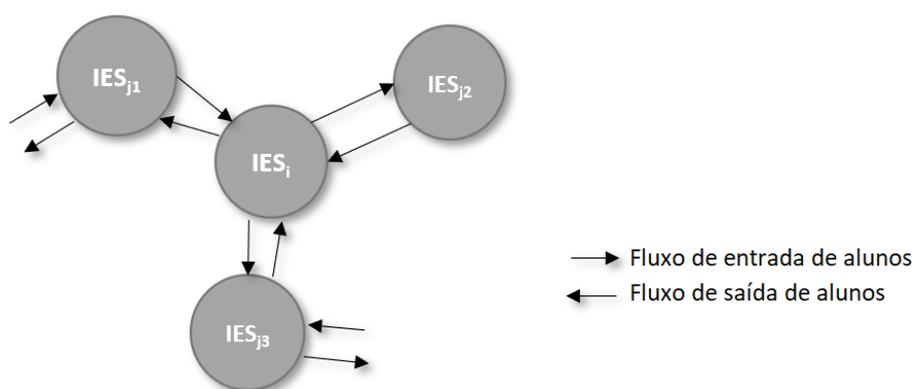
Conforme abordado anteriormente neste capítulo, a pesquisa busca explicar os efeitos na demanda de alunos pelo curso de Ciências Contábeis via transferência interinstitucional, alicerçada na concorrência entre as IES.

Desta forma, as variáveis explicativas estão associadas à estrutura de competição do mercado, representadas pela variação entre alunos entrantes e saíntes dos cursos, e a variação da variância do número de alunos por curso.

a) Variação da Concorrência (conexões)

A primeira variável explicativa que compõe o modelo diz respeito à variação do número de conexões entre os cursos, que é caracterizada pelo número de cursos onde houve ocorrência de entrada, subtraído do número de cursos onde houve ocorrência de saída de estudantes. À luz da teoria de redes complexas, a competição no modelo proposto se dá pelo número de nós adjacentes que se conectam com o nó principal (IES analisada) por meio de arestas (migração dos alunos), conforme ilustra a Figura 4.

Figura 4: Grau de competição do mercado



Fonte: elaborado pelo autor (2021).

De acordo com a métrica apresentada, o cálculo desta variável é representado pela Equação 3.8.

$$VAR_CONEX_i^t = \sum_j (k_{ij}^t - k_{ji}^t) \quad (3.7)$$

em que:

$VAR_CONEX_i^t$ = Variação do número de IES conectadas à IES I no tempo t ;

K_{ij}^t = Conexão de entrada entre as IES i e j no tempo t ;

K_{ji}^t = Conexão de saída entre as IES j e i no tempo t .

b) Variação da Variância de Estudantes nos Cursos

A segunda variável explicativa refere-se ao cálculo da variação da variância do número de alunos dos cursos, que foi definida para possibilitar determinar a influência da heterogeneidade dos canais de migração. Neste sentido, uma variância nula, implica em fluxo idêntico nos canais, ao passo que o contrário determina diferenças de magnitude das conexões. O cálculo desta variável é definido por:

$$VAR_VAR_i^t = \sum_j [var(E_{ij}^t) - var(E_{ji}^t)] \quad (3.8)$$

em que:

$VAR_VAR_i^t$ = Variação da variância de estudantes do curso i no tempo t ;

$var(E_{ij}^t)$ = Variância de estudantes que entraram no curso j oriundos dos cursos i no tempo t ;

$var(E_{ji}^t)$ = Variância de estudantes que saíram do curso j e migraram para os cursos i no tempo t ;

3.4.3 Variáveis de controle

As variáveis de controle da dimensão IES serão compreendidas por:

a) Tamanho da instituição (TAM):

O tamanho da instituição analisada foi definido pelo seu número de alunos ativos na instituição i , no tempo t , representado por TAM_{it} .

b) Modalidade de ensino (MOD):

Quanto à modalidade de ensino, esta variável é do tipo binária, assumindo o valor de 0 para modalidade EaD, e valor 1 para modalidade presencial.

c) Natureza da instituição (NAT)

Esta é uma variável do tipo binária, que representa a natureza jurídica da IES, assumindo valor igual a 1 (um) caso seja de natureza privada, ou igual a 0 (zero) caso seja de natureza pública;

3.5 Síntese das Variáveis

O Quadro 4 apresenta uma síntese de todas as variáveis utilizadas na pesquisa, com sua descrição, *proxy* e fontes de coleta.

3.6 Modelos

Nesta seção, apresenta-se os modelos econométricos responsáveis por relacionar as variáveis indicadas na seção anterior.

3.6.1 Dimensão Aluno

Para estimar a probabilidade dos alunos migrarem entre as IES, será utilizado um modelo de regressão logística (logit) com dados em painel.

Esta técnica estatística permite tratar a variável dependente como categórica, ou seja, assumindo posição dicotômica ou binária. A utilização de dados em painel será utilizada, tendo em vista a aplicação do modelo em uma base de informações ao longo de vários anos.

Modelos como este, têm sido aplicados com sucesso nas pesquisas, como por exemplo no artigo de Minussi, Damacena e Ness Jr (2002), em que os autores utilizaram a regressão logística para criar um modelo de previsão de insolvência no mercado financeiro. A regressão logística é atualmente uma técnica bastante utilizada nas pesquisas que envolvem variáveis medidas de forma dicotômica. Esta técnica apresenta algumas vantagens, como por exemplo uma maior robustez para uma grande variedade de distribuições, uma vez que não se necessita da suposição de normalidade multivariada, o que a torna mais genérica (BITTENCOURT; CLARKE, 2001). Em termos de desvantagens, pode se destacar que a regressão logística é bastante sensível à colinearidade, tendo por consequência resultados de erros padrão muito elevados, não permitindo a realização de testes de significância para os coeficientes (BITTENCOURT; CLARKE, 2001).

A Equação 3.9 expressa o modelo de regressão logística para cálculo da probabilidade do aluno migrar de IES.

$$P_{kij}^t = \frac{1}{1 + e^{-\Delta U_{kij}^t}} \quad (3.9)$$

Quadro 4: Variáveis de pesquisa

Variável	Descrição	Proxy	Fonte
DIMENSÃO ALUNO			
P	Probabilidade do aluno migrar de IES	$P_{kij}^t = \frac{1}{1+e^{-\Delta U_{kij}^t}}$ 0 = Não migrou 1 = Migrou	INEP
ΔU	Variação da função utilidade	Em função da condição econômico-financeira, fator social, distância e nota do curso	INEP
E	Fator Econômico-financ.	$E_{ki}^t = \frac{F_{kj}^t + B_{kj}^t + A_{kj}^t}{3}$ F = Financiamento (1 sim; 0 não) B = Bolsa (1 sim; 0 não) A = Auxílios (1 sim; 0 não)	INEP
S	Fator Social	$S_{kj}^t = \frac{Se_{kj}^t + Id_{kj}^t + De_{kj}^t}{3}$ Se = % de alunos do mesmo sexo Id = % de alunos na mesma idade De = % de alunos com deficiência	INEP
D	Fator Distância	$\Delta D_{kij}^t = D_{kj}^t - D_{ki}^{t-1}$ Distância entre a cidade do aluno e a IES	INEP IBGE
N	Fator Nota do curso	$\Delta N_{ij}^t = N_j^t - N_i^{t-1}$ Diferença entre as notas do curso nas IES de origem e destino	INEP ENADE
DIMENSÃO IES			
VAR_ENT	Demanda	Variação de alunos Dados em Painel	INEP
VAR_CONEX	Variação de concorrentes	$VAR_CONEX_i^t = \sum_j (k_{ij}^t - k_{ji}^t)$ Variação do número de IES conectadas (entrada - saída)	INEP
VAR_VAR	Variação de variâncias	$VAR_VAR_i^t = \sum_j [var(E_{ij}^t) - var(E_{ji}^t)]$ Variância dos cursos conectados (entrada - saída)	INEP
TAM	Tamanho da IES	Número de alunos ativos	INEP
MOD	Modalidade de ensino	0 para presencial; 1 para EaD.	INEP
NAT	Natureza jurídica da IES	0 para IES pública; 1 para IES privada.	INEP

em que:

P_{kij}^t = probabilidade do aluno k migrar da IES/curso i para j no tempo t ;

ΔU_{kij}^t = variação da função utilidade (modelo "dimensão aluno").

Para a variação da função utilidade ΔU_{kij}^t , será empregado o seguinte modelo econométrico:

$$\Delta U_{kij}^t = \alpha_0 + \alpha_1 \Delta E_{kij}^t + \alpha_2 \Delta S_{kij}^t + \alpha_3 \Delta D_{kij}^t + \alpha_4 \Delta N_{ij}^t + \varepsilon_k^t \quad (3.10)$$

em que:

ΔE_{kij}^t = variação da condição econômico-financeira do aluno k no tempo t entre as IES i e j ;

ΔS_{kij}^t = variação do fator social do aluno k no tempo t entre as IES i e j ;

ΔD_{kij}^t = variação das distâncias entre as cidades i dos alunos k e cidades j , no tempo t ;

ΔN_{ij}^t = variação das notas do curso de Ciências Contábeis entre as IES i e j no tempo t ;

α_0 = intercepto;

α_1 ao α_4 = coeficientes alfa associados à sensibilidade das variáveis;

ε = erro residual da regressão,

3.6.2 Dimensão IES

Para o tratamento dos dados referentes à demanda, foi utilizado o modelo de regressão linear múltipla de dados em painel. Ao combinar séries temporais a cortes transversais, os dados em painel "oferecem dados mais informativos, maior variabilidade, menos colinearidade entre variáveis, mais graus de liberdade e mais eficiência"(GUJARATI; PORTER, 2011, p. 588). Este tipo de regressão tem sido aplicado em diversos estudos similares recentes no campo da educação, tais como Souza et al. (2021), que estudou o desempenho de candidatos do ENEM na região Nordeste, utilizando um modelo de dados em painel com efeitos fixos; e Sallaberry et al. (2022), que também aplicou uma regressão com dados em painel com efeitos fixos, para investigar os determinantes do gasto público em Educação das cidades Sul Mato-Grossenses, no período de 2010-2018.

A fim de contrastar as hipóteses de pesquisa e alcançar os objetivos propostos, o modelo econométrico tomou por base as variáveis expostas, sendo formulada, para tanto, a Equação 3.11, a seguir apresentada:

$$VAR_ENT_i^t = \beta_0 + \beta_1 VAR_CONEX_i^t + \beta_2 VAR_VAR_i^t + \sum_{i>4} \beta_i C_i^t + \varepsilon_i^t \quad (3.11)$$

em que:

$VAR_ENT_i^t$ = variação da migração de alunos, pela diferença entre entrantes e saíntes da instituição i no tempo t ;

$VAR_CONEX_i^t$ = variação do número de instituições que possuem conexão com a instituição i no tempo t ;

$VAR_VAR_i^t$ = variação da variância do número de alunos do curso i no tempo t ;

C_i^t = variáveis de controle;

β_0 = coeficiente beta constante;

β_1 ao β_i = coeficientes betas que medem a sensibilidade das variáveis;

ε = erro residual da regressão.

3.7 Processamento dos Dados

Para o processamento dos dados, foram obtidos, em bases públicas na internet, os arquivos com os microdados dos censos da educação superior dos anos de 2009 a 2017, bem como os dados do IBGE contendo os PIBs dos municípios dos anos de 2009 a 2019 e os dados referentes às notas do ENADE aplicadas às IES durante o período pesquisado.

O processamento para tratamento e extração das informações necessárias à pesquisa se deu por etapas, conforme a necessidade de estratificação e uniformização dos dados. Por se tratarem de bases volumosas para cada ano, não foi possível a utilização de softwares que manipulam planilhas eletrônicas, visto que a capacidade destes normalmente se limita a pouco mais de 1 milhão de linhas, e as bases trabalhadas têm pelo menos oito vezes essa quantidade em cada ano pesquisado. Sendo assim, fez-se uso de rotinas desenvolvidas pelo autor na linguagem *python* para o processamento dos dados. Esta linguagem, além de suportar arquivos de grandes volumes, também proporcionou agilidade no processamento e foi utilizada também para a execução dos cálculos estatísticos exigidos pelos modelos econométricos da pesquisa.

Conforme definido na pesquisa, os modelos abrangem as dimensões dos alunos e das IES, se fazendo então necessário que as bases fossem processadas a fim de fornecerem as variáveis consideradas por ambos os modelos. Desta forma, o processamento se dividiu nas seguintes linhas:

3.7.1 Dimensão Aluno

Para o processamento das variáveis relacionadas à dimensão do aluno, foram utilizados principalmente os arquivos dos microdados do censo da educação superior, além de outras bases necessárias, tais como arquivos do IBGE e ENADE. As bases

constantes dos microdados se dividem basicamente em três partes (o nome dos arquivos pode variar, dependendo do ano do censo):

- DM_IES: arquivo formato CSV que contém o detalhamento dos dados das IES do censo;
- DM_ALUNO: arquivo formato CSV que contém o detalhamento dos dados dos alunos
- DM_CURSO: arquivo formato CSV que contém o detalhamento dos dados dos cursos

As etapas para obtenção das variáveis deste modelo foram as seguintes:

a) Primeira etapa: estratificação e homogeneização dos dados

Primeiramente, foram criadas bases secundárias derivadas das bases principais dos microdados do censo da educação superior, com a finalidade de segmentar os dados e assim constituir bases anuais contendo apenas as informações pertinentes aos cursos de ciências contábeis. Nesta etapa, foram agregadas as coordenadas geográficas dos municípios, ao ser feito o cruzamento dos códigos existentes nas bases dos microdados com a base dos municípios obtidas do IBGE. Tendo em vista que o INEP ao longo dos anos alterou diversas informações na estrutura dos microdados, principalmente no que diz respeito aos nomes dos campos, houve a necessidade de ser feita a homogeneização das suas nomenclaturas nos códigos de programação.

Além da estratificação pelos cursos de ciências contábeis e da agregação das informações referentes aos municípios (nomes e coordenadas geográficas), as bases secundárias criadas nesta etapa foram constituídas a partir de uma mescla de dados das três bases dos microdados do INEP (IES, aluno e curso). Os campos quantitativos e qualitativos que não eram de interesse da pesquisa foram descartados para fins de otimização do desempenho, remanescendo apenas os campos necessários à realização dos cálculos da pesquisa.

b) Segunda etapa: raspagem de dados de geolocalização

Em uma segunda etapa, houve a necessidade de se adequar as coordenadas geográficas das IES, uma vez que uma mesma IES poderia oferecer cursos em cidades diferentes, ou até em uma mesma cidade, porém em endereços diferentes. Desta forma, as coordenadas geográficas processadas na primeira etapa foram atualizadas e/ou corrigidas com base nos endereços reais das IES, ao invés da cidade. Para que isso fosse possível, foi utilizada uma técnica de raspagem de dados na internet ("web scrape"), com 7 níveis alternativos de busca

e varredura. Houve um total de 1.545 cursos mapeados com coordenadas validadas de forma automatizada. Após esta pesquisa automatizada, ainda foram feitos pequenos ajustes manuais para cursos que não foram alcançados pelos níveis de busca da raspagem de dados. Após a atribuição dos dados de geolocalização, foi então possível realizar o cálculo das distâncias entre as cidades de origem de cada aluno e as IES, ano a ano.

c) Terceira Etapa: cálculo das variáveis da pesquisa

A partir da base secundária constituída nas etapas anteriores, houve o processamento das informações para agregar o cálculo das variáveis independentes e de controle propostas pelos modelos da pesquisa. Inicialmente, foram calculados os componentes da variável de "condição econômico-financeira", aluno por aluno em cada IES e ano. Isto envolveu o levantamento e computação individual das variáveis relacionadas às condições de financiamento, bolsa de estudo e auxílio. Para os componentes "financiamento" e "auxílio", a atribuição dos valores "1" e "0", representando se o aluno possuía ou não os componentes, a relação foi simples e direta. Entretanto, para o componente "bolsa de estudo", houve a necessidade de proceder com uma composição no cálculo, visto que há dois campos na base de dados do INEP que determinam a presença ou não da bolsa de estudo, culminando na necessidade da verificação conjunta destes campos.

Na sequência, foram calculadas as variáveis para determinação da variável de "fator social" proposta pela pesquisa. Para tanto, foram calculados os percentuais relativos ao gênero declarado dos alunos em relação a cada IES, por ano. Este cálculo foi feito com base no cômputo dos percentuais dos sexos de cada IES, comparando-se cada aluno em relação à sua IES, por ano. Em seguida, de forma semelhante, foi calculado o percentual de alunos da mesma idade que o aluno na IES em cada ano. Entretanto, o componente de maior complexidade para a obtenção desta variável foi a determinação da deficiência dos alunos. Isto porque as bases do CES apresentou, no último ano pesquisado, 13 tipos distintos de deficiência, sendo que no início do período pesquisado o número de tipos de deficiência era inferior, e foram sendo agregados outros tipos ao longo dos anos. Isto exigiu, primeiramente, uma homogeneização desta variável para todos os anos. O processamento do cálculo foi complexo, pois houve a necessidade de se comparar a deficiência de cada aluno em relação ao total de alunos com o mesmo tipo de deficiência, em cada IES e ano.

Por fim, foi levantada a variável "nota do curso", por meio do cruzamento dos dados do CES com os dados resultantes do ENADE. Como o ENADE é aplicado a cada três anos, foram considerados os conceitos do ano de 2012 para o triênio 2009-2011, de 2015 para o triênio 2012-2014, e de 2018 para o triênio

2015-2017. Mais uma vez observou-se uma quebra de uniformidade na nomenclatura dos campos nas bases de dados públicas ao longo dos anos, ponto que exigiu a customização das rotinas programadas para que fosse possível processar corretamente as informações. Também destaca-se que alguns cursos não haviam sido ainda reconhecidos pelo MEC, à época dos anos pesquisados, e por conseguinte não apresentaram nota do ENADE no período.

d) Quarta Etapa: criação da base dos fluxos migratórios

Após o cálculo das variáveis desta dimensão, como descrito nas etapas anteriores, procedeu-se, por fim, o processamento do fluxo migratório dos alunos, que se constituiu na principal base de dados para a aplicação do modelo econométrico da pesquisa. Este processamento foi feito partindo-se do ano de 2009 como base, e criados, a partir deste, ano os fluxos de 2010 a 2017, comparando-se ano a ano os dados de origem do aluno no ano anterior com o ano seguinte. Foi gerada uma base apenas dos alunos que migraram de uma IES para outra entre os anos, contendo todas as variáveis do modelo para cada aluno. Estes fluxos se constituíram nas bases de dados principais para a viabilização do cálculo econométrico apresentado na pesquisa. Em termos de volume de processamento, o primeiro ano apresentou 76.226 ocorrências processadas, e o último ano, 305.105. O processamento foi automatizado e ocorreu comparando-se aluno a aluno (linha a linha) do ano do processamento com o ano anterior, originando o fluxo de movimentação de cada ano. Ao término deste processo, foi gerada uma base única contendo cerca de 2,2 milhões de registros referentes aos alunos ativos de ciências contábeis ao longo do período de análise.

A Figura 5 ilustra as etapas percorridas para a obtenção da base de dados final da dimensão "Aluno".

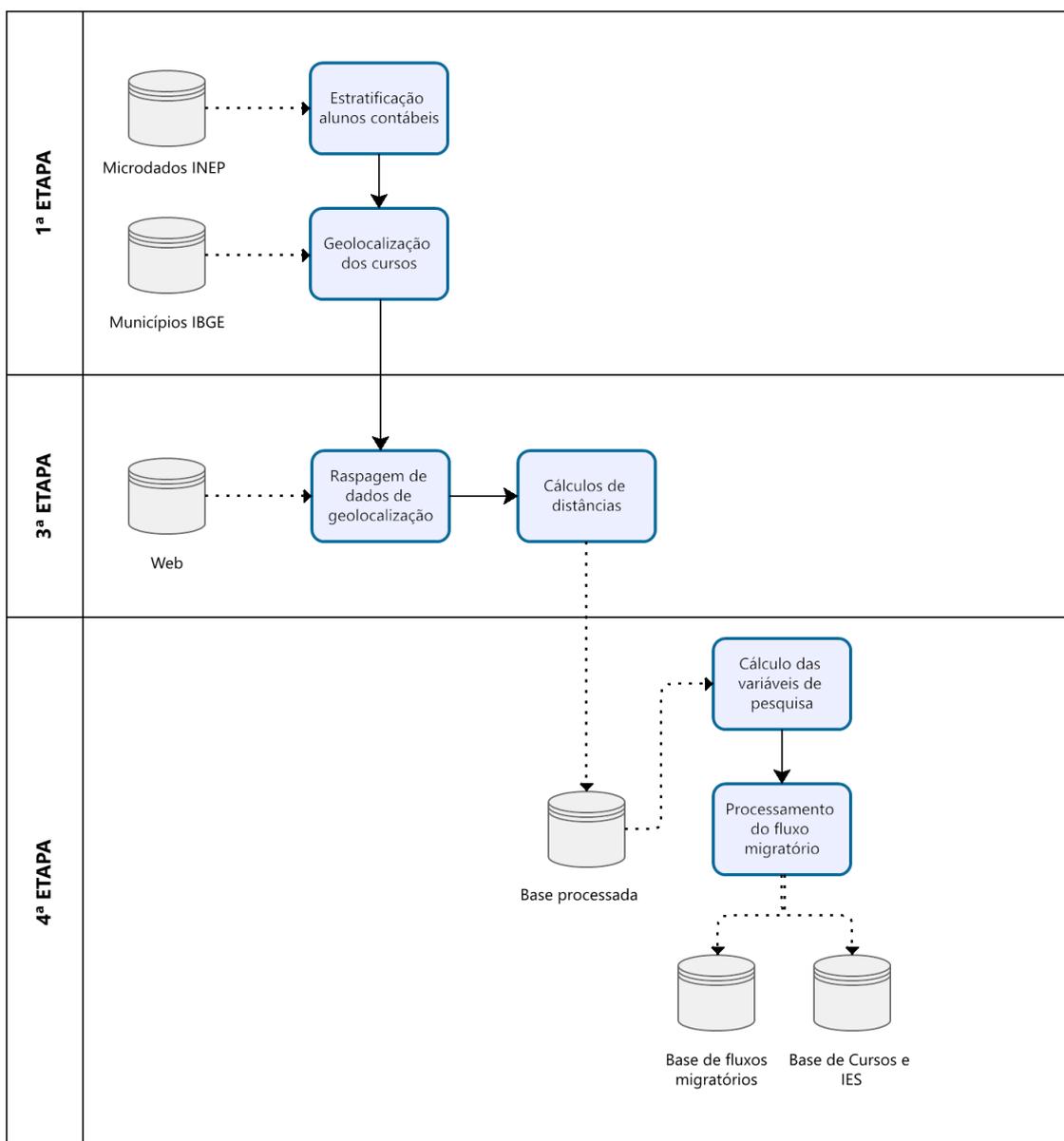
3.7.2 Dimensão IES

O preparo da base de dados para aplicação do modelo econométrico destinado à dimensão IES percorreu um caminho menos complexo em comparação à base dos alunos, tendo em vista a menor demanda de cruzamento de dados com informações paralelas (por exemplo, não exigiu raspagem de dados), bem como por aproveitar parte do processamento feito para a obtenção da primeira base (alunos). O preparo da base ocorreu de acordo com as seguintes etapas:

a) Primeira etapa: quantificação de alunos Presencial/EaD e natureza jurídica

O primeiro passo para viabilização da aplicação do modelo proposto, foi a totalização dos alunos segmentada pelas modalidades de ensino em cada IES, por

Figura 5: Diagrama do Processamento de Dados - Dimensão Aluno



Fonte: Elaborado pelo autor.

ano. Para tanto, se processou a contagem da quantidade de alunos em cada modalidade, utilizando-se da base secundária gerada na primeira etapa da dimensão aluno. Com isso, foram criadas 3 colunas na base de dados das IES, contendo, respectivamente, o total de alunos na modalidade presencial, o total de alunos na modalidade EaD e o somatório das duas modalidades (total de alunos). Na mesma etapa, foram processadas as naturezas jurídicas das IES. Segundo as bases originais dos microdados do INEP, as IES se dividem em públicas ou privadas, mas apresentam formatação multivariada, visto que ambas as modalidades possuem mais de um subtipo. A presente pesquisa não necessita do desdobramento previsto nos microdados, mas sim, apenas a caracterização dicotômica da natureza jurídica. Por este motivo, foi aplicado um algoritmo específico para converter este campo em uma variável categórica para cada IES, classificando-as como públicas ou privadas.

b) Segunda etapa: Conexões e migração de alunos entre os cursos

A seguir, foram computados o número de conexões existentes para cada curso (cursos que possuem relação com a IES pesquisada em termos de fluxo de alunos), bem como o quantitativo de alunos imigrantes e emigrantes, estratificados por ano e curso. Para tanto, utilizou-se um algoritmo que calculou as movimentações de cada aluno com base nos bancos de dados gerados no item anterior (Dimensão Aluno), mapeando os cursos de origem e destino, e assim computando os totais de cada item (conexão e migração) realizados. O mesmo algoritmo também foi usado para qualificar as conexões de entrada e saída, segmentadas por modalidade (presencial e EaD) e natureza jurídica (IES pública e privada). É importante destacar que o número de conexões pode ser menor ou igual ao número de alunos entrantes e saírem dos cursos. Isto se deve ao fato de que os alunos podem ser originários dos mesmos cursos, ou se destinarem a mesmos cursos. Por fim, nesta etapa foi calculada a variância das conexões de cada curso, em cada ano, para fins de composição do modelo econométrico.

c) Terceira etapa: PIB *per capita* dos municípios e conceitos dos cursos

Para agregar as informações de PIB *per capita* dos municípios e os conceitos dos cursos, foi desenvolvida uma rotina específica que cruzou os dados das bases do IBGE em relação ao PIB dos municípios, assinalando a cada curso de ciências contábeis em cada IES, o respectivo PIB *per capita* do município dos cursos.

De forma semelhante, as notas do ENADE foram computadas a cada IES, com base nos arquivos disponibilizados publicamente pelo INEP referente às avaliações das IES. Como o ENADE é realizado a cada três anos, utilizou-se a edição de 2012 para os anos de 2009 a 2011, a edição de 2015 para os anos de 2012 a 2014, e o ENADE de 2018 para os anos de 2015 a 2017.

- d) Quarta etapa: cálculo das variações de alunos, conexões e demais variáveis de controle

Por fim, foram agregadas na base de dados dos cursos, tanto as variações anuais de alunos (variável dependente) quanto às variações de conexões (variável independente). Nesta etapa, além das variáveis dependente e independente, também foram calculadas as variáveis relativas à renda média das cidades de origem dos alunos segmentada pelos cursos de ciências contábeis vinculados a cada IES, modalidades de ensino (EaD e presencial) e natureza jurídica das IES. Foi efetuado o levantamento do tamanho de cada IES, com base no número de alunos ativos de todos os cursos, para fins de normalização das variáveis utilizadas no modelo. Após completada esta etapa, o resultado foi a obtenção de uma base de dados com todas as variáveis necessárias para a execução dos cálculos econométricos e estatísticos que deram sustentação à pesquisa, no que tange à dimensão IES.

A Figura 6 ilustra as etapas percorridas para a obtenção da base de dados final da dimensão "Aluno".

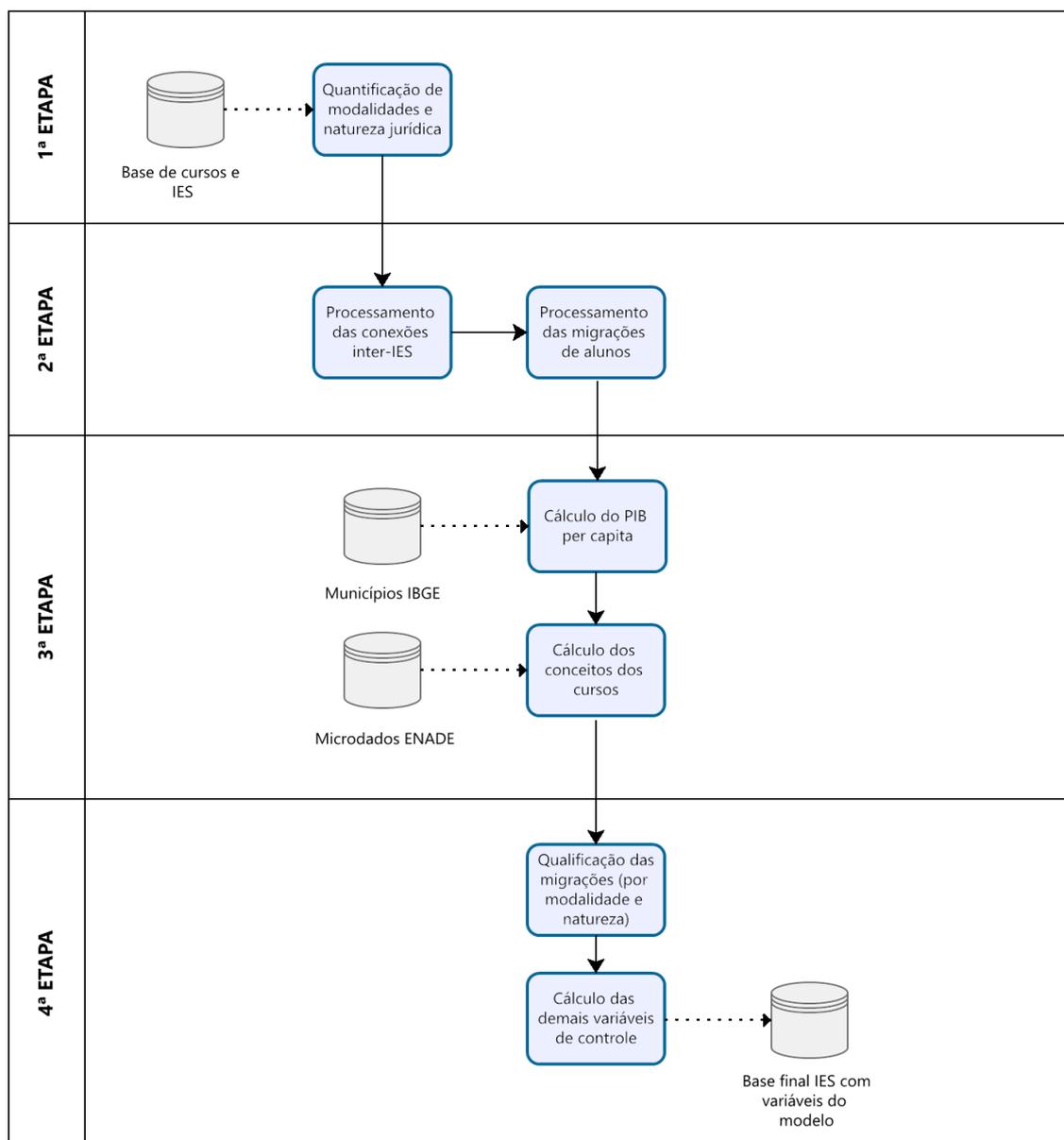
3.7.3 Limitações do Método

A presente pesquisa está limitada ao mercado brasileiro de educação superior, no que tange à migração interinstitucional dos alunos de Ciências Contábeis. Portanto, as migrações de alunos entre cursos na mesma IES não são analisadas, bem como migrações entre IES distintas relacionadas a outros cursos.

A pesquisa também apresenta uma limitação no que tange ao período de abrangência dos dados, o qual corresponde ao intervalo dos anos de 2009 a 2017. Esta limitação decorre de mudanças recentes na disponibilização pública das bases de dados pelo INEP. Os períodos anteriores a 2009 não possuíam o detalhamento necessário para a realização da pesquisa. Posteriormente, os anos de 2018 e 2019 sofreram nova modificação pelo INEP, a qual fez com que tornasse impossível o rastreamento dos alunos, inviabilizando o processamento do comportamento migratório. Recentemente no ano de 2022, o INEP retirou a disponibilidade pública de todos os dados pertinentes aos alunos, sob a justificativa de que a sua divulgação poderia vir a ter implicações com a nova LGPD (Lei Geral de Proteção de Dados).

Outro ponto a ser destacado nesta seção, diz respeito à informação disponível no tocante à cidade de origem dos alunos. Como esta informação só foi adicionada aos microdados do INEP a partir de 2011, somente foi possível considerá-la nos modelos econométricos deste ano em diante.

Figura 6: Diagrama do Processamento de Dados - Dimensão IES



Fonte: Elaborado pelo autor.

4 ANÁLISE DOS RESULTADOS

Este capítulo apresenta os resultados obtidos a partir da aplicação dos modelos propostos, bem como as respectivas análises pertinentes.

Seguindo a estrutura apresentada no capítulo de metodologia, inicialmente são apresentados os resultados obtidos para a dimensão aluno, que incluem a análise descritiva das variáveis e o resultado do modelo econométrico. Em seguida, as análises são apresentadas nos mesmos moldes para a dimensão IES. Por fim, o capítulo será finalizado com a análise global dos resultados.

Todos os dados demonstrados nas seções a seguir foram obtidos a partir das bases de dados derivadas dos microdados do CES/INEP, conforme descrito no capítulo 3, salvo citação explícita no conteúdo desta seção.

4.1 Dimensão Aluno

4.1.1 Estatística Descritiva das Variáveis

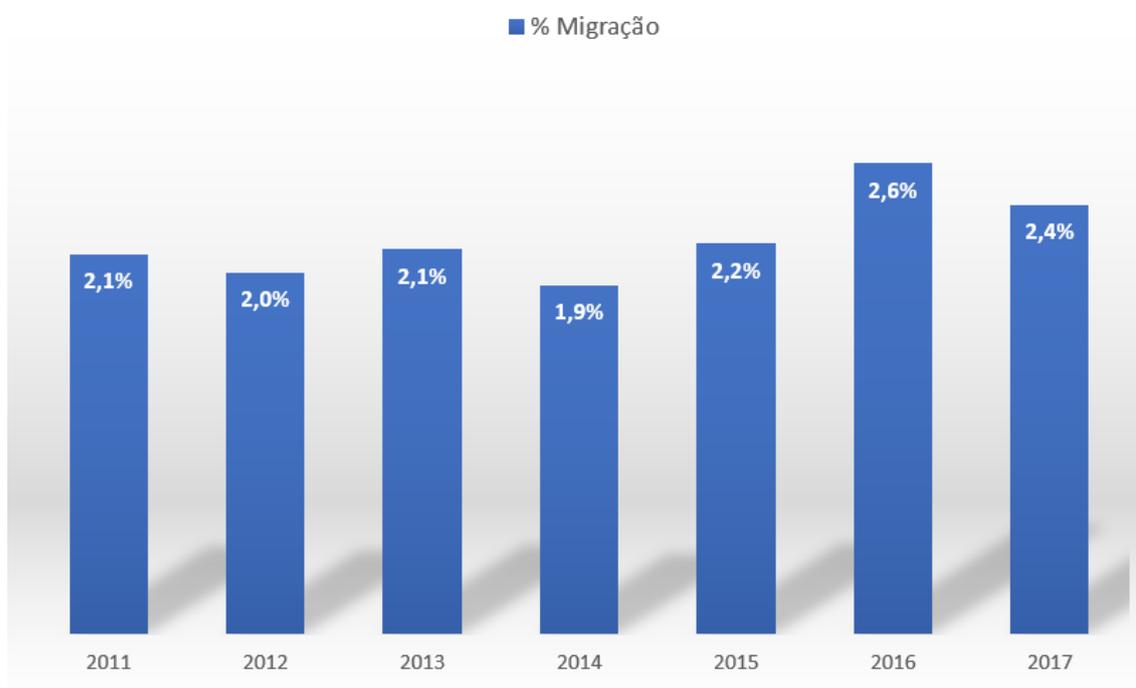
É possível constatar que no período de 2011 a 2017 houve um movimento migratório médio anual de 2,2% do total da base de alunos de ciências contábeis no Brasil. Conforme é possível observar na Figura 7, o período analisado apresenta um cenário de estabilidade em relação ao número de migrações.

Em relação ao total de alunos de ciências contábeis, houve um crescimento de cerca de 22% entre 2011 e 2014, e a partir de 2015 observou-se uma estabilização no número de alunos, conforme pode ser visualizado na Figura 8.

Considerando que as variáveis independentes do modelo dizem respeito às variações das características dos alunos entre os anos pesquisados (fatores econômicos, sociais, distâncias e notas ENADE), é possível inferir que, para aqueles alunos que não migraram de curso no período, os deltas destas variáveis tenderão a ficar próximos de zero. Isto foi comprovado nos dados analisados, uma vez que a grande maioria dos alunos não migrou (mais de 97% nos anos pesquisados). Houve uma grande concentração de incidência das variáveis independentes em torno de zero, conforme é possível observar nas médias das variáveis, constantes da Tabela 2, a qual traz a análise descritiva das variáveis independentes.

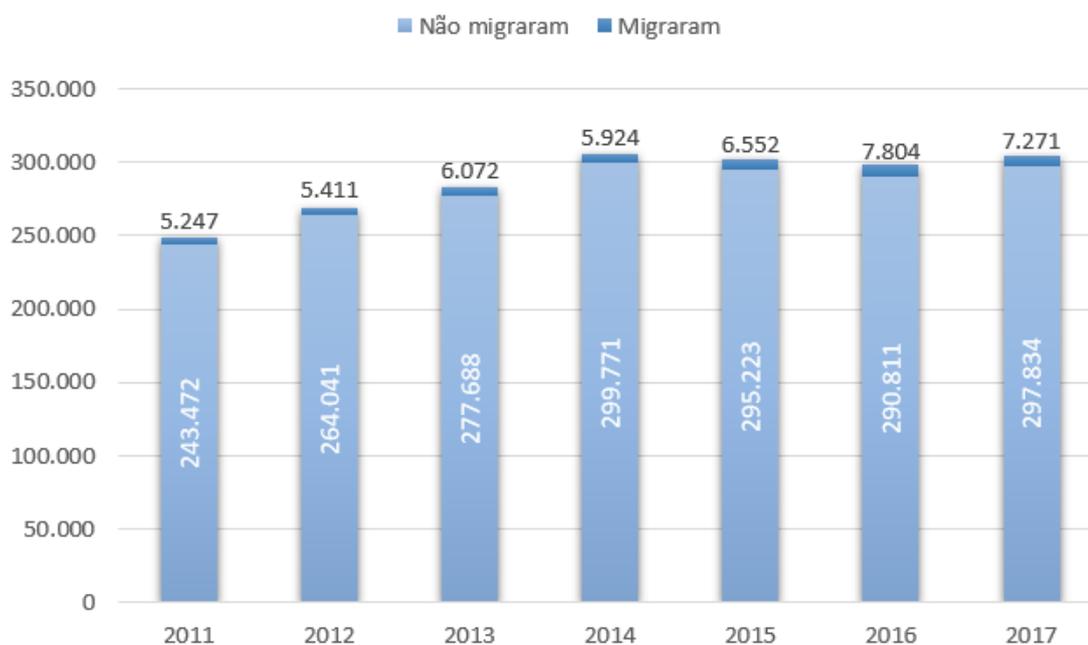
Ao analisar a Tabela 2, observa-se que em relação ao total da população, as variáveis que representam os fatores econômicos e sociais ("E" e "S") possuem médias muito baixas, entretanto, não nulas. Também apresentam uma alta dispersão ao se observar a razão entre os desvios-padrão e as médias. Conforme mencionado, estes resultados estão diretamente relacionados à ampla prevalência de alunos não migrantes, no período pesquisado.

Figura 7: Percentual de Migração dos Alunos - 2011 a 2017



Fonte: Elaborado pelo autor com base nos dados da pesquisa.

Figura 8: Número de Alunos de Ciências Contábeis - 2011 a 2017



Fonte: Elaborado pelo autor com base nos dados da pesquisa.

Tabela 2: Análise Descritiva das Variáveis Independentes - Dimensão Aluno

Variável	Migração	Média	D. Padrão	Mínimo	Máximo
E	Migrou	0,020	0,255	-0,833	0,833
	Total	0,002	0,104	-1,000	1,000
S	Migrou	0,002	0,036	-0,485	0,493
	Total	0,001	0,023	-0,667	0,836
D	Migrou	33,467	667,411	-3.625,009	3.770,588
	Total	-2,522	194,315	-3.800,083	3.770,588
N	Migrou	0,343	1,218	-4,179	4,530
	Total	0,035	0,414	-4,179	4,530

Fonte: elaborado pelo autor com base nos dados da pesquisa

Em relação à variável de distância entre as IES de destino e origem, houve uma média de 33,5km entre os alunos que migraram, também com uma alta dispersão da população analisada. O valor positivo observado na média desta variável significa que a distância pode não representar um incremento na utilidade esperada, uma vez que, na média da população, os alunos migraram para IES mais distantes em relação à IES de origem. Isto significa que a distância pode estar sendo sacrificada em prol de outros atributos com maior percepção de valor pelos alunos nas suas decisões de migração.

No que tange à variação dos conceitos das IES segundo avaliação do ENADE ("N"), naturalmente observou-se uma média baixa em relação à população total, condizente com a prevalência de alunos não-migrantes na população analisada, e também uma alta dispersão. A média positiva observada de 0,343 para os alunos que migraram, sinaliza que houve uma tendência majoritária de transferências para instituições com melhor nota no ENADE.

4.1.2 Análise do Modelo Econométrico

Conforme abordado na metodologia, utilizou-se um modelo de regressão logística com dados em painel para explicar a probabilidade dos alunos migrarem de IES.

O cálculo foi feito utilizando-se o método de dados empilhados, pelo qual se considera todo o período pesquisado, sem cortes temporais longitudinais, o que acarreta em um maior tamanho de amostra em relação aos modelos de efeitos fixos e efeitos aleatórios.

Inicialmente, as variáveis do modelo foram submetidas a um teste para detecção de multicolinearidade, para assegurar que as variáveis independentes não possuem

grau de correlação relevante entre si. Para tanto, foi utilizado o teste de fator de inflação da variância (VIF), o qual mede a quantidade de multicolinearidade de múltiplas variáveis de uma regressão.

Por meio da aplicação do teste VIF, observou-se que as variáveis independentes não apresentaram multicolinearidade. Os autores em geral consideram que uma variável possui alta colinearidade quando apresenta um valor VIF maior do que 10. Conforme demonstra a Tabela 3, as variáveis independentes do modelo apresentaram valores VIF muito baixos, próximos a 1, o que assegura a ausência deste efeito nas variáveis utilizadas na regressão.

Também foi elaborada uma matriz de correlação entre as variáveis independentes. A matriz indica os graus de correlação em uma escala de -1 a +1, onde os extremos denotam alto grau de correlação negativa ou positiva. A Figura 9 demonstra as correlações em um formato semelhante a um mapa de calor, onde fica evidenciada a baixa correlação existente entre as variáveis independentes. Pela matriz, é possível verificar que as variáveis possuem um grau de correlação próximo a zero, o que demonstra não haver correlação entre si.

Após testadas as variáveis para verificação de multicolinearidade, foi realizado o cálculo do modelo de regressão logística com dados em painel empilhados, abrangendo a totalidade das IES brasileiras.

No modelo analisado de regressão logística com dados em painel, a variável dependente se constitui em um valor binário representando a ocorrência de migração ou não dos alunos entre as IES, onde o valor 0 (zero) indicou a não ocorrência de migração para o aluno, e o valor 1 (um) indicou a migração realizada. Assim sendo, valores positivos do beta indicam uma relação direta entre a variável explicativa e a probabilidade de migração do aluno.

O resultado da regressão aplicada para a probabilidade dos alunos migrarem de IES pode ser observado na Tabela 4.

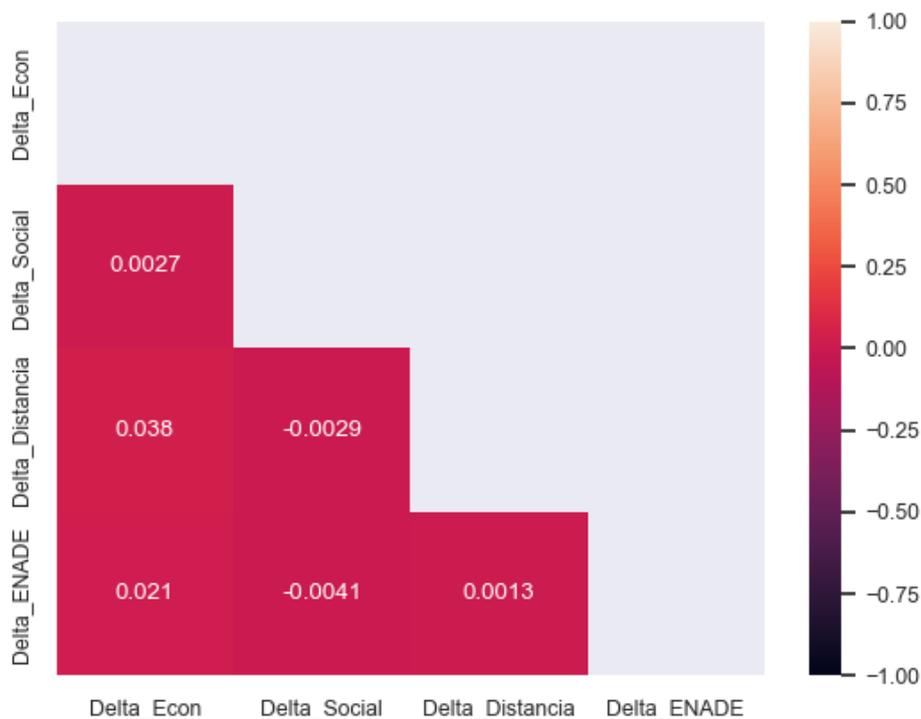
É possível constatar que todas as variáveis independentes do modelo p-valor inferior ao nível de significância de 0,01. Isto significa que a probabilidade dos alunos

Tabela 3: Teste VIF das variáveis independentes

Variável	VIF score
ΔE	1,001939
ΔS	1,000028
ΔD	1,001409
ΔN	1,000527

Fonte: elaborado pelo autor com base nos dados da pesquisa.

Figura 9: Matriz de Correlação de Variáveis Independentes



Fonte: Elaborado pelo autor com base nos dados da pesquisa.

Tabela 4: Resultado da Regressão Logística - Brasil

	coef	std err	z	P > z	[0.025	0.975]
Intercept	-3.9323	0.005	-753.971	0.000	-3.942	-3.922
ΔE	1.2537	0.041	30.249	0.000	1.172	1.335
ΔS	1.1541	0.211	5.457	0.000	0.740	1.569
ΔD	0.0007	1.89e-05	36.557	0.000	0.001	0.001
ΔN	0.8169	0.006	139.859	0.000	0.805	0.828

Fonte: elaborado pelo autor com base nos dados da pesquisa.

migrarem para outras IES encontra uma relação significativa com base nas condições econômicas, sociais, de distância, e também com base na qualificação das IES segundo a avaliação do ENADE.

Em relação ao conceito da IES no ENADE, é lógico esperar que uma IES melhor conceituada tenha uma maior probabilidade de receber alunos que migram de outras IES. Isto se confirmou na pesquisa, uma vez que a variação positiva no conceito entre as IES de origem e destino, impactou positivamente na probabilidade de migração do aluno.

Para que seja possível determinar quais as variáveis que causam mais impacto na probabilidade neste modelo, não se pode tão somente olhar para os coeficientes, eis que estes, em uma regressão logística, refletem uma taxa marginal em relação à probabilidade do modelo. A solução para contornar este efeito é calcular as chamadas *odds/ratio*, ou "razões de chance". As *odds/ratio* são usadas para comparar as chances relativas da ocorrência do resultado de interesse (no presente caso, a probabilidade de migração), dada a exposição às variáveis explicativas. Também pode ser usada para determinar se uma exposição em especial é um fator de risco para um determinado resultado e para comparar a magnitude de variáveis explicativas. A Tabela 5 exhibe, em ordem de importância, as razões de chance das variáveis deste modelo.

Os achados da aplicação deste modelo dizem respeito à constatação de que variações nos fatores econômicos e sociais (3,50 e 3,17) são fatores que influenciam mais na probabilidade de migração dos discentes em relação às variações nas notas de ENADE e distância (2,26 e 1,00). É importante destacar que o fator econômico, que se traduz em incremento da vantagem econômica do aluno, é o principal componente na probabilidade de migração. Em comparação com as demais variáveis, a variação nos fatores econômicos possui uma probabilidade 1,1 vez maior que a variação nos fatores sociais, 1,54 vezes maior que a nota da IES no ENADE, e pouco mais de 3,5 vezes maior do que a variação da distância. Isto sinaliza a confirmação da aplicação da teoria utilitarista, uma vez que a vantagem econômica no modelo significa que o agente decisor terá um atributo de custo menor na escolha alternativa. Desta forma, este atributo favorece a maximização da utilidade esperada, e, portanto, aumenta a probabilidade de tomada de decisão para a migração à outra IES.

Em termos de referenciais comparativos com base em outras pesquisas sobre o tema, após uma busca em bases públicas, verificou-se que há uma escassez de trabalhos com parâmetros similares. Entretanto, algumas pesquisas como a de LI e Chagas (2017), onde abordam alguns fatores que impactam na evasão, os autores concluíram que o auxílio social (fator social) contribui positivamente para reduzir a probabilidade de migração do aluno entre as IES, ainda que não no mesmo curso, o que é confirmado pela presente pesquisa.

Tabela 5: Tabela de *odd/ratio*

Variáveis	Odd/ratio
ΔE	3,5033
ΔS	3,1712
ΔN	2,2635
ΔD	1,0007

Fonte: elaborado pelo autor com base nos dados da pesquisa.

Outras pesquisas utilizaram-se de variáveis oriundas de bases internas de determinada IES, ou de dados genéricos populacionais. Destaca-se, dentre estas, a pesquisa de Furtado e Alves (2012), na qual os autores utilizaram variáveis oriundas de bases de dados internas de uma IES para analisar a evasão universitária. Dentre os demais achados da pesquisa, cabe destacar que o fator de distância entre a localidade de origem do aluno e a IES, diverge do presente estudo, uma vez que na referida pesquisa os autores concluíram que a distância é um fator de aumento da probabilidade de evasão.

4.2 Dimensão IES

A segunda modelagem desenvolvida na pesquisa envolveu o estudo da dimensão das IES, pela qual buscou-se analisar a demanda dos cursos de ciências contábeis em relação aos fatores ligados à migração dos alunos.

Para tanto, foi utilizado um modelo de regressão linear com dados em painel, e analisadas as variações dos modelos com dados empilhados, efeitos aleatórios e efeitos fixos.

4.2.1 Estatística Descritiva das Variáveis

Nesta seção é abordada a análise descritiva das variáveis do modelo pertinente à demanda das IES. A Tabela 6 apresenta o resumo estatístico das variáveis quantitativas independentes.

Conforme definido na pesquisa, a equação 3.11 representa o modelo econômico da dimensão IES, e é composto por:

- VAR_ENT: variação dos entrantes dos cursos (variável dependente)
- VAR_CONEX: variação dos cursos que possuem conexões com o curso pesquisado (variável explicativa)
- VAR_VAR: variação da variância do número de alunos que migraram e emigraram dos cursos (variável explicativa)
- TAM: tamanho do curso/IES definido pelo total de alunos (variável de controle)

A análise descritiva das variáveis quantitativas evidencia alta dispersão dos dados, conforme é possível visualizar na Tabela 6. Verificou-se elevados desvios-padrão em todas as variáveis, com destaque para as variáveis VAR_ENT e VAR_CONEX, ao se observar a razão entre os desvios-padrão e as médias. Isto se atribui ao tamanho da base de dados e principalmente pela variedade da população estudada, que abrange todos os cursos de ciências contábeis em suas modalidades presencial e remota.

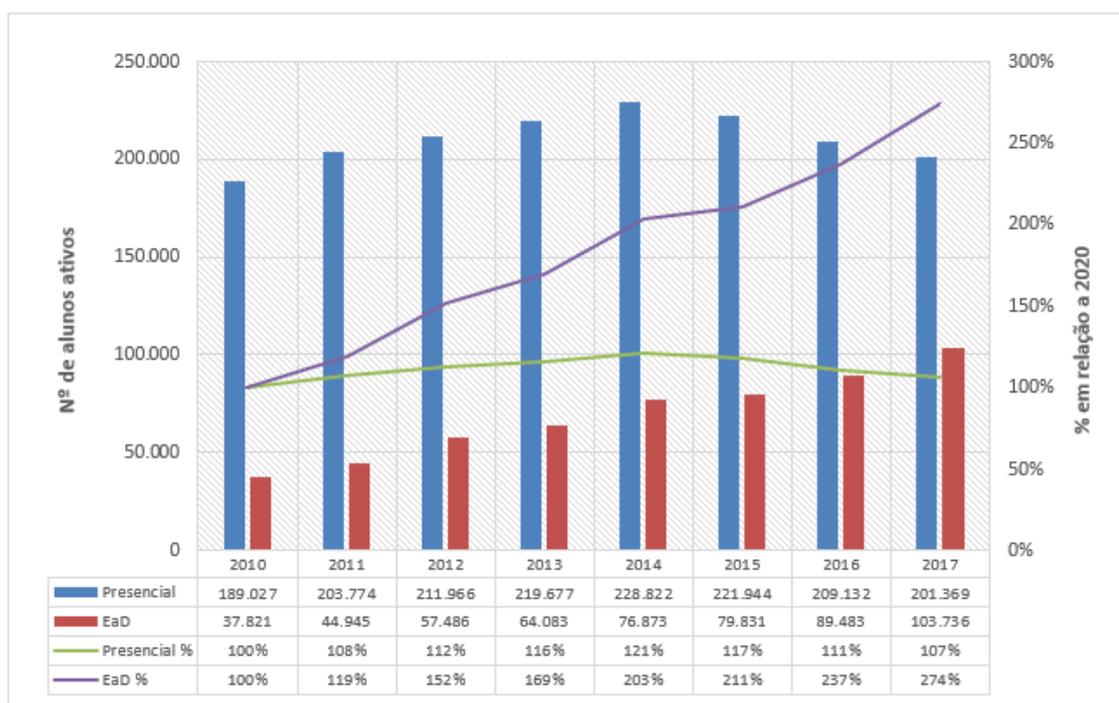
Tabela 6: Análise Descritiva das Variáveis Independentes - Dimensão IES

Variável	Média	D. Padrão	Mínimo	Máximo
VAR_ENT	0,407	12,902	-316,000	387,000
VAR_CONEX	0,029	3,307	-88,000	68,000
VAR_VAR	5,856	338,021	-9.657,810	22.615,667
TAM	235,000	819,000	1,000	29.234,000

Fonte: elaborado pelo autor com base nos dados da pesquisa

Em relação à variável qualitativa referente à modalidade (MODALIDADE) dos cursos, observou-se um crescimento acumulado significativo do número de alunos nos cursos de ciências contábeis na modalidade EaD, entre os anos de 2010 e 2017, conforme apresentado na Figura 10.

Figura 10: Número de alunos de ciências contábeis por modalidade de 2010 a 2017

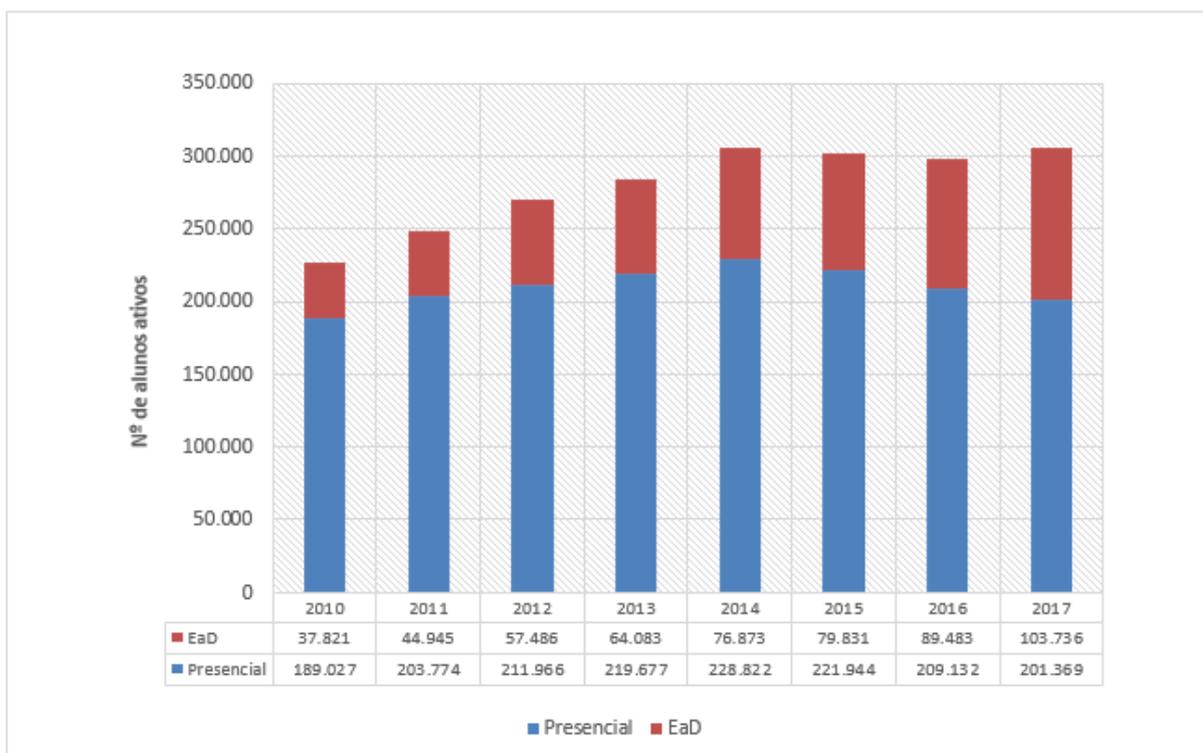


Fonte: Elaborado pelo autor com base nos dados da pesquisa.

O incremento foi da ordem de 174% de alunos ativos, ao passo que a modalidade presencial apresentou um crescimento de apenas 7%. Também é possível constatar uma curva decrescente do número de alunos em cursos presenciais a partir do ano de 2015, enquanto que os cursos à distância, no período pesquisado, apresentou crescimento em todos os anos. A Figura 11 apresenta o crescimento do número de alunos matriculados em ciências contábeis de 2010 a 2017, segmentados pela modalidade, onde evidência-se o gradativo aumento da modalidade EaD em detrimento

da modalidade presencial, em relação ao total de matriculados no curso de ciências contábeis. O percentual de alunos na modalidade EaD saltou de 16,7% em 2010, para 34,0% em 2017, ou seja, a modalidade EaD dobrou sua participação no total de alunos de ciências contábeis no período pesquisado.

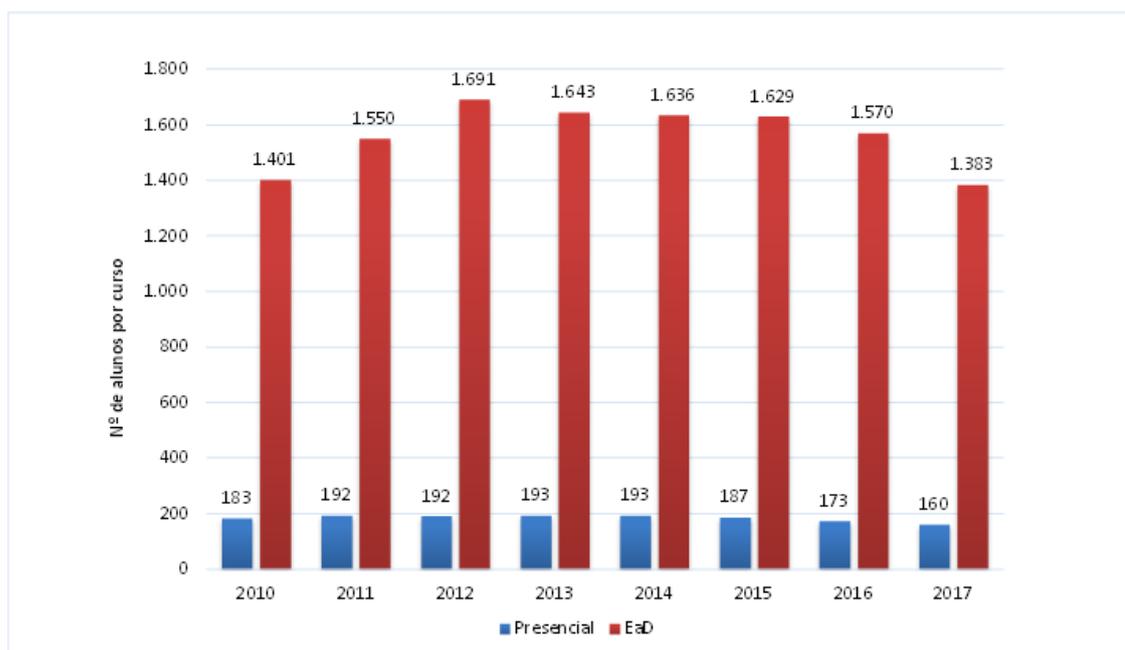
Figura 11: Crescimento do número de alunos de ciências contábeis de 2010 a 2017



Fonte: Elaborado pelo autor com base nos dados da pesquisa.

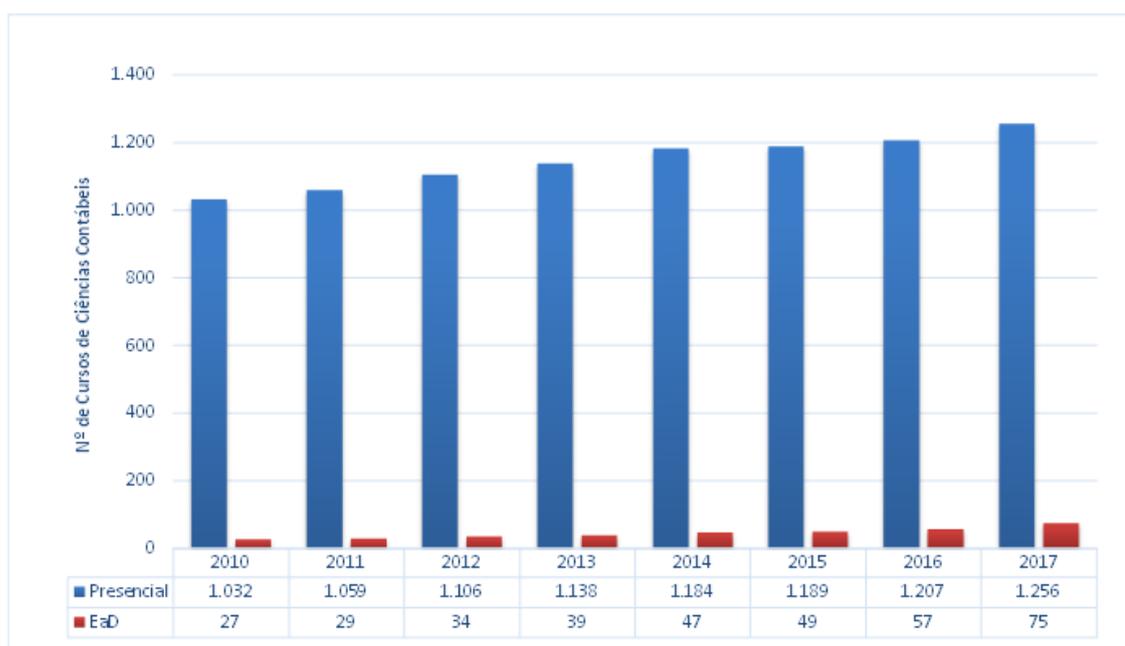
Outro aspecto que se destaca ao analisar a dinâmica das modalidades de ensino, é o fato de haver uma tendência de queda da densidade do número de alunos por curso, tanto nos cursos à distância quanto nos presenciais, conforme pode ser visualizado na Figura 12. Isto pode ser explicado pelo crescimento da oferta de cursos de ciências contábeis em ambas as modalidades, conforme evidenciado na Figura 13, uma vez que, conforme demonstrado na Figura 11, o número total de alunos apresentou um incremento praticamente constante, durante o período pesquisado.

Figura 12: Número de alunos por curso de ciências contábeis - 2010 a 2017



Fonte: Elaborado pelo autor com base nos dados da pesquisa.

Figura 13: Número de cursos de ciências contábeis por modalidade - 2010 a 2017



Fonte: Elaborado pelo autor com base nos dados da pesquisa.

4.2.2 Análise do Modelo Econométrico

Para o cálculo econométrico da dimensão IES, foi utilizado um modelo de regressão linear com dados em painel, conforme descrito no capítulo de metodologia.

Este modelo de regressão apresenta uma característica peculiar, por ser constituído de duas dimensões: uma temporal, e uma espacial.

Em relação à análise de regressão com dados em painel, três modelos podem ser aplicados: dados empilhados (“pooled”), efeitos fixos e efeitos aleatórios. No sub-tipo de dados empilhados, se considera que o intercepto do modelo e seus coeficientes angulares são constantes ao longo do tempo e espaço, e o termo de erro carrega a diferença no tempo entre os indivíduos; os modelos de efeitos fixos e efeitos aleatórios buscam calcular a estimativa de efeitos não observados. No modelo de efeitos fixos, a regressão considera que os coeficientes angulares são constantes e o intercepto varia entre os indivíduos; e em relação aos efeitos aleatórios, a regressão assume que o intercepto possui um valor médio comum entre os indivíduos e os coeficientes angulares variam ao longo do tempo (WOOLDRIDGE, 2015).

Para fins de comparação, foram apurados os resultados do modelo econométrico aplicando os três tipos de efeitos de dados em painel, dos quais obteve-se os resultados constantes da Tabela 7. As colunas rotuladas como (1), (2) e (3) referem-se, respectivamente, aos modelos de dados empilhados, efeitos fixos e efeitos aleatórios. Os cálculos foram feitos utilizando-se o tipo de covariância como “robusto” para contornar os efeitos indesejados de heteroscedasticidade da amostra.

Conforme é possível observar nos resultados obtidos, as variáveis quantitativas do modelo, representadas por VAR_CONEX, TAM e VAR_VAR, possuem significância estatística em um intervalo de confiança de 99% nos três modelos de dados em painel, e a variável TAM possui significância estatística em intervalo de confiança de 95% nos modelos de efeitos fixos e efeitos aleatórios. Por outro lado, tanto a modalidade de ensino como a natureza jurídica das IES não apresentaram significância estatística, e portanto, não se pode afirmar que impactaram na migração de alunos neste modelo.

Para selecionar o melhor modelo às características da amostra de dados, foram feitos testes de multicolinearidade e heteroscedasticidade. Para a verificação de multicolinearidade, foi aplicado o teste VIF, onde observou-se que os dados não apresentaram alto grau de correlação (Tabela 8). Com relação à heteroscedasticidade, verificou-se, por meio do teste de Breusch-Pagan, que a amostra assume uma condição de heteroscedástica, tendo em vista que o p-Valor resultou em um valor próximo a 0 (zero), o que indica a rejeição da hipótese nula (dados homoscedásticos). Em virtude da presença de heteroscedasticidade na base de dados, os cálculos foram feitos utilizando-se um parâmetro de robustez para estimativa da covariância.

Em seguida, foi aplicado o teste de Hausmann para avaliar a consistência entre os estimadores, e definir o modelo de dados em painel a ser utilizado. A hipótese nula do teste de Hausmann indica que o modelo apropriado para a análise é o de efeitos aleatórios, ao passo que a rejeição da hipótese nula aponta para a utilização de efeitos fixos. Conforme pode ser observado na Tabela 9, o p-valor do teste Hausmann

Tabela 7: Dados em painel da variação de entrantes nos cursos de ciências contábeis

Erros-padrão estão mostrados entre parênteses.

Variáveis	(1)	(2)	(3)
Dep. Variable	VAR_ENT	VAR_ENT	VAR_ENT
Estimator	PooledOLS	PanelOLS	RandomEffects
No. Observations	9528	9528	9528
Cov. Est.	Clustered	Robust	Robust
R-squared	0,6425	0,6382	0,6425
R-Squared (Within)	0,6371	0,6382	0,6371
R-Squared (Between)	0,6489	0,6275	0,6489
R-Squared (Overall)	0,6425	0,6407	0,6425
F-statistic	3.423,3000	2.962,2000	3.423,3000
P-value (F-stat)	0,0000	0,0000	0,0000
VAR_CONEX	1,9529 *** (0,0731)	1,9296 *** (0,0597)	1,9529 *** (0,0560)
VAR_VAR	0,0222 *** (0,0053)	0,0223 *** (0,0054)	0,0222 *** (0,0054)
TAM	0,0012 *** (0,0004)	0,0015 ** (0,0006)	0,0012 ** (0,0005)
MOD	-1,3036 (1,2531)	-3,4942 ** (1,5081)	-1,3036 (0,9955)
NAT	-0,2649 (0,1872)	0,4964 (0,4802)	-0,2649 (0,1743)
Efeitos	-	Entidade	-
AIC	6,598E+04	6,593E+07	6,598E+04
BIC	6,602E+04	6,596E+07	6,605E+04
Log-Likelihood	-3,299E+04	-3,296E+04	-3,298E+04
F-Statistic	289,7	269,40	169,9
Ombibus	8628,68	8617,784	8618,944
Prob(Omnibus)	0,000	0,000	0,000
Prob-(F-Statistic)	1,74E-289	2,85E-270	7,49E-300
Skew	2,855	2,833	2,849
Kurtosis	333,879	345,584	333,127
Durbin-Watson	1,972	1,958	1,975
Jarque-Bera	43476861,748	46606094,035	43279432,166
Prob(JB)	0,00	0,00	0,00
VIF médio	1,087303	1,087303	1,087303

*** *p*-valor inferior ao nível de significância de 0,01; ** *p*-valor inferior ao nível de significância de 0,05.

Fonte: Elaborado pelo autor com base nos dados da pesquisa.

Tabela 8: Teste VIF para multicolinearidade

Variável	VIF
VAR_CONEX	1,132336
VAR_VAR	1,004279
TAM	1,001208
MOD	1,033797
NAT	1,197184

Fonte: Elaborado pelo autor com base nos dados da pesquisa.

Tabela 9: Teste Hausmann

Qui-quadrado	9,041491
Graus de liberdade	6
p-Valor	0,171258

Fonte: Elaborado pelo autor.

é superior a 0,05, o que portanto não rejeita a hipótese nula e indica a utilização dos dados em painel com efeitos aleatórios como modelo mais apropriado.

Após definido o tipo de efeito para o modelo de dados em painel, e para uma melhor abordagem dos efeitos da dinâmica competitiva do mercado, utilizou-se a informação de região do território brasileiro como variável *dummy*, sendo a região sudeste escolhida como base de referência. Assim, é possível observar o efeito da aplicação do modelo por diferentes perspectivas regionais, conforme apresentado na Tabela 10.

É importante destacar que, como a pesquisa teve a abrangência dos dados de toda a população, e não apenas de uma amostra, isto possibilitou deixar em segundo plano a hipótese p-valor significativo, e dar ênfase nos valores dos coeficientes *beta*.

Em relação à variação dos entrantes, com base na região sudeste, é possível observar que praticamente todas as regiões apresentam coeficientes negativos, tanto em relação à região dos alunos, quanto em relação à variação das conexões e variação da variância. As exceções são as regiões Norte, para as variáveis de região e variação da variância, e a região Sul, para a variável de variação da variância. Estes coeficientes indicam que, enquanto a região sudeste registra a entrada de alunos, as demais regiões registram um movimento contrário. No caso da variável relacionada à região da IES, apenas a região Norte apresentou coeficiente positivo; já em relação à variável de variação de conexões, todas as regiões apresentaram coeficientes negativos; por fim, a variação da variância resultou nas regiões Sul e Norte com coeficientes positivos (Norte praticamente coeficiente nulo), e Centro-oeste e Nordeste com coeficientes negativos.

Em termos de concorrência e demanda, é possível constatar que a variação positiva da migração, que representa o movimento líquido de entrada de alunos nas IES,

Tabela 10: Dados em painel com efeitos aleatórios - Dimensão IES

Parâmetros	Valores	
Dependente	VAR_ENT	
Estimador	Efeitos Aleatórios	
Nº Observações	9.528	
Estimador Covariância	Robust	
R-squared	0,7297	
R-Squared (Within)	0,6891	
R-Squared (Between)	0,7283	
R-Squared (Overall)	0,7297	
F-statistic	1.510,1000	
P-value (F-stat)	0,0000	
Log-Likelihood	-3,165E+04	
F-Statistic (robust)	133,85	
Variáveis	Coeficiente	Erro Padrão
VAR_CONEX	1,9972***	0,1129
VAR_VAR	0,0435***	0,0087
MOD	-1,0888	0,7659
NAT	-0,0548	0,1591
TAM	0,0013***	0,0005
Região[T.SUL]	-0,1177	0,1315
Região[T.CENTRO-OESTE]	-0,3280	0,2523
Região[T.NORDESTE]	-0,0810	0,1967
Região[T.NORTE]	0,5014	0,2815
Região[T.SUL]:VAR_CONEX	-0,1504	0,1868
Região[T.CENTRO-OESTE]:VAR_CONEX	-0,1185	0,1410
Região[T.NORDESTE]:VAR_CONEX	-0,1465	0,1475
Região[T.NORTE]:VAR_CONEX	-0,1109	0,2098
Região[T.SUL]:VAR_VAR	0,0838***	0,0193
Região[T.CENTRO-OESTE]:VAR_VAR	-0,0294***	0,0092
Região[T.NORDESTE]:VAR_VAR	-0,0183*	0,0100
Região[T.NORTE]:VAR_VAR	0,0061	0,0101

*** p-valor inferior ao nível de significância de 0,01; ** p-valor inferior ao nível de significância de 0,05; * p-valor inferior ao nível de significância de 0,1

Fonte: Elaborado pelo autor.

possui alta relação com a variação positiva das conexões. Por outro lado, o tamanho das IES possui um baixo coeficiente *beta*, quase nulo, o que minimiza a importância deste fator na migração de alunos. Assim sendo, é possível afirmar que o nível de concorrência é mais afetado e determinado pelo maior número de conexões com outras IES, do que pelo tamanho dos seus concorrentes. Este achado vai ao encontro da hipótese H2 desta pesquisa, tendo em vista as características da estrutura de mer-

cado da educação superior no Brasil. O processo de mercantilização da educação no Brasil, e a consequente formação de oligopólios criados pela consolidação de IES privadas em poucos grupos de controladores, favoreceu o aumento da procura por cursos superiores, e, por conseguinte, uma dinâmica mais acentuada no movimento migratório de alunos. Ainda que a pesquisa trate cada IES separadamente pela sua geolocalização, independentemente se pertencem ao mesmo grupo econômico controlador, ou não, fica evidenciado que a migração possui maior grau de importância tanto maior forem as conexões interinstitucionais, em comparação com as migrações em regiões com maior concentração em torno de uma ou poucas IES, e que possuem menos conexões.

4.2.3 Síntese da Seção de Análise de Resultados

Após analisar os resultados obtidos nos dois modelos econométricos, é possível afirmar que as hipóteses de pesquisa H1 e H2 foram confirmadas, e, além disso, são complementares, à medida em que o grau de concentração do mercado de IES influencia no processo decisório do aluno no que tange à transferência institucional.

De um lado, o modelo confirma a hipótese H1, à medida em que as variações positivas nos benefícios geo-socioeconômicos proporcionam maior probabilidade do discente mudar de curso, ou seja, há uma relação positiva de tais benefícios na retenção dos estudantes. Dentre os fatores analisados, o incremento nos auxílios econômicos é o que representa maior utilidade esperada pelo aluno, influenciando de forma mais preponderante na sua decisão.

Pelo lado da concorrência entre as instituições, conclui-se que a quantidade de migrações possui uma relação direta com o número de competidores no corte regional, independentemente do tamanho dos concorrentes. Os achados confirmam a expectativa de H2, pois observa-se que o número de migrações é mais elevado nos mercados onde o número de concorrentes é maior. Mais do que isso, evidencia-se também em que medida este fenômeno ocorre. Esta conclusão é importante e pode ajudar os gestores das IES no seu processo de formulação de ações estratégicas. Ao analisarem o ambiente competitivo, é mais importante direcionar esforços nos locais onde há um maior número de concorrentes, em detrimento aos concorrentes de maior tamanho que estejam em um ambiente de maior concentração.

5 CONSIDERAÇÕES FINAIS

A presente pesquisa buscou abordar a relação entre as decisões dos alunos de ciências contábeis no Brasil e a concorrência entre as IES no mercado de ensino superior.

A pesquisa analisou a questão da migração de alunos e competitividade do mercado de educação superior, sob uma ótica pouco explorada nos trabalhos científicos sobre o mercado brasileiro publicados até o momento, pois foi integralmente feita a partir dos microdados do CES/INEP, e com um intervalo de tempo dilatado. Para que isso fosse concretizado, fez-se necessário o uso de ferramentas de processamento de dados tipo *big data*, visto o volume e a velocidade de processamento dos dados processados.

Para que o objetivo geral da pesquisa fosse alcançado, foi utilizado como base o fluxo de migração dos alunos de ciências contábeis no período entre 2009 e 2017, bem como os dados pertinentes às IES no mesmo período. Para se obter estas informações, foi utilizada a base dos microdados do CES/INEP e, após uma série de procedimentos para preparar a base de dados e adequá-la aos requisitos da pesquisa, foram aplicados dois modelos econométricos distintos para a obtenção dos resultados da pesquisa. Um dos modelos foi direcionado para o estudo da probabilidade de migração do aluno entre cursos de Ciências Contábeis (dimensão aluno), e outro modelo para abordar o nível de concorrência entre as IES (dimensão IES).

A primeira dimensão pesquisada, que diz respeito às probabilidades de migração dos discentes, foi submetida a um modelo econométrico de regressão logística, considerando como variáveis independentes as variações dos seguintes parâmetros: fatores sociais, fatores econômicos, distância e conceito da IES. Neste recorte, foi possível observar que variações nos fatores econômicos e sociais são mais relevantes na probabilidade dos alunos migrarem de curso, em detrimento da variação da nota ENADE e da variação da distância da cidade do aluno para a IES. A variação positiva no fator econômico dos alunos se mostrou como a variável que mais maximizou a utilidade esperada, potencializando a decisão do aluno pela transferência interinstitucional.

Com relação à segunda dimensão pesquisada, a qual trata da avaliação do nível de concorrência entre as IES, optou-se pela aplicação de um modelo de dados em painel com efeitos aleatórios. O principal achado desta análise, relaciona-se à constatação de que o nível de concorrência foi mais afetado pela quantidade de IES conectadas à IES pesquisada, do que pelo tamanho dos concorrentes. Este achado é importante, pois indica que a estratégia de crescimento de uma IES deveria passar pelo direcionamento de esforços com foco num leque mais amplo de IES concorrentes (novas conexões), por menores que sejam, em detrimento de focar em um ou em

poucos concorrentes que já estejam estabelecidos. Foi possível observar a evolução do mercado de educação superior ao longo do período pesquisado, tendo em vista os números crescentes de alunos, movimentos migratórios e a quantidade de IES. Também foi possível constatar a complexidade da rede de transferência interinstitucional, dada a quantidade de conexões existentes entre as IES pesquisadas, caracterizadas pelas arestas formadas por meio da malha de transferências interinstitucionais.

Outros autores também utilizaram mais recentemente os microdados do CES para estudar a movimentação discente, como é o caso de Carvalho e Jakob (2021), que propuseram apresentar o CES como fonte alternativa dos estudos da mobilidade estudantil ao nível de graduação. Porém, em relação a este estudo, os dados usados foram basicamente sintéticos caracterizados tão somente pela quantificação intra e interestadual, não detalhando ao nível individual dos discentes. Outro estudo que utilizou os microdados foi o de Vizzotto (2020), o qual pesquisou os cursos de licenciatura em Física no país. Entretanto, também neste caso, a pesquisa se ateve a quantificar o número de cursos existentes, segmentando-os por região e natureza jurídica, sem adentrar em detalhes pertinentes ao comportamento dos alunos na malha de IES. Por sua vez, Silva Filho et al. (2007) propôs um modelo preditivo para a evasão dos alunos de EaD da Universidade Federal de Santa Catarina. Este estudo não utilizou os microdados do CES, mas sim, os dados disponibilizados pela própria universidade. Mesmo tendo esta diferença na origem dos dados, o estudo se assemelha à presente pesquisa, pois também utilizou uma regressão logística para determinar a probabilidade de evasão dos alunos. No caso, os achados tiveram semelhanças limitadas apenas à idade e distância da cidade do aluno, uma vez que as demais variáveis são diferentes das que foram utilizadas neste estudo. Há vários outros estudos que pesquisaram principalmente sobre a evasão nos cursos superiores, mas invariavelmente se limitaram a dados sintéticos quantitativos e qualitativos, ou então valendo-se de outros métodos de coleta de dados, como questionários e bases de dados específicas de determinada IES.

Em termos de limitações da presente pesquisa, convém destacar alguns pontos que, em um momento futuro, podem ser desenvolvidos ou adaptados em outros estudos. O primeiro ponto a destacar, é a própria base de dados do INEP. Por exemplo, o período pesquisado ficou restrito ao ano de 2017, visto que os anos seguintes de 2018 e 2019, o INEP modificou completamente a codificação dos alunos, fazendo com que não mais fosse possível rastrear até a sua origem no ano anterior, e assim, impossibilitando mapear a sua migração. Posteriormente, o INEP suprimiu completamente os microdados relacionados aos alunos, inviabilizando completamente o processamento destes dados. O segundo ponto de destaque é a falta de alguns dados no CES que poderiam ter enriquecido mais a pesquisa: os locais de origem dos alunos poderiam ser substituídos pela cidade de última residência, ao invés da cidade de nascimento. Tam-

bém se destaca neste campo uma quantidade relevante de alunos que não possuem esta informação nos microdados. Por fim, a pesquisa limitou-se aos alunos do curso de ciências contábeis, tendo em vista a escolha do pesquisador e o pertencimento ao programa de pós-graduação ao qual pertence.

No que tange a sugestões para futuras pesquisas, destaca-se a possibilidade de incremento nas variáveis explicativas, incluindo uma variável que represente o valor de preço médio por aluno de cada IES. Esta variável ajudaria sobremaneira a enriquecer os modelos, tanto na visão de decisão dos alunos, quanto na concorrência entre as IES. Seria interessante haver o incremento dessa informação em uma nova versão do CES. Convém destacar que o CES é bastante dinâmico, e sofreu inúmeras modificações desde sua implantação. Também sugere-se a adoção de novas métricas com base nas ideias lançadas nesta pesquisa, agregando outras informações que possam estar ao alcance de pesquisadores, independentemente dos microdados do CES. Sugere-se também, a aplicação dos modelos aos demais cursos, além de ciências contábeis. Isto poderia servir para fins de comparação do comportamento entre os cursos, e até mesmo incluir situações envolvendo a evasão do aluno, e não somente a migração.

Ao encerrar a presente pesquisa, espera-se que haja a contribuição do que aqui foi exposto, tanto em termos acadêmicos, especialmente na linha de pesquisa de gestão e controle do PPG Contábeis, quanto em termos de sua aplicação prática na análise da competição no mercado da educação superior no Brasil. O entendimento desses comportamentos é um componente fundamental para uma gestão de excelência neste mercado.

REFERÊNCIAS

- ALSHMEMRI, M.; SHAHWAN-AKL, L.; MAUDE, P. Herzberg's two-factor theory. **Life Science Journal**, v. 14, n. 5, p. 12–16, 2017.
- AMANCIO, D. R. A complex network approach to stylometry. **PloS one**, v. 10, n. 8, p. e0136076, 2015.
- ARNOTT, D.; GAO, S. Behavioral economics for decision support systems researchers. **Decision Support Systems**, v. 122, p. 113063, 2019.
- BERGMANN, M.; BRÜCK, C.; KNAUER, T.; SCHWERING, A. Digitization of the budgeting process: determinants of the use of business analytics and its effect on satisfaction with the budgeting process. **Journal of Management Control**, p. 1–30, 2020.
- BITTENCOURT, H. R.; CLARKE, R. T. Um classificador baseado na Discriminação Logística: vantagens e desvantagens. **Anais X SBSR**, p. 1217–1223, 2001.
- BORGES, E. R. K. **Microdados Censo da Educação Superior - INEP**. 2021. Acesso em: 2021-10-05, Disponível em:< <https://github.com/erkborges/mestrado>>.
- BREITENBACH, M.; ALVES, T. W.; DIEHL, C. A. Indicadores financeiros aplicados à gestão de instituições de ensino de educação básica. **Contabilidade vista & revista**, v. 21, n. 3, p. 167–203, 2010.
- BUCHANAN, L.; O CONNELL, A. A brief history of decision making. **Harvard business review**, v. 84, n. 1, p. 32, 2006.
- CARVALHO, M. M. de; JAKOB, A. A. O uso do Censo da Educação Superior do MEC/INEP como fonte de dados alternativa para o estudo das migrações estudantis. **E-book**, p. 569–579, 2021.
- CHAVES, V. L. J.; REIS, L. F.; GUIMARÃES, A. R. Dívida pública e financiamento da educação superior no Brasil. **Acta Scientiarum. Education**, v. 40, n. 1, p. 1–12, 2018.
- CIMINI, G.; SQUARTINI, T.; SARACCO, F.; GARLASCHELLI, D.; GABRIELLI, A.; CALDARELLI, G. The statistical physics of real-world networks. **Nature Reviews Physics**, v. 1, n. 1, p. 58–71, 2019.
- CLARKE, J.; JANDIK, T.; MANDELKER, G. The efficient markets hypothesis. **Expert financial planning: Advice from industry leaders**, v. 7, n. 3/4, p. 126–141, 2001.
- COIMBRA, C. L.; SILVA, L. B.; COSTA, N. C. D. A evasão na educação superior: definições e trajetórias. **Educação e Pesquisa**, v. 47, 2021.
- CORBUCCI, P. R.; KUBOTA, L. C.; MEIRA, A. P. B. **Reconfiguração estrutural ou concentração do mercado da educação superior privada no Brasil?** 2016. Acesso em: 2010-09-30, Disponível em:< http://repositorio.ipea.gov.br/bitstream/11058/7080/1/Radar_n46_reconfiguracao.pdf>.

- CORDEIRO, C. M.; VOLDNES, G. An Integrated Economic Sociology Approach to Market-as-Network: the example of a shared business environment between norway and china. **Journal of East-West Business**, p. 1–25, 2021.
- DIEHL, C. A. **Proposta de um sistema de avaliação de custos intangíveis**. 1997. Acesso em: 2021-09-17, Disponível em:< <https://www.lume.ufrgs.br/bitstream/handle/10183/121788/000199236.pdf?sequenc>>.
- DINIZ, J. A. R.; OLIVEIRA, J. F. de; LIMA, D. d. C. B. P. A mercantilização da educação superior no Brasil: financeirização e oligopolização. **Revista Educação em Questão**, v. 59, n. 61, 2021.
- ERICSON, K. M.; LAIBSON, D. Intertemporal choice. In: **Handbook of Behavioral Economics: applications and foundations 1**. Elsevier, 2019. v. 2, p. 1–67.
- FURTADO, V. V. A.; ALVES, T. W. **Fatores determinantes da evasão universitária: uma análise com alunos da unisinos**. 2012. Acesso em: 2021-09-17, Disponível em:< http://repositorio.ufc.br/bitstream/riufc/7479/1/2012_art_vvafurtado.pdf>.
- GIL, A. C. et al. **Como elaborar projetos de pesquisa**. 2017. Acesso em: 2021-07-26, Disponível em:< <https://integrada.minhabiblioteca.com.br/#/books/9788597012934/>>.
- GUJARATI, D. N.; PORTER, D. C. **Econometria básica-5**. Amgh Editora, 2011.
- HOED, R. M. **Análise da evasão em cursos superiores: o caso da evasão em cursos superiores da área de computação**. 2016. Acesso em: 2021-10-05, Disponível em:< https://repositorio.unb.br/bitstream/10482/22575/1/2016_RaphaelMagalh%C3%A3esHoed.pdf>.
- IBGE, I. B. d. G. e. E. **Dados IBGE**. 2021. Acesso em: 2021-10-05, Disponível em:< <https://www.ibge.gov.br/estatisticas/downloads-estatisticas.html>>.
- INEP, I. N. d. E. e. P. E. A. T. **Microdados ENADE**. 2021. Acesso em: 2021-10-05, Disponível em:< <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enade>>.
- JALILI, M.; PERC, M. Information cascades in complex networks. **Journal of Complex Networks**, v. 5, n. 5, p. 665–693, 2017.
- KAHNEMAN, D.; TVERSKY, A. Tversky. A.(1979). Prospect theory: an analysis of decision under risk. **Econometrica**, v. 47, n. 2, p. 263, 1979.
- KAUARK, F. d. S.; MANHÃES, F. C.; MEDEIROS, C. H. **Metodologia da pesquisa: um guia prático**. Via Litterarum, 2010.
- KROHLING, R. A.; SOUZA, T. T. de. Combining prospect theory and fuzzy numbers to multi-criteria decision making. **Expert Systems with Applications**, v. 39, n. 13, p. 11487–11493, 2012.
- LI, L.; CHAGAS, A. L. S. Efeitos do SISU sobre a migração e a evasão estudantil. **XV ENCONTRO NACIONAL DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS REGIONAIS E URBANOS. Anais...** São Paulo: Aber, 2017.

LOURENÇO, D.; SÁ, C. Spatial competition for students: what does (not) matter? **The Annals of Regional Science**, v. 63, n. 1, p. 147–162, 2019.

MALKIEL, B. G. The efficient market hypothesis and its critics. **Journal of economic perspectives**, v. 17, n. 1, p. 59–82, 2003.

MARQUES, R. M.; XIMENES, S. B.; UGINO, C. K. Governos Lula e Dilma em matéria de seguridade social e acesso à educação superior. **Brazilian Journal of Political Economy**, v. 38, p. 526–547, 2018.

MARTINS, M. P.; MIGUEIS, V. L.; FONSECA, D.; GOUVEIA, P. D. Previsão do abandono acadêmico numa instituição de ensino superior com recurso a data mining. **Revista Ibérica de Sistemas e Tecnologias de Informação**, p. 188–203, 2020.

MATA, A. S. da. Complex Networks: a mini-review. **Brazilian Journal of Physics**, p. 1–15, 2020.

METZ, J.; CALVO, R.; SENO, E. R.; ROMERO, R. A. F.; LIANG, Z. et al. **Redes complexas: conceitos e aplicações**. 2007. Acesso em: 2010-08-31, Disponível em:<http://repositorio.icmc.usp.br/bitstream/handle/RIICMC/6720/Relatório%20Técnico_290_2007.pdf?sequence=3>.

MINUSSI, J. A.; DAMACENA, C.; NESS JR, W. L. Um modelo de previsão de solvência utilizando regressão logística. **Revista de Administração Contemporânea**, v. 6, p. 109–128, 2002.

OKASHA, S. On the interpretation of decision theory. **Economics and Philosophy**, v. 32, n. 3, p. 409, 2016.

PEREIRA, T. L.; BRITO, S. H. A. d. As aquisições e fusões no ensino superior privado no Brasil (2005–2013). In: CONGRESSO IBERO-AMERICANO DE POLÍTICAS E ADMINISTRAÇÃO DA EDUCAÇÃO; CONGRESSO LUSO-BRASILEIRO DE POLÍTICA E ADMINISTRAÇÃO DA EDUCAÇÃO, 2014. **Anais...** 2014. v. 6.

PILATTI, L. A. Qualidade de vida no trabalho e a teoria dos dois fatores de Herzberg: possibilidades-limite das organizações. **Qualidade de vida no ambiente corporativo**, v. 1, 2004.

PRATOLO, S.; SOFYANI, H.; ANWAR, M. Performance-based budgeting implementation in higher education institutions: determinants and impact on quality. **Cogent Business & Management**, v. 7, n. 1, p. 1786315, 2020.

READ, D.; MCDONALD, R.; HE, L. Intertemporal choice. **London School of Economics and Political Science**, 2004.

ROSSETTI, J. P. Introdução à economia. 21. **Edição. São Paulo: Atlas**, 2016.

RUGGERI, K.; ALÍ, S.; BERGE, M. L.; BERTOLDO, G.; BJØRNDAL, L. D.; CORTIJOS-BERNABEU, A.; DAVISON, C.; DEMIĆ, E.; ESTEBAN-SERNA, C.; FRIEDEMANN, M. et al. Replicating patterns of prospect theory for decision under risk. **Nature human behaviour**, v. 4, n. 6, p. 622–633, 2020.

SA, C.; FLORAX, R. J.; RIETVELD, P. Determinants of the regional demand for higher education in the Netherlands: a gravity model approach. **Regional Studies**, v. 38, n. 4, p. 375–392, 2004.

SALLABERRY, J. D.; SANTOS, E. A. d.; PEREIRA, P. H. d. S. M.; BRUM, D. L. Determinantes do gasto público em Educação nos municípios do estado do Mato Grosso do Sul. **Ensaio: Avaliação e Políticas Públicas em Educação**, 2022.

SALMAN, S. Market-as-Network versus Relationship Marketing: clarifying some crucial differences. **EDITORIAL**, v. 29, p. 43–55, 2020.

SAMPAIO, H. Setor privado de ensino superior no Brasil: crescimento, mercado e estado entre dois séculos. **Ensino Superior: Expansão e Democratização. Rio de Janeiro: Editora**, v. 7, p. 103–126, 2013.

SILVA FILHO, R. L. L.; MOTEJUNAS, P. R.; HIPÓLITO, O.; LOBO, M. B. d. C. M. A evasão no ensino superior brasileiro. **Cadernos de pesquisa**, v. 37, p. 641–659, 2007.

SILVA, G. P. d. Análise de evasão no ensino superior: uma proposta de diagnóstico de seus determinantes. **Avaliação: Revista da Avaliação da Educação Superior (Campinas)**, v. 18, p. 311–333, 2013.

SOUZA, A. E. de; SÁ SANTOS, L. M. de; MENEZES LARRUSCAIM, I. de; NOBREGA BESARRIA, C. da. Determinantes do Desempenho no ENEM na Região Nordeste: uma análise de dados em painel do período de 2015 a 2019. **Revista Brasileira de Estudos Regionais e Urbanos**, v. 15, n. 4, p. 690–711, 2021.

STEELE, K.; STEFÁNSSON, H. O. **Decision theory**. 2015.

THOMA, J. **Decision theory**. 2019.

TONTINI, G.; WALTER, S. A. Pode-se identificar a propensão e reduzir a evasão de alunos?: ações estratégicas e resultados táticos para instituições de ensino superior. **Avaliação: Revista da Avaliação da Educação Superior (Campinas)**, v. 19, p. 89–110, 2014.

VARELA, L. M.; ROTUNDO, G.; AUSLOOS, M.; CARRETE, J. Complex network analysis in socioeconomic models. In: **Complexity and Geographical Economics**. Springer, 2015. p. 209–245.

VIZZOTTO, P. A. Um panorama sobre as licenciaturas em Física do Brasil: análise descritiva dos microdados do censo da educação superior do inep. **Revista Brasileira de Ensino de Física**, v. 43, 2020.

WOOLDRIDGE, J. M. **Introductory econometrics: a modern approach**. Cengage learning, 2015.