

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE ECONOMIA
ECONOMIA DA INDÚSTRIA E DA TECNOLOGIA DESTA UNIVERSIDADE
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

THIAGO DRUMMOND DE MENDONÇA
GIUDICI

**ANÁLISE DA DESIGUALDADE
HORIZONTAL NO ENSINO
SUPERIOR BRASILEIRO**

Rio de Janeiro
2018

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE ECONOMIA
ECONOMIA DA INDÚSTRIA E DA TECNOLOGIA DESTA UNIVERSIDADE
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

THIAGO DRUMMOND DE MENDONÇA
GIUDICI

**ANÁLISE DA DESIGUALDADE
HORIZONTAL NO ENSINO
SUPERIOR BRASILEIRO**

Dissertação de Mestrado apresentado ao Programa de Pós-Graduação em Economia da Indústria e Tecnologia, Instituto de Economia, Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de Mestre em Economia

Rio de Janeiro
2018

FICHA CATALOGRÁFICA

G537 Giudici, Thiago Drummond de Mendonça
Análise da desigualdade horizontal no ensino superior brasileiro / Thiago Drummond de Mendonça Giudici. - 2018.
93 p. ; 31 cm.

Orientador: Valéria Lúcia Pero.

Dissertação (mestrado) – Universidade Federal do Rio de Janeiro, Instituto de Economia, Programa de Pós-Graduação em Economia da Indústria e da Tecnologia, 2018.

Bibliografia: f. 67 – 72.

1. Ensino superior. 2. Desigualdade horizontal. 3. Escolha do curso de graduação. I. Pero, Valéria Lúcia, orient. II. Universidade Federal do Rio de Janeiro. Instituto de Economia. III. Título.

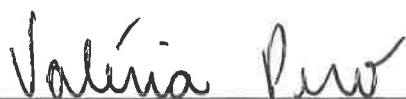
CDD 378.16

THIAGO DRUMMOND DE MENDONÇA GIUDICI

**Análise da desigualdade horizontal no Ensino Superior
Brasileiro**

Dissertação de Mestrado submetida ao Corpo Docente do Departamento Economia - PPGE da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários para obtenção do título de Mestre em Economia.

Aprovado em: Rio de Janeiro, 29 de AGOSTO de 2018.



Profa. Dr. Valéria Lúcia Pero (Orientador)



Profa. Dr. Danielle Carusi Machado



Prof. Dr. Romero Cavalcanti Barreto da Rocha

Rio de Janeiro
2018

AGRADECIMENTOS

Agradeço a minha orientadora, Profa. Dra. Valéria Pero pela atenção e paciência em me ensinar não apenas em relação a pesquisa mas também a como escrever uma dissertação.

Ao instituto de Economia da Ufrj que me abraçou e me ajudou a superar todas as dificuldades devido a minha graduação ser em engenharia. Aprendi muito e quero continuar aprendendo.

A minha namorada Letícia Corrêa que me acompanhou de perto e me ajudou em todos os momentos de dificuldade.

Por fim, aos meus pais que sempre me deram apoio e dividiram minhas alegrias e angústias nesses últimos anos de mestrado.

RESUMO

Giudici, Thiago Drummond de Mendonça. **Análise da desigualdade horizontal no Ensino Superior Brasileiro**. 2018. 93 f. Dissertação (Mestrado em Economia) - PPGE, Instituto de Economia, Instituto Economia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018.

Ao longo da última década o ensino superior no Brasil passou por uma grande etapa de expansão e diversificação. Em 2004 com um total de 4.278.133 matrículas saltou para 8.033.574 em 2015, um aumento de 83% em 11 anos (Censo Educação Superior, Inep). Programas governamentais como Fies, ProUni, Reuni e as políticas de cotas sociais vêm exercendo um impacto importante nessa expansão da educação superior e na democratização no campus brasileiro com maior participação de alunos de origens socioeconômicas menos privilegiadas.

No entanto, vê-se uma grande desigualdade horizontal no ensino superior, isto é, alunos de estratos sociais menos privilegiados tendem a se concentrar em cursos com menor perspectiva de retorno financeiro.

Essa dissertação quer verificar como esse viés de distribuição se comportou nos últimos anos. Para tanto, usa-se a base socioeconômica do Enade (de 2004 a 2015), afim de verificar como se distribuem os alunos com distintas características socioeconômicas nos cursos oferecidos pelo Ensino Superior Brasileiro e busca-se o melhor modelo de regressão logístico para avaliar a desigualdade horizontal na formação do curso do ensino superior. Para atingir esses objetivos, dividi-se os cursos em grupos com base no retorno financeiro esperado e estima-se a probabilidade dos alunos com distintas características socioeconômicas frequentarem determinado grupo de cursos.

Os resultados mostraram que as características socioeconômicas são significativas e que a elevada desigualdade horizontal no ensino superior brasileiro se manteve entre os anos de 2004 a 2015.

Palavras-chave: Desigualdade horizontal, Ensino superior, Escolha do curso, Distribuição dos alunos, Escolha do curso.

ABSTRACT

Giudici, Thiago Drummond de Mendonça. **Análise da desigualdade horizontal no Ensino Superior Brasileiro**. 2018. 93 f. Dissertação (Mestrado em Economia) - PPGE, Instituto de Economia, Insituto de Economia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018.

Over the last decade, higher education in Brazil has undergone a great stage of expansion and diversification. In 2004 with a total of 4.278.133 enrollments jumped to 8.033.574 in 2015, an increase of 83 % in 11 years (Higher Education Census, Inep). Government programs such as Fies, Prouni, Reuni and social quota policies have been having a major impact on this expansion of higher education and democratization on the Brazilian campus with a greater participation of students from less privileged socioeconomic backgrounds.

However, there is a great horizontal inequality in higher education, that is, students from less privileged social status tend to focus on courses with less prospect of financial return.

This dissertation wants to verify how this distribution bias has behaved in recent years. In order to do so, we use the socioeconomic data of Enade between 2004 and 2015, and we verify how students with different socioeconomic characteristics are distributed in the courses offered by the Brazilian Higher Education, and we search for the best logistic regression model to evaluate horizontal inequality in the course of higher education. To achieve these objectives, the courses were divided into groups based on the expected financial return and the probability of students with different socioeconomic characteristics is estimated to attend a certain group of courses.

The results showed that the socioeconomic characteristics are significant and the high horizontal inequality in college, has remained in recent years.

Keywords: Horizontal inequality in higher education, Choice of college major, college major gap, Heterogeneous human capital.

LISTA DE TABELAS

2.1	Fatores que influenciam a escolha do curso	13
3.1	Quantidade de alunos nos cursos selecionados por ciclo	21
3.2	Renda familiar, escolaridade do pai e origem escolar dos alunos nos campi do ensino superior brasileiro nos dois ciclos estudados .	24
3.3	Cor, gênero e situação do estudante quanto ao trabalho nos campi do ensino superior brasileiro nos dois ciclos estudados	25
3.4	Cursos com maior proporção de alunos com renda menor que 3 S.M., pai sem escolaridade e provenientes de escolas públicas nos dois ciclos	27
3.5	Cursos com menor proporção de alunos com renda menor que 3 S.M., com pai sem escolaridade e provenientes de escolas públicas nos dois ciclos	29
3.6	Cursos com menor proporção de alunos brancos/asiáticos; maior proporção de alunos que ajudam no sustento da família e menor proporção de homens nos dois ciclos	32
3.7	Cursos com maior proporção de alunos brancos/asiáticos; menor proporção de alunos que ajudam no sustento da família e maior proporção de homens nos dois ciclos	34
4.1	Divisão de cursos por faixas salariais	40
4.2	Distribuição da frequência da variável grupo_salarial dependente em 4 grupos	41
4.3	Descrição e distribuição das variáveis de interesse por faixas salariais dos cursos estudados entre os ciclos	42
4.4	Variance inflation factor test	43
4.5	Comparação da qualidade do ajuste dos três modelos	50
5.1	Efeitos marginais modelo logístico binomial primeiro ciclo	53
5.2	Efeitos marginais modelo logístico binomial segundo ciclo	55
5.3	Resultados em efeitos marginais do modelo multinomial para o primeiro ciclo Enade	59
5.4	Resultados em efeitos marginais do modelo multinomial para o segundo ciclo Enade	62
A.1	Renumeração por curso	73
A.1	Renumeração por curso	74

C.1	Descrição das variáveis controle	77
D.1	Resultados exponenciados modelo logístico primeiro ciclo	80
D.2	Resultados exponenciados modelo logístico segundo ciclo	81
F.1	Resultados em risco relativo modelo multinomial logit para o primeiro ciclo	86
F.2	Resultados em risco relativo modelo multinomial logit para o segundo ciclo	87
G.1	Resultados regressão odds proporcionais primeiro ciclo	88
G.2	Resultados regressão odds proporcionais segundo ciclo	89
H.1	Brant test	90
H.1	Brant test	91
I.1	Resultados exponenciados para o modelo de odds não proporcionais no primeiro ciclo	92
I.2	Resultados exponenciados para o modelo de odds não proporcionais no segundo ciclo	93

SUMÁRIO

1	INTRODUÇÃO	1
2	REVISÃO DA LITERATURA	5
2.1	Breve revisão da literatura sobre capital humano	5
2.2	Evidências empíricas sobre os fatores que impactam a escolha do curso superior	8
2.2.1	Evidências empíricas internacionais	8
2.2.2	Evidências empíricas nacionais	11
2.3	Desigualdade horizontal no ensino superior	14
3	CONTEXTO EMPÍRICO - PERFIL DOS ESTUDANTES ENTRE OS CURSOS DO ENSINO SUPERIOR BRASILEIRO	19
3.1	Uma breve explicação do Enade	19
3.2	Perfil socioeconômico dos alunos no ensino superior brasileiro	22
3.2.1	Renda, Origem escolar e educação do pai	22
3.2.2	Cor, gênero e situação do estudante	24
3.3	Principais cursos universitários por perfil socioeconômico do estudante	26
3.3.1	Renda familiar, educação do pai e origem escolar	26
3.3.2	Cor, gênero e situação do estudante	31
4	DADOS E ESTRATÉGIA EMPÍRICA	37
4.1	Variáveis dependentes e independentes	39
4.1.1	Variável dependente	39
4.1.2	Variáveis independentes de interesse e controle	41
4.1.3	Correlação entre as variáveis independentes	43
4.2	Regressões logísticas	43
4.2.1	Modelo logístico binomial	44
4.2.2	Modelo multinomial	45
4.2.3	Modelo odds proporcionais	46
4.2.4	Modelo odds não proporcionais	48
4.3	Seleção do modelo	48
5	RESULTADOS	51
5.1	Efeitos marginais modelo logístico binomial	52
5.1.1	Primeiro ciclo Enade	52
5.1.2	Segundo ciclo Enade	54

5.1.3	Comparação entre os ciclos	56
5.2	Efeitos marginais modelo logístico multinomial	56
5.2.1	Primeiro ciclo Enade	57
5.2.2	Segundo ciclo Enade	60
5.2.3	Comparação entre os ciclos	63
6	CONCLUSÃO	64
7	REFERÊNCIAS	67
APÊNDICE A	RETORNO SALARIAL POR CURSO	73
APÊNDICE B	PERGUNTAS SOCIOECONÔMICAS ENADE	75
APÊNDICE C	DESCRIÇÃO DAS VARIÁVEIS CONTROLE	77
APÊNDICE D	MODELO LOGÍSTICO BINOMIAL	80
APÊNDICE E	APÊNDICE E - REGRESSÃO PASSO A PASSO	82
E.0.1	Regressão passo-a-passo	82
APÊNDICE F	RESULTADOS EXPONENCIADOS MODELO MULTI- NOMIAL	86
APÊNDICE G	RESULTADOS EXPONENCIADOS MODELO ODDS PRO- PORCIONAIS	88
APÊNDICE H	BRANT TEST	90
APÊNDICE I	RESULTADOS DO MODELO DE ODDS NÃO PRO- PORCIONAIS	92

1 INTRODUÇÃO

Os estudos sobre retorno do capital humano indicam que há uma relação direta entre o investimento em educação e rendimento salarial, sendo Schultz(1973), Mincer(1974) e Becker(1983) os percursores sobre o tema. Mesmo dentro de um nível educacional há grandes diferenças no tocante ao valor atribuído a diferentes tipos de capital humano. Por exemplo, no tocante a cursos universitários, há diferenças significativas no retornos salariais entre os graduados em diferentes cursos do ensino superior devido à demanda pelo mercado da competência associada a cada carreira ligada ao curso (James et al. (1989), Rumberger e Thomas(1993), Finnie e Frenette (2003), Thorson(2005)).

Diferenças essas que podem contribuir e perpetuar desigualdades socioeconômicas. Portanto, não somente deve-se considerar a desigualdade socioeconômica no acesso ao ensino superior, o que se chama de estratificação ou desigualdade vertical (indivíduos de origens sociais mais privilegiadas tendem a estudar por mais tempo e possuem mais chances de superarem as transições dentro do sistema educacional), mas também dar atenção a um viés de distribuição dos alunos entre cursos e instituições, o que denomina-se de desigualdade ou estratificação horizontal. Essa dissertação quer contribuir sobre o tema da desigualdade horizontal verificando que cursos ligados a carreiras com maiores retornos financeiros esperados no mercado de trabalho concentram alunos de estratos sociais mais privilegiados enquanto o inverso acontece, cursos que levam a carreiras com menor expectativa de retorno financeiro concentram alunos de estratos sociais menos privilegiados.

Ao longo da última década o ensino superior no Brasil passou por uma grande etapa de expansão e diversificação do perfil dos estudantes. O total de matrículas

saltou de 4.278.133 em 2004 para 8.033.574 em 2015, um aumento de 83% em 11 anos (Censo educação Superior, Inep). Programas governamentais como Fies, ProUni, Reuni e as políticas de cotas sociais vêm exercendo um impacto importante nessa expansão da educação superior e na democratização nos campi brasileiros com maior participação de alunos de origens socioeconômicas menos privilegiadas. Porém, esses alunos tendem a se concentrar em cursos menos privilegiados e com menor perspectiva de retorno financeiro, ou seja, apesar da aparente queda da desigualdade vertical nos últimos anos do ensino superior brasileiro, há uma manutenção da desigualdade horizontal dentro do campus. "A estratificação verificada no curso universitário naturalmente tem impacto em termos salariais."(Ribeiro 2014, p.140)

A importância do tema sobre desigualdade horizontal na educação superior foi escrita por Bourdieu já em 1964 em *Les Heritages*, na qual o autor questiona a questão meritocrática da educação e afirma que a formação inicial em um ambiente social e familiar criaria um conjunto de disposições, pensamentos e percepções no indivíduo(o autor denomina de "habitus"), os quais passam a conduzir o indivíduo ao longo do tempo e que permite explicar as similares práticas, escolhas e formas de pensamento educacionais dos indivíduos de classes sociais semelhantes. A noção de habitus elaborada por Bourdieu apresentou uma explicação para os casos de auto seleção na escolha do curso superior, ou seja, para o fato dos estudantes tenderem a escolher cursos mais ou menos rentáveis ou prestigiosos de acordo com sua origem social, perfil escolar e cultural. Nogueira(2014)

Essa dissertação quer aprofundar sobre o tema da desigualdade horizontal no Ensino Superior brasileiro, analisando duas questões principais:

1. Como os alunos de diferentes perfis socioeconômicos estão distribuídos nos cursos universitários atualmente no Brasil?

2. Existe um viés nessa distribuição? Ele aumentou ou reduziu nos últimos 10 anos?

Para tanto, faz-se uma análise empírica com os dados do Exame Nacional de Desempenho de Estudantes (Enade) entre 2004 e 2015, com intuito de verificar como estudantes de diferentes origens socioeconômicas se distribuem nos diversos cursos oferecidos pelo ensino superior brasileiro. Além disso, busca-se o melhor modelo logístico para probabilidade de fazer cursos universitários classificados pelo retorno salarial esperado, considerando as diferenças por características socioeconômicas dos estudantes e sua evolução temporal.

Assim sendo, além dessa introdução, a dissertação está estruturada da seguinte forma. No capítulo 2 (Referencial teórico) tem-se uma breve revisão teórica sobre o tema do capital humano heterogêneo que envolve retorno do capital humano e escolha do curso de formação no ensino superior, assim como uma análise da literatura que aborda o tema de desigualdade horizontal no ensino superior, isto é, alunos de diferentes origens socioeconômicas se concentram de forma diferente entre os cursos universitários de acordo com seus prestígios e retornos salariais esperados.

No capítulo 3 (Contexto Empírico) apresenta uma breve explicação do Enade e uma análise socioeconômica da proporção de estudantes de diferentes origens sociais nos campi brasileiros como um todo e sua evolução temporal. Esse capítulo analisa também quais cursos universitários tem maior frequência de estudantes com origem socioeconômica mais e menos favorecidas.

O capítulo 4 (Base de dados e metodologia) apresenta as variáveis dependentes e independentes, de interesse e de controle e a estratégia empírica com base nos modelos regressivos logit, afim de encontrar o melhor modelo para analisar a desigualdade horizontal no ensino superior. Por fim verifica-se como a desigualdade

horizontal se comportou entre os anos de 2004 a 2015 a partir do modelo logístico selecionado.

No capítulo 5 (Resultados) tem-se a interpretação dos resultados e, por fim, no capítulo 6 apresenta-se a conclusão.

2 REVISÃO DA LITERATURA

Este capítulo apresenta a teoria do capital humano e diversos trabalhos que mostraram evidências empíricas sobre retorno salarial da educação, escolha do curso e desigualdade horizontal no ensino superior.

2.1 Breve revisão da literatura sobre capital humano

Desde a metade do século XX existe um amplo debate teórico sobre capital humano e a área da economia do trabalho. Os precursores a relacionar capital humano com retorno salarial foram Schultz(1973), Mincer(1974) e Becker(1983).

Schultz(1973) é um dos principais autores sobre o tema e procura mostrar que o investimento em capital humano traz benefícios individuais e para sociedade, pois melhora as competências e qualificações do indivíduo levando ao aumento de sua produtividade junto com os lucros dos capitalistas, impactando assim a economia como um todo. O autor preocupa-se com os benefícios indiretos do investimento em capital humano, isto é, os benefícios a sociedade em geral. Alguns exemplos clássicos desses benefícios são os acréscimos no desenvolvimento cultural e intelectual do indivíduo, consciência dos direitos civis, melhores hábitos alimentares, higiene e saúde em geral e maiores motivações profissionais.

O autor utiliza um conjunto de fatores com o objetivo de mensurar o acúmulo de capital humano como: educação formalmente organizada, treinamento no local de trabalho e programas de extensão ou migração das famílias e indivíduos.

Schultz(1973) também alerta para o problema da incerteza sobre o investimento em capital humano ligado a educação, pois os indivíduos não possuem informação perfeita sobre rendimentos futuros, isto é, os investimentos em educação situam-se num horizonte de longo prazo, logo, fica sujeito a incertezas e a imperfeições do mercado de capitais. O fato da incerteza e risco existirem não pode ser desprezado, pois afeta a capacidade cognitiva do indivíduo no processo de maximizar sua função de utilidade. A decisão educacional só será confrontada em um período de tempo futuro, no qual o cenário do mercado de trabalho se altera constantemente, portanto é inviável para o autor, avaliar o nível ótimo de investimento conforme o paradigma da escolha racional.

Já Becker(1983) fundamenta sua teoria nos pressupostos básicos da economia neoclássica de maximização da utilidade individual e no paradigma da escolha racional. Ao contrário de Schultz(1973), pode-se dizer que o investidor em capital humano pode prever todos os possíveis estados de mundo a partir de sua trajetória educativa, pois o indivíduo tem informações sobre o mercado de trabalho e assim a incerteza no investimento da educação é pequena.

Becker(1983) define o capital humano como um conjunto de capacidades produtivas que um individuo acumula por conhecimentos gerais ou específicos. A partir de evidências empíricas, o autor constatou que quanto maior a qualificação de um indivíduo, maiores suas chances de estar empregado e sua remuneração no mercado de trabalho.

O autor mostra que há uma assimetria de rendimento devido ao investimento em qualificação e, assim, pode-se modelar o nível ótimo de educação ou estoque de conhecimento que um individuo deseja adquirir, a fim de maximizar a sua utilidade segundo o paradigma da escolha racional.

Em seus trabalhos mais recentes, Becker (1996) busca ampliar as variáveis que influenciam o acúmulo de capital humano. Além da educação formal nas escolas ou emprego, ele incorpora outras variáveis, como o capital social e o capital da imaginação. O capital social está ligado ao contexto socioeconômico do indivíduo, como a importância do background familiar. O autor exemplifica os esforços dos pais, sendo os pais um modelo para os filhos, que absorvem seus valores. Ainda segundo o autor, esses valores absorvidos possuem uma influência muito maior na educação do que a renda dos pais.

Já o capital imaginação é a busca dos indivíduos em maximizar suas utilidades presentes e futuras. O emprego de recursos escassos no presente, como o tempo, no capital imaginação leva a uma melhor orientação ao futuro. O autor afirma que indivíduos com melhores orientações ao futuro criam preferências habituais que resultam em consequências benéficas a longo prazo. Isto é, agir no presente sendo orientado para o futuro implica que os indivíduos consigam antecipar consequências de suas escolhas no presente a fim de estarem em posições melhores no futuro.

Quanto as aplicações da teoria do capital humano, o modelo de Mincer (1974) tem sido o mais utilizado a fim de determinar empiricamente a relação entre o investimento em capital humano e o retorno da educação. O autor enfatiza que os rendimentos individuais de uma pessoa correspondem diretamente ao seu acúmulo de capital humano ao longo do tempo. Assim, o autor elabora a função "human capital earnings function", muito utilizada em trabalhos empíricos sobre retornos educacionais.

Simplificando, pode-se afirmar que a base central da teoria do capital humano é que o investimento pessoal em educação ou treinamento adquirido, aumenta a produtividade do indivíduo, o que resulta em maiores salários.

2.2 Evidências empíricas sobre os fatores que impactam a escolha do curso superior

2.2.1 Evidências empíricas internacionais

Outra literatura menos tradicional sobre capital humano vem ganhando espaço nos últimos anos, chamada de teoria do capital humano heterogêneo, busca compreender quais os fatores que impactam a demanda por tipos específicos de capital humano, por exemplo, no contexto da trajetória educacional, a escolha do curso superior pelos estudantes que terminam o ensino médio e querem seguir uma graduação.

Diversos estudos internacionais buscam compreender os fatores que impactam a decisão do estudante sobre qual curso seguir, entre eles Berger(1988), Paglin e Rufolo(1990) e Montmarquette, et al(2002), verificaram que o retorno esperado,¹ inicial ou ao longo prazo da carreira ligada a determinado curso, tem influência positiva na escolha dos estudantes.

Porém, Betts(1995) questiona se os indivíduos formam expectativas racionais sobre os salários futuros das diferentes carreiras no mercado de trabalho. O autor questiona a credibilidade das expectativas dos estudantes quanto ao rendimento futuro de suas carreiras e verifica como eles adquirem as informações sobre o mercado de trabalho. Segundo o autor, há uma grande heterogeneidade formada nessas expectativas entre os estudantes. Os estudantes não procuram prever suas expectativas de salário ao escolherem um curso superior, mas adquirem essas informações durante seu curso, logo, é muito difícil modelar a escolha do curso de ensino superior, pois os indivíduos não têm informação perfeita.

¹A variável área de conhecimento e curso do ensino superior possui uma significativa influência sobre o retorno financeiro. (Finnie e Frenette (2003), James et al. (1989), Thorson(2005).

Staniec(2004) verificou que o rendimento esperado não tinha impacto na escolha dos cursos. As variáveis que mais impactavam eram as socioeconômicas, como por exemplo a variável de controle ligada a habilidades do estudante e moradia. O aumento de um quantil na nota do vestibular de matemática aumentava em 4.8% a probabilidade de um estudante escolher cursos na área de ciência, engenharia e matemática. Além disso, estudantes que moravam com a mãe tinham 4% a menos de probabilidade de não escolherem esses cursos e 2.4% a mais a mais de probabilidade de escolherem cursos ligados a humanidades e artes, usando como referência os cursos de ciências sociais/outros.

Segundo a autora isto pode ser explicado pois a probabilidade de mulheres escolherem cursos mais técnicos é historicamente pequena, o que fazia com que seus filhos não tivessem incentivos para seguir cursos mais técnicos.

Seguindo a linha de Staniec(2004), a partir de regressões multinomiais, trabalhos como Porter e Umbach(2006), Dickon(2010) e Nores(2010) verificaram o impacto das variáveis socioeconômicas na escolha do curso. Os estudos mostraram diferenças significativas quanto a questões raciais, de gênero, habilidades(nota no vestibular) e renda na escolha da área de atuação no ensino superior.

No trabalho de Porter e Umbach(2006), por exemplo, os autores verificaram que mulheres tem menos chances de escolherem áreas ligadas a ciência comparadas a homens, enquanto brancos tem mais chances de escolherem áreas ligadas a ciências quando comparados a latinos e negros, e alunos provindos de escolas privadas tinham mais chances de estudarem artes. Os autores também verificaram que a nota do vestibular impactava a escolha do curso, isto é, alunos que foram melhores nos exames de linguáas, escolhiam áreas de conhecimento ligados a humanas e artes, enquanto alunos com melhores notas nos exames de matemática, escolhiam cursos ligados a ciências.

Bourdieu(1964), um dos autores mais importantes da sociologia, estudou a relação entre os estudantes e o tipo de curso frequentado, verificou que cursos com maiores retornos financeiros concentravam alunos de estratos sociais mais altos, enquanto o inverso ocorria para os alunos de estratos sociais mais baixos. Entretanto, para Bourdieu, a variável chave para escolha do curso não era o retorno esperado, mas sim as características socioeconômicas do estudante.

O autor criou o conceito de habitus, no qual defende que, a formação inicial em um ambiente social e familiar criaria um conjunto de disposições, pensamentos e percepções no indivíduo que possibilitariam explicar as escolhas, similaridades práticas e formas de pensamento educacionais dos indivíduos de classes sociais semelhantes. A noção de habitus tem como objetivo, apresentar uma explicação para os casos de auto seleção dos indivíduos ao tenderem escolher cursos mais ou menos seletivos e rentáveis, à medida que a idade, perfil social, escolar e cultural, se mostrem mais ou menos favoráveis. Nogueira(2004)

Segundo Bourdieu, outro fator importante para escolha do curso é o risco do investimento escolar, isto é, o nível econômico do estudante impacta o grau de audácia e disposição ao risco na escolha dos cursos a seguir.

Duru e Mingat(1979) abordou a temática do impacto do risco na escolha do curso superior. O autor verificou que indivíduos de classes sociais menos privilegiadas apresentam maior sensibilidade ao risco, os melhores estudantes no caso francês escolhiam o curso de medicina enquanto que, se eles viessem de uma classe social menos privilegiada, eles escolhiam cursos técnicos de curta duração. Isto porque na França, medicina é um curso com retorno financeiro alto porém com elevado risco de desistência ou fracasso no primeiro ano, já estudos técnicos de curta duração no caso francês possuem menor retorno financeiro e uma menor probabilidade de evasão. Assim, segundo os autores, estudantes de classes sociais mais baixas são

menos relutantes a tomar riscos, mesmo que o retorno financeiro no longo prazo seja maior.

2.2.2 Evidências empíricas nacionais

Casari(2006) verificou que a influência do retorno salarial na escolha do curso era quase nula. Porém, variáveis individuais do aluno, como acréscimo na nota do vestibular, aumentavam as chances do estudante cursar engenharia em 4,29 pontos percentuais, enquanto na área de educação, a probabilidade caía em 9,79 pontos percentuais. Número de TVs foi usado como proxy para renda e a autora verificou que, uma unidade a mais de TVs por domicílio aumentava a probabilidade de escolher engenharia em 1,41 pontos percentuais e caía na área de educação em 1,66 pontos percentuais.

Já Bartalotti e Menezes-Filho(2007) apontaram, por meio de estimações por mínimos quadrados ordinários generalizados, que o salário médio de cada carreira influencia positivamente na demanda de um curso no vestibular. Entretanto, a taxa de desemprego e dispersão salarial das carreiras não se mostraram variáveis importantes no número de inscritos no distintos cursos estudados ligados a essas carreiras. Os autores alertam para a importância de variáveis subjetivas como vocação ou habilidades individuais que podem ter influência na escolha do curso.

Soares(2007) procurou compreender os principais fatores que impactam na escolha do curso superior através de um estudo exploratório com 655 estudantes, ingressantes no ensino superior de Caxias do Sul/RS. Segundo o autor, os estudantes não se consideraram influenciados pela família na escolha do curso, porém essa influência pode ser subjetiva. Os fatores que se mostraram mais importantes foram a ascensão financeira, qualidade da instituição de ensino, oportunidades de trabalho,

vocação e aptidão dos estudantes a determinado curso.

Já Martins e Carusi(2015) verificaram que o rendimento esperado do curso somente é significativo nos quartis de renda familiar mais elevadas, sendo o rendimento no mercado de trabalho no início da carreira mais impactante na escolha do curso, quando comparado ao rendimento ao longo da carreira.

Os autores mostraram que as variáveis socioeconômicas tinham influência na escolha do curso superior, por exemplo, homens brancos com renda familiar mais elevada tinham probabilidade de 12,60% em cursar engenharia, enquanto para mulheres brancas com renda familiar menor, a probabilidade era de apenas 0,97%. Os autores constataram também o impacto da concorrência no vestibular nos indivíduos do quartil de renda familiar mais baixos, isto é, eles evitavam seguir cursos mais competitivos e optavam por cursos mais fáceis de passar no vestibular.

a fim de resumir o que foi visto até aqui, a tabela 2.1 abaixo mostra as principais variáveis que podem impactar a escolha do curso, dividas entre preferências, características socioeconômicas e habilidade do estudante.

As preferências segundo Nogueira(2004) estão relacionadas com as variáveis socioeconômicas, por exemplo, no estudo de Martins e Carusi(2015), o retorno salarial apenas era significativo nas famílias com renda familiar mais elevadas. Isto pode ser explicado devido o aluno querer maximizar sua utilidade na escolha do cursos superior, a qual depende do contexto socioeconômico a que o aluno esta inserido. O próprio ingresso no ensino superior para alunos de extratos sociais menos privilegiados e com pais sem educação formal, já seria um triunfo para família, independente do curso escolhido, enquanto que alunos de extratos sociais mais altos querem no mínimo manter o nível que seus pais já alcançaram, ingressando em cursos mais prestigiados.

Tabela 2.1: Fatores que influenciam a escolha do curso

Variáveis socioeconômica	Cor Renda familiar Educação dos pais Gênero Situação familiar
Habilidades individuais	Tipo de escola Nota no vestibular Aptidões Vocação pro curso Subjetivas
Preferências	Risco(capacidade de concluir o curso) Retorno esperado do curso Condições no mercado de trabalho do curso Concorrência do curso

Segundo Nogueira "No caso das famílias cujo patrimônio se concentra, sobretudo no capital cultural e cuja trajetória ascensional passada e perspectivas de ascensão futura dependem basicamente da escola, cria-se um habitus familiar completamente voltado para o sucesso escolar. A entrada no ensino superior e a escolha de cursos prestigiosos e rentáveis se tornam, então, elementos centrais das estratégias de reprodução dessas famílias." (Nogueira 2004, p.89)

Vale lembrar também que no Brasil, em sua maioria, alunos de baixa renda familiar estudam em escolas públicas, de pior qualidade em geral, com uma preparação ruim para o vestibular e isto pode impactar na escolha do curso, pois os alunos com uma formação escolar fraca evitariam de se inscrever em cursos com maior concorrência no vestibular, pois sabem que suas chances de ingressar em determinados cursos são menores comparados a outros. Além do mais, existe o risco de não passar no vestibular, atrasando a entrada no mercado de trabalho em um ano. Esse risco é menor para classes sociais mais altas, pois segundo Nogueira(2004), os estudantes mais privilegiados sabem que os pais estão dispostos a investir em mais um ano de cursinho e arcar com os custos do prolongamento da carreira escolar e adiantamento

na entrada no mercado de trabalho, já o atraso no mercado de trabalho tem um impacto muito maior nas famílias mais necessitadas, isto é, o custo de oportunidade é maior.

"As famílias dos diferentes níveis sociais são mais ou menos sensíveis ao risco porque os custos e os benefícios associados ao investimento escolar variam de acordo com a posição social."(Nogueira(2004), p.28)

O custo de oportunidade de um estudante de baixa renda cursar o ensino superior é maior, pois o mesmo poderia estar trabalhando e ajudando no sustento da família, assim suas opções são mais sensíveis, devendo optar por cursos menos desafiadores com baixa concorrência e taxas de evasão, os quais geralmente são cursos menos prestigiados e com menor expectativa de retorno salarial.

Portanto, entende-se que as variáveis socioeconômicas estão por traz das preferências dos indivíduos, mesmo subjetivamente e que elas podem perpetuar a desigualdade horizontal no ensino superior, isto é, alunos de origens sociais privilegiadas tendem a frequentar cursos de maior prestígio e retorno financeiro, enquanto alunos de origens menos privilegiadas se concentram em cursos com menor prestígio e retorno. A seguir tem-se uma revisão dos trabalhos que abordaram o tema da desigualdade horizontal no ensino superior brasileiro.

2.3 Desigualdade horizontal no ensino superior

Diversos estudos brasileiros verificaram uma grande discrepância nos rendimentos salariais entre grupos de área de formação superior(Entre eles, destaca-se, Fernandes e Narita(2001), Neri(2013), Reis e Carusi(2016)), que reflete uma heterogeneidade dentro do sistema educacional superior, e que pode contribuir para

o problema de desigualdade devido a influência das variáveis socioeconômicas na trajetória do estudante.

Trabalhos como Ribeiro e Klein(1982), Braga et al(2001), Ristoff(2013) e Ribeiro(2014), mostram que no Brasil, alunos de origens sociais menos privilegiadas se concentram em cursos menos prestigiados e com menor retorno financeiro no mercado de trabalho, enquanto o inverso ocorre para alunos de estratos sociais mais privilegiados.

Ribeiro(2014) constatou que, entre 1960 e 2010, houve uma grande expansão na oferta de vagas no ensino superior brasileiro, porém há duas tendências claras de sentidos opostos. A expansão levou a uma democratização com maior participação de minorias e classes sociais menos privilegiadas, apesar dessa expansão não representar acesso igual aos cursos universitários. Ele observou que a inclusão de estudantes das camadas menos privilegiadas, no ensino superior, ocorrem em cursos de menor prestígio e retorno financeiro como história, ciências sociais, filosofia, letras, serviço social e cursos ligados a educação, enquanto o inverso ocorre nas carreiras de odontologia, engenharia, economia e medicina.

Ribeiro(2014) argumentou em seu livro: "Sabemos que com a expansão universitária que ocorreu nas últimas décadas houve uma enorme diversificação do ensino superior. A hipótese da "desigualdade efetivamente mantida" (effectively maintained inequality) prevê que a desigualdade horizontal tende a aumentar quando há expansão dos diversos nível de ensino hierarquicamente organizados (Lucas, 2001), porque os privilegiados do ponto de vista socioeconômico seriam tentados a assegurar para si e seus filhos, sempre que possível, alguma vantagem. Se as diferenças quantitativas são comuns e escassas, essas pessoas perseguiriam vantagens quantitativas; quando elas se tornam maiores, passariam a visar vantagens qualitativas. Em termos educacionais, o mecanismo significaria que quando o acesso a dado nível de

escolaridade ainda é muito restrito os privilegiados usariam sua posição para assegurar sua aquisição. Quando, no entanto, certo nível escolar se torna mais difundido, a estratégia seria procurar garantir o acesso a diferenças qualitativas."(Ribeiro, 2014, p.139)

Ribeiro e Klein (1982) estudaram a desigualdade horizontal no meio da Fundação Cesgranrio e identificaram que quanto maior o prestígio social de um curso, mais elevado os pontos no vestibular para a classificação. Alunos de substratos sociais mais baixos, já no pré vestibular, optavam por cursos de menor concorrência, prestígio e retorno social. Portanto, no exame em si, a seleção é meritocraticamente deturpada pois os concorrentes a uma determinada carreira estão disputando entre concorrentes com estratos sociais semelhantes.

Braga et al. (2001) foi outro autor brasileiro que analisou a demanda pela graduação no âmbito da Universidade Federal de Minas Gerias (UFMG), estudando o perfil socioeconômico dos candidatos e suas preferências por áreas e cursos. Ele observou que, a partir de 1995, houve um grande aumento na demanda de candidatos provindos de escolas públicas e provenientes de famílias com baixo poder aquisitivo. Essa demanda se concentrou em cursos de baixo prestígio social e menos competitivos, com notas de corte baixas no vestibular (enfermagem, geografia, biblioteconomia, ciências biológicas, letras, pedagogia, a maioria cursos que oferecem licenciatura).

Segundo Braga, isso é explicado pela seletividade social associada ao sucesso no vestibular. Estudantes de origem social menos privilegiada já reconhecem não ter condições e preparo para os cursos com maior grau de concorrência no vestibular, optando por cursos de menor concorrência e prestígio. Estudantes que desconheciam a realidade pagavam um custo elevado pela falta de informação. O autor verificou que no período de 1992/1999, por exemplo, o curso de direito da UFMG era um

dos mais concorridos e, de cada oito candidatos de origem social mais privilegiada, um era aprovado, enquanto que candidatos de origens sociais mais baixas, apenas um em cada cem era aprovado. Nos cursos de enfermagem esses dois grupos de estudantes tinham chances de sucesso similares, a cada treze concorrentes um foi aprovado.

Ristoff(2013) estudou a inclusão de alunos de classes menos privilegiadas no ensino superior brasileiro e como essa inclusão se deu entre os diferentes cursos universitários, usando a base do Enade entre 2004 e 2011. Segundo o autor, há uma lado otimista, "O campus brasileiro como um todo torna-se menos branco; o percentual de pretos começa a se aproximar mais dos percentuais da sociedade; mais pardos frequentam os cursos universitários; e mais pessoas de baixa renda conseguem chegar à educação superior."(Ristoff(2013), p.21) e um pessimista, "Fica evidente que há uma forte correlação entre os indicadores socioeconômicos analisados e os cursos nos quais os estudantes estão matriculados."(Ristoff(2013, p.21)

Ainda segundo Ristoff(2013), cursos como licenciatura são onde estão concentrados os alunos de baixas classes sociais, enquanto cursos mais privilegiados e com maior perspectiva de retorno financeiro, como medicina, concentram os estudantes provindos de classes sociais mais privilegiadas. "Os dados do Inep demonstram que somente em sete cursos, seis de licenciatura e no Serviço Social, os percentuais, na faixa de até três salários mínimos são superiores a 50%. Diante da comprovação oferecida pelos dados, de que a origem social e a situação econômica da família são fatores determinantes na trajetória do jovem pela educação superior. "(Ristoff(2013), p.22)

O estudo até aqui evidenciou que a escolha do curso superior é complexa, envolvendo fatores socioeconômicos, habilidades individuais e preferências subjetivas do aluno. Viu-se que os fatores socioeconômicos tem um impacto muito forte nessa

escolha e que parece haver um viés na distribuição dos alunos entre os cursos quanto a esses fatores. Esse viés perpetua desigualdades sociais já existentes, isto é, sabe-se que há uma diferença de retornos esperados entre as carreiras no mercado de trabalho ligadas aos cursos e que, estudantes de origens sociais privilegiadas, tendem a se concentrar em cursos com maiores expectativas de retorno financeiro, enquanto que estudantes de origens menos privilegiadas, tem maior participação em cursos menos prestigiados e com menor expectativa de retorno financeiro.

No capítulo seguinte é feita uma análise descritiva desse fato, usando a base do questionário socioeconômico do Enade², entre os anos de 2004 a 2015, a fim de explorar a proporção de alunos de diferentes backgrounds socioeconômicos nos campi brasileiros e como esses alunos se distribuem entre os cursos.

²A base socioeconômica do Enade é muito rica e pouco utilizada em trabalhos sobre educação, essa dissertação utiliza os mais recentes dados disponibilizados pelo Inep.

3 CONTEXTO EMPÍRICO - PERFIL DOS ESTUDANTES ENTRE OS CURSOS DO ENSINO SUPERIOR BRASILEIRO

3.1 Uma breve explicação do Enade

A Lei nº 10.861, de 14/04/2004, criou o Sistema Nacional de Avaliação da Educação Superior (Sinaes), instituindo nacionalmente a avaliação das instituições, dos cursos e dos estudantes. O Exame Nacional de Desempenho de Estudantes (Enade), criado pela lei do Sinaes e regulamentado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), tem como objetivo avaliar os estudantes do ensino superior e a qualidade das instituições. Tal avaliação é feita através de exames específicos por curso e perguntas sobre a infraestrutura da instituição a qual o aluno pertence. A realização do exame é obrigatória para os estudantes que possuem pelo menos 75% do curso completo e que visam obter o diploma da instituição de ensino superior (IES). No entanto, a nota do aluno no Enade não tem impacto em sua formação, apenas na avaliação do instituto ao que o aluno pertence.

A fim de realizar a prova do Enade, é necessário que o estudante responda um conjunto de questionários sobre sua situação socioeconômica. Esse questionário será a principal base de dados para estudar como os alunos de diferentes origens socioeconômicas estão distribuídos nos campi brasileiros.

A aplicação do Enade é feita anualmente por grupos de áreas de conhecimento e cada grupo é submetido ao exame de três em três anos. Assim, o primeiro

grupo(cursos das áreas de Ciências Agrárias e da Saúde) realizou o exame nos anos de 2004, 2007, 2010 e 2013, o segundo (cursos das áreas das engenharias e licenciaturas) nos anos de 2005, 2008, 2011 e 2014 e o terceiro grupo (principalmente cursos das áreas de Ciências Sociais Aplicadas e Jurídicas) em 2006, 2009, 2012 e 2015. Os cursos avaliados em cada grupo são os mesmos com alguns acréscimos de cursos ligados a tecnologia (p.ex: Automação Industrial, Fabricação Mecânica, Gestão da Produção Industrial, Manutenção Industrial, Processos Químicos, Fabricação Mecânica e etc.) e a exclusão de outros, como certas engenharias muito específicas sem representação nacional devido a pouca quantidade de alunos, conforme pode ser visto em Fernandes(2014).

Optou-se, nessa dissertação, realizar uma análise por ciclos completos de três anos do Enade. O primeiro ciclo¹ a ser estudado nessa dissertação será o de 2004 a 2006, que abrange os cursos da tabela 3.1 escolhidos pelo autor e, o segundo, será o ciclo mais atual disponível até então, de 2013 a 2015, os mesmos escolhidos para o primeiro ciclo:

¹2004 – Agronomia, Educação Física, Enfermagem, Farmácia, Fisioterapia, Fonoaudiologia, Medicina, Medicina Veterinária, Nutrição, Odontologia, Serviço Social, Terapia Ocupacional e Zootecnia; 2005 – Arquitetura e Urbanismo, Biologia, Ciências Sociais, Computação, Engenharia (em oito grupos), Filosofia, Física, Geografia, História, Letras, Matemática, Pedagogia e Química; • 2006 – Administração, Arquivologia, Biblioteconomia, Biomedicina, Ciências Contábeis, Ciências Econômicas, Comunicação Social, Design, Direito, Formação de Professores (Normal Superior), Música, Psicologia, Secretariado Executivo, Teatro e Turismo. Assim o próximos ciclos (2007 a 2009, 2010 a 2012 e 2013 a 2015 se repetem com os mesmos cursos do ciclo anterior

Tabela 3.1: Quantidade de alunos nos cursos selecionados por ciclo

Cursos	1º Ciclo		2º Ciclo	
	N	%	N	%
Administração	153,509	18.88	151,997	16.37
Arquitetura	7,417	0.91	15,739	1.70
Biologia	9,481	1.17	7,422	0.80
Biologia(Lic)	18,022	2.22	19,704	2.12
C.Sociais	3,987	0.49	3,806	0.41
C.Sociais(Lic)	1,073	0.13	4,098	0.44
Com Social	53,942	6.63	34,634	3.73
Computação	36,087	4.44	9,822	1.06
Computação(Lic)	3,248	0.40	2,226	0.24
Contabilidade	49,814	6.13	65,483	7.05
Direito	74,476	9.16	124,924	13.45
Economia	15,075	1.85	9,57	1.03
Enfermagem	23,11	2.84	30,289	3.26
Engenharia Ambiental	4,083	0.50	7,329	0.79
Engenharia Civil	10,229	1.26	23,539	2.54
Engenharia De Produção	8,355	1.03	17,307	1.86
Engenharia Florestal	2,744	0.34	2,081	0.22
Engenharia Grupo 2	20,264	2.49	18,065	1.95
Engenharia Grupo 4	5,91	0.73	6,448	0.69
Engenharia Mecânica	5,295	0.65	12,328	1.33
Farmácia	15,423	1.90	14,061	1.51
Filosofia	3,718	0.46	1,813	0.20
Filosofia(Lic)	15,306	1.88	4,92	0.53
Física	2,602	0.32	702	0.08
Física(Lic)	2,841	0.35	3,498	0.38
Geografia	3,98	0.49	3,439	0.37
Geografia(Lic)	9,415	1.16	12,773	1.38
História	6,349	0.78	2,602	0.28
História(Lic)	15,49	1.91	21,823	2.35
Letras	13,561	1.67	1,028	0.11
Letras(Lic)	38,819	4.77	33,408	3.60
Matemática	14,354	1.77	522	0.06
Matemática(Lic)	9,298	1.14	16,825	1.81
Medicina	8,943	1.10	16,264	1.75
Odontologia	8,394	1.03	10,476	1.13

Pedagogia	55,238	6.79	131,146	14.13
Psicologia	26,706	3.28	26,544	2.86
Química	4,155	0.51	3,813	0.41
Química(Lic)	4,012	0.49	6,433	0.69
Serv. Social	10,642	1.31	41,716	4.49
Veterinária	6,227	0.77	7,841	0.84
Total	813,104		928,458	

Fonte: Inep Enade, elaborado pelo autor

Nessa seção optou-se analisar as seguintes dimensões socioeconômicas presentes no questionário socioeconômico do Enade: renda mensal da família do estudante; origem escolar do estudante; cor do estudante; escolaridade dos pai do estudante; situação de trabalho do estudante e gênero. (As perguntas e respostas estão no apêndice B).

As duas próximas seções contêm as tabelas que mostram a porcentagem de respostas predominantemente selecionadas em duas etapas: a primeira nos campi como um todo e a seguinte por curso do ensino superior selecionados pelo autor, com o intuito de ilustrar a desigualdade vertical e horizontal, respectivamente.

3.2 Perfil socioeconômico dos alunos no ensino superior brasileiro

3.2.1 Renda, Origem escolar e educação do pai

A tabela 3.2 mostra a proporção de estudantes no ensino superior brasileiro no que se refere a: renda familiar do estudante, escolaridade do pai do estudante e origem escolar (estudantes que estudaram maior parte de sua instrução em escolas

públicas/privadas).

Na variável renda familiar, vê-se que no segundo ciclo (2013-2015) mais de um terço dos estudantes brasileiros (42,68%) estão na faixa de renda mensal familiar de até três salários mínimos. Houve um aumento significativo de 15,66% na presença de alunos nesta faixa de renda comparada ao primeiro ciclo e uma diminuição bastante significativa nas faixas de renda mais altas, principalmente entre as faixas de 10 a 30 salários mínimos, uma queda de 11,65% nessa faixa entre os 10 anos estudados. É perceptível que, na variável socioeconômica renda familiar do estudante, os resultados se mostram positivos, havendo uma maior democratização nos campi e queda da desigualdade vertical.

Vê-se na tabela 3.2 que a proporção de estudantes com pai sem escolaridade subiu de 4,92% para 7,45% entre os dois ciclos estudados, e a proporção de estudantes com pai com escolaridade até 4^o série subiu de 28,65% para 30,45%. Entre os dois ciclos houve uma queda de 8,26% na proporção de estudantes de pai com ensino superior. Portanto, é perceptível uma democratização entre os dois ciclos quanto a variável educação do pai do estudante.

Na questão sobre origem escolar dos estudantes, estudantes de escolas públicas já eram a maioria no primeiro ciclo e houve um aumento de cerca de 9,91% entre os dois ciclos da proporção desses estudantes, chegando a 71,57% no segundo ciclo.

Tabela 3.2: Renda familiar, escolaridade do pai e origem escolar dos alunos nos campi do ensino superior brasileiro nos dois ciclos estudados

Variável socioeconômica	1º Ciclo		2º Ciclo		Variação	
	N	%	N	%	N	%
Renda Familiar						
Até 3 SM	179.867	27,02	455.096	42,68	275.229	15,66
De 3 até 10 SM	326.474	49,05	496.369	46,55	169.895	-2,5
De 10 a 30 SM	136.953	20,58	95.18	8,93	-41.773	-11,65
Mais que 30 SM	22.307	3,35	19.619	1,84	-2.688	-1,51
Total	665,601		1066,264			
Escolaridade do pai						
Sem escolaridade	32.807	4,92	79.388	7,45	46.581	2,53
Fundamental: 4ª série	191.015	28,65	324.932	30,47	133,917	1,82
Fundamental: 8ª série	109.662	16,45	155.488	14,58	45.826	-1,87
Ensino médio	185.838	27,87	307.881	28,87	122.043	1
Ensino superior	147.367	22,1	147.597	13,84	23	-8,26
Total	519.322		1015.286			
Origem escolar						
Pública	397.3	61,65	760.07	71,57	362.807	9,91
Privada	247.145	38,35	302.08	28,43	54.863	-9,92
Total	644.445		1062.115			

Fonte: Inep Enade, elaborado pelo autor

3.2.2 Cor, gênero e situação do estudante

A tabelas 3.3 mostra a proporção de estudantes no ensino superior brasileiro quanto as variáveis cor, gênero e situação do estudante no quesito trabalho e sustento da família. Vê-se que, na variável cor, brancos seguem sendo a maioria: 56,18% no segundo ciclo. Porém, a proporção de brancos vem cedendo espaço entre os dois ciclos estudados, já que houve uma queda de 13,35% e um aumento de negros e pardos de 3,62% e 3,64%, respectivamente.

A análise por gênero revela o que já era conhecido na literatura educacional,

mulheres são a maioria no ensino superior e sua participação segue aumentando entre os dois ciclos estudados.

Ainda na tabela 3.3, encontram-se informações sobre a situação do estudante em relação ao trabalho e família. Vemos que a maioria dos estudantes trabalham e que houve uma queda significativa dos estudantes que não trabalham entre os dois ciclos. Tal fato pode ser explicado com a maior participação de estudantes com renda familiar menor que três salários mínimos, já que grande parte desses estudantes precisam trabalhar para contribuir com as finanças da família.

Tabela 3.3: Cor, gênero e situação do estudante quanto ao trabalho nos campi do ensino superior brasileiro nos dois ciclos estudados

Variável socioeconômica	1º Ciclo		2º Ciclo		Variação	
Cor	N	%	N	%	N	%
Branco	460.688	69,53	593.359	56,18	132.671	-13,35
Negro	35.257	5,32	94.409	8,94	59.152	3,62
Pardo	153.516	23,17	283.149	26,81	129.633	3,64
Asiático	13.154	1,99	85.307	8,08	72.153	6,09
Total	662.615		1.056.224			
Gênero	N	%	N	%	N	%
Mulher	545.79	57,93	736.106	59,95	190.316	2,02
Homem	396.34	42,07	491.719	40,05	95.379	-2,02
Total	942.13		1227.825			
Situação do estudante	N	%	N	%	N	%
Não Trabalho	227.478	33,94	259.537	25,72	32.06	-8,22
Trabalho e recebo ajuda da família	190.973	28,5	291.485	28,88	100.512	0,38
Trabalho e me Sustento	78.655	11,74	137.347	13,61	58.69	1,87
Trabalho e contribuo com a família	125.671	18,75	223.341	22,13	97.67	3,38
Trabalho e sustento a Família	47.366	7,07	97.435	9,66	50.07	2,59
Total	670.143		1009.145			

Fonte: Inep Enade, elaborado pelo autor

Vê-se um aumento da participação de estudantes de grupos socioeconômicos menos favorecidos, indicando uma queda da desigualdade vertical. Essa mudança no perfil dos estudantes do ensino superior brasileiro é efeito provável das políticas

públicas aplicadas no ensino superior, como os conhecidos programas de financiamento governamentais como FIES e PROUNI, além das políticas de cotas. No entanto, como esses alunos se distribuíram entre os cursos? Busca-se nas seções seguintes analisar, através de tabelas, a distribuição dos alunos por curso quanto à variável socioeconômica.

3.3 Principais cursos universitários por perfil socioeconômico do estudante

Essa seção apresenta uma análise dos dez cursos universitários com a maior ou menor proporção de estudantes de acordo com as variáveis socioeconômicas selecionadas. As tabelas 3.4 e 3.5 informam a proporção de estudantes com renda familiar até três salários mínimos, proporção de estudantes de pai sem nenhuma escolaridade e a proporção de estudantes que estudaram a maior parte de suas vidas em escolas públicas. Já, as tabelas 3.6 e 3.7 informam a proporção de estudantes brancos ou asiáticos, proporção de homens e, finalmente, proporção de estudantes que trabalham e ajudam no sustento da família.

3.3.1 Renda familiar, educação do pai e origem escolar

Observa-se, na tabela 3.4 abaixo, os dez cursos com maior proporção de alunos com renda familiar menor que três salários mínimos, com pai sem escolaridade e provenientes de escolas públicas. As licenciaturas são os cursos que se destacam nesses três fatores socioeconômicos, junto com os cursos de serviço social e pedagogia, mostrando que esses cursos concentram grande parte da parcela dos alunos de origens socioeconômicas menos privilegiadas.

Tabela 3.4: Cursos com maior proporção de alunos com renda menor que 3 S.M., pai sem escolaridade e provenientes de escolas públicas nos dois ciclos

Cursos	1º Ciclo (%)	2º Ciclo(%)	Variação(%)
Maior proporção de alunos com Renda menor que 3 S.M.			
Biologia(Lic)	32,7	61,25	20,08
Filosofia(Lic)	19,88	58,44	20,63
Física(Lic)	36,13	56,22	22,52
Geografia	35,17	62,12	21,5
História(Lic)	41,16	60,23	38,55
Letras(Lic)	41,15	61,78	28,56
Matemática(Lic)	28,69	59,24	30,55
Pedagogia	42,22	64,74	26,95
Química(Lic)	34,76	56,26	32,82
Serv. Social	33,54	66,36	19,07
Média	34,54	60,66	26,12
Maior proporção de alunos com pai sem escolaridade			
C.Sociais(Lic)	6,84	12,33	5,49
Computação(Lic)	2,57	14,97	12,4
Filosofia(Lic)	4,42	15,48	11,06
Física(Lic)	6,82	12,5	5,68
Geografia(Lic)	10,31	14,5	4,19
História(Lic)	9,8	14,89	5,09
Letras(Lic)	9,3	14,69	5,39
Matemática(Lic)	8,01	15,26	7,25
Pedagogia	11,09	16,47	5,38
Serv, Social	7,24	14,75	7,5
Média	7,64	14,58	6,94
Maior proporção de alunos de escolas públicas			
Biologia(Lic)	68,46	79,45	10,99
Contabilidade	76,24	82,53	6,29
Filosofia(Lic)	53,52	81,14	27,61
Física(Lic)	73,04	79,43	6,38
Geografia(Lic)	79,7	85,89	6,19
História(Lic)	75,18	84,12	8,93
Letras(Lic)	79,42	84,14	4,72
Matemática(Lic)	81,99	86,69	4,7
Pedagogia	81,59	90,15	8,55
Serv. Social	73,02	84,28	11,26
Média	74,21	83,78	9,56

Fonte: Inep Enade, elaborado pelo autor

Nota-se, na tabela 3.4, que nos cursos de licenciatura, pedagogia e serviço social, a proporção de alunos com renda familiar de até três salários mínimos no segundo ciclo é muito maior que a proporção no campus geral (42,68%, tabela 3.3), a maioria dos cursos tem uma proporção com média acima de 60%. Além disso, nota-se a variação média de 26,12% na proporção de alunos dessa faixa de renda entre os dois ciclos, é maior que a registrada no ensino superior como um todo, 15,6%, visto na tabela 3.3.

Quanto a variável estudantes de pai sem nenhuma escolaridade, percebe-se que os dez cursos com a maior proporção desses estudantes são de licenciatura, serviço social e pedagogia. A variação média de estudantes de pai sem escolaridade entre os ciclos dos cursos da tabela 3.4 foi de 6,94%, mais que o dobro da variação no ensino superior como um todo, de 2,53%, como mostrado na tabela 3.3.

As licenciaturas, pedagogia e serviço social também são os cursos que possuem maior proporção de alunos provenientes de escolas públicas, chegando a ter 90% no curso de pedagogia. Vale lembrar que, segundo ranking dos cursos de acordo com o salário esperado elaborado Neri(2013), apêndice A, licenciatura e pedagogia são cursos ligados a educação e formação de professores, com o salário médio mais baixo do ranking, R\$1.844,29, sendo um índice da grande desigualdade horizontal presente no ensino superior.

A tabela 3.5, ao contrário da tabela 3.4, mostra os dez cursos com menor proporção de alunos com renda familiar menor que três salários mínimos; com menor proporção de alunos com pai sem escolaridade e provenientes de escolas públicas nos dois ciclos do Enade estudados.

Tabela 3.5: Cursos com menor proporção de alunos com renda menor que 3 S.M., com pai sem escolaridade e provenientes de escolas públicas nos dois ciclos

Cursos	1º Ciclo(%)	2º Ciclo(%)	Variação(%)
Menor proporção de alunos com Renda menor que 3 S.M.			
Arquitetura	8,47	18,01	9,54
Economia	21,86	22,62	0,76
Engenharia Ambiental	6,85	27,48	20,63
Engenharia Civil	11,18	19,28	8,09
Engenharia De Produção	6,11	17,38	11,26
Engenharia Grupo 2	9,06	17,62	8,56
Engenharia Grupo 4	8,94	21,47	12,53
Engenharia Mecânica	9,47	15,62	6,15
Medicina	9,35	8,54	-0,81
Odontologia	9,83	20,2	10,37
Média	10,11	18,82	8,71
Menor proporção de alunos com pai sem escolaridade			
Arquitetura	1,24	1,85	0,61
Biologia	2,65	2,68	0,03
Com Social	1,75	2,33	0,58
Computação	2,06	2,57	0,51
Engenharia Grupo 2	1,7	2,27	0,57
Engenharia Grupo 4	1,27	2,05	0,78
Engenharia Mecânica	1,63	2,00	0,37
Medicina	0,55	0,84	0,29
Odontologia	0,73	1,68	0,95
Veterinária	1,07	1,97	0,9
Média	1,46	2,02	0,56
Menor proporção de alunos de escolas públicas			
Arquitetura	35,66	41,29	5,63
C.Sociais	54,32	54,98	0,66
Com Social	43,49	46,63	3,14
Direito	55,22	55,38	0,16
Economia	58,09	50,58	-7,52
Engenharia Civil	48,43	51,38	2,95
Engenharia Mecânica	50,56	55,08	4,52
Medicina	13,01	15,6	2,59
Odontologia	25,42	40,76	15,34
Veterinária	33,51	43,94	10,43
Média	41,77	45,56	3,79

Fonte: Inep Enade, elaborado pelo autor

Na questão da renda familiar, medicina é o curso com menor proporção de alunos com renda menor que três salários mínimos, apenas 9,35% e 8,54% no primeiro e segundo ciclo respectivamente, sendo o único curso no qual a proporção de alunos dessa faixa de renda diminuiu entre os ciclos. Ademais, Medicina é o curso com menor proporção de alunos com pai sem nenhuma escolaridade, apenas 0,84% no segundo ciclo, e também o curso com menor proporção de alunos de escolas públicas, apenas 15,6% no segundo ciclo.

Nota-se, na tabela 3.5, a pequena variação média entre os cursos que possuem alunos com renda menor que três salários mínimos: apenas 8,71%, na proporção de alunos com renda familiar menor que três salários mínimos (nos campi como um todo a variação foi 15,6%, tabela 3.3)

Houve uma variação muito pequena na proporção de alunos com pai sem nenhuma escolaridade nos cursos da tabela 3.5, apenas 0,56%, evidenciando que, apesar da melhora nos campi em geral quanto a esse quesito, uma parcela muito pequena desses alunos ingressaram nos cursos de arquitetura, engenharia grupo 2, engenharia grupo 4, engenharia mecânica, medicina e odontologia.

Houve uma significativa melhora no ingresso de alunos de escolas públicas nos cursos de arquitetura(5,63%), engenharia mecânica(4,52%) e odontologia(15,34%). Porém, verifica-se uma queda no curso de economia(-7,52%).

Na próxima seção, faz-se a mesma análise realizada nessa seção, entretanto, considerando as variáveis cor, gênero e proporção dos estudantes que trabalham e ajudam no sustento da família.

3.3.2 Cor, gênero e situação do estudante

A tabela 3.6 abaixo mostra os dez cursos com a menor proporção de brancos/asiáticos, cursos com a maior proporção de alunos que sustentam a família e cursos com menor proporção de homens, enquanto a tabela 3.7 mostra o inverso, ou seja, os cursos com maior proporção de alunos brancos/asiáticos, cursos com menor proporção de alunos que sustentam a família e cursos com maior proporção de homens.

Tabela 3.6: Cursos com menor proporção de alunos brancos/asiáticos; maior proporção de alunos que ajudam no sustento da família e menor proporção de homens nos dois ciclos

Cursos	1º Ciclo(%)	2º Ciclo(%)	Variação(%)
Menor proporção de alunos brancos/asiáticos			
Biologia(Lic)	67,8	47,26	-20,55
C,Sociais(Lic)	58,41	45,31	-13,09
Física(Lic)	62,76	42,62	-20,14
Geografia(Lic)	56,64	42,66	-13,99
História(Lic)	58,28	42,82	-15,46
Letras(Lic)	61,82	43,13	-18,69
Matemática	66,53	41,86	-24,68
Pedagogia	61,84	48,02	-13,81
Química(Lic)	65,95	43,69	-22,27
Serv. Social	62,29	40,49	-21,81
Média	62,23	43,79	-18,45
Maior proporção de alunos que ajudam no sustento família			
C.Sociais(Lic)	30,86	43,92	13,06
Computação(Lic)	22,69	53,05	30,35
Contabilidade	38,95	41,59	2,64
Filosofia(Lic)	29,05	41,48	12,43
Física(Lic)	28,32	38,10	9,78
Geografia(Lic)	35,71	42,19	6,48
História(Lic)	33,97	42,30	8,32
Letras(Lic)	34,47	40,56	6,09
Matemática(Lic)	36,25	48,47	12,23
Pedagogia	40,75	47,67	6,93
Média	33,11	43,94	10,83
Menor proporção de alunos homens			
Biologia	33,34	30,26	33,61
Biologia(Lic)	27,86	27,78	43,44
Enfermagem	15,24	14,38	22,78
Letras	23,3	25,88	25,24
Letras(Lic)	17,22	22,04	25,17
Odontologia	36,10	31,41	33,00
Pedagogia	8,18	6,79	19,90
Psicologia	16,81	17,42	34,50
Serv. Social	6,62	9,10	28,74
Média	21,66	21,19	29,28

Fonte: Inep Enade, elaborado pelo autor

Na tabela 3.6 tem-se os dez cursos com menor proporção de brancos/asiáticos e, conseqüentemente, com maior proporção de pardos/negros, sendo esses majoritariamente de licenciatura, junto com pedagogia, serviço social e matemática. Nota-se que a maioria desses cursos são voltados para a educação ou formação de professores. Esses cursos tiveram uma redução na proporção de brancos/asiáticos e o curso de pedagogia se tornou o curso com maior proporção de pardos/negros, 51,98%(100%-48,02%) no segundo ciclo, número muito maior que a taxa geral de pardos/negros nos campi no segundo ciclo, 33,55% segundo a tabela 3.3.

A tabela 3.6 mostra que os dez cursos com maior proporção de alunos que sustentam a família são todos de licenciatura e um de pedagogia. Houve uma variação média de 10,83% entre os dois ciclos na proporção desses alunos nesses cursos. Por fim, vê-se na tabela 3.6 a pequena proporção de homens nos cursos de serviço social e pedagogia, apenas 6,79% e 9,10% no segundo ciclo.

A tabela 3.7, ao contrário da tabela 3.6, mostra os dez cursos com maior proporção de brancos/asiáticos, os dez cursos com menor proporção de alunos que ajudam no sustento da família e maior proporção de alunos homens.

Tabela 3.7: Cursos com maior proporção de alunos brancos/asiáticos; menor proporção de alunos que ajudam no sustento da família e maior proporção de homens nos dois ciclos

Cursos	1º Ciclo(%)	2º Ciclo(%)	Varição(%)
Maior proporção de alunos brancos/asiáticos			
Arquitetura	78,68	73,49	-5,19
Com Social	73,99	68,41	-5,57
Economia	70,34	66,95	-3,38
Engenharia de produção	77,18	66,41	-10,77
Engenharia Grupo 4	76,98	70,1	-6,88
Engenharia Mecânica	76,7	69,1	-7,6
Medicina	79,85	73,86	-5,99
Odontologia	80,9	71,8	-9,1
Química	72,9	67,69	-5,21
Veterinária	81,34	70,45	-10,89
Média	76,89	69,83	-7,06
Menor proporção de alunos que ajudam no sustento família			
Arquitetura	11,03	12,25	1,22
Biologia	12,83	12,23	-0,6
Engenharia Florestal	7,95	7,39	-0,56
Engenharia Grupo 4	11,27	11,3	0,03
Física	8,72	12,80	4,08
Matemática	27,95	12,36	-15,59
Medicina	1,41	1,19	-0,22
Nutrição	7,52	8,64	1,12
Odontologia	2,2	3,83	1,63
Veterinária	2,71	3,84	1,13
Média	9,35	8,58	-0,78
Maior proporção de alunos homens			
Computação	81,05	87,58	6,52
Computação(lic)	86,01	60,64	0,74
Engenharia Civil	79,04	71,85	-7,19
Engenharia De Produção	79,8	69,88	-9,92
Engenharia Grupo 2	90,99	88,48	-2,51
Engenharia Mecânica	94,37	91,22	-3,15
Filosofia	68,8	68,89	0,09
Física	77,13	70,51	-6,62
Física(Lic)	71,35	71,18	-0,16
Matemática	50,18	70,69	20,51
Média	77,87	75,09	-0,169

Fonte: Inep Enade, elaborado pelo autor

. O curso que possui maior proporção de brancos/asiáticos no segundo ciclo é medicina, com 73,86% , seguido por arquitetura, 73,49% e odontologia 71,8%. Verificou-se que, em todos os cursos, a proporção de brancos diminuiu. A média dessa queda, entre esses dez cursos, foi de apenas 7,06%, menos que a metade da queda na proporção de brancos/asiáticos nos dez cursos com menor proporção de brancos/asiáticos, 18,45% , visto na tabela 3.6.

A tabela 3.7 também mostra os dez cursos com menor proporção de alunos que sustentam a família. Medicina é o curso com a menor proporção no segundo ciclo, apenas 1,19%, seguida por odontologia, 3,83%. Esses indicadores apontam que esses cursos não permitem uma jornada de estudo acompanhada de trabalho, o que impossibilita o ingresso de alunos de origens sociais que precisam trabalhar e contribuir para o sustento da família. Fica nítida a discrepância desses cursos comparada aos cursos voltados a educação, como pedagogia com 47,67%, na proporção de alunos que ajudam no sustento da família.

Por fim, na tabela 3.7, tem-se os dez cursos com maior proporção de homens, sendo majoritariamente ligados a ciências exatas, com exceção de filosofia.

No geral, a análise mostra uma grande discrepância entre os cursos quanto as variáveis socioeconômicas selecionadas para indicar estudantes de origem social menos e mais privilegiada. Os cursos de licenciaturas, pedagogia e serviço social, por exemplo, possuem grande proporção de alunos de origem menos privilegiadas, enquanto que cursos como medicina, odontologia, determinadas engenharias, economia e arquitetura se mostraram menos inclusivos. Portanto, fica claro que a desigualdade horizontal ainda é elevada no ensino superior brasileiro apesar da queda da desigualdade vertical.

Na próxima seção, busca-se através de regressões logísticas analisar o viés de

distribuição dos alunos entre cursos com distintos retornos salariais usando a base socioeconômica do Enade e indicadores de retorno salarial por curso.²

² Usa-se o ranking salarial por curso acadêmico realizada por Neri(2013), apêndice A, como proxy para prestígio do curso, pois acredita-se que o retorno salarial da carreira ligada ao curso no mercado de trabalho é um indicador eficiente e prático para analisar a desigualdade horizontal dentro dos campi do ensino superior brasileiro.

4 DADOS E ESTRATÉGIA EMPÍRICA

Verificou-se que, entre os dois ciclos aqui analisados do Enade, houve uma maior participação de estudantes de origens sociais menos privilegiadas no ensino superior, porém essa maior participação foi desigual entre os cursos. Alunos de classes sociais menos privilegiadas parecem estar se concentrando em cursos com menor retorno financeiro, enquanto que aluno de classes sociais mais privilegiadas, nos cursos de maior retorno financeiro.

Neste capítulo, elabora-se uma estratégia empírica a fim de averiguar por modelos regressivos a desigualdade horizontal no ensino superior.

Dividi-se os cursos de acordo com seu retorno salarial esperado e a partir da base do Enade, verifica-se a influência das características socioeconômicas sobre a probabilidade dos alunos frequentarem cursos com distintos retornos salariais esperados.

Os modelos logísticos são usados em escolhas categóricas e segundo Greene (1997) o indivíduo escolhe uma alternativa que maximiza sua utilidade. Sendo a probabilidade do indivíduo i escolher a alternativa m dada como: $Prob(U_{im}) > U_{ik}$ para todo $j \neq k$ e a utilidade para escolha da alternativa m dada: $U_{im} = \beta x_{im} + \epsilon_{im}$: sendo ϵ_{im} o termo aleatório.

Segundo Greene(1997) se e somente se os termos aleatórios forem independentes e identicamente distribuídos, tem-se o modelo logístico da probabilidade do

indivíduo i escolher alternativa m dada na equação 4.1 abaixo:

$$Pr(Y_i = m) = \frac{(\beta'_m X_i)}{\sum_{k=1}^M (\beta'_k X_i)} \quad (4.1)$$

Nesta dissertação m representa o grupo de cursos frequentado por um indivíduo i . Os cursos foram agrupados de acordo com o salário esperado, isto é, nos modelos com 4 categorias $m=0$ representa o grupo de cursos de baixos salários, $m=1$ para médio-baixos salários, $m=2$ para médio-altos salários e $m=3$ para altos salários. X_i é o vetor das características do estudante i que incluem variáveis socioeconômicas como raça, renda, educação dos pais; variáveis quanto a habilidade do estudante como horas de estudo, livros lidos, dificuldade da prova e por fim características da instituição do estudante como privada ou pública, federal ou estadual e região da instituição. Já o vetor β é o parâmetro relevante que identifica a influência de determinada característica de um estudante i , nas chances de frequentar determinado grupo de cursos.

Usa-se 4 modelos logísticos(binomial, multinomial, ordinal proporcional e não proporcional) a fim de escolher o modelo com os melhores indicadores de ajuste.

Assume-se que esses estudantes estão no processo de conclusão do curso devido a obrigatoriedade de realizar o Enade com no mínimo 75% do curso concluído e a fim de isolar o efeito das variáveis de interesse, usa-se variáveis de controle que determinam as habilidades do estudante, características da instituição e características individuais do aluno.(Ver apêndice C)

No modelo logístico binomial a variável dependente é dividida em apenas 2 categorias de curso (cursos com salário esperado até 3000R\$ e acima de 3000R\$), ver tabela 4.1, enquanto que nos outros modelos logísticos, desmembra-se a variável dependente em 4 categorias de cursos(baixos, médio-baixos, médio-altos e altos salários).

4.1 Variáveis dependentes e independentes

4.1.1 Variável dependente

A variável dependente é a probabilidade de estudantes de nível superior estarem matriculados em cursos classificados pela faixa salarial. Essa classificação foi feita com base no ranking salarial elaborado por Neri(2013), que usou os dados censo de 2010 para encontrar o retorno esperado dos cursos universitário, conforme pode ser visto no Apêndice A.

A tabela 4.1 ilustra a divisão dos cursos de acordo com faixas salariais. Vê-se que cursos ligados a exatas, como diversas engenharias, e na área de saúde, como medicina e odontologia, possuem maiores retornos salariais enquanto que, cursos ligados a humanas, especialmente na área educacional como pedagogia ou licenciaturas, possuem baixo retorno salarial.

Tabela 4.1: Divisão de cursos por faixas salariais

Cursos com salários até 3000R\$		Cursos com salários acima de 3000R\$	
Baixos até 2000R\$	Médio-baixos de 2000 a 3000R\$	Médio-altos de 3000 a 4000R\$	Altos Acima 4000R\$
Pedagogia	Administração	C.Sociais	Medicina
Biologia(Lic)	Biologia	Com Social	Engenharia Civil
C.Sociais(Lic)	Computação	Economia	Engenharia Grupo 4
Computação(Lic)	Contabilidade	Engenharia Grupo 2	Engenharia Mecânica
Filosofia(Lic)	Enfermagem	Psicologia	Odontologia
Física(Lic)	Engenharia De Produção	Veterinária	Arquitetura
Geografia(Lic)	Farmácia		Direito
Historia(Lic)	Filosofia		Engenharia Ambiental
Letras(Lic)	Física		
Matemática(Lic)	Geografia		
Química(Lic)	Historia		
	Letras		
	Matemática		
	Química		
	Serv. Social		

Fonte: Elaborado pelo autor de acordo com o ranking salarial de Neri(2013) no âmbito do Ipea

Os cursos de licenciaturas e pedagogia estão como educação e formação de professores no ranking do Ipea(2013)

A tabela 4.2 abaixo, apresenta a distribuição dos alunos nos grupos de cursos por faixas salariais nos dois ciclos estudados do Enade. Verifica-se a maior proporção de alunos nos cursos de baixos e médio-baixos salários, isto é nos cursos até 3000R\$. No primeiro ciclo a proporção era de 67,13% e no segundo de 65,54%. Porém houve um grande aumento na proporção de alunos nos cursos de baixos salários(até 2000R\$), de 20,62% para 27,34% e uma queda na proporção de alunos no médio-baixos salários(de 2000\$ até 3000R\$), de 46,51% para 38,2%.

Já nos cursos acima de 3000R\$, houve uma queda entre os dois ciclos nos cursos de médio-altos salários(de 3000R\$ a 4000R\$), de 16,53% para 10,9% e um aumento nos cursos de altos salários(acima de 4000R\$), de 16,34% para 23,55%.

Tabela 4.2: Distribuição da frequência da variável grupo_salarial dependente em 4 grupos

Grupo cursos	1º Ciclo		2º Ciclo	
	N	%	N	%
Baixos salários	157,456	20.62	251,934	27.34
Médio-baixos salários	355,14	46.51	352,016	38.20
Cursos médio-altos salários	126,201	16.53	100,46	10.90
Cursos altos salários	124,747	16.34	217,047	23.55
Total	763.544	100,00	921.457	100,00

4.1.2 Variáveis independentes de interesse e controle

Na tabela 4.3, tem-se descrição e distribuição das variáveis de interesse por faixa salariais dos cursos estudados. Essas variáveis de interesse são as variáveis socioeconômicas escolhidas a fim de analisar a desigualdade horizontal no ensino superior.

Além disso, de acordo com a revisão da literatura empírica feita anteriormente, variáveis que determinam as preferências e o ambiente universitário também influenciam a escolha do curso. Considerando isso, foram selecionadas variáveis de controle divididas em habilidades individuais, características do aluno e características da instituição que o aluno frequenta, conforme podem ser vistas no apêndice C.

Tabela 4.3: Descrição e distribuição das variáveis de interesse por faixas salariais dos cursos estudados entre os ciclos

Descrição	Baixos salários		Médio-baixos		Médio-altos		Altos salários	
	1º Ciclo (%)	2º Ciclo (%)	1º Ciclo (%)	2º Ciclo (%)	1º Ciclo (%)	2º Ciclo (%)	1º Ciclo (%)	2º Ciclo (%)
(0) Renda até 3 S.M.	38,63	62,62	26,08	40,25	18,17	27,36	16,31	23,77
(1) Renda de 3 até 10 SM.	51,02	34,8	52,73	49,86	47,16	52,09	43,54	51,3
(2) Renda de 10 até 30 SM	9,49	2,39	18,74	8,23	29,33	16,57	32,22	19,84
(3) Renda maior que 30 SM	0,86	0,20	2,45	1,66	5,33	3,98	7,92	5,08
(0) negros/pardos	36,23	52,84	27,5	41,25	24,01	32,29	23,81	32,83
(1) brancos/asiáticos	63,77	47,16	72,5	58,75	75,99	67,71	76,19	67,17
(0) Escola Pública	77,41	86,96	66,6	74,16	47,41	53,2	45,02	49,58
(1) Escola Privada	22,59	13,04	33,4	25,84	52,59	46,8	54,98	50,42
(0) Pai sem educação	9,01	15,47	4,26	6,86	2,18	3,62	3,04	3,69
(1) Pai estudou até 4º série	41,12	44,97	30,67	33,28	18,05	20,24	19,59	20,64
(2) Pai estudou até 8º série	17,29	13,5	18,85	16,65	14,06	14,21	13,000	13,11
(3) Pai com ensino médio	21,97	19,42	29,03	31,0	31,96	37,38	28,41	35,79
(4) Pai com ensino superior	10,61	6,64	17,2	12,22	33,75	24,55	35,96	26,77

4.1.3 Correlação entre as variáveis independentes

Para verificar problemas de multicolinearidade nas variáveis categóricas, utiliza-se o teste Variance inflation factor test (VIF), indicado por Willians(2015), aplicando o comando Collin no stata. O teste foi aplicado nas variáveis da tabela 4.4, que são para o autor as mais sucessíveis a serem correlacionadas.

O teste mostrou uma alta taxa de tolerância a multicolinearidade entre elas, isto é, a colinearidade foi considerada baixa.

1

Tabela 4.4: Variance inflation factor test

Variable	VIF	SORT VIF	Tolerance	Squared
renda	1.27	1.12	0.7902	0.2098
cor	1.06	1.03	0.9417	0.0583
escola_privada	1.27	1.13	0.7889	0.2111
educa_pai	1.78	1.33	0.5622	0.4378
educa_mae	1.71	1.31	0.5844	0.4156
caso_trabalho	1.20	1.91	0.8358	0.1642
in _n oturno	1.11	1.05	0.8994	0.1006
sexo	1.02	1.01	0.9768	0.0232

Fonte:Inep Enade

Aplicação do comando collin no stata nas respectivas variáveis

4.2 Regressões logísticas

Nas próximas seções são apresentados os modelos logísticos.

¹"Allison says he gets concerned when the VIF is over 2.5 and the tolerance is under .40. In Stata you can use the vif command after running a regression, or you can use the collin command (Willians,2015 p. 04)."

4.2.1 Modelo logístico binomial

Quando uma variável resposta é dicotômica (0 ou 1, cursos até 3000R\$ ou cursos acima de 3000R\$), ver tabela 4.1, o modelo por meio de regressão logística é adequado. (Greene, 1997)

Sua probabilidade de sucesso é explicada pelos dados segundo a equação 4.2

$$P(Y = 1|x_i) = \frac{\exp(\beta_i x_i)}{1 + \exp(\beta_i x_i)} \quad (4.2)$$

Na qual $P(Y = 1|x_i)$ representa a probabilidade do aluno i estar no grupo de cursos com retorno salarial acima de 3000R\$(contra a probabilidade do aluno frequentar cursos com retorno salarial menor que 3000R\$), onde x'_i é um vetor com as observações das variáveis explicativas para a i -ésima observação da variável resposta e β é o vetor com os coeficientes de regressão (efeitos dos parâmetros do modelo).² Os resultados exponenciados da equação 4.2 se encontram no apêndice D.

A função de log-verossimilhança do modelo que deve ser maximizada é:

$$l(\pi; \underline{y}) = \log \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i} = \sum_{i=1}^n y_i \log(\pi_i) + (1 - y_i) \log(1 - \pi_i) \quad (4.3)$$

A interpretação mais prática dos resultados de modelos logísticos é feita a

²Para modelos em que a variável resposta segue uma distribuição Bernoulli a logit é vista como a função canônica de ligação.

partir de efeitos marginais seguindo a equação:

$$\frac{\partial P_m}{\partial x_i} = P_m \beta_{im} (1 - P_m) - P_m \sum_{h \neq k}^K P_k \beta_{ik} X_{ik} \quad (4.4)$$

4.2.2 Modelo multinomial

O modelo de regressão multinomial é adequado quando há no mínimo 3 variáveis categóricas na variável dependente e é usado em diversos trabalhos que buscavam a influência das características socioeconômicas na probabilidade dos alunos frequentarem distintos cursos ou área de conhecimento, destaca-se os trabalhos de Turner e Bowen(1997), Staniec(2004) e Dickson(2010).

O modelo é uma expansão do modelo logístico binomial(Greene, 1997) e é ilustrado na equação abaixo:

$$(Y_i = m | x_i) = \frac{\exp(Z_{mi})}{1 + \sum_{m=1}^3 \exp(Z_{mi})} \quad (4.5)$$

sendo

$$Z_{mi} = \alpha_m + \sum_{k=0}^K \beta_{mk} X_{ik}$$

Onde o vetor x_i representa as características de um indivíduo i e β_j o efeito dessas características na probabilidade do indivíduo i escolher o grupo de cursos m , sendo m dividido em quatro grupos de cursos de acordo com o retorno salarial($m=0$ cursos até 2000R\$, $m=1$ cursos entre 2000R\$ e 3000R\$, $m=2$ cursos entre 3000R\$ e 4000R\$ e $m=3$ cursos acima de 4000R\$), ver tabela 4.1.

Os resultados da equação 4.5³ estão exponenciados e apresentados no apêndice F.

4.2.3 Modelo odds proporcionais

O modelo de Odds Proporcionais pode ser aplicado quando a variável dependente é proveniente de uma variável contínua que foi agrupada (McCullagh(1980), Hosmer and Lemeshow(2000)). No caso desse trabalho, o retorno médio de cada carreira ligada ao curso é uma variável contínua que foi agrupada em quatro faixas salariais, como visto na tabela 4.1.

O modelo é baseado na função de distribuição cumulativa e seu pressuposto básico de ordenação é considerar que os coeficientes das covariáveis são os mesmos em todas as categorias de resposta, ou seja, as relações entre x e y são independentes dos níveis de resposta. Isto faz com que os coeficientes β sejam iguais para cada valor do corte da variável dependente.

A variável dependente Y consiste de M variáveis categóricas ordinais:

$$(Y \geq m) = \frac{(Y \geq m)}{(Y \leq m)} \quad (4.6)$$

Uma categoria é escolhida como referência(grupo de cursos de baixos salários) e $M-1$ cortes são estimados. Esses estimadores dão informação para cada categoria consecutiva sobre a probabilidade acumulada. A probabilidade de estar na categoria dada, ou numa menor, são tomadas em conta juntas neste tipo de probabilidade acumulativa. Segue o exemplo a seguir:

$$(Y < m | x) = \frac{\exp(\alpha_m + \eta)}{1 + \exp(\alpha_m + \eta)} \quad (4.7)$$

³Usa-se todas as variáveis de controle do apêndice C devido o teste stepwise selection, aplicado no stata com o comando xi:stepwise, mostrarem ser dignas de fazerem parte do modelo final. A explicação do teste é explicada no apêndice E

sendo

$$\eta = (\beta_0 + \sum_{k=1}^M \beta_k X_k)$$

O modelo pode ser transformado em linear elevando exponencialmente os resultados logaritmos (odds ratios).

Os resultados da regressão (equação 4.7)⁴ do modelo de odds proporcionais se encontram no apêndice G.

4.2.3.1 *Teste de brant*

Segundo Willians(2006), os resultados do modelo de odds proporcionais podem ser melhorados relaxando a hipótese que os coeficientes β são constantes entre os cortes da variável dependentes. Portanto, deve-se usar o teste de Brant a fim de verificar se β varia entre os cortes. A hipótese nula do teste é que os coeficientes são paralelos entre os cortes, isto é, são iguais entre os grupos de cursos da variável dependente. O teste de brant foi realizado pelo comando `brant` no programa `stata`. A hipótese foi rejeitada, ver apêndice H, logo os coeficientes das variáveis independentes não são constantes entre os cortes da variável dependente.

Ainda segundo Willians(2006), nesse caso, deve-se usar modelos mais complexos que permitem a variação dos β entre os cortes, como o modelo de odds não proporcionais (Non-proportional Odds Model) proposto por Fu's (1998) ou o modelo multinomial. Por fim, testa-se o modelo de odds não proporcionais.

⁴ Usa-se novamente o processo de regressão passo-a-passo automatizado de ajuste de modelos de regressão, a partir do qual são escolhidas as variáveis de controle que serão usadas no modelo. Segundo o programa `stata`, nenhuma das variáveis explicativas testadas devem ser excluídas do modelo.

4.2.4 Modelo odds não proporcionais

O modelo é uma extensão do modelo de odds proporcionais. A diferença entre o modelo de odds proporcionais é a adição de coeficientes γ nas variáveis independentes, que podem variar de acordo com o corte da variável dependente; Ao adicionar γ_{mk} , faz com que os coeficientes das variáveis independentes não sejam mais constantes entre os cortes. Segue a equação: $\eta = (\beta_0 + \sum_{k=1}(\beta_k + \gamma_{mk})X_k)$

O comando `gologit2`, criado por Willians(2006) para o programa `stata`, permite a aplicação do modelo de odds não proporcionais. Os resultados exponenciados se encontram no apêndice I.

4.3 Seleção do modelo

A estratégia agora é encontrar o modelo que melhor explica a variável dependente, que é uma medida categórica organizada de acordo com o retorno financeiro esperado de cada curso. Para verificação da qualidade de ajuste do modelo, usa-se como principal indicador o pseudo R^2_{MF} ou pseudo R^2 McFadden, os critérios AIC e BIC, além de uma medida de desvio do modelo(deviance) sugerido por Bellocco(2013) e Ari e Yildiz(2014).

O pseudo R^2_{MF} ⁵ desenvolvido por McFadden e Domencich(1974), é um dos indicadores de ajuste de qualidade mais usado em regressões logísticas e chama-se

⁵ O R^2_{MF} de McFadden é definido como

$$R^2 = 1 - \frac{\log(L_{vr})}{\log(L_r)}$$

onde L_{vr} é o máximo da função de verossimilhança da regressão não restrita e L_r é o máximo da função de verossimilhança da regressão restrita quando $\beta_1 = \beta_2 \dots = \beta_k$.

de "pseudo" pois tem o objetivo de ser análogo ao R^2 de regressões lineares.

Em modelos de regressão logística, os valores dessa estatística não costumam ser altos, pois ela é uma medida da variabilidade das previsões individuais, em vez das previsões para cada categoria.

As três estatísticas podem ajudar muito na seleção de modelos, especialmente se usadas em conjunto. De forma geral, quanto menor os valores de desvio do modelo (Deviance), AIC e BIC para um modelo, melhor é o ajuste do mesmo. Já para o pseudo R^2 , quanto maior forem seus valores, maior é a variabilidade dos dados explicada pelo modelo, o que aponta para um bom ajuste.

Vê-se na tabela 4.5 que o modelo multinomial possui os melhores indicadores entre os modelos não binomiais. Quando compara-se o modelo multinomial com o binomial, vê-se que o multinomial possui maior pseudo R^2_{MF} , já os indicadores AIC, BiC e Deviance são melhores no modelo logístico. Portanto, usa-se esses dois modelos a fim de interpretar os resultados na próxima seção, a partir de efeitos marginais, que informam a influência de dada característica socioeconômica na probabilidade de frequentar determinado grupo de cursos.

Tabela 4.5: Comparação da qualidade do ajuste dos três modelos

Logístico binomial		
Indicadores	Ciclo1	Ciclo2
Mac Fadden R2	0.1311	0.193
AIC	366.690	606.053
BIC	367.192	606.595
Deviance	759.516	1.212.886
Odds não proporcionais		
Indicadores	Ciclo1	Ciclo2
Mac Fadden R2	0.141	0,202
AIC	769.570	1.235.215
BIC	770.985	1.237.707
Deviance	769.306	1.235.951
Odds proporcinais		
Indicadores	Ciclo1	Ciclo2
Mac Fadden R2	0.087	0,145
AIC	821.090	1.331.228
BIC	821.583	1.331.748
Deviance	820.998	1.331.136
Logístico multinomial model		
Indicadores	Ciclo1	Ciclo2
Mac Fadden R2	0.153	0.212
AIC	760.170	1.213.474
BIC	763.675	1.216.795
Deviance	759.516	1.212.886

5 RESULTADOS

Deseja-se saber a probabilidade de cada categoria da variável de interesse estar em determinado grupo de cursos com distintas faixas salariais e como essa probabilidade varia por ciclo. Para isso, deve-se interpretar os resultados por efeitos marginais.

Os efeitos marginais (equação 4.4) nos diz a probabilidade de uma categoria da variável de interesse estar em determinado grupo da variável dependente, controlando pelos efeitos das demais variáveis independentes. O efeito marginal é facilmente encontrado aplicando o comando `margins` após os modelos de regressão escolhidos serem aplicados no programa Stata. (Williams, 2012) O comando `margins` calcula as estatísticas marginais, que provêm das probabilidades previstas pelo modelo que se ajustou. Sua interpretação é mais fácil de compreender empiricamente.

Segue, nas próximas tabelas os resultados dos efeitos marginais das variáveis de interesse para os dois ciclos estudados do Enade, a partir dos modelos logísticos binomial e multinomial.

A variável dependente são os grupos de cursos divididos por faixas salariais, apresentados no capítulo anterior.

5.1 Efeitos marginais modelo logístico binomial

5.1.1 Primeiro ciclo Enade

Vê-se na tabela 5.1 que quanto maior a renda, maior a probabilidade de estar nos cursos de altos salários (no modelo binomial acima de 3000R\$), por exemplo, alunos com renda familiar até 3 salários mínimos ($\text{renda}=0$) possuem 30,4% de probabilidade de frequentarem cursos de altos salários, já alunos com renda familiar maior que 30 salários mínimos ($\text{renda}=3$) possuem 48,1%.

O mesmo acontece para educação dos pais, quanto maior a educação dos pais, maior a probabilidade dos alunos estarem em cursos de altos salários, por exemplo, filhos de pais sem educação formal possuem 29,4% de probabilidade de frequentarem cursos de altos salários, já filhos de pais com ensino superior ($\text{educa_pai}=4$) possuem 42,3%

Quanto a variável tipo de escola, alunos de escolas privadas ($\text{escola_privada}=1$) possuem 39,9% de probabilidade de estarem nos cursos de alto salários, já alunos de escolas públicas, 32,2%. Já na variável cor, alunos brancos/asiáticos ($\text{cor}=1$) possuem 36,67% de probabilidade de frequentarem cursos de altos salários, já negros/pardos possuem 34,3%.

Tabela 5.1: Efeitos marginais modelo logístico binomial primeiro ciclo

Observations	322,557
Pseudo R2	0.1309
VARIABLES	(1)
renda=0	0.304*** (0.00175)
renda=1	0.344*** (0.00112)
renda=2	0.431*** (0.00183)
renda=3	0.481*** (0.00462)
cor=0	0.343*** (0.00157)
cor=1	0.367*** (0.000905)
escola_privada=0	0.332*** (0.00111)
escola_privada=1	0.399*** (0.00136)
educa_pai=0	0.294*** (0.00473)
educa_pai=1	0.314*** (0.00185)
educa_pai=2	0.334*** (0.00202)
educa_pai=3	0.370*** (0.00154)
educa_pai=4	0.423*** (0.00198)
Additional controls	Yes

Standard errors in parentheses: *** $p \leq 0,01$, ** $p \leq 0,05$, * $p \leq 0,1$

Fonte: Inep Enade, elaborado pelo autor

(1) Cursos retorno acima de 3000R\$

5.1.2 Segundo ciclo Enade

No segundo ciclo vê-se algo parecido com o ocorrido no primeiro ciclo. A renda influencia na probabilidade de estar em cursos de altos salários (no modelo binomial acima de 3000R\$), alunos com renda menor que 3 salários mínimos (renda=0) possuem 25% de probabilidade de frequentarem cursos de alto salários, já alunos com renda maior que 30 salários mínimos (renda=3) possuem 40,5% de probabilidade.

Para variável educação dos pais, quanto maior a educação, maior a probabilidade dos alunos frequentarem cursos de altos salários, por exemplo, 27,4% para alunos de pais sem educação formal e 37,6% para alunos de pais com ensino superior.

Quanto a variável tipo de escola, alunos de escolas privadas possuem 38,6% de probabilidade de frequentarem cursos de alto salários, número maior que alunos de escolas públicas, 30%. Já para variável cor, alunos brancos/asiáticos possuem 33,4% de probabilidade de frequentarem cursos de alto salários e negros/pardos 31,3%.

Tabela 5.2: Efeitos marginais modelo logístico binomial segundo ciclo

Observations	594.766
Pseudo R2	0.19
VARIABLES	(1)
renda=0	0.254*** (0.000892)
renda=1	0.356*** (0.000810)
renda=2	0.447*** (0.00213)
renda=3	0.405*** (0.00498)
cor=0	0.313*** (0.000866)
cor=1	0.334*** (0.000701)
escola_privada=0	0.300*** (0.000683)
escola_privada=1	0.386*** (0.00119)
educa_pai=0	0.274*** (0.00243)
educa_pai=1	0.295*** (0.00110)
educa_pai=2	0.316*** (0.00141)
educa_pai=3	0.343*** (0.00103)
educa_pai=4	0.376*** (0.00162)
Additional controls	Yes

Standard errors in parentheses: *** $p \leq 0,01$, ** $p \leq 0,05$, * $p \leq 0,1$

Fonte: Inep Enade, elaborado pelo autor

(1) Cursos retorno acima de 3000R\$

5.1.3 Comparação entre os ciclos

Quando compara-se as diferenças de probabilidade entre os ciclos vemos que a probabilidade em frequentar cursos com retorno acima de 3000R\$, caiu independente das variáveis socioeconômicas(exceto para renda=1).

Na variável renda por exemplo, alunos com renda familiar menor que 3 salários mínimos(renda=0), tinham 30,4% de probabilidade de estudar em cursos de altos salários no primeiro ciclo, já no segundo ciclo essa probabilidade é 25,4%, entretanto a probabilidade de frequentar cursos de alto salários também caiu para os alunos com renda familiar maior que 30 salários mínimos(renda=3), de 48,1% para 40,5%. Quanto aos alunos com renda familiar entre 3 a 10 salários mínimos(renda=1) ou entre 10 a 30 salários(renda=2), houve diferenças muito pequena na probabilidade de frequentar cursos de altos salários entre os dois ciclos.

Quanto as variáveis educação dos pais, cor e tipo de escola, houve uma queda na probabilidade de frequentar cursos de altos salários entre todas as categorias, porém a desigualdade horizontal se manteve. Isto é, alunos de contexto socioeconômicos mais privilegiados possuem maior probabilidade de frequentar cursos de altos salários.

5.2 Efeitos marginais modelo logístico multinomial

No modelo multinomial, os cursos são divididos em 4 categorias como visto na tabela 4.1(Baixos salários até 2000R\$, Médio-baixos salários(de 2000R\$ a 3000R\$, Médio-altos salários(de 3000R\$ a 4000R\$, altos salários(acima de 4000R\$).

5.2.1 Primeiro ciclo Enade

Vê-se, na tabela 5.3, que alunos com renda familiar até 3 salários mínimos($renda=0$) têm 34,5% de probabilidade de estarem no grupo de cursos de baixos salários($grupo=0$). Essa probabilidade cai quanto maior a renda familiar: alunos de renda familiar entre 10 a 30 salários mínimos($renda=2$) e maior que 30 salários mínimos($renda=3$) possuem, respectivamente, 20% e 16,7,1% de probabilidade em estarem no grupo de cursos de baixos salários.

Já quando observamos o grupo de cursos com altos salários($grupo=3$), percebe-se que alunos com renda familiar menor que 3 salários mínimos possuem 14,1% de probabilidade de estarem nesse grupo, enquanto que alunos com renda familiar entre 10 e 30 e maior que 30 salários mínimos possuem, respectivamente, 21,2% e 27,2% de probabilidade de estarem nesse grupo.

Alunos de pais sem nenhuma escolaridade($educa_pai=0$) possuem 34,7% de probabilidade de estarem no grupo de cursos de baixos salários. Essa probabilidade cai quanto maior a educação do pai, chegando a 26,6% para alunos de pais com ensino superior($educa_pai=4$).

Já no grupo de cursos de altos salários, filhos de pais sem nenhuma educação formal($educa_pai=0$) possuem apenas 14,9% de probabilidade de estarem nesse grupo, enquanto que alunos de pais com ensino superior possuem 20,1% de probabilidade.

Quanto a variável cor dos alunos, vemos uma discrepância nos cursos de baixos salários: Brancos/asiáticos($cor=1$) possuem 27,8% de probabilidade de estarem no grupo de cursos de baixos salários, enquanto que negros/pardos($cor=0$) possuem 31,3% de probabilidade de estarem nesse grupo. Já alunos de escolas priva-

das(escola_privada=1) possuem 25,1% de probabilidade de estarem no grupo de cursos de baixos salários, enquanto que alunos de escolas públicas(escola_privada=0) possuem 30,7% de probabilidade. No grupo de médio-altos salários, alunos de escolas privadas possuem 20,5% de probabilidade de estarem nesse grupo de cursos e, para alunos de escolas públicas, esta probabilidade é de 17%.

Tabela 5.3: Resultados em efeitos marginais do modelo multinomial para o primeiro ciclo Enade

Observations	322.557			
Pseudo R2	0.153			
Variables	(0)	(1)	(2)	(3)
renda=0	0.345*** (0.00153)	0.347*** (0.00170)	0.167*** (0.00150)	0.141*** (0.00144)
renda=1	0.294*** (0.000994)	0.362*** (0.00111)	0.185*** (0.000952)	0.159*** (0.000909)
renda=2	0.200*** (0.00166)	0.366*** (0.00188)	0.222*** (0.00155)	0.212*** (0.00146)
renda=3	0.167*** (0.00447)	0.344*** (0.00478)	0.218*** (0.00368)	0.272*** (0.00380)
cor=0	0.313*** (0.00135)	0.341*** (0.00157)	0.183*** (0.00138)	0.163*** (0.00132)
cor=1	0.278*** (0.000826)	0.356*** (0.000918)	0.190*** (0.000765)	0.176*** (0.000734)
escola_privada=0	0.307*** (0.000917)	0.360*** (0.00109)	0.178*** (0.000972)	0.156*** (0.000946)
escola_privada=1	0.251*** (0.00131)	0.350*** (0.00141)	0.205*** (0.00113)	0.194*** (0.00106)
educa_pai=0	0.347*** (0.00620)	0.344*** (0.00784)	0.160*** (0.00741)	0.149*** (0.00724)
educa_pai=1	0.322*** (0.00177)	0.363*** (0.00221)	0.168*** (0.00207)	0.147*** (0.00199)
educa_pai=2	0.295*** (0.00188)	0.374*** (0.00220)	0.174*** (0.00189)	0.156*** (0.00188)
educa_pai=3	0.274*** (0.00173)	0.358*** (0.00180)	0.195*** (0.00142)	0.172*** (0.00141)
educa_pai=4	0.266*** (0.00316)	0.321*** (0.00297)	0.211*** (0.00228)	0.201*** (0.00220)
Additional controls	Yes			

Standard errors in parentheses: *** $p \leq 0,01$, ** $p \leq 0,05$, * $p \leq 0,1$

Fonte: Inep Enade, elaborado pelo autor

(0) Cursos baixo salários(até 2000R\$)

(1) Cursos médio-baixos salários(de 2000R\$ a 3000\$)

(2) Cursos médio-altos salários(de 3000R\$ a 4000R\$)

(3) Cursos altos salários(acima de 4000R\$)

5.2.2 Segundo ciclo Enade

Na tabela 5.4 temos os resultados para o segundo ciclo. Vê-se que alunos com renda familiar até 3 salários mínimos (renda=0) têm 35,8% de probabilidade de estarem no grupo de cursos de baixos salários. Essa probabilidade, assim como no primeiro ciclo, cai quanto maior a renda familiar. Para alunos de renda familiar entre 10 a 30 salários mínimos, a probabilidade é de 14,8%, chegando a 9,47% para alunos com renda familiar maior que 30 salários mínimos.

No grupo de cursos de altos salários, nota-se que alunos com renda familiar menor que 3 salários mínimos possuem 17,1% de probabilidade de estarem nesse grupo, enquanto que alunos com renda familiar entre 10 e 30 e maior que 30 salários mínimos possuem, respectivamente, 32,9% e 31,9% de probabilidade de estarem nesse grupo.

No segundo ciclo, alunos de pais sem nenhuma escolaridade (educa_pai=0) possuem 32,7% de probabilidade de estarem no grupo de cursos de baixos salários. Assim como no primeiro ciclo, essa probabilidade cai quanto maior a educação dos pais, chegando a 26,6% para alunos de pais com ensino médio (educa_pai=4) e 27,5% para alunos de pais com ensino superior.

Já alunos de pais sem nenhuma educação formal possuem 19% de probabilidade de estarem no grupo de cursos de alto salários, enquanto que, alunos de pais com ensino superior possuem 11,2% e 25,7% de estarem nos grupos de médio-altos salários e altos salários, respectivamente.

Branco/asiáticos possuem uma pequena diferença de probabilidade de estarem no grupo de baixos salários, sendo 28,8% para branco/asiáticos e 29,9% para negro/pardos. Já nos grupos de médio-altos salários e altos salários essa diferença

é ainda menor.

Quanto a variável alunos de escolas privadas(`escola_privada=1`), vê-se que eles possuem 23,9% de probabilidade de estarem no grupo de baixos salários, enquanto alunos de escolas públicas(`escola_privada=0`) possuem 30,1% de probabilidade. Já no grupo de cursos de altos salários, alunos de escolas privadas possuem 27% de probabilidade de estarem nessa categoria enquanto que alunos de escolas públicas possuem 20%.

Tabela 5.4: Resultados em efeitos marginais do modelo multinomial para o segundo ciclo Enade

Observations	594.766			
Pseudo R2	0.212			
Variables	0	1	2	3
renda=0	0.358*** (0.000813)	0.383*** (0.000957)	0.0877*** (0.000650)	0.171*** (0.000817)
renda=1	0.241*** (0.000692)	0.403*** (0.000846)	0.112*** (0.000573)	0.243*** (0.000747)
renda=2	0.148*** (0.00182)	0.393*** (0.00226)	0.131*** (0.00144)	0.329*** (0.00199)
renda=3	0.0947*** (0.00500)	0.463*** (0.00591)	0.123*** (0.00318)	0.319*** (0.00464)
cor=0	0.299*** (0.000715)	0.387*** (0.000917)	0.0957*** (0.000630)	0.218*** (0.000825)
cor=1	0.280*** (0.000638)	0.387*** (0.000756)	0.107*** (0.000498)	0.227*** (0.000637)
escola_privada=0	0.301*** (0.000546)	0.399*** (0.000712)	0.0973*** (0.000512)	0.203*** (0.000645)
escola_privada=1	0.239*** (0.00115)	0.373*** (0.00129)	0.118*** (0.000787)	0.270*** (0.00108)
educa_pai=0	0.327*** (0.00239)	0.390*** (0.00351)	0.0927*** (0.00278)	0.190*** (0.00340)
educa_pai=1	0.309*** (0.000913)	0.395*** (0.00128)	0.0918*** (0.000978)	0.205*** (0.00124)
educa_pai=2	0.281*** (0.00136)	0.403*** (0.00163)	0.103*** (0.00110)	0.213*** (0.00141)
educa_pai=3	0.266*** (0.00124)	0.392*** (0.00135)	0.111*** (0.000813)	0.231*** (0.00108)
educa_pai=4	0.275*** (0.00224)	0.356*** (0.00238)	0.112*** (0.00137)	0.257*** (0.00192)
Additional controls	Yes			

Standard errors in parentheses: *** $p \leq 0,01$, ** $p \leq 0,05$, * $p \leq 0,1$

Fonte: Inep Enade, elaborado pelo autor

(0) Cursos baixo salários(até 2000R\$)

(1) Cursos médio-baixos salários(de 2000R\$ a 3000\$)

(2) Cursos médio-altos salários(de 3000R\$ a 4000R\$)

(3) Cursos altos salários(acima de 4000R\$)

5.2.3 Comparação entre os ciclos

Ao comparamos os dois ciclos, vemos uma queda expressiva na probabilidade entre todas as variáveis de interesse em frequentar os cursos de médio-altos(grupo 2) salários e o aumento na probabilidade entre todas as variáveis de interesse em frequentar cursos de altos(grupo 3) salários. Houve também um leve aumento em frequentar cursos de médio-baixos salários. Isso esta coerente com o que foi visto na tabela 4.2, na qual a proporção de alunos nos cursos de médio-altos salários caiu entre os dois ciclos e subiu nos cursos de médio-baixos e nos de altos salários.

A probabilidade de alunos com renda familiar menor que 3 salários frequentarem o grupo de baixos salários praticamente se manteve entre os dois ciclos(34,5% para 35,8%), porém, caiu a probabilidade dos demais alunos(com renda familiar maiores que 3 salários mínimos) frequentarem o grupo de cursos de baixos salários. Por exemplo, para alunos com renda familiar maior que 30 salários mínimos, a probabilidade de frequentarem cursos de baixos salários caiu de 16,7% para 9,47%. Ainda no grupo de baixo salários, houveram pequenas variações na probabilidade entre os dois ciclos quanto a variável educação dos pais.

Para alunos de escola pública, a diferença só foi expressiva na queda da probabilidade de frequentarem cursos de médio-baixos salários(de 17,8% para 9,73% e um aumento na probabilidade de frequentarem cursos de altos salários(de 15,6% para 20,3%).

Já quanto a variável cor, as maiores mudanças foram no grupo de curso de médio-altos e altos salários. Negros/pardos tinham 18,3% de frequentarem cursos de médio-altos salários, essa probabilidade caiu para 9,57%, já nos cursos de altos salários, a probabilidade subiu de 16,3% para 21,8%.

6 CONCLUSÃO

A análise sobre o perfil dos estudantes no ensino superior mostrou que, apesar do aumento na proporção de estudantes de origens sociais menos favorecidas, ainda persiste um viés na distribuição desses alunos por cursos. Cursos ligados à educação como licenciatura e pedagogia, por exemplo, concentram alunos de origens socioeconômicas menos privilegiadas, enquanto que cursos como medicina, odontologia e engenharias, concentram alunos de origens sociais mais privilegiadas.

A fim de verificar o viés de distribuição entre os cursos, dividiu-se os cursos entre grupos de acordo com suas expectativas de retornos salariais e buscou-se o melhor modelo logístico regressivo (entre binomial, multinomial, ordinal proporcional e ordinal não proporcional), afim de identificar a influência das variáveis socioeconômicas dos alunos nas probabilidades de frequentarem cursos com distintos retornos salariais.

Os modelos logísticos binomial e multinomial foram escolhidos com o objetivo de analisar a desigualdade horizontal. A partir de efeitos marginais, obteve-se as probabilidades dos estudantes, com distintas características socioeconômicas, em frequentarem distintos grupos de cursos, de acordo com seus respectivos retornos salariais esperados.

A respeito das variáveis socioeconômicas, pode-se dizer que nos dois ciclos temporais estudados, a renda familiar e educação dos pais tem um grande impacto nos grupos de cursos com distintos retornos salariais esperados. Por exemplo, no modelo multinomial quanto maior a renda e a educação dos pais, menor a probabilidade dos estudantes frequentarem o grupo de baixos salários(até 2000R\$). Já

no grupo de cursos de altos salários(acima de 4000R\$), quanto maior a renda ou educação dos pais, maiores as probabilidades dos estudantes frequentarem.

Em respeito ao tipo de escola dos alunos, viu-se que alunos de escolas privadas possuem menor probabilidade de estarem no grupo de cursos de baixos salários(até 2000R\$). Já no grupo de cursos de altos salários(acima de 4000R\$), alunos de escolas privadas possuem maior probabilidade de frequentarem esse grupo quando comparados a alunos de escolas públicas.

Quanto à variável cor, viu-se que alunos brancos/asiático possuem menor probabilidade de estarem no grupo de cursos de baixos salários(até 2000R\$), quando comparados aos alunos negros/pardos, e maior probabilidade de estarem no grupo de cursos de altos salários(acima de 4000R\$), porém essa diferença parece ser pequena.

Pode-se afirmar, com os resultados apresentados, que existe uma desigualdade horizontal no ensino superior e essa desigualdade se manteve entre os dois ciclos.

No entanto, não é possível dizer se a desigualdade horizontal aumentou ou diminuiu entre os ciclos. Isto pode ser analisado mais facilmente no modelo logístico binomial, no qual mostra que entre os dois ciclos, as diferenças nas probabilidades entre as variáveis de interesse de frequentarem cursos acima de 3000R\$ continuaram, isto é, alunos de estratos sociais mais privilegiados continuam possuindo maior probabilidade de frequentarem esses cursos quando comparados aos alunos de estratos sociais menos privilegiados.

Conclui-se que a desigualdade horizontal é um problema no qual políticas públicas precisam estar mais atentas e que as mesmas não devem se basear apenas na inclusão de grupos historicamente excluídos no ensino superior, mas também na inclusão desses grupos em cursos de maiores prestígios e retorno financeiro esperados,

a fim de uma maior democratização nos campi brasileiros.

Em pesquisas futuras pretende-se focar na questão das políticas sociais e como os alunos ingressantes a partir de programas de inclusão governamentais com financiamentos ou por políticas de cotas, estão se distribuindo nos cursos do ensino superior brasileiro.

7 REFERÊNCIAS

ARI, E.; YILDIZ, Z. Parallel lines assumption in ordinal logistic regression and analysis approaches. **International Interdisciplinary Journal of Scientific Research**,v.1 ,n.1,p.8–23, 2014.

BARTALOTTI, O.; MENEZES-FILHO, N. A. A Relação entre o desempenho da carreira no mercado de trabalho e as escolhas profissionais dos jovens. **Revista de Economia Aplicada**. v.11, p.487-506, 2007.

BECKER, G. S. *The human capital*. Chicago: Chicago University, 1983

_____ *Accounting for tastes*. Cambridge, London: Harvard University Press, 1996

BELLOCO R., ALGERI S. Goodness-of-Fit Tests for categorical data.**Stata Journal**,v.13, n.2, p.356–365, 2013.

BETTS, J. R. What do students know about wages? Evidence from a survey of undergraduates. **Journal of Human Resources**, v.31, n.1, p.27-56, 1996

BOURDIEU, P.; PASSERON, J. C. Os herdeiros: os estudantes e a cultura. Florianópolis: Editora da UFSC, 2014.

BRAGA, M. M. ; PEIXOTO, M. C. L.; BOGUTCHI, T. F. Tendências da demanda pelo ensino superior: estudo de caso da UFMG. **Cadernos de Pesquisa**. São Paulo

, n.113, p.129-152, 2001.

BROWN, T. . Confirmatory factor analysis for applied research. Nova York: Editora Guilford, 2006.

CASARI, P., *Retorno esperado e escolha profissional: fatores associados à escolha da carreira dos alunos da Universidade de São Paulo*. 2006. 68 f. Dissertação (Mestrado em Economia) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo

DICKSON, L. Race and gender differences in college major choice. **The Annals of the American Academy of Political and Social Sciences**, v.627, p.108–124, 2010

DISTEFANO C.; ZHU M.; Mîndrilă D. Understanding and Using Factor Scores: Considerations for the Applied Researcher. *textbfPractical Assessment, Research & Evaluation*, v.14 ,n.20, 2009.

DRUKKER, D. M. Using margins to estimate partial effects. Italian Stata Users Group Meeting, Bologna, Italy, 2010.

DURU, B. M.; MINGAT, A. Comportement des bacheliers: modèle de choix de disciplines. **Consommation**, n.34, 1979

FERNANDES, R.; NARITA, R. Instrução superior e o mercado de trabalho no brasil., **Economia Aplicada**, v.5, n.1, 2001

FINNIE, R.;FRENETTE, M. Earnin differences by major field of study: Evidence from the cohorts of recent Canadian graduates. **Economics of Education Review** v.22,p. 179–192, 2003.

FU, V., K. Estimating Generalized Ordered Logit Models. **Stata Technical Bulletin**. StataCorp LP, vol. 8, n. 44, 1998.

GOUVEIA, A J. Democratização do ensino superior. **Revista Brasileira de Estudos Pedagógicos**, Rio de Janeiro, v.50, n.112, p.232-244, out./dez. 1968.

GREENE, W. H. *Econometric Analysis*. 5ed. Nova York:Editora Prentice Hall. 2003

HOSMER, D. W.; LEMESHOW, S. *Applied Logistic Regression*.**Second International Interdisciplinary Journal of Scientific Research** v.1, n.3, Dec/2014.

JAMES, E.; NABEEL, A.; CONATY, J.; TO, D. College quality and future earnings: where should you send your child to college? **American Economic Review** v.79, n.2, p.247–252, 1989.

MARTINS, F. S; CARUSI D. M. Determinantes do acesso ao ensino superior no Brasil e a probabilidade de sucesso por curso. Textos para discussão UFF/ECONOMIA. n.317, Dez/2015

MCCULLAGH P., *Regression Models for Ordinal Data*, **Journal of the Royal Statistical Society, Series B (Methodological)**, v.42, n.2, p. 109-142, 1980.

MCFADDEN, D. Conditional logit analysis of qualitative choice behavior. **Frontiers in Econometrics**, Academic Press, New York, p.105-142, 1974.

MINCER, J. *Schooling, experience and earnings*. Nova Iorque. National Bureau of Economic Research, 1974.

MONTMARQUETTE, C; CANNINGS, K; MAHSEREDJAN, S. How do people choose college majors? **Economics of Education Review**. v.21, p.543-556, 2002

NERI, M. C. . Escolhas Universitárias e Performance Trabalhista. Brasília: IPEA. Radar 27, p. 7-20, 2013 .

NOGUEIRA, C. M. M. *Dilemas na análise sociológica de um momento crucial das trajetórias escolares: o processo de escolha do curso superior*. 2004. 185f. Tese (Doutorado em Educação) – Universidade Federal de Minas Gerais, Belo Horizonte, MG.

NORES, M. Differences in college major choice by citizenship status. **The Annals of the American Academy of Political and Social Science**, v.627, n.1, p.125-141. 2010

PORTER, S. R;UMBACH, P. D. College Major Choice: An Analysis of Person-Environment Fit. **Research in Higher Education**, v.47, n.4, p.429-449, Jun/2006

REIS, M. MACHADO; CARUSI, D. Uma análise dos rendimentos do trabalho entre indivíduos com ensino superior no Brasil. Brasília: Ipea, n.2110, jul/2015.

RESENDE, M.; WYLLIE, R. Retornos para educação no Brasil: evidências empíricas adicionais. **Econ. Apl**, Ribeirão Preto , v.10, n.3, p. 349-365, 2006.

RIBEIRO, C. A. C. Juventudes e educação: escola e transições para a vida adulta no Brasil. Rio de Janeiro: Coleção Pronex-Juventude, 2014

RIBEIRO, S. COSTA; KLEIN, R. A divisão interna da universidade: posição social das carreiras. **Educação e Seleção**, n.5, p.29-43, jan./jun, 1982.

RISTOFF, D. O perfil socioeconômico do estudante de graduação: uma análise de dois ciclos completos do Enade (2004 a 2009). **Cadernos do GEA**, n.4, jul./dez, 2013.

RUMBERGER, R. W.; THOMAS S.L. The economic returns to college major, quality and performance: A multilevel analysis of recent graduates. **Economics of Education Review**. v.12, n.1, p.1-19, mar/1993.

SCHULTZ, T. O capital humano: investimento em educação e pesquisa. Rio de Janeiro: Zahar, 1973

SOARES, F. L. B., A escolha no ensino superior: fatores de decisão. 2007. 126 f. Dissertação (Mestrado em Economia) – Faculdade de Ciências Econômicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007.

STANIEC, J. F. O. The effects of race, sex, and expected returns on the choice of college major. **Eastern Economic Journal** v.30, n.4, p.549-69, 2004.

TAFNER, P.; CARVALHO, M. Educação superior no Brasil. Rio de Janeiro, 2008. Mimeografado.

THORSON, A. The Effect of College Major on Wages. **The Park Place Economist**. v.13,p.45-57, 2005 University: Senior Honors Project, 2005.

TURNER, S. E., BOWEN G. W. Choice of major: The changing (unchanging) gender gap. **Industrial and Labor Relations Review**, v.52 ,n.2, p.289-313, jan/1993

WILLIAMS, R. Generalized Ordered Logit/ Partial Proportional Odds Models for ordinal dependent variables, **Stata Journal** v. 6, n. 1, 2006

_____. (2012) Using the margins command to estimate and interpret adjusted predictions and marginal effects. **Stata Journal** v. 12, n. 2, p.308–331, 2012.

_____. Logistic Regression, Part III: Hypothesis Testing, Comparisons to OLS, 2015, < <https://www3.nd.edu/~rwilliam/stats2/Logistic-Stata.pdf> > Acesso em 5 jun, 2018.

_____. Multicollinearity. 2015, University of Notre Dame, < <https://www3.nd.edu/~rwilliam/stats2/l11.pdf> >. Acesso em 9 jun, 2018.

_____. Scalar Measures of Fit: Pseudo R2 and Information Measures (AIC & BIC), 2018, < <https://www3.nd.edu/~rwilliam/stats3/L05.pdf> > Acesso em 25 jun, 2018

APÊNDICE A RETORNO SALARIAL POR CURSO

Tabela A.1: Renumeração por curso

Cursos	R\$
Medicina	6940,12
Estatística	5416,1
Engenharia civil	4604,41
Engenharia química	4549,12
Serviços de transportes	4460,89
Setor militar e de defesa	4433,67
Engenharia mecânica e metalúrgica	4258,69
Odontologia	4238,65
Arquitetura e urbanismo	4206,01
Engenharia (outros)	4168,05
Direito	4104,84
Engenharia naval e aeronáutica	4089,29
Economia	3914,51
Veterinária	3782,46
Engenharia elétrica e automação	3734,1
Sociologia e ciência política	3672,37
Comunicação social (jornalistas)	3655,01
Psicologia	3529,77
Marketing e publicidade	3517,69
Artes	3229,36
Biblioteconomia	3140,29
Outras ciências físicas (gerais)	3124,56
Outras ciências sociais	3102,22
Administração e diversos	3057,31
Agronomia pecuária e pesca	2964,62
Farmácia	2964,38
Física	2961,39
Computação	2886,57
Turismo viagens e lazer	2884,15
Administração e secretariado	2866,27
Engenharia produção e processamento	2785,79
Serviços de segurança	2785,47

Tabela A.1: Renumeração por curso

Cursos	R\$
Outras saúde (reabilitação)	2609,16
Química	2606,01
Contabilidade e atuariais	2588,04
Biologia e ciências da vida	2454,9
Serviço social	2428,14
Educação física e esportes	2423,49
Historia e arqueologia	2304,05
Enfermagem	2282,46
Filosofia e ética	2261,28
Letras linguas e culturas	2228,33
Matemática	2216,96
Outros serviços pessoais (beleza e domésticos)	2205,83
Proteção ambiental	2187,21
Geologia e ciências da terra	2181,34
Educação e formação de professores	1844,29
Religião	1413,01

Fonte: Neri(2013)

APÊNDICE B PERGUNTAS SOCIOECONÔMICAS ENADE

“Qual a faixa de renda mensal da sua família?” As opções são:

- a) até 3 salários mínimos;
- b) mais de 3 até 10 salários mínimos;
- c) mais de 10 até 30 salários mínimos;
- d) mais de 30 salários mínimos

“Em que tipo de escola você cursou o ensino médio?” As opções são:

- a) todo em escola pública;
- b) todo em escola privada (particular);
- c) a maior parte em escola pública;
- d) a maior parte em escola privada (particular);
- e) metade em escola pública e metade em escola privada (particular)

“Como você se considera?” As opções são::

- a) branco(a);
- b) negro(a);
- c) pardo(a)/mulato(a);
- d) amarelo(a) (asiático);

Aos estudantes foram ainda feitas as perguntas: “Qual o grau de escolaridade de seu pai?” As opções são::

- a) nenhuma escolaridade;
- b) ensino fundamental: de 1^a a 4^a série;
- c) ensino fundamental de 5^a a 8^a série;

- d) ensino médio;
- e) ensino superior”.

"Qual seu gênero"?:

- a)Feminino
- b)Masculino

E, por último, afim de identificar os estudantes trabalhadores que contribuem com o sustento da família: “Assinale a situação abaixo que melhor descreve seu caso”.Com as seguintes opções:

- a) não trabalho e meus gastos são financiados pela família;
- b) trabalho e recebo ajuda da família;
- c) trabalho e me sustento;
- d) trabalho e contribuo com o sustento da família;
- e) trabalho e sou o principal responsável pelo sustento da família”.

APÊNDICE C DESCRIÇÃO DAS VARIÁVEIS CONTROLE

Tabela C.1: Descrição das variáveis controle

Nome	Descrição
Características da instituição	
tipo_ies	Categoria administrativa da IES 1-Federal 2-Estadual 3-Municipal 4-Privada
organacad	Organização acadêmica da IES 1-Centro federal de educação tecnológica 2-Centro Universitário 3-Faculdade 4-Universidade
região	Região do curso 1 = Norte 2 = Nordeste 3 = Sudeste 4 = Sul 5 = Centro-Oeste
in_noturno	Indicador de turno noturno 0 = Não 1 = Sim
Características do aluno	
idade	Idade do inscrito na prova
sexo	Sexo do inscrito 0 = Feminino 1 = Masculino
est_civil	Estado civil? 0= Solteiro(a) 1 = Casado(a) 2 = Separado(a)/Desquitado(a)/Divorciado(a) 3 = Viúvo(a) 4 = Outro
moradia	Mora atualmente?

	0 = Com os pais e(ou) com outros parentes
	1 = Com o(a) esposo(a) e(ou) com o(s) filho(s)
	2 = Com amigos (compartilhando despesas ou de favor)
	3 = Com colegas, em alojamento universitário
	4 = Sozinho(a)
caso_trabalho	Situação que melhor descreve seu caso.
	0 = Não trabalho e meus gastos são financiados pela família
	1 = Trabalho e recebo ajuda da família
	2 = Trabalho e me sustento
	3 = Trabalho e contribuo com o sustento da família
	4 = Trabalho e sou o principal responsável pelo sustento da família
educa_mae	Grau de escolaridade do pai/mae?
	0- Nenhuma escolaridade.
	1- Ensino fundamental: de 1.a a 4. a serie.
	2- Ensino fundamental: de 5. a 8. a serie.
	3- Ensino médio.
	4- Ensino superior.

Habilidade

livros_lidos	Livros lidos por ano
	0- Nenhum.
	1-No máximo dois.
	2-Entre três e cinco.
	3-Entre seis e oito.
	4- Mais de oito.
horas_estudo	Horas estudo por semana
	0- Nenhuma, apenas assisto as aulas.
	1- Uma a duas.
	2- Três a cinco.
	3- Seis a oito.
	4- Mais de oito.
dificuldadeFG	Dificuldade na Formação Geral?
	0= Muito fácil.
	1= Fácil.
	2 = Médio.
	3 = Difícil.
	4 = Muito difícil.
dificuldadeCE	Dificuldade Componente Especifico?
	0= Muito fácil.
	1= Fácil.

2 = Médio.
3 = Difícil.
4 = Muito difícil.

APÊNDICE D MODELO LOGÍSTICO BINOMIAL

Tabela D.1: Resultados exponenciados modelo logístico primeiro ciclo

Observations	322,557
Pseudo R2	0.1309
VARIABLES	(1)
1.renda	1.238*** (0.0134)
2.renda	1.884*** (0.0248)
3.renda	2.380*** (0.0565)
1.cor4	1.130*** (0.0109)
1.escola_privada2	1.397*** (0.0130)
1.educa_pai	1.115*** (0.0295)
2.educa_pai	1.238*** (0.0345)
3.educa_pai	1.482*** (0.0409)
4.educa_pai	1.916*** (0.0546)
Additional controls	Yes

Standard errors in parentheses: *** $p \leq 0,01$, ** $p \leq 0,05$, * $p \leq 0,1$

Fonte: Inep Enade, elaborado pelo autor

(1) Cursos altos salários(Acima de 3000R\$)

Tabela D.2: Resultados exponenciados modelo logístico segundo ciclo

Observations	594.74
Pseudo R2	0.1925
VARIABLES	(1)
1.renda	1.238*** (0.0134)
2.renda	1.884*** (0.0248)
3.renda	2.380*** (0.0565)
1.cor4	1.130*** (0.0109)
1.escola_privada2	1.397*** (0.0130)
1.educa_pai	1.115*** (0.0295)
2.educa_pai	1.238*** (0.0345)
3.educa_pai	1.482*** (0.0409)
4.educa_pai	1.916*** (0.0546)
Addition controls	Yes

Standard errors in parentheses: *** $p \leq 0,01$, ** $p \leq 0,05$, * $p \leq 0,1$

Fonte: Inep Enade, elaborado pelo autor

(1) Cursos altos salários(Acima de 3000R\$)

APÊNDICE E APÊNDICE E - REGRESSÃO PASSO A PASSO

E.0.1 Regressão passo-a-passo

A regressão passo-a-passo é um processo automatizado de ajuste de modelos de regressão, a partir do qual são escolhidas as variáveis explicativas que serão usadas no modelo final. Uma regressão passo-a-passo bidirecional, que será utilizada nesse caso, consiste da adição e retirada - constante e sequencialmente - de variáveis explicativas.

O uso dessa técnica de seleção de variáveis é muito criticada, e algumas das críticas são:

- Ela testa um espaço grande de modelos possíveis, o que aumenta a possibilidade de um sobreajuste dos dados.
- Não lida bem com a colinearidade dos dados.
- As estimativas de parâmetros e os coeficientes de determinação de seus modelos tendem a ser altos demais.

Todavia, acredita-se que a técnica, se usada tendo em mente seus pontos fracos, pode dar informações importantes sobre as possíveis variáveis explicativas, especialmente se usada numa análise exploratória.

E.0.1.1 Algoritmo

Como processo automatizado, a regressão passo-a-passo segue uma série de passos, listados a seguir.

Passo 1

O primeiro passo do algoritmo da regressão passo-a-passo consiste em selecionar os dois níveis de significância que serão usados. O primeiro (α_{inc}) define o limiar para que uma variável seja incluída no modelo. O segundo (α_{exc}) define o limiar para que uma variável seja excluída.

Passo 2

O segundo passo consiste em ajustar modelos com apenas uma variável explicativa, um para cada possível variável explicativa. Caso nenhum dos p-valores desses modelos seja menor que α_{inc} , para-se o processo. Caso pelo menos um p-valor seja menor, escolhe-se o menor, e a variável explicativa de seu modelo é incluída ao modelo final.

Passo 3

No terceiro passo, pega-se todos os modelos com duas variáveis explicativas em que uma das variáveis explicativas é a variável incluída no passo anterior. Caso nenhum dos p-valores desses modelos com duas variáveis seja menor que α_{inc} , para-se o processo. Caso pelo menos um p-valor seja menor, escolhe-se o menor, e a variável explicativa de seu modelo é incluída ao modelo final. Em seguida, testa-se se a inclusão da segunda variável ao modelo final tornou a primeira irrelevante, e se o p-valor para o teste de significância - no modelo com as duas variáveis - da primeira

variável for maior que α_{exc} ela será excluída do modelo final, e refaz-se o passo 3, levando em conta que a única variável incluída no modelo final seria a segunda.

Passo 4

Nesse passo, pega-se todos os modelos com três variáveis explicativas em que duas delas são as variáveis incluídas no modelo final. Caso nenhum dos p-valores desses modelos com três variáveis seja menor que α_{inc} , para-se o processo. Caso pelo menos um p-valor seja menor, escolhe-se o menor, e a variável explicativa de seu modelo é incluída ao modelo final. Em seguida, testa-se se a inclusão da terceira variável ao modelo final tornou as duas primeiras irrelevantes, e se pelo menos um dos p-valores para os teste de significância - no modelo com as três variáveis - das duas primeiras variáveis for maior que α_{exc} sua variável será excluída do modelo final, e refaz-se o passo 4 - ou volta-se para o passo 3-.

Passos subsequentes

Segue-se esse padrão de passos até o momento em que o processo tiver de ser parado porque nenhum p-valor foi menor que α_{inc} ou até o momento em que não há mais variáveis a serem testadas.

E.0.1.2 Dados

Para esse conjunto de dados, a regressão passo-a-passo foi executada tanto no R quanto no Stata. E as respostas de ambos os programas foram semelhantes: nenhuma das variáveis explicativas testadas (*idade*, *in_noturno*, *regiao*, *privada*, *tipo_ies*, *est_civil*, *sexo*, *renda*, *escola_privada*, *livros_lidos*, *moradia2*, *caso_trabalho*, *horas_estudo*, *organacad*, *dificuldadeFG*, *dificuldadeCE*, *educa_pai*, *educa_mae* e

cor) deveria ser excluída do modelo, mesmo usando-se uma correção de Bonferroni para os níveis de significância do processo de seleção (isso foi feito para tentar diminuir alguns dos problemas típicos da regressão passo-a-passo). Ou seja, de um ponto de vista de seleção de variáveis, todas elas parecem ser dignas de fazer parte do modelo final.

APÊNDICE F RESULTADOS EXPONENCIADOS MODELO MULTINOMIAL

Tabela F.1: Resultados em risco relativo modelo multinomial logit para o primeiro ciclo

Observations	333.963		
Pseudo R2	0.153		
VARIABLES	(1)	(2)	(3)
1.renda	1.301*** (0.0153)	1.361*** (0.0201)	1.433*** (0.0234)
2.renda	2.190*** (0.0388)	2.780*** (0.0554)	3.321*** (0.0696)
3.renda	2.596*** (0.117)	3.500*** (0.157)	5.595*** (0.248)
1.cor	1.185*** (0.0138)	1.201*** (0.0167)	1.216*** (0.0181)
1.escola_privada	1.310*** (0.0164)	1.589*** (0.0221)	1.747*** (0.0253)
1.educa_pai	1.213*** (0.0308)	1.216*** (0.0452)	1.129*** (0.0417)
2.educa_pai	1.395*** (0.0386)	1.438*** (0.0565)	1.335*** (0.0527)
3.educa_pai	1.504*** (0.0416)	1.803*** (0.0701)	1.680*** (0.0657)
4.educa_pai	1.514*** (0.0460)	2.245*** (0.0916)	2.292*** (0.0939)
Additional controls	Yes		

Standard errors in parentheses:*** $p \leq 0,01$,** $p \leq 0,05$,* $p \leq 0,1$

Fonte: Inep Enade, elaborado pelo autor

(0) Cursos baixo salários(até 2000R\$)

(1) Cursos médio-baixos salários(de 2000R\$ a 3000\$)

(2) Cursos médio-altos salários(de 3000R\$ a 4000R\$)

(3) Cursos altos salários(acima de 4000R\$)

Tabela F.2: Resultados em risco relativo modelo multinomial logit para o segundo ciclo

Observations	594.766		
Pseudo R2	0.212		
VARIABLES	(1)	(2)	(3)
1.renda	1.948*** (0.0161)	2.450*** (0.0294)	2.741*** (0.0272)
2.renda	3.879*** (0.0861)	6.149*** (0.154)	8.155*** (0.184)
3.renda	8.592*** (0.683)	10.66*** (0.874)	14.84*** (1.171)
1.cor	1.127*** (0.00892)	1.293*** (0.0144)	1.213*** (0.0112)
1.escola_privada	1.418*** (0.0162)	1.960*** (0.0271)	2.196*** (0.0265)
1.educa_pai	1.144*** (0.0357)	1.173*** (0.0720)	1.084* (0.0476)
2.educa_pai	1.432*** (0.104)	1.706*** (0.209)	1.228** (0.125)
3.educa_pai	1.549*** (0.112)	1.715*** (0.209)	1.483*** (0.142)
4.educa_pai	1.098 (0.129)	1.420* (0.284)	1.557*** (0.222)
Additional controls	Yes		

Standard errors in parentheses:*** $p \leq 0,01$,** $p \leq 0,05$,* $p \leq 0,1$

Fonte: Inep Enade, elaborado pelo autor

(0) Cursos baixo salários(até 2000R\$)

(1) Cursos médio-baixos salários(de 2000R\$ a 3000\$)

(2) Cursos médio-altos salários(de 3000R\$ a 4000R\$)

(3) Cursos altos salários(acima de 4000R\$)

APÊNDICE G RESULTADOS EXPONENCIADOS MO- DELO ODDS PROPORCINAIS

Tabela G.1: Resultados regressão odds proporcionais primeiro ciclo

Observations	333.963
Pseudo R2	0.089
Variables	(1)
1.renda	1.946*** (0.0113)
2.renda	3.078*** (0.0336)
3.renda	2.598*** (0.0611)
1.cor	1.142*** (0.00627)
1.escola_privada	1.505*** (0.0278)
1.educa_pai	1.162*** (0.0277)
2.educa_pai	1.431*** (0.0767)
3.educa_pai	1.588*** (0.0855)
4.educa_pai	1.370*** (0.128)
Additional controls:	Yes

Standard errors in parentheses:*** $p \leq 0,01$, ** $p \leq 0,05$, * $p \leq 0,1$

Fonte: Inep Enade, elaborado pelo autor

Tabela G.2: Resultados regressão odds proporcinais segundo ciclo

Observations	595.740
Pseudo R2	0.1458
Variables	(1)
1.renda	0.837*** (0.0268)
2.renda	2.295*** (0.0885)
3.renda	2.415*** (0.274)
1.cor	-0.228*** (0.00819)
1.escola_privada	0.568*** (0.0126)
1.educa_pai	0.156*** (0.0144)
2.educa_pai	0.387*** (0.0170)
3.educa_pai	0.524*** (0.0165)
4.educa_pai	0.687*** (0.0220)
Additional controls:	Yes

Standard errors in parentheses:*** $p \leq 0,01$,** $p \leq 0,05$,* $p \leq 0,1$

Fonte: Inep Enade, elaborado pelo autor

APÊNDICE H BRANT TEST

Tabela H.1: Brant test

Variables	chi2	p>chi2	df
All	96572.95	0.000	86
2.organacad	401.48	0.000	2
3.organacad	1317.51	0.000	2
4.organacad	648.59	0.000	2
2.tipo_ies	1516.44	0.000	2
3.tipo_ies	301.71	0.000	2
4.tipo_ies	2255.61	0.000	2
1.privada	348.56	0.000	2
1.educa_pai	3.77	0.152	2
2.educa_pai	34.00	0.000	2
3.educa_pai	43.55	0.000	2
4.educa_pai	50.48	0.000	2
1.educa_mae	12.68	0.002	2
2.educa_mae	24.27	0.000	2
3.educa_mae	11.62	0.003	2
4.educa_mae	25.11	0.000	2
1.sexo	3877.91	0.000	2
1.cor3	147.94	0.000	2
1.renda	938.70	0.000	2
2.renda	1283.83	0.000	2
3.renda	459.00	0.000	2
1.escola_privada	203.50	0.000	2
1.in_noturno	3100.12	0.000	2
idade	11955.98	0.000	2
1.caso_trabalho	1385.80	0.000	2
2.caso_trabalho	4020.10	0.000	2
3.caso_trabalho	3751.69	0.000	2
4.caso_trabalho	2291.49	0.000	2
1.livros_lidos	161.15	0.000	2
2.livros_lidos	1352.85	0.000	2
3.livros_lidos	1689.90	0.000	2
4.livros_lidos	2420.69	0.000	2

Tabela H.1: Brant test

Variables	chi2	p>chi2	df
1.horas_estudo	216.74	0.000	2
2.horas_estudo	431.37	0.000	2
3.horas_estudo	748.19	0.000	2
4.horas_estudo	886.91	0.000	2
1.dificuldadeFG	18739.32	0.000	2
2.dificuldadeFG	28330.15	0.000	2
3.dificuldadeFG	29389.11	0.000	2
4.dificuldadeFG	20760.67	0.000	2
1.dificuldadeCE	2124.46	0.000	2
2.dificuldadeCE	3540.71	0.000	2
3.dificuldadeCE	3724.18	0.000	2
4.dificuldadeCE	2621.25	0.000	2

APÊNDICE I RESULTADOS DO MODELO DE ODDS NÃO PROPORCIONAIS

Tabela I.1: Resultados exponenciados para o modelo de odds não proporcionais no primeiro ciclo

Observations	334.798		
Pseudo R2	0.144		
VARIABLES	(1)	(2)	(3)
1.renda	1.261*** (0.0124)	1.251*** (0.0127)	1.235*** (0.0168)
2.renda	2.265*** (0.0337)	1.910*** (0.0240)	1.745*** (0.0270)
3.renda	3.332*** (0.124)	2.364*** (0.0542)	2.367*** (0.0567)
1.cor4	1.248*** (0.0114)	1.149*** (0.0103)	1.156*** (0.0127)
1.escola_privada2	1.429*** (0.0142)	1.378*** (0.0122)	1.347*** (0.0144)
1.educa_pai	1.218*** (0.0263)	1.121*** (0.0280)	1.074** (0.0318)
2.educa_pai	1.434*** (0.0336)	1.238*** (0.0326)	1.186*** (0.0369)
3.educa_pai	1.635*** (0.0383)	1.462*** (0.0381)	1.366*** (0.0417)
4.educa_pai	2.071*** (0.0522)	1.847*** (0.0498)	1.654*** (0.0518)
Additional controls	Yes		

Standard errors in parentheses:*** $p \leq 0,01$, ** $p \leq 0,05$, * $p \leq 0,1$

Fonte: Inep Enade, elaborado pelo autor

(0) Cursos baixo salários(até 2000R\$)

(1) Cursos médio-baixos salários(de 2000R\$ a 3000\$)

(2) Cursos médio-altos salários(de 3000R\$ a 4000R\$)

(3) Cursos altos salários(acima de 4000R\$)

Tabela I.2: Resultados exponenciados para o modelo de odds não proporcionais no segundo ciclo

Observations	598,48		
Pseudo R2	0.202		
VARIABLES	(1)	(2)	(3)
1.renda	2.136*** (0.0725)	1.726*** (0.0464)	1.644*** (0.0361)
2.renda	4.503*** (0.0958)	2.746*** (0.0343)	2.493*** (0.0306)
3.renda	8.757*** (0.653)	2.294*** (0.0639)	2.155*** (0.0560)
1.cor4	1.183*** (0.0407)	1.122*** (0.0307)	1.072*** (0.0272)
1.escola_privada2	1.604*** (0.0522)	1.662*** (0.0496)	1.561*** (0.0383)
1.educa_pai	1.177*** (0.0389)	1.175*** (0.0341)	1.187 (0)
2.educa_pai	1.438 (0)	1.327*** (0.0207)	1.306*** (0.0168)
3.educa_pai	1.574*** (0.0527)	1.556*** (0.0324)	1.498*** (0.0143)
4.educa_pai	1.758*** (0.0402)	1.852 (0)	1.733*** (0.0344)
Additional controls	Yes		

Standard errors in parentheses:*** $p \leq 0,01$,** $p \leq 0,05$,* $p \leq 0,1$

Fonte: Inep Enade, elaborado pelo autor

(0) Cursos baixo salários(até 2000R\$)

(1) Cursos médio-baixos salários(de 2000R\$ a 3000\$)

(2) Cursos médio-altos salários(de 3000R\$ a 4000R\$)

(3) Cursos altos salários(acima de 4000R\$)