



Universidade de Brasília – UnB
Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas – FACE
Departamento de Economia – ECO
Programa de Pós-Graduação em Economia – PPGE
Doutorado em Economia

**RETORNOS ECONÔMICOS À EDUCAÇÃO SUPERIOR:
Cursos, qualidade da instituição, desempenho acadêmico
e implicações sobre a escolha de carreiras**

Anapaula Regina Lobo Cavalcanti

Brasília
2017

ANAPAULA REGINA LOBO CAVALCANTI

**RETORNOS ECONÔMICOS À EDUCAÇÃO SUPERIOR:
Cursos, qualidade da instituição, desempenho acadêmico
e implicações sobre a escolha de carreiras**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Economia da Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas da Universidade de Brasília como requisito parcial para obtenção do título de Doutor em Economia.

Orientadora: Profa. Dra. Maria Eduarda Tannuri-Pianto.

Brasília

2017

Universidade de Brasília – UnB
Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas – FACE
Departamento de Economia – ECO
Programa de Pós-Graduação em Economia
Doutorado em Economia

ANAPAULA REGINA LOBO CAVALCANTI

**RETORNOS ECONÔMICOS À EDUCAÇÃO SUPERIOR:
Cursos, qualidade da instituição, desempenho acadêmico
e implicações sobre a escolha de carreiras**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Economia da Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas da Universidade de Brasília como requisito parcial para obtenção do título de Doutor em Economia.

Orientadora: Profa. Dra. Maria Eduarda Tannuri-Pianto.

Banca Examinadora:

Profa. Dra. Maria Eduarda Tannuri-Pianto – Orientadora
PPGE/ECO/FACE/UnB

Profa. Dra. Ana Carolina Zoghbi – Membro Interno
PPGE/ECO/FACE/UnB

Prof. Dr. Donald Matthew Pianto – Membro Externo
Departamento de Estatística/UnB

Prof. Dr. Mário Rubem do Coutto Bastos – Membro Externo
Banco Central do Brasil/BCB

Prof. Dr. Rodrigo Leandro de Moura – Membro Externo
Ministério da Fazenda/STN

Prof. Dr. Rafael Terra de Menezes – Suplente
PPGE/ECO/FACE/UnB

Brasília, 16 de Maio de 2017.

Aos meus filhos: Pedro e Ana

AGRADECIMENTOS

Aos meus amigos: Fabio Araujo, José Ricardo da Costa e Silva e Wagner Piazza Gaglianone, que me incentivaram a ingressar no Programa de Pós-Graduação em Economia da Universidade de Brasília (PPGE/UnB).

A todos os meus mestres e colegas de curso, em especial, à professora orientadora Dra. Maria Eduarda Tannuri-Pianto, pela ajuda inestimável para a confecção do presente estudo, bem como por toda atenção, paciência e apoio durante todo o curso de Doutorado.

À professora Dra. Ana Carolina Zoghbi, pela ajuda com a base de dados e por todas as considerações feitas no exame de qualificação.

Ao professor Dr. Donald Matthew Pianto, pelas considerações feitas no exame de qualificação.

À minha família e aos amigos, em especial, minha mãe – Lenita Barreto de Lemos Lobo –, minha irmã – Nara Lenita de Lemos Lobo – e minha amiga – Rosa Cândida de Nascimento Miranda.

Em especial, agradeço ao meu esposo André Gustavo César Cavalcanti, por ter vivido e compartilhado todos os momentos de alegrias e angústias do curso de Doutorado.

“Desistir... eu já pensei seriamente nisso, mas nunca me levei realmente a sério”

Cora Coralina

RESUMO

O presente estudo, dividido em duas partes, tem por objetivo: a) separar os impactos de três diferentes fatores relacionados à formação universitária dos alunos sobre os salários dos recém-formados, a saber: 1) o curso universitário escolhido; 2) a qualidade da Instituição de Ensino Superior; e, 3) o desempenho escolar do recém-graduado, com base em um modelo hierárquico linear que distingue com maior grau de precisão os efeitos de fatores individuais e fatores institucionais sobre os ganhos financeiros dos recém-graduados. Os resultados obtidos neste aspecto apontam que os três tipos de fatores qualitativos supramencionados influenciam os ganhos iniciais e que os efeitos da qualidade da instituição e desempenho escolar não são uniformes para os graduados entre os diferentes cursos universitários. E ainda, a segunda parte do estudo tem por objetivo: b) analisar como os universitários escolhem suas carreiras, ou seja, questionar sobre os fatores relevantes para a decisão – o retorno esperado, a probabilidade de ser aprovado para um determinado curso e as preferências, as habilidades, a preparação e as características individuais e familiares dos candidatos que influenciam na sua escolha da carreira. A análise aqui tem por base os modelos multinomiais – adequados para as situações onde a variável de interesse discreta pode assumir diversos resultados possíveis, geralmente mutuamente excludentes, como é o caso na escolha de carreira. Nessa etapa valeu-se o modelo para a primeira parte do estudo, a fim de estimar as remunerações esperadas futuras dos candidatos de modo individual. Os resultados obtidos demonstram que o retorno esperado e a probabilidade de aprovação para o ingresso no curso são importantes na decisão de escolha da carreira dos estudantes, além de indicar que as variáveis de habilidade, preferências e preparação, bem como o nível de escolaridade dos pais, o gênero do candidato e se este optou pela seleção via cota racial foram aspectos significativos para a construção da probabilidade de ingresso. Assim, como as cotas raciais na Instituição de Ensino Superior analisada foram relevantes para estimar a probabilidade de passar e esta influenciam os estudantes nas escolhas de carreiras, é possível afirmar que as políticas afirmativas via cotas raciais, de fato, altera as escolhas dos candidatos a uma vaga na instituição em questão.

Palavras-chave: Escolha de carreiras; *Mixed logit*; Rendimentos; Retornos esperados; Probabilidade de admissão; Probabilidade de ingresso.

ABSTRACT

The present study, divided into two parts, has as its objective: a) to separate the impacts of three different factors related to the university formation of students on the salaries of the newly graduated ones, namely: 1) college major; 2) the quality of the Institution of Higher Education; And (3) the recent graduate school performance based on a linear hierarchical model that distinguishes more accurately the effects of individual factors and institutional factors on recent graduates' financial gains. The results obtained in this aspect indicate that the three types of qualitative factors mentioned above influence the initial gains and that the effects of institution quality and school performance are not uniform for the graduates between the different university courses. Moreover, the second part of the study aims to: b) analyze how university students choose their college major, that is, question the factors relevant to the decision – the expected return, the probabilities of being approved for a particular course and the preferences, the skills, preparation, and individual and family characteristics of candidates that influence their career choice. The analysis here is based on multinomial models – suitable for situations where the variable of discrete interest may assume several possible, usually mutually exclusive, outcomes, as is the case in career choice. In this stage, the model for the first part of the study was used, in order to estimate the expected future earnings of the candidates individually. The results show that the expected return and the probability of passing the course are important in deciding the career of the students, besides indicating that the variables of ability, preferences and preparation, as well as the level of schooling of the parents, the gender of the candidate and whether the candidate chose the racial quota selection were significant aspects for the construction of the probability of entry. Thus, as the racial quotas in the Institution of Higher Education analyzed were relevant to estimate the probability of passing and this influences the students in the career choices, it is possible to affirm that the affirmative policies via racial quotas, in fact, alter the choices of the candidates to place in the institution concerned.

Keywords: Choice of careers. Mixed logit. Income. Expected returns. Probability of admission. Probability of entry. College major.

LISTA DE TABELAS

Tabela 1 – Estatísticas descritivas das variáveis utilizadas	11
Tabela 2 – Percentual de recém-formados por grupos de cursos e por gênero.....	16
Tabela 3 – Remuneração dos recém-formados por grupos de cursos e por gênero.....	17
Tabela 4 – Coeficientes da regressão MQO para o log da remuneração dos recém-formados.....	18
Tabela 5 – Modelos HLM em três níveis.....	23
Tabela 6 – Modelo nulo em dois níveis.....	27
Tabela 7 – Modelos com variáveis explicativas em nível do aluno.....	28
Tabela 8 – Modelos com variáveis explicativas em nível da instituição de ensino superior.....	31
Tabela 9 – Estatísticas descritivas.....	45
Tabela 10 – Coeficientes do modelo <i>Logit</i> para a probabilidade de conclusão do curso.....	47
Tabela 11 – Probabilidade de conclusão do curso por grupos.....	48
Tabela 12 – Probabilidade de conclusão do curso por grupos (projetada).....	49
Tabela 13 – Coeficientes do modelo HLM para e_{ij} dos recém-formados.....	51
Tabela 14 – Modelo HLM para a remuneração futura $e_{ij} - u_{0j}$ para a UnB	52
Tabela 15 – Média das remunerações (e_{ij}) projetadas por grupos de cursos.....	53
Tabela 16 – Modelo MQO para a remuneração futura e_{i0}	55
Tabela 17 – Modelo - 3 níveis para a remuneração futura e_{i0}	56
Tabela 18 – Média das remunerações (e_{i0}) projetadas por grupos de cursos	57
Tabela 19 – Média das remunerações esperadas (Y) por grupos de cursos.....	58
Tabela 20 – Coeficientes do modelo <i>logit</i> para a probabilidade de passar.....	60
Tabela 21 – Probabilidade de passar no curso por grupos.....	61
Tabela 22 – Probabilidade de passar no curso por grupos (projetada).....	62
Tabela 23 – Resultado <i>Mixed Logit Model</i>	64
Tabela 24 – Efeitos marginais sobre a probabilidade de escolha.....	66
Tabela 25 – Resultado <i>Random Parameters Logit Model</i>	68
Tabela A1 – Correlação entre as variáveis em nível da instituição.....	82
Tabela A2 – Coeficientes da regressão MQO com as variáveis de nível de escolaridade dos pais.....	82
Tabela A3 – Modelos HLM em três níveis com as variáveis de nível de escolaridade dos pais	84
Tabela A4 – Modelo nulo em dois níveis com as variáveis de nível de escolaridade dos pais	86
Tabela A5 – Modelo 1 em dois níveis com as variáveis de nível de escolaridade dos pais	87
Tabela A6 – Modelo 2 em dois níveis com as variáveis de nível de escolaridade dos pais	88
Tabela A7 – Modelos com os efeitos randômicos.....	89
Tabela A8 – Modelos com os efeitos randômicos e interações – parte 1.....	90
Tabela A9 – Modelos com os efeitos randômicos e interações – parte 2.....	93
Tabela A10 – Resultado <i>Mixed Logit Model</i> com todas as variáveis explicativas.....	95

LISTA DE ABREVIATURAS E SIGLAS

ARUM	-	Additive Random Utility Model
CONAES	-	Comissão Nacional de Avaliação da Educação Superior
CPC	-	Conceito Preliminar de Curso.
ENADE	-	Exame Nacional de Desempenho de Estudantes
EUA	-	Estados Unidos da América
FGTS	-	Fundo de Garantia por Tempo de Serviço
HLM	-	Hierarchical Linear Model
ICC	-	Intraclass Correlation Coefficient
IES	-	Instituição de Ensino Superior
MQO	-	Mínimos Quadrados Ordinários
MTE	-	Ministério do Trabalho e Emprego
PASEP	-	Programa de Formação do Patrimônio do Servidor Público
PIS	-	Programa de Integração Social
PPGE	-	Programa de Pós-Graduação em Economia
PPI	-	Pretos, Pardos e Indígenas
RAIS	-	Relação Anual de Informações Sociais
RPL	-	Random Parameters Logit
SINAES	-	Sistema Nacional de Avaliação da Educação Superior
UnB	-	Universidade de Brasília

SUMÁRIO

INTRODUÇÃO	01
2 CURSO UNIVERSITÁRIO ESCOLHIDO, QUALIDADE DA INSTITUIÇÃO DE ENSINO SUPERIOR E DESEMPENHO ACADÊMICO DO ALUNO: QUAIS OS EFEITOS DESSES TRÊS FATORES SOBRE AS REMUNERAÇÕES DOS RECÉM-FORMADOS?	03
2.1 INTRODUÇÃO	03
2.2 METODOLOGIA	04
2.3 RESULTADOS	18
2.4 SUMÁRIO E CONCLUSÕES.....	33
3 COMO OS UNIVERSITÁRIOS ESCOLHEM SUAS CARREIRAS? QUAIS OS FATORES RELEVANTES: RETORNO ESPERADO, PROBABILIDADE DE PASSAR NO CURSO, HABILIDADES, INFLUÊNCIAS SOCIOECONOMICAS DOS CANDIDATOS?	35
3.1 INTRODUÇÃO.....	35
3.2 METODOLOGIA	37
3.3 RESULTADOS	45
3.4 SUMÁRIO E CONCLUSÕES.....	69
4 CONCLUSÃO	70
REFERÊNCIAS.....	72
ANEXOS	75
APÊNDICES.....	82

INTRODUÇÃO

O presente estudo, dividido em duas partes, tem por objetivo: a) separar os impactos de três diferentes fatores relacionados à formação universitária dos alunos sobre os salários dos recém-formados. Analisam-se os impactos do curso universitário graduado, da qualidade da Instituição de Ensino Superior (IES) e do desempenho escolar do recém-graduado em sua remuneração no mercado de trabalho. Aqui se fez uso de um Modelo Hierárquico Linear (em inglês, *Hierarchical Linear Model* – HLM), que distingue com maior grau de precisão os efeitos de fatores individuais e fatores institucionais sobre os ganhos financeiros dos recém-formados. Embora a literatura americana, em especial, o artigo de Rumberger (1993), trate detalhadamente à temática, a literatura brasileira parece não ter explorado exaustivamente tais influências. Conhecer como o mercado de trabalho valoriza a formação como um todo dos trabalhadores com o nível superior pode ajudar na escolha de quais tipos de políticas afirmativas devem ser implementadas para minimizar distorções no mercado de trabalho, na maioria das vezes, decorrentes de motivos raciais, étnicos, religiosos, de gênero e outros.

Diante do exposto, a pesquisa, na análise em questão, se propõe a responder basicamente quatro questionamentos, a saber:

- 1) Quais são os efeitos do curso, a qualidade da instituição e o desempenho sobre os ganhos dos recém-formados?
- 2) Como tais efeitos são diferentes entre homens e mulheres e por grupos raciais?
- 3) Quais são os efeitos de diferentes medidas de qualidade da universidade sobre os ganhos dos recém-formados com o mesmo curso?
- 4) Existem diferenças nos impactos das variáveis entre os cursos universitários?

Na segunda parte da pesquisa tem-se por objetivo: b) analisar como os jovens escolhem suas carreiras universitárias. Aqui se verificam quais os fatores relevantes para as escolhas dos cursos dos candidatos a um diploma universitário. As literaturas nacionais e internacionais apontam que o retorno esperado é relevante para a escolha, mas não há consenso quanto ao cálculo do retorno e a intensidade de seu efeito.

Alguns estudos definem o retorno esperado como o valor presente esperado de todos os ganhos futuros líquidos devido aos cursos frequentados pelo estudante (ALTONJI, 2012; ARCIDIACOMO, 2004; BERGER, 1988). Porém, estimar essa trajetória de retornos é algo complicado e, na prática, os artigos quantitativos optam por utilizar, em grande parte, o retorno esperado logo após o estudante encerrar o curso universitário (MONTMARQUETTE, 2002).

Os estudos nacionais utilizam a média da remuneração dada pelo mercado para cada curso universitário (CASARI, 2006). Alguns adicionam a remuneração, dada pelo mercado, o desvio-padrão e uma medida de necessidade do profissional no mercado (taxa de desemprego por área de atuação e percentual de participação no total de ocupados) para capturar um pouco da incerteza da

remuneração existente na escolha dos cursos pelos candidatos (COSTA DE SILVA; SILVEIRA NETO, 2015; BARTALOTTI; MENEZES, 2007).

Diante do exposto, a presente pesquisa utiliza o trabalho da primeira parte do curso para estimar a remuneração futura dos candidatos a uma vaga na Universidade de Brasília (UnB). Neste sentido, é preciso considerar, para estimar o retorno do curso almejado, a qualidade da instituição de ensino superior, seu desempenho acadêmico, além de todas as características individuais, socioeconômicas e de habilidades, ou seja, o retorno financeiro é individual de acordo com as escolhas e características dos candidatos.

Além de a remuneração ser individualizada, as linhas que se seguem captam a incerteza do alcance dessa remuneração no futuro estimando uma probabilidade de conclusão do curso, dado que existe a possibilidade do candidato não concluir o mesmo. Assim, de modo semelhante ao artigo de Montmarquette *et al.* (2002), o retorno esperado é uma média ponderada pela probabilidade de conclusão do curso via dois retornos financeiros individualizados, ou seja, caso o candidato conclua o curso e caso não alcance o diploma, ambos estimados e utilizando a metodologia e a base de dados da primeira parte do estudo. E ainda, além dos retornos esperados, tem-se que quase todas as políticas afirmativas recentemente implementadas no Brasil interferem diretamente na probabilidade de ingresso no curso pelos estudantes. Logo, a pesquisa analisa se a probabilidade de aprovação para o ingresso no curso é relevante para a escolha de carreiras dos candidatos, bem como quais os fatores que a influenciam.

Neste íterim, a segunda parte do curso tenta responder, em prol do universo dos candidatos a uma vaga na IES analisada as seguintes questões:

- 1) O Retorno esperado é relevante para a escolha de carreiras na UnB?
- 2) A escolha do curso pelos candidatos depende da probabilidade deste em ser aprovado no ingresso do curso em relação aos demais cursos?
- 3) Quais fatores influenciam a probabilidade de um candidato ser aprovado no ingresso em um dado curso?

A análise é baseada nos modelos multinomiais, que são adequados para situações onde a variável de interesse discreta – escolha do curso – pode assumir diversos resultados possíveis, em geral, mutuamente excludentes – como é o caso do problema proposto.

2 CURSO UNIVERSITÁRIO ESCOLHIDO, QUALIDADE DA INSTITUIÇÃO DE ENSINO SUPERIOR E DESEMPENHO ACADÊMICO DO ALUNO: QUAIS OS EFEITOS DESSES TRÊS FATORES SOBRE AS REMUNERAÇÕES DOS RECÉM-FORMADOS?

2.1 Introdução

De acordo com a literatura internacional, os graduados em cursos superiores ganham substancialmente mais que os graduados do Ensino Médio (MURPHY; WELCH, 1989). Algumas pesquisas apontam ainda que existem fatores qualitativos que afetam os retornos econômicos dos recém-formados (RUMBERGER, 1993). Assim, três fatores qualitativos foram identificados, quais sejam: 1) diferenças na escolha do curso universitário; 2) diferenças na qualidade da Instituição de Ensino Superior (IES) frequentada; e, 3) diferenças no desempenho do graduado na instituição.

Outros estudos apontam que grande parte da diferença dos ganhos dos graduados universitários é devida a sua escolha de curso universitário (SILVA; SILVEIRA NETO, 2015; BERGER, 1988a, 1988b; RUMBERGER, 1984). Em geral, os graduados que se formam em Engenharia, por exemplo, tanto no Brasil quanto nos Estados Unidos da América (EUA), tendem a ter salários mais elevados do que aqueles que se formam em outras áreas. Além disso, as diferenças salariais tendem a aumentar ao longo do tempo (BERGER, 1988a, 1988b).

A qualidade da instituição aparece influenciando os ganhos dos formados nos estudos de Solmon (1973, 1975), Trusheim e Crouse (1981), Smart (1988), Mueller (1988), James *et al.* (1989) e Rumberger (1993). Entretanto, os efeitos parecem explicar apenas uma pequena proporção da variância dos ganhos. Já Karabel e McClelland (1987) apontam que são mais fortes a partir de estudantes de classe social mais alta do que aqueles de classe mais baixa.

Em relação ao desempenho escolar, os resultados de alguns estudos demonstram que tal variável é relevante para a remuneração dos graduados, como, por exemplo, um estudo de James *et al.* (1989), onde se observa que o desempenho escolar interfere na remuneração mesmo quase 10 anos após os estudantes terem concluído o curso universitário.

Embora a literatura americana, em especial, o estudo de Rumberger (1993), demonstre com sucesso que as diferenças qualitativas nas experiências dos graduados na universidade têm influência importante sobre a remuneração dos formados, a literatura brasileira parece não ter explorado de modo exaustivo tais influências.

Neste íterim, o presente estudo analisa se os efeitos das diferenças qualitativas apontadas pela literatura americana também influenciam a remuneração laboral no Brasil. Este inclui dados de recém-formados de todas as regiões do Brasil. Teve-se o cuidado de analisar uma amostra que inclui diferença de gênero, raça e características sociais, a fim de estudar os retornos econômicos dos recém-formados, bem como fazer uso de um modelo robusto para estimar os ganhos dos recém-formados, que inclui várias dimensões da qualidade institucional, curso superior e de desempenho dos universitários. Por último, fez-se uso de uma técnica estatística – *Hierarchical Linear Model* (HLM) – para estimar adequadamente os efeitos, uma vez que se tem aqui um conjunto de unidades

agrupadas em diferentes níveis e de forma hierárquica. No caso, têm-se alunos agrupados por cursos em IESs e também agrupados por cursos universitários semelhantes.

De modo específico, o estudo tenta responder aos seguintes questionamentos:

- 1) Quais são os efeitos do curso, da qualidade da instituição e do desempenho sobre os ganhos dos recém-formados?
- 2) Como tais efeitos são diferentes entre homens e mulheres e por grupos raciais?
- 3) Quais são os efeitos de diferentes medidas de qualidade da universidade sobre os ganhos dos recém-formados com o mesmo curso?
- 4) Existe diferença nos impactos das variáveis entre os cursos universitários?

2.2 Metodologia

2.2.1 Modelos

A fim de analisar os impactos da qualidade da instituição de ensino, do curso escolhido e do desempenho dos recém-formados no curso, dois tipos de abordagem econométrica e procedimentos de estimação foram aqui empregados. O primeiro, muito utilizado na literatura internacional, envolveu estimar os ganhos salariais individuais usando regressão dos Mínimos Quadrados Ordinários (MQO), que é a abordagem mais comum para estimar equações de rendimentos.

O MQO é apropriado para estimar as contribuições únicas de uma grande variedade de variáveis de características a nível individual, do aluno, sobre os resultados individuais. No entanto, modelos em nível individual desse tipo pede por hipóteses que os indivíduos sejam independentes, o que pode não ser verdade quando se tem grupos de alunos agrupados por cursos e por IESs, como no caso em questão. Assim, pode-se pensar que alunos de um curso de uma dada instituição podem ser mais semelhantes aos olhos do mercado de trabalho que os demais de outra instituição. Logo, as observações individuais não são independentes, o que pode resultar em erros de estimativas utilizando técnicas MQO (RAUDENBUSH, 1988).

Outro problema em utilizar MQO é que tal abordagem não oferece um meio conveniente para testar modelos explícitos em dois ou mais níveis de interesse: No presente estudo têm-se três níveis de interesse, quais sejam: 1) os impactos do aluno com suas características e desempenho escolar na sua remuneração; 2) os impactos da qualidade da IES; e, 3) o curso em que se graduou na remuneração do recém-formado.

A segunda técnica aqui utilizada foi a HLM. A técnica foi desenvolvida especificamente para modelar fenômenos dentro e entre escolas. Esta também tem diversos sinônimos, sendo também conhecida como regressão multinível, modelo multinível, modelo de coeficiente randômico, modelo de componente de variância e modelo linear misto (*mixed model*). Para o estudo em questão, é a técnica apropriada; porém, existe uma limitação no modelo em questão: como ele possui vários níveis, o número de variáveis utilizadas nos modelos a nível individual é limitado pelo número de observações dentro de unidades institucionais que, em grande parte dos estudos disponíveis, é muito pequena. Neste sentido, no estudo em questão, como a base de dados é grande, alcançou-se o número de 721

IESs com o mínimo 30 alunos em cada instituição para todos os agrupamentos de cursos universitários. Assim, foram utilizadas técnicas de MQO para obter uma imagem mais robusta das influências individuais sobre os ganhos e, em seguida, técnicas HLM para obter uma imagem mais robusta de influências institucionais sobre os ganhos.

2.2.1.1 Modelos Mínimos Quadros Ordinários

Primeiramente, fez-se uso de uma função de ganho relacionando de forma linear à remuneração do recém-formado com características individuais, de antecedentes familiares, de característica da educação e do mercado de trabalho. O modelo básico pode ser escrito conforme o que se segue:

$$\ln(Y_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon_1 \quad (1)$$

onde:

Y_i = remuneração do recém-formado;

X_1 = vetor de variáveis de características individuais, tais como: gênero, raça, idade, estado civil e se possui filhos;

X_2 = vetor de variáveis da base familiar, tais como: renda da família (alta, média ou baixa), nível de escolaridades dos pais (se possui superior completo, médio completo ou menos que nível médio completo);

X_3 = vetor de variáveis de características da educação, tais como: grupos de cursos graduado (no presente estudo, se agrupam os cursos semelhantes em oito grupos, a saber: 1) Engenharias; 2) Ciências; 3) Saúde; 4) Medicina/Direito; 5) Ciências Sociais; 6) Negócios; 7) Educação; e, 8) Outros); desempenho escolar do aluno (nota na prova de conhecimento específico do Exame Nacional de Desempenho de Estudantes – ENADE¹); medidas da qualidade da instituição de ensino superior (se o curso graduado foi em uma escola particular, o percentual de alunos Pretos, Pardos ou Indígenas – PPI) no curso no qual se graduou, o percentual de pais de alunos com nível superior e médio no curso, média obtida pelos calouros do curso na prova do ENADE, conceito do curso e conceito preliminar de curso do curso no qual se graduou; e

X_4 = vetor de variáveis do mercado de trabalho, tais como: jornada de trabalho (se parcial, integral ou mais de 44h por semana) e tipo de emprego (se o emprego é no setor público e se o diploma é necessário para o emprego).

¹ Reconhece que a variável nota Enade possa carregar algum viés, uma vez que as provas de conhecimento específico para os cursos são diferentes.

2.2.1.2 Hierarchical Linear Model

Para a análise de regressão multinível tem-se o método proposto por Hox (2010), o qual implica em cinco passos de análise, quais sejam:

- Passo 1 – Análise do modelo sem nenhuma variável explicativa – modelo nulo ou modelo 0. Tal estimação serve de base de comparação para todos os demais modelos subsequentes, bem como para o Coeficiente de Correlação Intraclasse (*Intraclass Correlation Coefficient - ICC*), para a estimativa de *deviance*, uma medida de desajuste do modelo, e para a análise da necessidade da técnica.

O ICC é uma correlação que pode variar entre 0 e 1, e indica a necessidade do uso de uma análise em multinível. Quanto mais próximo de zero for o ICC, mais homogêneas são as instituições de Ensino Superior e, portanto, a remuneração dos recém-formados é independente da escola a qual ele se formou. Entretanto, quando o ICC for cerca de 10% ou mais, o método multinível deve ser considerado (LEE, 2008), indicando que existem diferenças significativas entre as instituições de nível superior.

O *deviance*, $-2 \cdot \log \text{likelihood}$, é uma estimativa de desajuste do modelo em relação aos dados que deve ser comparada com modelos posteriores. Assim, o valor do *deviance* do modelo nulo pode ser utilizado para comparação com os modelos posteriores e, caso a diferença seja significativa, o modelo com menor *deviance* é considerado melhor (KREFT; DELEEUW, 1998). Para critério de comparação entre os modelos também é possível utilizar as estatísticas Akaike's Information Criterion (AIC) e Bayesian Information Criterion (BIC), além do teste χ^2 . O AIC e o BIC são os mais populares para comparar modelos estimados por máxima verossimilhança, desde que os modelos sejam aninhados. Os critérios AIC e BIC são assim definidos:

$$\text{AIC} = -2 \cdot \ln(\text{likelihood}) + 2 \cdot k \quad (2)$$

$$\text{BIC} = -2 \cdot \ln(\text{likelihood}) + \ln(N) \cdot k \quad (3)$$

onde:

k = número de parâmetros estimados; e

N = número de observações.

O modelo nulo, também conhecido como modelo incondicional de apenas intercepto, pode ser assim exposto:

$$\begin{aligned} \ln(Y_{ij}) &= \beta_{0j} + \varepsilon_{ij} \\ \beta_{0j} &= \gamma_{00} + u_{0j} \end{aligned} \quad (4)$$

onde:

γ_{00} = média;

u_{0j} = termo aleatório que captura as diferenças entre as IESs

Nessa especificação está se dividindo o ganho dentro e entre as escolas. Podem-se reescrever as equações supramencionadas de modo equivalente em outro formato conhecido como *Mixed Model*.

$$\ln(Y_{ij}) = \gamma_{00} + u_{0j} + \varepsilon_{ij} \quad (5)$$

- Passo 2 – Inclusão no modelo anterior de variáveis explicativas fixas do nível mais baixo. No presente caso, o nível mais baixo é o aluno e as variáveis explicativas são, por exemplo, as de características dos alunos, conforme se segue:

$$\begin{aligned} \ln(Y_{ij}) &= \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + \beta_{4j}X_{4ij} + \varepsilon_{ij} \\ \beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} \\ \beta_{2j} &= \gamma_{20} \\ \beta_{3j} &= \gamma_{30} \\ \beta_{4j} &= \gamma_{40} \end{aligned} \quad (6)$$

No formato *Mixed Model* tem-se:

$$\ln(Y_{ij}) = \gamma_{00} + \gamma_{10}X_{1ij} + \gamma_{20}X_{2ij} + \gamma_{30}X_{3ij} + \gamma_{40}X_{4ij} + u_{0j} + \varepsilon_{ij} \quad (7)$$

Na especificação supramencionada tem-se um modelo, dentro das IESs, que tenta explicar diferenças de ganhos com base em variáveis individuais dentro de cada IESs.

- Passo 3 – Inclusão no modelo anterior das variáveis fixas do nível mais alto (Z_j) (aqui, em nível da instituição), conforme se segue:

$$\begin{aligned} \ln(Y_{ij}) &= \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + \beta_{4j}X_{4ij} + \varepsilon_{ij} \\ \beta_{0j} &= \gamma_{00} + \gamma_{01}Z_{1j} + \gamma_{02}Z_{2j} + \gamma_{03}Z_{3j} + \gamma_{04}Z_{4j} + \gamma_{05}Z_{5j} + \gamma_{06}Z_{6j} + u_{0j} \\ \beta_{1j} &= \gamma_{10} \\ \beta_{2j} &= \gamma_{20} \\ \beta_{3j} &= \gamma_{30} \\ \beta_{4j} &= \gamma_{40} \end{aligned} \quad (8)$$

No formato *Mixed Model* tem-se:

$$\begin{aligned} \ln(Y_{ij}) &= \gamma_{00} + \gamma_{01}Z_{1j} + \gamma_{02}Z_{2j} + \gamma_{03}Z_{3j} + \gamma_{04}Z_{4j} + \gamma_{05}Z_{5j} + \gamma_{06}Z_{6j} + \gamma_{10}X_{1ij} + \gamma_{20}X_{2ij} + \gamma_{30}X_{3ij} + \\ &\gamma_{40}X_{4ij} + u_{0j} + \varepsilon_{ij} \end{aligned} \quad (9)$$

Na especificação supramencionada tem-se um “modelo condicional”, utilizando variáveis de nível de escola para explicar as diferenças no intercepto da regressão que varia significativamente entre as IESs

- Passo 4 – Avalia-se a existência de coeficientes de inclinação das variáveis explicativas do primeiro nível que possuem componentes de variância significativos entre as IESs , conforme se segue:

$$\begin{aligned}
 \ln(Y_{ij}) &= \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + \beta_{4j}X_{4ij} + \varepsilon_{ij} \\
 \beta_{0j} &= \gamma_{00} + \gamma_{01}Z_{1j} + \gamma_{02}Z_{2j} + \gamma_{03}Z_{3j} + \gamma_{04}Z_{4j} + \gamma_{05}Z_{5j} + \gamma_{06}Z_{6j} + u_{0j} \\
 \beta_{1j} &= \gamma_{10} + u_{1j} \\
 \beta_{2j} &= \gamma_{20} + u_{2j} \\
 \beta_{3j} &= \gamma_{30} + u_{3j} \\
 \beta_{4j} &= \gamma_{40} + u_{4j}
 \end{aligned} \tag{10}$$

No formato *Mixed Model* tem-se:

$$\begin{aligned}
 \ln(Y_{ij}) &= \gamma_{00} + \gamma_{01}Z_{1j} + \gamma_{02}Z_{2j} + \gamma_{03}Z_{3j} + \gamma_{04}Z_{4j} + \gamma_{05}Z_{5j} + \gamma_{06}Z_{6j} + \gamma_{10}X_{1ij} + \gamma_{20}X_{2ij} + \gamma_{30}X_{3ij} + \\
 &\gamma_{40}X_{4ij} + u_{0j} + u_{1j}X_{1ij} + u_{2j}X_{2ij} + u_{3j}X_{3ij} + u_{4j}X_{4ij} + \varepsilon_{ij}
 \end{aligned} \tag{11}$$

Na especificação supramencionada qualquer variável explicativa em nível do aluno pode variar seu impacto entre as IESs.

- Passo 5 – Adicionam-se as interações entre níveis (efeito *cross-level*): relação entre as variáveis explicativas em nível da IESs e aquelas em nível do aluno, conforme se segue:

$$\begin{aligned}
 \ln(Y_{ij}) &= \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + \beta_{4j}X_{4ij} + \varepsilon_{ij} \\
 \beta_{0j} &= \gamma_{00} + \gamma_{01}Z_{1j} + \gamma_{02}Z_{2j} + \gamma_{03}Z_{3j} + \gamma_{04}Z_{4j} + \gamma_{05}Z_{5j} + \gamma_{06}Z_{6j} + u_{0j} \\
 \beta_{1j} &= \gamma_{10} + \gamma_{11}Z_{1j} + \gamma_{12}Z_{2j} + \gamma_{13}Z_{3j} + \gamma_{14}Z_{4j} + \gamma_{15}Z_{5j} + \gamma_{16}Z_{6j} + u_{1j} \\
 \beta_{2j} &= \gamma_{20} + \gamma_{21}Z_{1j} + \gamma_{22}Z_{2j} + \gamma_{23}Z_{3j} + \gamma_{24}Z_{4j} + \gamma_{25}Z_{5j} + \gamma_{26}Z_{6j} + u_{2j} \\
 \beta_{3j} &= \gamma_{30} + \gamma_{31}Z_{1j} + \gamma_{32}Z_{2j} + \gamma_{33}Z_{3j} + \gamma_{34}Z_{4j} + \gamma_{35}Z_{5j} + \gamma_{36}Z_{6j} + u_{3j} \\
 \beta_{4j} &= \gamma_{40} + \gamma_{41}Z_{1j} + \gamma_{42}Z_{2j} + \gamma_{43}Z_{3j} + \gamma_{44}Z_{4j} + \gamma_{45}Z_{5j} + \gamma_{46}Z_{6j} + u_{4j}
 \end{aligned} \tag{12}$$

No formato *Mixed Model* tem-se:

$$\begin{aligned}
 \ln(Y_{ij}) &= \gamma_{00} + \gamma_{01}Z_{1j} + \gamma_{02}Z_{2j} + \gamma_{03}Z_{3j} + \gamma_{04}Z_{4j} + \gamma_{05}Z_{5j} + \gamma_{06}Z_{6j} + \gamma_{10}X_{1ij} + \gamma_{11}Z_{1j}X_{1ij} + \gamma_{12}Z_{2j}X_{1ij} + \\
 &\gamma_{13}Z_{3j}X_{1ij} + \gamma_{14}Z_{4j}X_{1ij} + \gamma_{15}Z_{5j}X_{1ij} + \gamma_{16}Z_{6j}X_{1ij} + \gamma_{20}X_{2ij} + \gamma_{21}Z_{1j}X_{2ij} + \gamma_{22}Z_{2j}X_{2ij} + \gamma_{23}Z_{3j}X_{2ij} + \\
 &\gamma_{24}Z_{4j}X_{2ij} + \gamma_{25}Z_{5j}X_{2ij} + \gamma_{26}Z_{6j}X_{2ij} + \gamma_{30}X_{3ij} + \gamma_{31}Z_{1j}X_{3ij} + \gamma_{32}Z_{2j}X_{3ij} + \gamma_{33}Z_{3j}X_{3ij} + \gamma_{34}Z_{4j}X_{3ij} +
 \end{aligned}$$

$$\gamma_{35}Z_{5j}X_{3ij} + \gamma_{36}Z_{6j}X_{3ij} + \gamma_{40}X_{4ij} + \gamma_{41}Z_{1j}X_{4ij} + \gamma_{42}Z_{2j}X_{4ij} + \gamma_{43}Z_{3j}X_{4ij} + \gamma_{44}Z_{4j}X_{4ij} + \gamma_{45}Z_{5j}X_{4ij} + \gamma_{46}Z_{6j}X_{4ij} + u_{0j} + u_{1j}X_{1ij} + u_{2j}X_{2ij} + u_{3j}X_{3ij} + u_{4j}X_{4ij} + \varepsilon_{ij} \quad (13)$$

Nesse passo final na análise HLM tem-se um “modelo condicional”, utilizando variáveis em nível de IES para explicar as diferenças em todos os coeficientes da regressão em nível do aluno que variam significativamente entre as instituições.

2.2.2 Dados

O presente estudo fez uso de dois tipos de dados, a saber: 1) os dados do ENADE; e, os dados da Relação Anual de Informações Sociais (RAIS).

O ENADE foi instituído pela Lei n. 10.861, de 14 de abril de 2004, que definiu a obrigatoriedade do Sistema Nacional de Avaliação da Educação Superior (SINAES), que estabelece no § 1º do art. 1º, *in verbis*:

Art. 1º Fica instituído o Sistema Nacional de Avaliação da Educação Superior - SINAES, com o objetivo de assegurar processo nacional de avaliação das instituições de educação superior, dos cursos de graduação e do desempenho acadêmico de seus estudantes, nos termos do art. 9º, VI, VIII e IX da Lei n. 9.394, de 20 de dezembro de 1996.

§ 1º O SINAES tem por finalidades a melhoria da qualidade da educação superior, a orientação da expansão da sua oferta, o aumento permanente da sua eficácia institucional e efetividade acadêmica e social e, especialmente, a promoção do aprofundamento dos compromissos e responsabilidades sociais das instituições de educação superior, por meio da valorização de sua missão pública, da promoção dos valores democráticos, do respeito à diferença e à diversidade, da afirmação da autonomia e da identidade institucional.

[...].

O SINAES avalia a qualidade da oferta do Ensino Superior de graduação nas instituições de educação superior em nível de graduação e o desempenho acadêmico de seus estudantes, no início e no final dos cursos – o ENADE.

O ENADE avalia o rendimento dos concluintes dos cursos de graduação, em relação aos conteúdos programáticos, habilidades e competências adquiridas em sua formação. O exame é obrigatório e a situação de regularidade do estudante neste deve constar em seu histórico escolar.

A primeira aplicação do exame em questão ocorreu em 2004, e a periodicidade máxima da avaliação é trienal para cada área do conhecimento.

No presente estudo fizeram-se uso os dados dos exames realizados nos anos de 2004, 2005 e 2006 em todas as regiões do País. Em 2004, foram avaliados 143.170 estudantes (85.056 ingressantes e 58.114 concluintes), de 2.184 cursos de graduação das seguintes áreas do conhecimento, a saber: Agronomia, Educação Física, Enfermagem, Farmácia, Fisioterapia, Fonoaudiologia, Medicina, Medicina Veterinária, Nutrição, Odontologia, Serviço Social, Terapia Ocupacional e Zootecnia.

No ano de 2005, foram as seguintes áreas de conhecimento: Arquitetura e Urbanismo, Biologia, Ciências Sociais, Computação, Engenharia – Grupos I a VIII, Filosofia, Física, Geografia, História, Letras, Matemática, Pedagogia e Química.

No ano de 2006, os seguintes cursos: Administração, Arquivologia, Biblioteconomia, Biomedicina, Ciências Contábeis, Ciências Econômicas, Comunicação Social, Design, Direito, Formação de Professores, Música, Psicologia, Secretariado Executivo, Teatro e Turismo.

No período 2004-2006 foram avaliados ao todo 13.396 cursos em 48 diferentes áreas, através da aplicação de provas a 807.170 alunos, dos quais aproximadamente 60% foram ingressantes e 40% concluintes. Com a adoção de processos de amostragem, cerca de 50% dos alunos-alvo participaram a cada ano. No caso de cursos pequenos, com menos do que 20 candidatos potenciais, a participação dos alunos foi censitária. Assim, para um número pequeno de áreas, compostas apenas de cursos pequenos, o sistema amostral não foi aplicado.

A RAIS é um relatório de informações socioeconômicas solicitadas pelo Ministério do Trabalho e Emprego (MTE) às pessoas jurídicas e outros empregadores anualmente. Foi instituída pelo Decreto n. 76.900, de 23 de dezembro de 1975, e regulamentada, em seguida, pela Portaria MTE n. 651, de 28 de dezembro de 2007. Tem por objetivo gerar as informações necessárias ao controle da atividade laboral no País, o provimento de dados para a elaboração de estatísticas do trabalho e a disponibilização de informações do mercado de trabalho às entidades governamentais.

Os dados coletados pela RAIS² constituem expressivos insumos para atendimento das necessidades da legislação da nacionalização do trabalho, de controle dos registros do Fundo de Garantia por Tempo de Serviço (FGTS), dos Sistemas de Arrecadação e de Concessão e Benefícios Previdenciários, de estudos técnicos de natureza estatística e atuarial e de identificação do trabalhador com direito ao abono salarial do Programa de Integração Social/Programa de Formação do Patrimônio do Servidor Público (PIS/PASEP).

A RAIS, com o transcorrer do tempo, foi se tornando uma das fontes estatísticas mais confiáveis sobre o mercado de trabalho formal. Tal registro administrativo constitui referência nacional e internacional, sendo considerado um verdadeiro censo.

No estudo presente fez-se uso da RAIS de 2007 para abastecer o modelo com as variáveis do mercado de trabalho. Foram utilizadas informações sobre o salário, horas trabalhadas, tipo de vínculo empregatício e se o emprego exige o nível superior.

A partir das amostras originais, excluíram-se os indivíduos sem os dados de nível escolar, que não estavam trabalhando no ano de referência e, portanto, não tinham rendimentos ou foram matriculados novamente em outro curso superior nos anos seguintes. Nesta última condição, reclassificou-se o indivíduo para o atual curso superior que ele passou a frequentar, bem como para sua nova situação de conclusão do curso naquele ano. Após a conclusão deste processo de ajustamento da base de dados, o tamanho final da amostra foi de 201.363 recém-formados distribuídos em 721 instituições de ensino superior públicas e privadas localizadas em todas as regiões do Brasil.

² Reconhece que a base de dados da RAIS não contempla os trabalhadores que estão no mercado informal, desempregados ou profissionais liberais e isto pode gerar algum viés no estudo em questão.

2.2.3 Variáveis

Uma variedade de variáveis em nível do aluno e em nível da instituição de ensino superior foi utilizada no presente estudo. Seus nomes, médias, desvios padrão e uma breve descrição são evidenciados na Tabela 1, a seguir.

Tabela 1 – Estatísticas descritivas das variáveis utilizadas (continua).

Variáveis	Obs.	Média	DP	Min.	Max.	Descrição
Nível do Aluno (N = 201363)						
Características Individuais						
Mulher	201363	0,59	0,49	0	1	Variável <i>dummy</i> , mulher = 1 e homem = 0
PPI	149376	0,29	0,46	0	1	Variável <i>dummy</i> , PPI = 1 e brancos ou asiáticos = 0
Idade	201363	29,99	7,25	20	74	Idade do recém-formado em dezembro de 2007
Casado	149549	0,25	0,44	0	1	Variável <i>dummy</i> , casado = 1 e não casado = 0
Filhos	149456	0,28	0,45	0	1	Variável <i>dummy</i> , filhos = 1 e não filhos = 0
Família						
Renda Baixa	148686	0,23	0,42	0	1	Variável <i>dummy</i> , renda baixa = 1 e renda média ou alta = 0
Renda Média	148686	0,70	0,46	0	1	Variável <i>dummy</i> , renda média = 1 e renda baixa ou alta = 0
Renda Alta	148686	0,07	0,26	0	1	Variável <i>dummy</i> , renda alta = 1 e renda baixa ou média = 0
Pai_edu_NS	148840	0,19	0,39	0	1	Variável <i>dummy</i> , pai com nível superior = 1 e pai sem nível superior = 0
Pai_educ_NM	148840	0,25	0,43	0	1	Variável <i>dummy</i> , pai com nível médio = 1 e pai com nível superior ou menos de nível médio = 0
Mae_educ_NS	149309	0,19	0,39	0	1	Variável <i>dummy</i> , mãe com nível superior = 1 e mãe sem nível superior = 0
Mae_educ_NM	149309	0,28	0,45	0	1	Variável <i>dummy</i> , mãe com nível médio = 1 e mãe com nível superior ou menos de nível médio = 0

Tabela 1 – Estatísticas descritivas das variáveis utilizadas (continuação).

Variáveis	Obs.	Média	DP	Min.	Max.	Descrição
Nível do Aluno (N = 201363)						
Educação						
Nota no ENADE	171787	39,09	16,17	0	98,7	Nota na prova de conhecimento específico do ENADE
Engenharias	201363	0,07	0,26	0	1	Variável <i>dummy</i> , Engenharias = 1 e demais grupos = 0
Ciências	201363	0,05	0,21	0	1	Variável <i>dummy</i> , Ciências = 1 e demais grupos = 0
Saúde	201363	0,06	0,24	0	1	Variável <i>dummy</i> , Saúde = 1 e demais grupos = 0
Medicina/Direito	201363	0,05	0,22	0	1	Variável <i>dummy</i> , Medicina/Direito = 1 e demais grupos = 0
Ciências Sociais	201363	0,15	0,36	0	1	Variável <i>dummy</i> , Ciências Sociais = 1 e demais grupos = 0
Negócios	201363	0,28	0,45	0	1	Variável <i>dummy</i> , Negócios = 1 e demais grupos = 0
Educação	201363	0,22	0,42	0	1	Variável <i>dummy</i> , Educação = 1 e demais grupos = 0
Outros	201363	0,12	0,32	0	1	Variável <i>dummy</i> , Outros = 1 e demais grupos = 0
Mercado de Trabalho						
Trabalho Parcial	201363	0,22	0,42	0	1	Variável <i>dummy</i> , trabalho parcial = 1 e integral ou +44/sem = 0
Trabalho Integral	201363	0,72	0,45	0	1	Variável <i>dummy</i> , trabalho integral = 1 e parcial ou +44/sem = 0
Trabalho +44/sem	201363	0,06	0,24	0	1	Variável <i>dummy</i> , trabalho +44/sem = 1 e parcial ou integral = 0
Estatutário	201363	0,29	0,46	0	1	Variável <i>dummy</i> , emprego com vínculo estatutário = 1 e demais vínculos = 0
Diploma não Necessário	201363	0,48	0,50	0	1	Variável <i>dummy</i> , emprego onde o diploma não é necessário = 1 e necessário = 0

Tabela 1 – Estatísticas descritivas das variáveis utilizadas (conclusão).

Variáveis	Obs.	Média	DP	Min.	Max.	Descrição
Nível da Instituição (N = 721)						
Particular	201363	0,66	0,47	0	1	Variável <i>dummy</i> , Instituição particular = 1 e pública = 0
%_PPI_na_inst	201363	0,31	0,19	0	0,82	Percentual de membros que se autodeclararam PPI por grupo e instituição
%_pai_educ_NS	201363	0,21	0,15	0	0,86	Percentual de pais de alunos com nível superior por grupo e instituição
%_pai_educ_NM	201363	0,27	0,08	0	0,70	Percentual de pais de alunos com nível médio por grupo e instituição
%_mae_educ_NS	201363	0,21	0,14	0	0,83	Percentual de mães de alunos com nível superior por grupo e instituição
%_mae_educ_NM	201363	0,30	0,08	0,05	0,60	Percentual de mães de alunos com nível médio por grupo e instituição
Media_ingressante	200957	30,09	5,58	0	55,07	Média da nota da prova de conhecimento específico do ENADE dos calouros por grupo e instituição
Conceito	199969	3,11	0,66	1	5	Média dos conceitos dos cursos para cada grupo e instituição
CPC	201363	2,43	0,57	0,67	4,43	Média dos CPC dos cursos para cada grupo e instituição

Fonte: Da autora.

2.2.3.1 Variáveis em nível do aluno

As variáveis no nível do aluno consistem em quatro tipos, a saber: 1) características individuais; 2) antecedentes familiares; 3) educação; e, 4) mercado de trabalho.

As variáveis de características individuais consistem de variáveis para o gênero (mulher), para raça (PPI), para idade (idade), para estado civil (casado) e para se possuem um ou mais filhos (filhos), uma vez que cada um destes grupos geralmente tem experiências no mercado de trabalho diferentes das dos homens brancos jovens solteiros e sem filhos, que no estudo proposto é à base de referência. Para a variável mulher, a amostra apresenta 59,31% de recém-formadas e 40,69% de recém-formados. A variável PPI indica o valor 1 para os recém-formados que se consideram pretos, pardos ou de origem indígenas, e 0 caso branco ou asiático. Na amostra, 29,30% se consideram PPI, enquanto 70,70% se consideram brancos ou asiáticos. A média da idade dos recém-formados é de 30 anos na amostra, onde os mais novos têm 20 anos, e os mais velhos 74 anos. Para a variável de estado civil casado tem-se 25,45% de graduados casados, contra 74,55% de solteiros, divorciado, viúvo ou outros. E, finalmente, para a variável filhos tem-se 27,78% para os recém-formados que possuem um ou mais filhos, contra 72,22% que não possuem filhos.

As variáveis de antecedentes familiares consistem de variáveis que classificam os alunos quanto à renda familiar (baixa, média e alta) e quanto à escolaridade dos pais (com nível superior, nível médio e inferior ao nível médio). Uma família é considerada de baixa renda quando a renda familiar é inferior a três salários mínimos, de média renda quando a renda familiar é de três a 20 salários mínimos e de alta renda quando superior a 20 salários mínimos. Na amostra final, a distribuição ficou baixa renda com 23%, média renda com 70% e alta renda com 7%. Para as variáveis de escolaridade dos pais tem-se pai com nível superior – pai_educ_NS –, com 19%; pai com nível médio – pai_educ_NM –, com 25%; e, pai com menos de nível médio – pai_educ_inf_NM –, com 56%. Para as mães tem-se 19% para mães com nível superior – mae_educ_NS; com 28%, mães com nível médio – mae_educ_NM; e, com 53%, mães com menos de nível médio – mae_educ_inf_NM.

As variáveis de educação consistem em duas áreas de interesse, a saber: 1) desempenho na instituição de ensino superior; e, 2) curso escolhido. O desempenho na instituição foi medido pela nota no ENADE para a prova de conhecimento específico – variável nota ENADE. A nota varia de 0 a 100, e a média se encontra em torno de 39 pontos (vide Tabela 1). Os cursos escolhidos por cada aluno são agrupados nas seguintes áreas de concentração (grupos) (a ideia é seguir o padrão internacional sem se afastar muito dos artigos da literatura brasileira): Engenharias (grupo 1), Ciências (grupo 2), Saúde (grupo 3), Medicina/Direito (grupo 4), Ciências Sociais (grupo 5), Negócios (grupo 6), Educação (grupo 7) e Outros (grupo 8). O grupo de Engenharias é formado por todas as Engenharias (Civil, Elétrica, Mecânica, Química, de Materiais, de Produção, Ambiental, Florestal/pesca e Agronomia). O grupo Ciências é composto dos cursos de Matemática, Física, Química, Biologia, Computação, Biomedicina e retirado às licenciaturas desses cursos para serem incluídas no grupo de Educação. O grupo de Saúde é formado pelos cursos Medicina Veterinária, Odontologia, Farmácia, Enfermagem, Fisioterapia, Fonoaudiologia, Nutrição, Zootecnia, Terapia

Ocupacional. O grupo Medicina/Direito é formado apenas pelos cursos de Medicina e Direito. O grupo de Ciências Sociais compreende os cursos de Jornalismo, Ciências Sociais, Filosofia, História, Geografia, Serviço Social, Psicologia, e retiradas às licenciaturas desses cursos para serem adicionadas no grupo Educação. O grupo de Negócios é constituído pelos cursos de Administração, Ciências Contábeis e Economia. Já o grupo de Educação é formado pelos cursos Pedagogia, Educação Física, Magistério (ensino fundamental e infantil) e todas as licenciaturas dos demais cursos. Finalmente, os demais cursos estão no grupo Outros. Este é constituído basicamente de Arquivologia, Biblioteconomia, Arquitetura, Turismo, Secretariado, Design, Letras, Teatro, Música, e retiradas às licenciaturas desses cursos incluídas no grupo Educação.

Na amostra, os grupos foram assim distribuídos: Engenharias, com 7,16%; Ciências, com 4,63%; Saúde, com 6,21%; Medicina/Direito, com 5,25%; Ciências Sociais, com 15,08%; Negócios, com 27,57%; Educação, com 22,29%; e, Outros, com 11,82%.

As variáveis do mercado de trabalho incluem aquelas relacionadas ao número de horas trabalhadas durante uma semana e ao tipo de emprego. Têm-se três variáveis indicadoras associadas ao número de horas trabalhadas por semana, a saber: 1) trabalho parcial, caso o número de horas trabalhadas seja inferior às 40h semanais; 2) trabalho integral, caso as horas trabalhadas sejam de 40h às 44h semanais; e, 3) trabalho +44h/sem, se mais de 44h semanais são as horas trabalhadas pelo recém-formado. Para as variáveis relacionadas ao tipo de emprego tem-se a variável estatutário, que identifica se o recém-formado tem vínculo estatutário com o setor público, e a variável diploma não necessário, onde o emprego não exige o diploma recém alcançado com a formatura. Na amostra, as variáveis estão assim distribuídas: para as variáveis relacionadas ao tempo trabalhado tem-se trabalho parcial – 22,16%, trabalho integral – 71,63% e trabalho +44h/sem – 6,22%, e para as variáveis estatutário e diploma não necessário tem-se 29,32% e 47,66%, respectivamente.

2.2.3.2 Variáveis em nível da instituição de Ensino Superior

As variáveis de nível da instituição cobrem uma vasta gama de características das universidades que refletem a qualidade das instituições. Em primeiro lugar, foi incluída uma variável para indicar se a instituição é particular ou pública: variável particular. Em segundo lugar, tem-se uma série de variáveis indicando a composição social dos alunos, a saber: a variável %_PPI_na_inst representa o percentual de estudantes que se autodeclararam PPI para cada grupo de curso na instituição; e, a variável %_pai_educ_NS – o percentual de pais de alunos com o nível superior completo também por grupo de curso na instituição. Em terceiro lugar, são incluídas três medidas de avaliação da instituição, a saber: 1) média ingressante, que é a média dos calouros na prova do ENADE; 2) conceito, que é a média dos concluintes na prova do ENADE; e, 3) CPC, que é o Conceito Preliminar de Curso. Este se configura em uma média de oito quesitos que avaliam a instituição como um todo; é um indicador de qualidade dos cursos superiores. Ele é calculado com base na avaliação de desempenho de estudantes, corpo docente, infraestrutura, recursos didático-pedagógicos e

demais insumos, conforme orientação técnica aprovada pela Comissão Nacional de Avaliação da Educação Superior (CONAES). O CPC é constituído de oito componentes, agrupados em três dimensões que se destinam a avaliar a qualidade dos cursos de graduação, a saber: 1) desempenho dos estudantes; 2) corpo docente; e, 3) condições oferecidas para o desenvolvimento do processo formativo. A dimensão desempenho dos estudantes é constituída pelos seguintes componentes: Nota dos Concluintes no Enade (NCc) e Nota do Indicador da Diferença entre os Desempenhos Observado e Esperado (NIDDc). A dimensão corpo docente é constituída pelos seguintes componentes: Nota de Proporção de Mestres (NMc), Nota de Proporção de Doutores (NDc) e Nota de Regime de Trabalho (NRc). A dimensão percepção discente sobre as condições do processo formativo é constituída pelos componentes: Nota referente à Organização Didático-Pedagógica (NOc), Nota referente à Infraestrutura e Instalações Básicas (NFc) e Nota referente às Oportunidades de Ampliação da Formação Acadêmica e Profissional (NAc). As estatísticas descritivas para as medidas de avaliação da instituição também se encontram na Tabela 1, e as correlações para todas as medidas de nível da instituição são mostradas no apêndice tabela A1.

Quando se observa separadamente cada grupo, é possível perceber que existem diferenças entre graduados por grupos, tanto em sua escolha do curso como nos ganhos associados a essa escolha. Tais diferenças, discriminadas por gênero, são ilustrados nas Tabelas 2 e 3, a seguir.

Tabela 2 – Percentual de recém-formados por grupos de cursos e por gênero.

Grupos	Homens (%)	Mulheres (%)
Engenharias	13,7	2,7
Ciências	7,9	2,4
Saúde	5,0	7,1
Medicina/Direito	6,8	4,2
Ciências Sociais	14,3	15,6
Negócios	33,0	23,8
Educação	12,7	28,8
Outros	6,6	15,4
Total	100	100
Observações	81941	119422

Fonte: Adaptado de ENADE (2004; 2005; 2006) e RAIS(2007).

Nessa tabela apresenta-se os percentuais dos recém-formados por grupos de cursos e por gênero da base de dados final do trabalho

Tabela 3 – Remuneração dos recém-formados por grupos de cursos e por gênero.

Grupos	Homens (R\$)	Mulheres (R\$)
Engenharias	3.663,04	2.609,58
Ciências	2.714,69	1.941,32
Saúde	1.774,10	1.544,10
Medicina/Direito	3.075,25	2.196,57
Ciências Sociais	1.801,38	1.445,70
Negócios	2.277,66	1.571,85
Educação	1.879,21	1.316,12
Outros	1.687,15	1.339,83
Média	2.373,37	1.503,34

Fonte: Adaptado de ENADE (2004; 2005; 2006) e RAIS(2007).

Nessa tabela apresenta-se as médias das remunerações dos recém-formados por grupos de cursos e por gênero da base de dados final do trabalho

Os dados na Tabela 2 revelam que existem grandes diferenças na escolha do curso superior entre alunos do sexo masculino e feminino. Aproximadamente 30% das mulheres escolhem cursos superiores em áreas tradicionalmente femininas, como, por exemplo, os cursos do grupo Educação, enquanto que cerca de 13% dos homens escolhem cursos nestes domínios. Em contraste, os homens são muito mais propensos a se formar em cursos do grupo Negócios (33%). Porém, como a distribuição dos graduados entre os grupos não é uniforme, vale a pena ressaltar as maiores diferenças entre as escolhas de homens e mulheres nos grupos. Neste sentido, vale destacar que 14% dos homens escolhem o grupo Engenharias, enquanto apenas 3% das mulheres optam por esses cursos. E ainda, o grupo Ciência é a opção de cerca de 8% dos homens, enquanto apenas 2,5% das mulheres escolhem um dos cursos desse grupo. Por outro lado, as mulheres escolhem os grupos de Educação e Saúde em percentual maior que os homens.

Os salários iniciais dos recém-formados em universidades variam muito de acordo com o curso superior escolhido (vide Tabela 3). Analogamente, na literatura internacional, os graduados em Engenharia de qualquer instituição ganham os salários mais altos independente do sexo. Recém-graduados em Engenharia do sexo masculino ganham cerca de R\$ 1.300,00 (um mil e trezentos reais), ou 54% a mais que a média de todos os recém-graduados (em torno de R\$ 2 373,37 – dois mil e trezentos e setenta e três reais e trinta e sete centavos). O mesmo ocorre quando se verificam as diferenças entre as mulheres: as graduadas em Engenharia do sexo feminino ganham quase R\$ 1.100,00 (um mil e cem reais), ou 73% a mais que a média de graduadas do sexo feminino em todas as áreas (cerca de R\$ 1.503,34 – um mil e quinhentos e três reais e trinta e quatro centavos). Os segundos maiores salários tanto para os homens quanto para as mulheres estão no grupo Medicina/Direito, seguidos dos grupos de Ciência e Negócios – áreas tradicionalmente masculinas. No comparativo entre os gêneros, os homens ganham salários maiores em todos os grupos, mesmo no grupo de Educação, área predominantemente feminina, cerca de 40% a mais que as mulheres. A menor diferença salarial está no grupo de Saúde, onde homens ganham em torno de 15% a mais que

as mulheres; porém, é o grupo com menor salário para os homens, quando se exclui da análise o grupo Outros.

2.3 Resultados

2.3.1 Modelos Mínimos Quadros Ordinários

O modelo MQO é estimado para várias amostras de estudantes. Inicialmente estima-se na amostra geral, com variáveis binárias para diferenciar gênero e raça, primeiramente, excluindo variáveis do mercado de trabalho e, posteriormente, incluindo variáveis do mercado de trabalho. Este último modelo é, então, estimado separadamente para homens e mulheres e para brancos/asiáticos e PPI, a fim de capturar os efeitos distintos entre tais grupos. Os coeficientes de regressão para estes modelos são apresentados na Tabela 4, a seguir. Os resultados indicam que uma série de variáveis de características individuais, de família, da instituição de ensino, do curso e do mercado de trabalho tiveram efeitos significativos sobre a remuneração dos alunos recém-graduados.

Tabela 4 - Coeficientes da regressão MQO para o log da remuneração dos recém-formados (continua).

Modelos MQO (coluna 1)	Sem Mercado de Trabalho (coluna 2)	Com Mercado de Trabalho (coluna 3)	Homens (coluna 4)	Mulheres (coluna 5)	Brancos / Asiáticos (coluna 6)	PPI (coluna 7)
<i>Variáveis</i>	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
Mulher	-0,2098 *** (0,004)	-0,2102 *** (0,004)			-0,2073 *** (0,004)	-0,2126 *** (0,007)
PPI	-0,0229 *** (0,004)	-0,0242 *** (0,004)	-0,0331 *** (0,007)	-0,0224 *** (0,005)		
Idade	0,0194 *** (0,0005)	0,0170 *** (0,0004)	0,0191 *** (0,0008)	0,0157 *** (0,0005)	0,0180 *** (0,0006)	0,0150 *** (0,0007)
Casado	0,0864 *** (0,005)	0,0860 *** (0,005)	0,1873 *** (0,008)	0,0143 ** (0,005)	0,0961 *** (0,006)	0,0653 *** (0,008)
Filhos	0,0307 *** (0,005)	0,0214 *** (0,005)	0,0683 *** (0,009)	-0,0155 ** (0,006)	0,0040 (0,006)	0,0506 *** (0,008)
Renda Baixa	-0,2942 *** (0,004)	-0,2921 *** (0,004)	-0,3468 *** (0,008)	-0,2605 *** (0,005)	-0,2809 *** (0,005)	-0,3123 *** (0,007)
Renda Alta	0,1868 *** (0,007)	0,1799 *** (0,007)	0,2185 *** (0,010)	0,1422 *** (0,009)	0,1710 *** (0,007)	0,2155 *** (0,017)
Nota no Enade	0,0033 *** (0,0001)	0,0030 *** (0,0001)	0,0036 *** (0,0002)	0,0027 *** (0,0001)	0,0030 *** (0,0001)	0,0030 *** (0,0002)
Trabalho Parcial	-	-0,0718 *** (0,004)	-0,0996 *** (0,008)	-0,0482 *** (0,005)	-0,0731 *** (0,005)	-0,0694 *** (0,007)
Trabalho +44h/sem	-	0,4463 *** (0,007)	0,3631 *** (0,013)	0,4922 *** (0,008)	0,4466 *** (0,009)	0,4501 *** (0,012)
Diploma ã Necessario	-	-0,2081 *** (0,003)	-0,2074 *** (0,006)	-0,2067 *** (0,004)	-0,2271 *** (0,004)	-0,1625 *** (0,006)
Estatutario	-	0,1611 *** (0,004)	0,1387 *** (0,007)	0,1892 *** (0,005)	0,1425 *** (0,005)	0,1939 *** (0,007)

Tabela 4 - Coeficientes da regressão MQO para o log da remuneração dos recém-formados (conclusão).

Modelos MQO (coluna 1)	Sem Mercado de Trabalho (coluna 2)	Com Mercado de Trabalho (coluna 3)	Homens (coluna 4)	Mulheres (coluna 5)	Brancos / Asiáticos (coluna 6)	PPI (coluna 7)
Particular	-0,1051 *** (0,016)	-0,0840 *** (0,015)	-0,1278 *** (0,028)	-0,0330 (0,018)	-0,0620 ** (0,020)	-0,1535 *** (0,026)
Particular * idade	0,0056 *** (0,0005)	0,0058 *** (0,0005)	0,0074 *** (0,0009)	0,0040 *** (0,0006)	0,0056 *** (0,0006)	0,0069 *** (0,0008)
%_pai_educ_NS	0,4948 *** (0,016)	0,5218 *** (0,016)	0,4635 *** (0,025)	0,5820 *** (0,020)	0,5589 *** (0,018)	0,3632 *** (0,033)
%_pai_educ_NM	0,1602 *** (0,025)	0,3010 *** (0,024)	0,1431 *** (0,042)	0,3504 *** (0,030)	0,3647 *** (0,031)	0,2923 *** (0,043)
%_PPI_na_inst	-0,0838 *** (0,012)	-0,1321 *** (0,012)	-0,1228 *** (0,020)	-0,1148 *** (0,014)	-0,0904 *** (0,014)	-0,2655 *** (0,021)
Media_ingressantes	-0,0016 *** (0,0005)	-0,0003 (0,0004)	0,0019 * (0,0008)	-0,0006 (0,0005)	0,0007 (0,0005)	-0,0032 *** (0,0008)
Conceito	0,0181 *** (0,004)	0,0061 (0,004)	0,0250 *** (0,007)	-0,0117 * (0,005)	0,0097 * (0,005)	-0,0010 (0,007)
CPC	0,0141 *** (0,004)	0,0087 * (0,004)	-0,0144 * (0,007)	0,0262 *** (0,005)	-0,0040 (0,005)	0,0451 *** (0,008)
Engenharias	0,6413 *** (0,009)	0,6358 *** (0,009)	0,7366 *** (0,014)	0,4635 *** (0,015)	0,6557 *** (0,010)	0,5851 *** (0,018)
Ciências	0,3441 *** (0,012)	0,3590 *** (0,012)	0,4854 *** (0,019)	0,2570 *** (0,018)	0,3813 *** (0,014)	0,2966 *** (0,024)
Saúde	-0,0828 *** (0,009)	-0,0695 *** (0,009)	-0,0735 *** (0,018)	-0,0503 *** (0,010)	-0,0614 *** (0,010)	-0,0604 *** (0,019)
Medicina /Direito	0,1799 *** (0,009)	0,1975 *** (0,009)	0,2435 *** (0,015)	0,1564 *** (0,011)	0,1977 *** (0,010)	0,2253 *** (0,016)
Ciências Sociais	0,0039 (0,007)	0,0157 * (0,006)	0,0187 (0,013)	0,0238 *** (0,007)	0,0282 *** (0,008)	-0,0091 (0,011)
Negócios	0,1586 *** (0,006)	0,2367 *** (0,006)	0,2597 *** (0,012)	0,2292 *** (0,007)	0,2552 *** (0,007)	0,1984 *** (0,010)
Educação	0,0212 *** (0,006)	0,0107 (0,006)	0,1216 *** (0,014)	-0,0046 (0,007)	0,0203 ** (0,008)	-0,0076 (0,010)
Constante	7,2170 *** (0,002)	7,2166 *** (0,002)	7,3388 *** (0,003)	7,1180 *** (0,002)	7,2238 *** (0,002)	7,2056 *** (0,004)

OBS: *** significante a menos de 0,1%, ** significante a 1% e * significante a 5% e erros-padrão entre parênteses. Nessa tabela se apresentam os coeficientes das estimações dos modelos MQO para as diversas amostras do estudo. Na coluna 2 se apresentam os resultados das estimações para a amostra geral sem as variáveis de mercado de trabalho, na coluna 3 os coeficientes com as variáveis de mercado de trabalho e nas colunas 4 , 5 , 6 e 7 os resultados para as amostras separadas por gênero e raça

2.3.1.1 Variáveis de características individuais

Para a amostra geral (colunas 2 e 3), as estimativas revelam diferenças significativas nos salários iniciais para todas as variáveis. O maior efeito está na diferença de gênero, ou seja, as mulheres recebem salários em média 21% menor em comparação com os homens e quando se analisa a amostra separada por raça (colunas 6 e 7), é possível observar que as mulheres que se declaram PPI recebem ainda um pouco menos que as mulheres brancas e asiáticas. Os membros do grupo PPI recebem cerca de 2,5% menos que os brancos e asiáticos (colunas 2, 3 e 5), sendo esta

diferença maior (-3,3%) quando se compara aos homens PPI com os homens brancos e asiáticos (coluna 4).

Em relação ao estado civil e à presença de filhos na família, tem-se uma diferença interessante em relação ao gênero. Em geral, os casados recebem cerca de 8% a mais que os solteiros, e 2% a mais se possuem filhos. Entretanto, ao olhar o modelo separado por gênero (colunas 4 e 5), é possível observar que os homens casados recebem bem mais que seus pares solteiros – cerca de 18,7% –, enquanto que as mulheres casadas recebem praticamente o mesmo salário das mulheres solteiras. No que tange à presença de filhos na família tem-se uma inversão de sinal, ou seja, os homens com filhos recebem em torno de 7% a mais que seus pares sem filhos, e as mulheres com filhos recebem 1,6% a menos que as mulheres sem filhos.

2.3.1.2 Variáveis de base familiar

As variáveis de base familiar possuem uma correlação forte e, quando estão juntas no modelo, é possível perceber que as variáveis indicadoras de renda absorvem os efeitos das variáveis indicadoras de escolaridade do pai e da mãe. Assim, têm-se como resultado principal os modelos que possuem apenas as variáveis de renda familiar e, nos apêndices, os resultados dos mesmos – apenas alterando as variáveis explicativas de renda familiar pelas variáveis de escolaridade do pai e da mãe (Tabelas A2, A3, A4, A5 e A6).

Os modelos apontam que ser de uma classe social de baixa renda (até três salários mínimos) reduz drasticamente – cerca de quase 30% – o salário inicial dos recém-formados em relação a ser da classe média (de três a vinte salários mínimos) (diferença que aumenta para 35% se comparada à renda de homens da classe baixa com homens da classe média). Por outro lado, se a família do recém-formado possuir renda acima de 20 salários mínimos, este receberá, em média, 18% a mais que os membros da classe média – diferença que aumenta para 22% se comparada entre os homens de classes altas e médias e os membros que se autodeclararam PPI de renda alta com os membros do grupo PPI de renda média.

2.3.1.3 Variáveis de educação

Os resultados apontam que os tipos de medidas qualitativas – curso superior, qualidade da universidade e desempenho na universidade – afetam significativamente os salários iniciais da maioria dos recém-formados. Os efeitos latentes estão entres os cursos superiores, ou seja, têm-se diferenças substanciais nos salários iniciais para os graduados universitários com diferentes cursos superiores, mesmo após o controle de uma série de variáveis de características individuais, de família e de mercado de trabalho. Os cursos dos grupos Engenharias, Ciências e Medicina/Direito recebem salários iniciais bem maiores que os graduados no grupo omitido no modelo Outros (basicamente, Arquivologia, Biblioteconomia, Arquitetura, Turismo, Secretariado, Design, Letras, Teatro e Música). No grupo Engenharias, o ganho é de 64% a mais que no grupo Outros, chegando a ser 74% a mais quando comparados apenas os homens dos dois grupos. Para o grupo de Ciências, o ganho é de

36%, chegando a ser 49% quando comparados apenas os homens. E no grupo de Medicina/Direito, o ganho é de 20%, chegando a 24% na amostra apenas para os estudantes do sexo masculino. Por outro lado, os grupos de Ciências Sociais e Educação não se distinguiram em relação ao grupo Outros, e o de Saúde recebeu um salário ligeiramente inferior – cerca de 7% a menos.

O controle com as variáveis do mercado de trabalho afetou o efeito no grupo Negócios, não afetando, de modo relevante, os demais grupos, sugerindo que os grupos de cursos universitários tenham acesso a diferentes mercados de trabalho que influenciam os ganhos dos recém-formados. A introdução das variáveis de mercado alterou os ganhos no grupo Negócios de 16% para 24%, ou seja, muito provavelmente, alunos nessa área ingressaram no mercado em trabalhos que não exigem o diploma de nível superior.

As diferenças relativas entre os cursos superiores variaram para homens e mulheres e membros do grupo PPI e não PPI. Têm-se praticamente os mesmos sinais nos ganhos salariais em relação ao grupo Outros que os observados na amostra com todos os estudantes, porém, a amplitude do impacto é atenuada quando se analisa a amostra somente com estudantes do sexo feminino e membros do grupo PPI. Por exemplo, para o grupo de Engenharia, o ganho é de 64% na amostra com todos os estudantes, porém, na amostra somente com as alunas, é possível observar um ganho atenuado de 47%, e de 59% na amostra apenas com os membros do grupo PPI – o que indica que fatores além do capital humano podem afetar os rendimentos. Diante do exposto, tem-se apenas uma exceção observada no grupo Medicina/Direito para os membros do grupo PPI, os quais recebem um ganho salarial cerca de 23% a mais que os membros PPI do grupo Outros – na amostra geral, o ganho salarial é de 20%.

A medida de desempenho escolar no modelo é a variável nota no ENADE, que se refere à nota do aluno na prova de conhecimento específico no exame em questão. Para todas as amostras, o impacto foi significativo e semelhante – cerca de 0,3%. Como as notas variam de 0 a 100, para cada aumento de 10 pontos na nota do ENADE, o salário do recém-formado aumenta em cerca de 3%. Entretanto, tal ganho salarial parece ser irrelevante, em valores, quando comparado aos demais.

Neste íterim, grande parte das medidas de qualidade da instituição foi significativa. A primeira medida verifica se o fato da instituição de nível superior ser uma universidade particular³ interfere nas remunerações dos recém-formados. Tal fato apontou para a redução, em cerca de 8%, do ganho salarial dos recém-formados – redução acentuada quando se observa a amostra apenas com os recém-formados do sexo masculino, e com a amostra apenas de recém-formados PPI: queda de 13% e 15%, respectivamente. A segunda e terceira medidas de qualidade da instituição verifica se o fato dos pais dos alunos terem mais escolaridade (percentual de pais de alunos com nível superior completo e percentual de pais com apenas nível médio completo) interfere nas remunerações dos recém-formados. Aqui, a escolaridade dos pais acarreta em um impacto forte e positivo nos ganhos dos graduados. Entretanto, a quarta medida, que verifica se o fato do curso apresentar um percentual alto de alunos de PPI impacta as remunerações de todos os recém-formados, gera um impacto

³ Utilizou-se a interação das variáveis particular e idade, Particular * idade, para expurgar o efeito da variável idade na variável particular, uma vez que a média dos alunos de universidades particulares é maior que a média dos estudantes das universidades públicas.

negativo na remuneração. E neste sentido, é possível observar que o ganho reduz cerca de 13%, e é ainda mais acentuado (27%) na amostra onde se analisa apenas os recém-formados PPI – o que sugere que de, algum modo, estudar em um curso com uma maior concentração de alunos brancos e asiáticos ajuda na remuneração dos recém-formados PPI que estudaram no mesmo curso.

As três últimas medidas de qualidade da instituição utilizadas no estudo estão relacionadas com três índices de avaliação calculados pelo SINAES. Os índices média dos alunos ingressantes – conceito –, que é a média dos alunos concluintes, e CPC impactam a remuneração dos recém-formados. Entretanto, são nas amostras separadas por gênero que se observa um maior efeito. O conceito da instituição impacta fortemente a remuneração dos recém-formados do sexo masculino, enquanto que o índice CPC impacta a remuneração das recém-formadas. Tem-se aí uma forte correlação entre tais índices, sendo importante observar que o impactos dos três juntos é algo positivo para todas as amostras.

2.3.1.4 Variáveis do mercado de trabalho

No mercado de trabalho têm-se dois tipos de variáveis *dummies*: quanto à jornada de trabalho e quanto ao tipo de emprego. Todas as variáveis do mercado de trabalho foram significativas e impactam fortemente as remunerações dos recém-formados. O fato de o graduado trabalhar além das 44h semanais acarreta em um aumento de aproximadamente 45% na sua remuneração inicial, e quando este opta por trabalhar menos de 40h semanais, o impacto é negativo, reduzindo em cerca de 7% sua remuneração. Graduados empregados em trabalhos onde um diploma universitário não é necessário, ou exigido, ganham substancialmente menos que outros graduados universitários – aproximadamente 21%, independente do sexo, e 16% a menos se for PPI. Entretanto, se o recém-formado se empregar no setor público com vínculo estatutário tem-se um aumento, em média, de 16% na sua remuneração, que é ainda maior (cerca de 19%) quando se analisam as amostras com apenas mulheres e apenas membros do grupo PPI.

Comparando os resultados supramencionados com a literatura americana, é possível observar algumas semelhanças, apesar do ensino e do mercado de trabalho distintos estadunidenses. Sobre a questão, Rumberger (1993) evidencia que as mulheres recebem salários significativos menores que seus pares do gênero masculino e que existem grandes diferenças na remuneração dos cursos de nível superior, sendo os alunos dos cursos de Engenharias os mais bem sucedidos no mercado de trabalho. Aquele autor ainda revela que graduados empregados em trabalhos que não exigem o diploma universitário ou que o emprego não está relacionado com a formação alcançada possuem uma queda substancial nas suas remunerações em relação aos demais e, contrariamente ao resultado desse trabalho, os graduados norte-americanos sofrem uma penalidade se empregados no setor público.

Solnick (1990) assevera que os graduados universitários em engenharia negros em uma empresa receberam salários iniciais mais elevados que os demais graduados em outros cursos, bem como um dos resultados aqui apontados. Porém, Berger (1988), em um estudo apenas com graduados, de nível superior completo, não encontrou nenhuma diferença salarial entre brancos e

não brancos. Neste ínterim, a diferença salarial entre raças aparece na literatura americana quando se observam artigos mais antigos onde a amostra abrange todos os grupos de educação, em conjunto com os artigos do Gwartney (1978), Tienda (1987) e Meisenheimer(1990). Para a variável de desempenho escolar, o impacto no presente estudo parece ser inferior aos impactos encontrados na literatura internacional, porém, pode ser uma questão de escala, uma vez que a variável nota no ENADE varia de zero a 100 e, em Rumberger (1993), por exemplo, a variável GPA varia de 0 a 4.

Em geral, os resultados da regressão são consistentes com estudos anteriores internacionais, evidenciando que os aspectos de qualidade educacional – o curso universitário, a qualidade da instituição e o desempenho na instituição – afetam os salários dos recém-graduados universitários.

2.3.2 Modelos Hierárquicos

2.3.2.1 Modelo em três níveis

O primeiro passo para a realização da análise HLM é o modelo em três níveis, onde o nível mais alto é o curso superior, “grupos” (nesse trabalho os cursos superiores foram agrupados em grupos de cursos superiores semelhantes), seguido da instituição de ensino e finalmente, por último, os alunos (os recém-formados) como nível mais baixo do modelo. A tabela 5 mostra os resultados da análise em três níveis segundo Hox (2010).

Tabela 5 – Modelos HLM em três níveis (continua)

Modelos - 3 níveis (coluna 1)	Modelo 0 (coluna 2)	Modelo 1 (coluna 3)	Modelo 2 (coluna 4)
Numero de alunos	201363	143423	142398
Numero de instituições	1713	1713	1697
Numero de grupos	8	8	8
Intercepto	R\$ 1.472,75	R\$ 1.463,92	R\$ 989,91
Log likelihood	-203952,31	-128197,17	-127212,39
Deviance	407904,62	256394,34	254424,78
AIC	407912,60	256426,30	254470,80
BIC	407953,50	256584,30	254697,70
Efeito Aleatório			
Variância do Nível 1 - Aluno	0,4342 (0,001)	0,3393 (0,001)	0,3395 (0,001)
Variância do Nível 2 - Instituição	0,0575 (0,002)	0,0596 (0,002)	0,0512 (0,002)
Variância do Nível 3 - Grupo	0,0816 (0,041)	0,0811 (0,041)	0,0719 (0,036)
ICC -grupo	14,24% (0,061)	16,90% (0,071)	15,53% (0,066)
ICC instituição grupo	24,27% (0,054)	29,32% (0,060)	26,60% (0,058)

Tabela 5 – Modelos HLM em três níveis (conclusão)

Modelos - 3 níveis (coluna 1)	Modelo 0 (coluna 2)	Modelo 1 (coluna 3)	Modelo 2 (coluna 4)
Efeito Fixo - Variáveis do nível Aluno			
Mulher	-	-0,2022 *** (0,003)	-0,2010 *** (0,003)
PPI	-	0,0257 *** (0,004)	0,0261 *** (0,004)
Idade	-	0,0205 *** (0,0003)	0,0154 *** (0,0004)
Casado	-	0,0877 *** (0,004)	0,0859 *** (0,004)
Filhos	-	0,0281 *** (0,005)	0,0273 *** (0,005)
Renda Baixa	-	0,2275 *** (0,004)	0,2270 *** (0,004)
Renda Alta	-	0,1477 *** (0,006)	0,1476 *** (0,006)
Nota no Enade	-	0,0033 *** (0,0001)	0,0032 *** (0,0001)
Trabalho Parcial	-	0,0530 *** (0,004)	0,0527 *** (0,004)
Trabalho +44h/sem	-	0,4607 *** (0,007)	0,4614 *** (0,007)
Diploma ã Necessario	-	0,1974 *** (0,003)	0,1966 *** (0,003)
Estatutario	-	0,1812 *** (0,004)	0,1824 *** (0,004)
Efeito Fixo - Variáveis do nível Instituição			
Particular	-	-	-0,1278 *** (0,022)
Particular * idade	-	-	0,0073 *** (0,0005)
%_PPI_na_inst	-	-	0,0793 ** (0,037)
%_pai_educ_NS	-	-	0,5239 *** (0,051)
Media_ingressante	-	-	0,0023 (0,001)
Conceito	-	-	0,0296 * (0,013)
CPC	-	-	0,0060 (0,014)
Constante	7,2949 *** (0,101)	7,2889 *** (0,101)	6,8976 *** (0,107)
Variância explicada			
Nível Aluno	-	27,98%	27,89%
Nível Instituição	-	-3,52%	12,38%
Nível Grupo	-	0,60%	13,59%

OBS : *** significãnte a menos de 0,1%, ** significãnte a 1% e * significãnte a 5% e erros-padrão entre parênteses. Nessa tabela se apresentam os modelos HLM em três níveis, na coluna 2, o modelo nulo, sem as variáveis explicativas, na coluna 3, o modelo com as variáveis explicativas em nível do aluno e na coluna 4, o modelo com todas as variáveis explicativas (em nível do aluno e da instituição de ensino superior)

Na segunda coluna tem-se o modelo nulo, ou seja, sem nenhuma variável explicativa, na terceira coluna o modelo 1 com as variáveis explicativas no nível mais baixo do aluno e, por último, na quarta coluna o modelo 2, onde são acrescentadas as variáveis em nível da instituição. O modelo nulo, que é definido pelo intercepto e os erros em nível do aluno, instituição e grupo, é utilizado como base para a estimação da variância explicada *versus* variância não explicada em comparação com os modelos posteriores. Destaca-se o intercepto representando o salário médio de R\$1472,75, o que hoje seria cerca de R\$ 2574,81 ajustado apenas pela inflação do período (dez 2007 a dez 2016, cerca de 75% ap) e as variâncias de 0,4342 em nível do aluno, 0,0575 em nível da instituição e 0,0816 em nível do curso superior (Grupo).

A coluna 2 da tabela mostra que a maior parte da variância dos rendimentos dos recém-formados ocorre entre os alunos. Embora os parâmetros estimados de 0,0575 e 0,0816 para o nível instituição e o grupo, respectivamente, pareçam pequenos, eles representam desvios padrão de 0,24 e 0,29, respectivamente. Ao verificar o ICC – *interclass correlation coefficient* observa-se que 14,24% da variância da remuneração está associado ao grupo e quase 25% é devido a instituição e o grupo conjuntamente. Dessa forma, como o percentual é superior a 10%, deve se considerar o método multinível (Lee,2008) e estas são as diferenças que os modelos HLM tentam explicar.

Na coluna 3 da tabela 5, modelo 1, são inseridas variáveis explicativas fixas em nível do aluno. Todas as variáveis são significativas a 0,1% e a variância residual – nível do aluno reduziu de 0,4342 para 0,3393, ou seja, as variáveis explicativas em nível do aluno explicam 27,98% da variância em nível do aluno. Neste modelo são utilizadas as mesmas variáveis do modelo MQO, os impactos são na mesma direção, porém de intensidades um pouco diferente, vale ressaltar que a técnica HLM é mais apropriada, uma vez que as observações individuais não são independentes, o que é um problema para o modelo MQO.

No modelo 1 observa-se para as variáveis de características individuais uma queda de 20% se o recém graduado for do sexo feminino, queda de 2,5% se for membro do grupo PPI e aumento de 2%, 9% e 3% pela idade, se casado e se tem filhos, respectivamente. Para as variáveis de características familiares, observa-se uma queda de 23% se o graduado pertence a uma família de renda familiar inferior a 3 salários mínimos e um aumento de 14% se pertence a uma família de renda superior a 20 salários mínimos. Para a variável de desempenho escolar, observa-se um aumento de 0,3% para cada ponto a mais na prova do Enade e finalmente para as variáveis de mercado de trabalho observa-se, quanto as horas trabalhadas pelo recém-formado, uma queda de 5% se este trabalha menos de 40h semanais contra um aumento de 46% se o mesmo trabalha acima das 44h semanais e para o tipo de emprego observa-se uma queda de 20% se este não exige o grau de escolaridade obtido e um aumento de 18% se o emprego é no setor público (estatutário).

No modelo 2, são adicionadas ao modelo anterior variáveis explicativas em nível da instituição, tipo: se a instituição é particular, percentual de alunos membros do grupo PPI na instituição, percentual de pais de alunos com nível superior e com nível médio, média da nota do Enade para os ingressantes no curso, média dos concluintes e o Conceito Preliminar de Curso para a instituição em um determinado grupo. As variáveis particular, percentuais de PPI e de pais com nível superior em uma instituição em um determinado grupo e a nota dos concluintes no Enade para uma

dada instituição em um determinado grupo (*conceito_inst_grupo*) são significativas e explicam 12,38% e 13,59% da variância em nível da instituição e grupo respectivamente.

Assim como nos modelos anteriores, impactos são na mesma direção, porém de intensidades um pouco diferente, exceto para a variável percentual de PPI na instituição. Nesse caso, o modelo HLM de três níveis, que distingue com maior grau de precisão os efeitos de fatores individuais e fatores institucionais sobre os ganhos financeiros dos recém-graduados, contrariamente ao modelo MQO, apresenta um coeficiente positivo, indicando que quanto maior a concentração de alunos PPI em um curso maior o ganho salarial para esses alunos quando eles se formarem, cerca de 8%. Dessa forma, o que temos de fato é que uma maior concentração de alunos PPI em um dado curso aumenta a remuneração de todos os recém-formados desse curso. O modelo de três níveis explicou cerca de 12% das variâncias do nível da instituição e do grupo, porém a variância em nível do grupo é cerca de 50% maior que a em nível da instituição.

Na literatura norte americana um dos indicadores mais poderosos dos ganhos iniciais para graduados universitários é o curso universitário e, nesse estudo se observa também no modelo de três níveis a variância no nível do grupo (cursos universitários) maior que a em nível da instituição. Dessa forma existe ganho em analisar separadamente por grupo para os dados brasileiros assim como Rumberger(1993) analisou para os dados norte americano.

2.3.2.2 Modelo em dois níveis

Nesse tópico se analisam separadamente os oito grupos, utilizando a metodologia de Hox(2010) para cada um. Os grupos, formados por cursos semelhantes e já definidos anteriormente, são grupo 1 - Engenharias, grupo 2 - Ciências, grupo 3 - Saúde, grupo 4 - Medicina/Direito, grupo 5 - Ciências sociais, grupo 6 - Negócios, grupo 7 - Educação e por fim grupo 8 - Outros. Na tabela 6 tem-se a primeira parte utilizando a metodologia de Hox(2010) que é a estimação do modelo nulo para cada grupo, destaca-se nessa tabela a média salarial do grupo Engenharias que bem maior que os demais chegando a ser aproximadamente o dobro das médias salariais dos grupos Saúde, Ciências Sociais, Negócios, Educação e Outros.

É observado, igualmente ao modelo de três níveis, que a maior parte da variância nos rendimentos é devido à diferença entre os alunos dentro das escolas para todos os grupos. Entretanto ao verificar o ICC – *interclass correlation coefficient* observa-se que, exceto para os grupos Saúde e Outros, o valor é maior ou aproximadamente 10% o que justifica se considerar o método multinível (Lee,2008).

Tabela 6 – Modelo nulo em dois níveis

Modelo 0	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8								
(Coluna 1)	Engenharias	Ciências	Saúde	Medicina /Direito	Ciências Sociais	Negócios	Educação	Outros								
(coluna 1)	(coluna 2)	(coluna3)	(coluna 4)	(coluna 5)	(coluna 6)	(coluna7)	(coluna 8)	(coluna 9)								
Intercepto	R\$ 2.621,93	R\$ 1.876,78	R\$ 1.216,60	R\$ 1.784,50	R\$ 1.183,19	R\$ 1.360,77	R\$ 1.165,27	R\$ 1.108,78								
Variância do Nivel 1 - Aluno	0,3760	0,4426	0,4855	0,6786	0,4261	0,4546	0,3838	0,3895								
Variância do Nivel 2 - Instituição	0,0999	0,0646	0,0347	0,0826	0,0466	0,0641	0,0505	0,0361								
ICC	21,0%	12,7%	6,7%	10,8%	9,9%	12,4%	11,6%	8,5%								
Numero de alunos	14408	9323	12502	10568	30366	55517	44877	23802								
Numero de instituições	109	60	144	127	227	539	281	226								
Log likelihood	-13.582,23	-9.517,72	-13.356,60	-13.085,43	-30.421,01	-57.578,08	-42.575,44	-22.796,53								
Deviance	27.164,46	19.035,45	26.713,20	26.170,87	60.842,02	115.156,16	85.150,88	45.593,05								
AIC	27.170,46	19.041,45	26.719,09	26.176,87	60.848,02	115.162,20	85.156,88	45.599,05								
BIC	27.193,19	19.062,87	26.741,40	26.198,66	60.872,98	115.188,90	85.183,00	45.623,28								
	Var(β_x)	% explicado	Var(β_x)	% explicado	Var(β_x)	% explicado	Var(β_x)	% explicado	Var(β_x)	% explicado	Var(β_x)	% explicado	Var(β_x)	% explicado	Var(β_x)	% explicado
modelo 1	0,105	-5,18%	0,068	-5,19%	0,034	2,65%	0,084	-2,24%	0,050	-6,39%	0,067	-3,99%	0,052	-2,93%	0,038	-4,48%
modelo 2	0,069	31,26%	0,043	33,37%	0,032	7,57%	0,057	31,06%	0,040	14,40%	0,050	22,46%	0,044	12,06%	0,035	1,93%
modelo 3	0,069	30,90%	0,042	34,77%	0,033	5,46%	0,056	31,87%	0,040	13,99%	0,049	23,23%	0,044	12,97%	0,036	1,12%
modelo 4	0,068	31,55%	0,043	33,83%	0,032	6,76%	0,056	31,85%	0,040	14,52%	0,049	22,96%	0,044	12,97%	0,035	2,97%

OBS : Os valores são nominais no ano de 2007 e a tabela apresenta o modelo nulo para cada grupo

A tabela 7 mostra o modelo 1 (segunda parte da metodologia). Nesse modelo são inseridas variáveis explicativas fixas em nível do aluno, as mesmas variáveis do modelo MQO e HLM de três níveis. A grande maioria das variáveis é significativa a 0,1% e a variância residual – nível do aluno reduziu consideravelmente para todos os grupos. A introdução dessas variáveis conseguiu explicar cerca de 38% da variância para o grupo de Medicina/Direito, 28% para Negócios e Saúde, 23% para Educação, 22% para Ciências Sociais e Outros, 20% para Ciências e por fim 17% para Engenharias.

Tabela 7 – Modelos com variáveis explicativas em nível do aluno (continua)

Modelo 1 (coluna 1)	Engenharias (coluna 2)	Ciências (coluna 3)	Saúde (coluna 4)	Medicina /Direito (coluna 5)	Ciências Sociais (coluna 6)	Negócios (coluna 7)	Educação (coluna 8)	Outros (coluna 9)
Numero de alunos	9.150	3.986	9.497	8.765	21.170	42.299	30.020	18.501
Numero de instituições	109	60	144	127	227	539	281	226
Wald chi2(12)	1683,47	950,47	3493,90	5151,53	4969,69	14158,46	8016,07	4545,04
Log likelihood	-7878,92	-3661,74	-8644,46	-8843,55	18716,24	37191,67	-24675,88	15494,86
Deviance	15757,84	7323,48	17288,92	17687,11	37432,48	74383,34	49351,75	30989,72
AIC	15787,84	7353,48	17318,92	17717,11	37462,48	74413,34	49381,75	31019,72
BIC	15894,72	7447,84	17426,30	17823,29	37581,89	74543,13	49506,40	31137,10
Efeito Fixo - Variaveis do nivel Aluno								
Mulher	-0,2650 *** (0,015)	-0,2440 *** (0,021)	-0,0973 *** (0,013)	-0,1959 *** (0,015)	-0,1325 *** (0,008)	-0,1882 *** (0,006)	-0,2483 *** (0,008)	-0,1377 *** (0,010)
PPI	-0,0312 * (0,016)	-0,0600 ** (0,025)	0,0094 (0,018)	-0,0492 ** (0,017)	-0,0344 *** (0,010)	-0,0439 *** (0,007)	0,0062 (0,007)	-0,0266 ** (0,010)
Idade	0,0142 *** (0,0017)	0,0248 *** (0,0024)	0,0022 (0,0016)	0,0296 *** (0,0011)	0,0197 *** (0,0007)	0,0293 *** (0,0006)	0,0137 *** (0,0005)	0,0189 *** (0,0007)
Casado	0,1325 *** (0,022)	0,1781 *** (0,034)	-0,0460 * (0,024)	0,1224 *** (0,019)	0,0960 *** (0,012)	0,1496 *** (0,008)	0,0160 * (0,008)	0,0451 *** (0,011)
Filhos	0,0208 (0,023)	0,0368 (0,036)	-0,0033 (0,023)	0,0445 * (0,021)	0,0011 (0,013)	0,0596 *** (0,009)	0,0158 (0,009)	0,0079 (0,012)
Renda Baixa	-0,1867 *** (0,025)	-0,3125 *** (0,037)	0,0000 (0,018)	-0,3280 *** (0,022)	-0,2304 *** (0,010)	-0,2845 *** (0,008)	-0,1702 *** (0,007)	-0,2111 *** (0,010)
Renda Alta	0,1067 *** (0,016)	0,1805 *** (0,031)	0,0192 (0,018)	0,1667 *** (0,021)	0,1490 *** (0,016)	0,2225 *** (0,013)	0,1226 *** (0,022)	0,1239 *** (0,021)
Nota no Enade	0,0016 *** (0,0004)	0,0050 *** (0,0006)	0,0008 (0,0005)	0,0067 *** (0,0005)	0,0030 *** (0,0003)	0,0049 *** (0,0002)	0,0017 *** (0,0002)	0,0027 *** (0,0003)
Trabalho Parcial	-0,1730 *** (0,022)	-0,2443 *** (0,029)	-0,0723 *** (0,014)	-0,0772 *** (0,018)	-0,0739 *** (0,010)	0,0406 *** (0,009)	-0,0701 *** (0,007)	0,0003 (0,010)
Trabalho +44h/sem	0,2779 *** (0,038)	0,3436 *** (0,050)	0,6155 *** (0,023)	0,3430 *** (0,039)	0,4002 *** (0,017)	0,5088 *** (0,024)	0,4818 *** (0,011)	0,4778 *** (0,015)
Diploma ã Necessario	-0,3788 *** (0,015)	-0,1995 *** (0,020)	-0,5295 *** (0,013)	-0,3043 *** (0,015)	-0,1701 *** (0,008)	-0,1298 *** (0,006)	-0,1512 *** (0,007)	-0,1318 *** (0,009)
Estatutario	-0,2008 *** (0,023)	-0,0368 (0,032)	0,2747 *** (0,015)	0,4349 *** (0,016)	0,1807 *** (0,010)	0,0800 *** (0,011)	0,1798 *** (0,007)	0,1823 *** (0,010)
constante	7,8681 *** (0,032)	7,5239 *** (0,035)	7,0968 *** (0,017)	7,4824 *** (0,027)	7,0772 *** (0,016)	7,2081 *** (0,012)	7,0572 *** (0,014)	7,0058 *** (0,014)

Tabela 7 – Modelos com variáveis explicativas em nível do aluno (conclusão)

Modelo 1 (coluna 1)	Engenharias (coluna 2)	Ciências (coluna 3)	Saúde (coluna 4)	Medicina /Direito (coluna 5)	Ciências Sociais (coluna 6)	Negócios (coluna 7)	Educação (coluna 8)	Outros (coluna 9)
Efeito Aleatório								
var(constante)	0,1051 (0,015)	0,0680 (0,014)	0,0338 (0,005)	0,0844 (0,012)	0,0496 (0,005)	0,0667 (0,004)	0,0520 (0,005)	0,0377 (0,004)
var(Residual)	0,3134 (0,005)	0,3541 (0,008)	0,3513 (0,005)	0,4250 (0,006)	0,3340 (0,003)	0,3286 (0,002)	0,2955 (0,002)	0,3042 (0,003)
teste LR - chibar2(01)	1841,75	497,67	539,45	908,08	2381,52	6451,60	4407,03	1718,54
Variância explicada								
Nível aluno	16,65%	20,00%	27,64%	37,37%	21,61%	27,72%	23,01%	21,89%
Nível Instituição	-5,23%	-5,25%	2,72%	-2,21%	-6,47%	-3,96%	-2,96%	-4,54%

OBS : *** significante a menos de 0,1%, ** significante a 1% e * significante a 5% e erros-padrão entre parênteses. A tabela apresenta os modelos com as variáveis explicativas para o primeiro nível - aluno

Os impactos de cada variável são praticamente na mesma direção dos demais modelos, porém as intensidades variaram bastante tanto entre os modelos (modelos MQO e modelos HLM em três níveis) quanto entre os grupos nos modelos HLM de dois níveis. Utiliza-se mulher, PPI, idade, casado e filhos como variáveis de características individuais, observa-se que a variável indicadora do sexo feminino, mulher, impacta negativamente a remuneração dos recém-formados. O impacto mais acentuado se observa quando os recém-formados são do grupo de engenharias, uma queda de 27%, seguido de uma queda de aproximadamente 25% para os grupos de ciências e educação e as menores quedas nos grupos de saúde, ciências sociais e outros cerca de 10%, 13% e 14% respectivamente. Quanto aos grupos de medicina/direito e negócios a tabela apresentou o mesmo resultado dos modelos MQO e HLM de três níveis, uma queda de cerca de 20%.

Para a variável de raça PPI observa-se que não há diferença salarial para os grupos de saúde e educação, porém os demais apresentaram uma queda de 3% a 6%. Para a variável idade o impacto varia de aproximadamente 1,5% a 3%, sendo os grupos ciências, medicina/direito e negócios os que o mercado de trabalho mais valoriza a experiência alcançada com o passar do tempo. Para a variável casado o impacto é positivo para todos os grupos exceto para o grupo saúde, talvez porque o grupo esteja mais exposto a trabalhos noturnos. Para a variável Filhos o impacto é praticamente nulo para todos os grupos exceto para os grupos medicina/direito e negócios onde é aproximadamente 5%, talvez o impacto não seja nulo se a amostra for separada por gênero, como ocorreu no modelo MQO.

Para as variáveis de características familiares têm-se as indicadoras de renda familiar baixa e de renda familiar alta. Observa-se um forte impacto negativo se o recém-graduado pertencer a uma família de renda baixa, sendo mais acentuado se pertencer aos grupos medicina/direito e ciências, cerca de - 33% e -31% respectivamente e um forte impacto positivo se pertencer a uma família de renda familiar alta. O maior impacto é observado no grupo de negócios, cerca de 22%. Vale ressaltar

que não foi observado impacto algum no salário dos recém-formados se estes fizeram cursos no grupo saúde.

A variável de desempenho escolar, nota Enade, foi significativa, exceto para o grupo de saúde, entretanto, assim como no modelo MQO, seu impacto na remuneração dos recém-formados é muito pequeno de 1% a 6% na remuneração para cada 10 pontos na nota do Enade. Todavia maior que o apresentado por Rumberger(1993) que mostrou o impacto na remuneração dos recém-graduados norte americanos de aproximadamente 10% para os grupos negócios, educação e ciências, dado um aumento na nota do GPA que varia de 0 a 4. Porém essa diferença de impacto pode ser devido à escala, uma vez que a nota do Enade varia de 0 a 100.

Com relação ao mercado de trabalho têm-se dois grupos de variáveis: quanto ao número de horas trabalhadas por semana e quanto ao tipo de emprego. Para a variável *dummy* trabalho parcial ou trabalho inferior às 40h semanais o impacto é significativo e negativo para quase todos os grupos, sendo mais acentuado nos grupos ciências e engenharias, uma queda de 24% e 17% respectivamente. Essa queda é bem maior que a observada no modelo de três níveis e no MQO, cerca de 5% e 7%. Para os grupos saúde, medicina/direito, ciências sociais e educação a queda foi muito próxima dos 7% que aparece no modelo MQO, entretanto para o grupo de negócios o resultado surpreende em sendo positivo de 4% em relação ao trabalho de 44h semanais. Com relação ao trabalho acima de 44h semanais em todos os grupos o efeito foi positivo e significativo, sendo o grupo de saúde o que sofre um maior impacto na remuneração se o formado trabalhar além das 44h semanais.

Para as variáveis de tipo de emprego, têm-se duas: diploma não necessário, onde o recém-formado é empregado numa atividade onde o grau obtido não é necessário e Estatutário, onde o formado trabalha no setor público com vinculação estatutário. Para a variável diploma não necessário o resultado é significativo a menos de 0,1% e negativo para todos os grupos, sendo os grupos saúde, engenharias e medicina/direito os com maior redução salarial (-53%, -38% e -30%). Este resultado é bem acima dos -20% apontado pelos modelos HLM em três níveis e MQO, em especial para o grupo saúde. Para a variável Estatutário tem-se nenhum impacto para o grupo ciência e negativo para o grupo de engenharias (-20%). Os demais o impacto é positivo significativo com destaque para medicina/direito e saúde, 44% e 28%, respectivamente.

O modelo 2 , terceira parte da metodologia de Hox, é apresentado na tabela 8, neste modelo é inserido as variáveis explicativas fixas em nível da instituição: Particular, indicador para escola particular, %_PPI_na_inst, percentual de alunos membros do grupo PPI na instituição por grupo e %_pai_educ_NS, percentual dos pais dos alunos com nível de escolaridade maior ou igual ao nível superior e variáveis de avaliação do curso como média ingressante, conceito do grupo e CPC do grupo , que são as médias das notas do Enade dos alunos ingressantes para cada instituição e grupo , as médias das notas do Enade dos concluintes das instituições em cada grupo e o conceito preliminar de curso que leva em consideração as notas dos ingressantes, dos concluintes e avaliações com relação a infraestrutura e professores da instituição para cada instituição e grupo.

Tabela 8 - Modelos com variáveis explicativas em nível da Instituição de ensino superior (continua)

Modelo 2	Engenharias	Ciências	Saúde	Medicina /Direito	Ciências Sociais	Negócios	Educação	Outros
(coluna 1)	(coluna 2)	(coluna 3)	(coluna 4)	(coluna 5)	(coluna 6)	(coluna 7)	(coluna 8)	(coluna 9)
Numero de alunos	8655	3986	9497	8719	21123	42036	29918	18464
Numero de instituições	106	60	144	126	226	532	278	225
Wald chi2(19)	1772,6	979,47	3502,82	5166,42	5021,34	14145,21	8012,1	4559,2
Log likelihood	-7465,41	-3649,79	-8640,48	-8776,64	18637,81	-36896,88	-24580,22	-15458,74
Deviance	14930,83	7299,57	17280,96	17553,28	37275,62	73793,77	49160,43	30917,48
AIC	14974,83	7343,57	17324,96	17597,28	37319,62	73837,77	49204,43	30961,48
BIC	15130,28	7481,96	17482,45	17752,90	37494,70	74027,99	49387,17	31133,60
Efeito Fixo - Variáveis do nível Aluno								
Mulher	-0,2691 *** (0,016)	-0,2444 *** (0,021)	-0,0976 *** (0,013)	-0,1957 *** (0,015)	-0,1327 *** (0,008)	-0,1883 *** (0,006)	-0,2480 *** (0,008)	-0,1374 *** (0,010)
PPI	-0,0251 (0,016)	-0,0601 * (0,025)	0,0095 (0,018)	-0,0496 *** (0,018)	-0,0344 *** (0,010)	-0,0443 *** (0,007)	0,0060 (0,007)	-0,0267 ** (0,010)
Idade	-0,0043 (0,0026)	0,0246 *** (0,0042)	-0,0024 (0,0041)	0,0300 *** (0,0028)	0,0157 *** (0,0010)	0,0247 *** (0,0012)	0,0139 *** (0,0007)	0,0177 *** (0,0010)
Casado	0,1165 *** (0,023)	0,1774 *** (0,034)	-0,0467 * (0,024)	0,1211 *** (0,019)	0,0965 *** (0,012)	0,1496 *** (0,008)	0,0154 (0,008)	0,0433 *** (0,011)
Filhos	0,0131 (0,023)	0,0367 (0,036)	-0,0038 (0,023)	0,0438 * (0,021)	0,0017 (0,013)	0,0594 *** (0,009)	0,0169 (0,009)	0,0079 (0,012)
Renda Baixa	-0,1776 *** (0,026)	-0,3132 *** (0,037)	0,0000 (0,018)	-0,3304 *** (0,023)	-0,2308 *** (0,010)	-0,2852 *** (0,008)	-0,1697 *** (0,007)	-0,2111 *** (0,010)
Renda Alta	0,1020 *** (0,017)	0,1803 *** (0,031)	0,0191 (0,018)	0,1674 *** (0,021)	0,1496 *** (0,016)	0,2212 *** (0,013)	0,1226 *** (0,022)	0,1239 *** (0,021)
Nota no Enade	0,0015 *** (0,0005)	0,0050 *** (0,0006)	0,0008 (0,0005)	0,0067 *** (0,0005)	0,0029 *** (0,0003)	0,0048 *** (0,0002)	0,0018 *** (0,0002)	0,0027 *** (0,0003)
Trabalho Parcial	-0,1736 *** (0,023)	-0,2443 *** (0,029)	-0,0721 *** (0,014)	-0,0739 *** (0,018)	-0,0732 *** (0,010)	0,0416 *** (0,009)	-0,0699 *** (0,007)	0,0007 (0,010)
Trabalho +44h/sem	0,2744 *** (0,039)	0,3440 *** (0,050)	0,6156 *** (0,023)	0,3439 *** (0,039)	0,3994 *** (0,017)	0,5053 *** (0,025)	0,4816 *** (0,011)	0,4778 *** (0,015)
Diploma ã Necessario	-0,3800 *** (0,015)	-0,1997 *** (0,020)	-0,5294 *** (0,013)	-0,3047 *** (0,015)	-0,1702 *** (0,008)	-0,1295 *** (0,006)	-0,1512 *** (0,007)	-0,1317 *** (0,009)
Estatutario	-0,1826 *** (0,024)	-0,0376 (0,032)	0,2747 *** (0,015)	0,4356 *** (0,016)	0,1798 *** (0,010)	0,0812 *** (0,011)	0,1797 *** (0,007)	0,1826 *** (0,010)

Tabela 8 - Modelos com variáveis explicativas em nível da Instituição de ensino superior (conclusão)

Modelo 2	Engenharias	Ciências	Saúde	Medicina /Direito	Ciências Sociais	Negócios	Educação	Outros
(coluna 1)	(coluna 2)	(coluna 3)	(coluna 4)	(coluna 5)	(coluna 6)	(coluna 7)	(coluna 8)	(coluna 9)
Efeito Fixo - Variáveis do nível Instituição								
Particular	-0,4397 *** (0,119)	0,0777 (0,148)	-0,1586 (0,126)	0,0765 (0,114)	-0,1211 * (0,050)	0,0049 (0,049)	0,0285 (0,042)	-0,0062 (0,049)
Particular * idade	0,0280 *** (0,0030)	0,0004 (0,0046)	0,0052 (0,0042)	-0,0003 (0,0028)	0,0062 *** (0,0012)	0,0054 *** (0,0012)	-0,0003 (0,0008)	0,0020 (0,0012)
%_PPI_na_inst	-0,0043 (0,185)	0,3445 (0,244)	0,0465 (0,134)	0,8458 *** (0,164)	0,1382 (0,095)	-0,0095 (0,062)	0,0522 (0,081)	0,0418 (0,086)
%_pai_educ_NS	0,2298 (0,221)	0,9994 *** (0,284)	0,3599 * (0,175)	0,7488 *** (0,202)	0,5361 *** (0,096)	0,5895 *** (0,104)	0,8534 *** (0,183)	0,2191 (0,125)
Media_ingressantes	0,0153 (0,009)	-0,0146 (0,010)	0,0040 (0,005)	0,0156 * (0,007)	0,0070 (0,004)	0,0096 ** (0,004)	-0,0001 (0,003)	-0,0007 (0,003)
Conceito	0,1470 * (0,066)	0,1537 * (0,072)	0,0048 (0,046)	-0,0808 (0,049)	-0,0839 (0,041)	0,0742 ** (0,024)	-0,0349 (0,034)	-0,0185 (0,033)
CPC	-0,0975 (0,065)	-0,0059 (0,081)	-0,0601 (0,041)	-0,0415 (0,049)	0,0402 (0,034)	-0,0306 (0,025)	0,0731 * (0,034)	0,0579 (0,032)
constante	6,9766 *** (0,257)	6,9520 *** (0,288)	7,0107 *** (0,189)	6,8268 *** (0,197)	6,8020 *** (0,117)	6,5391 *** (0,097)	6,8600 *** (0,123)	6,8562 *** (0,127)
Efeito Aleatório								
var(constante)	0,0687 (0,010)	0,0431 (0,009)	0,0321 (0,005)	0,0569 (0,008)	0,0399 (0,004)	0,0497 (0,003)	0,0444 (0,004)	0,0354 (0,004)
var(Residual)	0,3176 (0,005)	0,3541 (0,008)	0,3512 (0,005)	0,4252 (0,006)	0,3336 (0,003)	0,3287 (0,002)	0,2957 (0,002)	0,3043 (0,003)
teste LR - chibar2(01)	1088,36	266	507,58	595,38	1609,38	4251,05	3018,15	1416,83
Variância explicada								
Nível aluno	15,52%	19,99%	27,66%	37,34%	21,72%	27,68%	22,96%	21,88%
Nível Instituição	31,26%	33,37%	7,57%	31,06%	14,40%	22,46%	12,06%	1,93%

OBS : *** significante a menos de 0,1%, ** significante a 1% e * significante a 5% e erros-padrão entre parênteses . A tabela apresenta os modelos com as variáveis explicativas tanto para o primeiro nível – aluno quanto para o ultimo nível – Instituição.

Pelo menos uma variável é significativa para cada grupo analisado e a introdução dessas variáveis explicou boa parte da variação no nível da instituição, como mostra a tabela 8. Para os grupos de engenharia, ciências, medicina/direito e negócios as novas variáveis explicam cerca de 31%, 33%, 31% e 22% respectivamente. Para a variável indicadora de a instituição ser particular, observa-se que é relevante para os grupos engenharias e ciências sociais, apresentando uma queda de 44% e 12% na remuneração dos recém-formados se a instituição for particular. Para a variável que mede o percentual de alunos membros do grupo PPI, observa-se que a variável é significativa apenas para o grupo medicina/direto e o impacto é positivo de 0,85 para cada aumento de 1 ponto percentual na variável. Para a variável que mede o percentual de pais com escolaridade acima do nível superior observa-se um impacto positivo para todos os grupos, entretanto para engenharia e outros a variável não foi significativa. Os maiores impactos foram nos grupos de ciências, educação e medicina/direito, sendo 1 , 0,85 , 0,75 para cada aumento de 1 ponto percentual na variável respectivamente. Para as variáveis que mede de alguma forma a qualidade do ensino da instituição tem-se um impacto positivo para os grupos onde foram significativas.

Para os passos 4 e 5 da metodologia de Hox(2010) se apresentam as tabelas A7, A8, A9 nos apêndices. O passo 4 deixa os coeficientes das variáveis explicativas em nível do Aluno variar entre as instituições de ensino superior e é utilizado para investigar se existem diferenças significativas no impacto de uma dada variável explicativa na remuneração dos recém-formado devido as instituições serem diferentes. O passo 5 tenta explicar com variáveis em nível da instituições o motivo da diferença dos coeficientes que variam entre as instituições no passo 4. Na tabela A2 dos apêndices se apresenta o resultado do passo 4, onde se permite que todas as 12 variáveis no nível do aluno possam variar seus impactos entre as instituições de ensino superior. Esses resultados de acordo com os critérios de *Deviance*, AIC e BIC não trouxeram ganhos em relação ao modelo anterior (modelo 2). Observa-se que a queda no *Deviance*, AIC e BIC diminui drasticamente ao longo dos passos indicados por Hox(2010) tabela 8 e também o número de parâmetros aumenta consideravelmente quando se caminha nessa direção o mesmo ocorre para o passo 5 apresentados nas tabelas A3 e A4 dos apêndices

Dado que o modelo HLM tem uma restrição quanto ao número máximo de parâmetros a ser estimado, que deve ser inferior ao número de alunos em uma dada instituição, o passo 5 fica praticamente inviável devido ao excesso de variáveis. Nesse passo se adiciona as interações entre níveis (efeito *cross-level*) e têm-se 12 variáveis no nível mais baixo (aluno) e 7 variáveis no nível superior (instituição) o que possibilitaria 54 opções de interações para cada grupo.

2.4 Sumário e conclusões

Este estudo investigou três fontes de influências qualitativas sobre os ganhos iniciais dos recém-formados universitários: o curso universitário, a qualidade da instituição de ensino superior e o desempenho do formado na instituição de ensino superior. Cada um destes fatores, como tem sido demonstrado em pesquisas anteriores influencia os salários dos recém-graduados. Aqui foram analisados os impactos dos três fatores para dados que abrangem todas as regiões do Brasil. Os dados utilizados neste estudo forneceram informações sobre os ganhos dos recém-formados no Brasil, a partir de uma amostra de dados da RAIS em 2007, SINAES/ENADE nos anos de 2004 a 2006 independentes do gênero, raça e demais características sociais e familiares e regionais. O estudo também empregou técnicas HLM, a fim de melhor testar a influência das medidas de nível institucional de qualidade da universidade.

Os resultados apresentados mostram que as três fontes são relevantes para influenciar as remunerações dos recém-formados também aqui no Brasil. Destaca-se que os cursos escolhidos pelos universitários é a fonte mais importante na remuneração dos recém-formados e que quando foram analisados separadamente os grupos observou-se que os impactos dos demais fatores (qualidade da instituição e desempenho escolar do formado) não se apresentaram uniformes entre os cursos.

Os formados do grupo de engenharias sofrem a maior queda salarial se o diploma do curso for de uma instituição de ensino superior particular, os de ciências possui o maior impacto positivo na

sua remuneração quanto melhor for o conceito da instituição de ensino superior cursada e o percentual de escolaridade dos pais com no mínimo nível superior, medicina/direito é o grupo que o mercado valoriza mais se o recém-formado vem de uma instituição concorrida para ingresso no curso, representado aqui pela média da nota do Enade dos alunos ingressantes no curso e o grupo de educação o que possui o maior impacto positivo na remuneração dos seus formados quanto ao conceito preliminar de curso que avalia a instituição como um todo. Já para grupo de saúde não foi relevante o desempenho escolar do formado e como também quase todas as medidas de qualidade da instituição de ensino superior exceto apenas o percentual dos alunos com escolaridade dos pais com no mínimo nível superior. Talvez a justificativa para esse resultado seja que o mercado nesse período estava bastante aquecido e o importante para emprego fosse o diploma universitário em um dos cursos do grupo independente da instituição de ensino.

Para finalizar os grupos de medicina/direito, seguido de negócios e ciências sociais são os grupos que o mercado mais valoriza o desempenho escolar. De maneira geral, pode-se dizer que o estudo mostra para o Brasil que a remuneração dos recém-formados cai cerca de 20% se o graduado for do sexo feminino, com destaque de maior queda se esta escolher o grupo de Engenharias, cerca de 27% em relação aos engenheiros. Entretanto é o grupo de maior remuneração tanto para os homens quanto para as mulheres. Pode-se também recém-formados pretos, pardos ou indígena sofrem uma redução de cerca de 3% na remuneração com destaque para os grupos de ciências e medicina/direito onde a queda é quase o dobro, cerca de 6% e 5%, respectivamente. Quanto à classe social do recém-formado pode-se dizer que se este pertence a uma família de baixa renda, então sofre uma queda salarial de aproximadamente 23%, onde também nos grupos de ciências e medicina/direito o impacto é mais acentuado, sendo um pouco mais de 30%.

Este estudo apenas analisou o impacto de fatores qualitativos sobre os salários iniciais. Desde que, pelo menos, algumas pesquisas sugerem que diferentes fatores influenciam os ganhos posteriores ou de crescimento dos salários, também é importante examinar os fatores que influenciam salários posteriores.

3 COMO OS UNIVERSITÁRIOS ESCOLHEM SUAS CARREIRAS? QUAIS OS FATORES RELEVANTES: RETORNO ESPERADO, PROBABILIDADE DE PASSAR NO CURSO, HABILIDADES, INFLUÊNCIAS SOCIOECONÔMICAS DOS CANDIDATOS?

3.1 Introdução

O objetivo desse estudo é analisar como é o processo de escolha de um curso universitário, quais os fatores relevantes: retorno esperado, probabilidade de passar no curso, habilidades, preferências, características socioeconômicas, gênero? O que realmente importa na decisão da escolha de carreiras dos estudantes?

Segundo a teoria microeconômica os indivíduos investem em sua qualificação profissional se o retorno desse investimento for maior que seu custo e, portanto, o retorno esperado deve ser importante na decisão dos estudantes. A literatura internacional e nacional possuem inúmeros artigos que indicam a importância do retorno esperado na escolha das carreiras. Entretanto eles diferem quanto ao cálculo desses retornos e a intensidade do seu efeito. Rumberger e Thomas(1993) mostram que há diferenças relevantes entre os retornos de cursos universitários e apontaram três fontes principais para justificar as diferenças salariais no mercado de trabalho para pessoas com nível superior: o desempenho do aluno na graduação, a qualidade da universidade e a área de atuação do curso universitário escolhido.

Berger(1988) baseou seu modelo de escolha de carreiras em expectativas racionais e modelou a utilidade do curso j como consumo do curso j mais o valor presente esperado de todos os ganhos futuros devidos ao curso na vida inteira do estudante. A escolha é resultado de um problema de maximização, onde se compara todas as opções de cursos e se escolhe o que maximiza a utilidade do estudante. Com essa abordagem Berger dá um passo à frente dos modelos que incorporam apenas os ganhos nos primeiros anos de formados. Entretanto, na prática, é extremamente difícil estimar todos os ganhos futuros associado ao curso. E também, o trabalho de Berger(1988) não menciona nada a respeito de incertezas, preferências e habilidades dos estudantes. Dessa forma, como justificar escolhas de cursos que são menos atrativos financeiramente em detrimento a outros mais atrativos?

Montmarquette et al(2002) insere incerteza ao modelo, o autor levanta a possibilidade de o estudante não concluir o curso. Desta forma, a escolha do curso continua sendo o resultado de um problema de maximização, porém é introduzida uma probabilidade de sucesso para concluir o curso universitário. Essa probabilidade de sucesso no curso está fortemente relacionada com as preferências, habilidades e características socioeconômicas dos estudantes. O retorno passa a ser definido como uma combinação do retorno dos ganhos futuros devido ao curso j e do retorno dos ganhos futuros, caso não tenha alcançado o nível superior ponderado pelas respectivas probabilidades de sucesso e fracasso no curso j .

Arcidiacomo(2004) estima um modelo dinâmico onde a cada período o estudante, de posse de mais informação sobre suas reais habilidades e preferências, escolhe entre continuar o curso ou entrar no mercado de trabalho. Arcidiacomo(2010) estima o modelo para a universidade de Duke

onde ele consegue as informações do contra factual através de uma pesquisa entre os alunos da universidade. Os alunos respondem um questionário bastante completo, onde, entre outras informações, revelam para todas as áreas de concentração quanto eles acreditam que ganhariam de rendimentos futuros se concluíssem um curso de certa área de concentração.

Saks & Shore (2005) inovam introduzindo aversão ao risco. Agora os estudantes não só estão preocupados com o retorno médio esperado para um determinado curso como também a sua variação. Carreiras consideradas mais “arriscadas” são as com maior variação no retorno esperado. Os autores encontram que estudantes mais ricos tendem a escolher carreiras mais arriscadas, em particular, na área de negócios, que os menos favorecidos. Observe que nesse modelo, carreiras mais arriscadas são as que possuem uma maior variação salarial no mercado de trabalho e não são, necessariamente, as consideradas mais difíceis e que a maioria dos estudantes atribuiria uma baixa probabilidade de conclusão do curso.

Altonji(2012) combina uma série de artigos para criar seu modelo dinâmico para escolha de carreiras, levando em consideração preferências, habilidades, incertezas, preparação, retorno do mercado de trabalho e outros relacionados a educação. O modelo teórico é bem geral e o próprio autor relata a dificuldade de estimá-lo, porém a ideia principal é colocar todas as informações listadas na literatura como relevantes sob o mesmo arcabouço.

Para o Brasil, Casari(2006) estudou a escolha de carreiras para alunos da Universidade de São Paulo - USP, o estudo utiliza como retorno esperado os rendimentos médios dos cursos . O estudo apontou vários resultados interessante: mostrou que o retorno esperado não tinha efeito na escolha da carreira dos universitários da USP e, observando algumas variáveis isoladamente, conclui que as áreas de negócios, engenharias e matemática/ciências são mais masculinas; olhando para as habilidades, as carreiras de ciências humanas, saúde e engenharias exigem notas mais altas e as áreas de educação e matemática/ciências notas mais baixa. Entretanto Bartalotti e Menezes(2007) com um modelo em que além do rendimento médio, considera a variabilidade do rendimento e a taxa de desemprego, concluem que o retorno médio influenciam as escolhas dos estudantes da USP. Costa de Silva e Silveira Neto(2015) apresentam um modelo semelhante com rendimento médio, desvio padrão dos rendimentos e participação no total de ocupados (em substituição a taxa de desemprego) para os candidatos a um curso na Universidade Federal de Pernambuco – UFPE, obtendo também resultados significativos para o retorno médio.

Além dos retornos esperados observa-se também que quase todas as políticas afirmativas implementadas no Brasil recentemente interferem diretamente na probabilidade de admissão no curso dos estudantes. As políticas afirmativas na educação com o intuito de eliminar desigualdades historicamente acumuladas, promover a igualdade de oportunidades e tratamento, bem como de compensar perdas provocadas pela discriminação e marginalização, decorrentes de motivos raciais, étnicos, religiosos, de gênero e outros em sua grande maioria separa os candidatos em grupos e reserva cotas para ingresso nas universidades. Dessa forma, é importante saber para o Brasil se esta probabilidade de passar no curso é relevante na escolha de carreiras e quais fatores a influência. Um estudo detalhado da escolha de carreiras, abordando gênero, raça e situação socioeconômica dos candidatos, pode ser importante para uma futura avaliação das políticas afirmativas no Brasil.

O estudo tem como objetivo estimar alguns modelos de escolha de carreiras para os estudantes da UnB, informações sobre gênero, raça, família, probabilidade de passar no curso, tipo de preparação dos candidatos são observados e assim como o retorno esperado, ou remuneração futura esperada caso o candidato alcance o diploma universitário. A metodologia do trabalho é semelhante a do artigo de Montmarquette et al(2002), entretanto adiciona-se a probabilidade de passar no curso como uma variável relevante no processo de escolha de carreiras.

Para este fim, se utiliza dados da base de dados da Cespe/UnB para todos os alunos que fizeram vestibular no período de 2004:1 a 2005:2. Nos dados, têm-se informações sobre os dados pessoais dos alunos, a sua escolha de curso, as respostas ao questionário sociocultural, as notas dos alunos no vestibular e se o aluno, anos após o ingresso, completou o curso ou não. A Universidade de Brasília foi a pioneira em implementar o sistemas de cotas no Brasil, bem antes da lei 12.711 de 29 de agosto de 2012, em julho de 2004 a UnB reservou 20% das vagas de seus cursos para alunos que se autodeclaravam “negros”. Desta forma, nessa base de dados têm-se informações sobre as cotas raciais da UnB. Também a fim de obter informações do mercado de trabalho para estimar o retorno esperado é utilizada a base de dados da Rais 2007 e o estudo do capítulo anterior.

3.2 Metodologia

3.2.1 Modelos

Os modelos multinomiais são adequados para situações onde a variável de interesse discreta pode assumir diversos resultados possíveis, geralmente mutuamente excludentes. No problema proposto se estuda como é feita à escolha de carreira dos estudantes, estes se deparam com uma grande variedade de cursos universitários e optam por um deles. É natural tentar modelar a escolha de carreiras por modelos multinomiais, já que os estudantes escolhem uma opção dentre várias possíveis.

Diferentes modelos multinomiais resultam de diferentes formas funcionais para as probabilidades da distribuição multinomial. Uma distinção entre os modelos também ocorre quando os regressores variam entre as alternativas para um determinado indivíduo ou quando regressores são constantes entre as alternativas.

O modelo multinomial mais simples é o modelo logit multinomial, proposto por Luce(1959). Este tem uma forma funcional específica para as probabilidades da distribuição multinomial e variantes desse modelo estão relacionadas com como os regressores variam entre as alternativas. Se estes variam são conhecidos como o modelo Logit Condicional e se não variam como Logit Multinomial. Ainda é possível ter uma combinação dos dois, onde uma parte dos regressores varia com as alternativas e outra parte não. Esse último modelo é conhecido como Mixed Logit Model.

No modelo Logit Condicional tem-se a probabilidade de um estudante i optar por um curso j dentre os J possíveis cursos descrita como:

$$p_{ij} = \frac{e^{x_{ij}\beta}}{\sum_{l=1}^J e^{x_{il}\beta}}, \text{ para todo } j= 1,2,\dots,J \quad (14)$$

Observe que os regressores assumem valores diferentes dependendo das J alternativas (x_{ij}) e o vetor de parâmetro β é único através das alternativas. Como p_{ij} é uma probabilidade, ou seja, $\sum_{j=1}^J p_{ij} = 1$ e, portanto, pode-se reescrever o modelo de forma que x_{ij} seja o desvio do valor do regressor para a alternativa j em relação a valor do regressor para a alternativa 1 e como consequência $x_{i1} = 0$. A interpretação dos resultados será sempre em relação aos regressores da alternativa 1.

No modelo Logit Multinomial tem-se a probabilidade de um estudante i optar por um curso j dentre os J possíveis cursos descrita como:

$$p_{ij} = \frac{e^{x_{ij}\beta_j}}{\sum_{l=1}^J e^{x_{il}\beta_l}}, \text{ para todo } j= 1,2,\dots,J \quad (15)$$

Observe que os regressores não variam dentre das J alternativas (x_{ij}) e o vetor de parâmetro β depende das J alternativas. Como p_{ij} é uma probabilidade, ou seja, $\sum_{j=1}^J p_{ij} = 1$, faz-se $\beta_1 = 0$ para que o modelo seja identificado e de forma análoga ao modelo Logit condicional a interpretação dos resultados será sempre em relação aos regressores da alternativa 1.

Porém na maioria dos casos têm-se variáveis explicativas que variam com as alternativas, como salário e variáveis de características individuais que não variam com as opções de escolha do curso, como gênero e raça. Para problemas com esse tipo de combinação utiliza-se o modelo *Mixed Logit*. Nesse modelo tem-se a probabilidade de um estudante i optar por um curso j dentre os J possíveis cursos descrita como:

$$p_{ij} = \frac{e^{x_{ij}\beta + w_i\gamma_j}}{\sum_{l=1}^J e^{x_{il}\beta + w_i\gamma_l}}, \text{ para todo } j= 1,2,\dots,J \quad (16)$$

Onde os regressores x_{ij} variam dentre das J alternativas e os regressores w_i não variam com as alternativas. Observe que pode-se reescrever os regressores inserindo *dummies* de forma a criar regressores que tragam a mesma informação, mas que também variem dentre as J alternativas. Dessa forma se reescreve os modelos Logit Multinomial e *Mixed Logit* como Logit Condicional. Seja d_{ijk} uma variável *dummy* que é igual a 1 se $j = k$ para cada indivíduo i e zero caso contrário, então $dw_{ijk} = d_{ijk}w_i$ é igual ao valor do regressor w_i se $j = k$ e zero caso contrário. Criam-se as variáveis *dummies* para as J alternativas, ou seja, k varia de 1 até J. Dessa forma podemos reescrever os modelos acima como logit condicional.

$$p_{ij} = \frac{e^{x_{ij}\beta + \sum_{k=1}^J dw_{ijk}\gamma_{jk}}}{\sum_{l=1}^J e^{x_{il}\beta + \sum_{k=1}^J dw_{ilk}\gamma_{jk}}}, \text{ para todo } j = 1,2,\dots,J \quad (17)$$

Estes modelos são casos particulares para uma dada forma funcional dos modelos conhecidos como *Additive Random Utility Model – ARUM*. Este é definido da seguinte forma: Os indivíduos i , no nosso caso os estudantes, escolhem o curso j entre as J alternativas possíveis de forma a maximizar sua função utilidade descrita abaixo.

$$U_j = V_j + \varepsilon_j \quad , \text{ para todo } j= 1,2,\dots, J \quad (18)$$

Onde V_j é o componente determinístico da utilidade e ε_j o componente aleatório. Os estudantes escolhem o curso $y = j$ se

$$\begin{aligned} \Pr[y = j] &= \Pr[U_j \geq U_k] \quad , \quad \forall k \neq j \\ \Pr[y = j] &= \Pr[\varepsilon_k - \varepsilon_j \leq V_j - V_k] \quad , \quad \forall k \neq j \\ \Pr[y = j] &= \Pr[\tilde{\varepsilon}_{kj} \leq -\tilde{V}_{kj}] \quad , \quad \forall k \neq j \end{aligned} \quad (19)$$

Dessa forma a escolha depende da probabilidade conjunta dos termos de erros

$$\begin{aligned} \Pr[y = j] &= \Pr[\tilde{\varepsilon}_{kj} \leq -\tilde{V}_{kj}] \quad , \quad \forall k \neq j \\ \Pr[y = j] &= \Pr[\tilde{\varepsilon}_{1j} \leq -\tilde{V}_{1j}, \dots, \tilde{\varepsilon}_{j-1j} \leq -\tilde{V}_{j-1j}, \tilde{\varepsilon}_{j+1j} \leq -\tilde{V}_{j+1j}, \dots, \tilde{\varepsilon}_{Jj} \leq -\tilde{V}_{Jj}] \\ &= \int_{-\infty}^{-\tilde{V}_{1j}} \dots \int_{-\infty}^{-\tilde{V}_{j-1j}} \int_{-\infty}^{-\tilde{V}_{j+1j}} \dots \int_{-\infty}^{-\tilde{V}_{Jj}} f(\tilde{\varepsilon}_{1j}, \dots, \tilde{\varepsilon}_{j-1j}, \tilde{\varepsilon}_{j+1j}, \dots, \tilde{\varepsilon}_{Jj}) d\tilde{\varepsilon}_{1j} \dots d\tilde{\varepsilon}_{j-1j} d\tilde{\varepsilon}_{j+1j} \dots d\tilde{\varepsilon}_{Jj} \end{aligned} \quad (20)$$

Em geral esse problema não possui solução fechada e é necessário utilizar métodos numéricos para estimar os parâmetros. Assumindo algumas hipóteses para o termo de erro, podemos encontrar uma solução fechada para a equação acima.

Seja ε_j iid conforme Weibull:

$$f(\varepsilon_{ij}) = \exp(-\varepsilon_{ij} - \exp(-\varepsilon_{ij})) \quad (21)$$

Como o erro, por hipótese, é independente a densidade conjunta torna-se o produto das densidades individuais e devido à forma funcional ser do tipo exponencial com os mesmos parâmetros esse produto nos fornecerá uma distribuição conjunta dos erros do tipo Gamma apenas reescrevendo a variável aleatória como função dos termos de erros. Como a variável de interesse é $\tilde{\varepsilon}_{kj} = \varepsilon_k - \varepsilon_j$ tem-se uma distribuição Logit e definindo $V_{ij} = x'_i \beta_j$ e $V_{ij} = x'_{ij} \beta$, tem-se Logit Multinomial e Condicional respectivamente. Dessa forma a probabilidade dos universitários escolherem o curso j é dada por:

$$p_{ij} = \frac{e^{V_j}}{\sum_{l=1}^J e^{V_l}} \quad (22)$$

O modelo a ser estimado vai depender do conjunto de variáveis explicativas X_i do problema proposto, se estas variam em relação as J alternativas deve-se utiliza Condicional Logit, se não variam em relação as J alternativas Multinomial Logit e ainda *Mixed Logit model* se parte dos regressores variam e outra parte não. Nesse estudo têm se ganhos futuros esperados, características socioeconômicas, variáveis que identificam as preferências, variáveis relacionadas às habilidades e preparação dos universitários compondo o conjunto de variáveis em X_i . Desta forma se utiliza o *Mixed Logit model*.

Uma variação dos modelos proposto acima é o modelo apresentado por Montmarquette et al.(2002), onde ele introduz no modelo a probabilidade de sucesso no curso j escolhido pelo estudante. O problema é descrito da seguinte maneira:

$$U_{ij} = \beta' y_{ij}^* + \alpha_j' w_i + \varepsilon_{ij} \quad \text{com} \quad y_{ij}^* = p_{ij}(x) e_{ij}(z) + (1 - p_{ij}(x)) e_{i0}(z) \quad (23)$$

U_{ij} é o nível da utilidade indireta esperada do indivíduo i no curso j, que não é observável. Porém observamos a escolha C_{ij} que assume os seguintes valores com relação à U_{ij} :

$$\begin{cases} C_{ij} = 1 & \text{se } U_{ij} \geq U_{ik}, \text{ para todo } k \neq j \\ C_{ij} = 0, & \text{caso contrário} \end{cases} \quad (24)$$

Onde,

y_{ij}^* é a remuneração futura esperada

$p_{ij}(x)$ é probabilidade de sucesso no curso escolhido j do indivíduo i

$e_{ij}(z)$ é o valor presente de todos os ganhos futuros esperado caso o indivíduo i conclua o curso j

$e_{i0}(z)$ é o valor presente de todos os ganhos futuros esperado caso o indivíduo i não conclua o curso j

w_i são as características individuais dos estudantes

ε_{ij} o termo aleatório

Assumindo que os resíduos (ε_{ij}) são independentes e identicamente distribuídos conforme Weibull, pode-se derivar um *Mixed Logit model* para a escolha discreta. P_{ij} é a probabilidade do indivíduo i escolher o curso j e ela pode ser escrita como:

$$P_{ij} = Prob(C_{ij} = 1) = \frac{e^{\beta y_{ij}^* + \alpha_j' w_i}}{\sum_{k=1}^J e^{\beta y_{ik}^* + \alpha_k' w_i}} \quad (25)$$

O problema pode ser resolvido em duas etapas onde primeiro se estima as probabilidades de sucesso nos cursos $p_{ij}(x)$ e os retornos dos ganhos futuros esperado $e_{ij}(z)$ e $e_{i0}(z)$ e depois os vetores de parâmetros α e β do *Mixed Logit model*.

Para esse estudo adiciona-se mais um termo, além da probabilidade de sucesso proposta por Montmarquette et al.(2002), a probabilidade de passar no vestibular $pp_{ij}(x)$ no curso desejado, que é uma das variáveis de interesse desse estudo. Como a escolha no Brasil, no período analisado, é feita antes das provas, esse componente de incerteza deve ter um peso grande na escolha do curso, uma vez que se o candidato não passar para o curso escolhido ele não pode optar por outro curso universitário nessa mesma instituição.

Então, nesse estudo, a probabilidade de escolha do indivíduo i pelo curso j é dada pela seguinte equação:

$$P_{ij} = Prob(C_{ij} = 1) = \frac{e^{\beta y_{ij}^* + \alpha_j w_i + \gamma pp_{ij}}}{\sum_{k=1}^J e^{\beta y_{ik}^* + \alpha_k w_i + \gamma pp_{ik}}}$$

Com
$$y_{ij}^* = pc_{ij}(x)e_{ij}(z) + (1 - pc_{ij}(x))e_{i0}(z) \quad (26)$$

Onde,

$pp_{ij}(x)$ é probabilidade de passar no curso escolhido j do indivíduo i

y_{ij}^* é a remuneração futura esperada

$pc_{ij}(x)$ é probabilidade de concluir o curso escolhido j do indivíduo i (sucesso)

$e_{ij}(z)$ é o valor presente de todos os ganhos futuros esperados caso o indivíduo i conclua o curso j

$e_{i0}(z)$ é o valor presente de todos os ganhos futuros esperados caso o indivíduo i não conclua o curso j

w_i são as características individuais dos estudantes

ε_{ij} o termo aleatório

A probabilidade do indivíduo i de passar no curso j foi estimada pelo modelo Logit, descrito da seguinte maneira.

$$pp_{ij}(x) = \Pr(P_p \neq 0 | \tilde{x}_{ij}) = \frac{e^{\tilde{x}_{ij}\beta}}{1 + e^{\tilde{x}_{ij}\beta}}, \text{ para cada grupo de cursos } J \quad (27)$$

Onde, $pp_{ij}^*(x)$ é uma variável não observada, porém observa-se a variável *dummy* P_p para todo indivíduo i

$$\begin{cases} P_p = 1, \text{ se passou no vestibular para o curso escolhido } J \\ P_p = 0, \text{ se foi reprovado no curso escolhido } J \end{cases} \quad (28)$$

\tilde{x}_{ij} são as variáveis de características individuais, socioeconômicas e de preparação dos candidatos para um dos cursos do grupo J .

A probabilidade do indivíduo i de concluir o curso escolhido e alcançado J foi estimada também pelo modelo Logit, descrito da seguinte maneira.

$$pc_{ij}(x) = \Pr(P_c \neq 0 | \hat{x}_{ij}) = \frac{e^{\hat{x}_{ij}\beta}}{1+e^{\hat{x}_{ij}\beta}}, \text{ para cada grupo de cursos } J \quad (29)$$

Onde, $pc_{ij}^*(x)$ é uma variável não observada, porém observa-se a variável *dummy* P_c para todo indivíduo i

$$\begin{cases} P_c = 1, & \text{se se concluiu o curso } \textit{escolhido e alcançado } J \\ P_c = 0, & \text{se desistiu ou foi reprovado no curso escolhido e alcançado } J \end{cases} \quad (30)$$

\hat{x}_{ij} são as variáveis de características individuais, socioeconômicas e de habilidade dos alunos universitários.

Para o valor presente dos ganhos futuros é utilizada apenas a remuneração esperada logo após a conclusão do curso como no modelo do capítulo 1. Um modelo muito semelhante ao do capítulo 1 é estimado e utilizam-se os coeficientes estimados da instituição de ensino UnB para calcular os ganhos futuros $e_{ij}(z)$ dos candidatos para cada opção de grupo j e também é estimado um modelo utilizando a mesma abordagem e base de dados para obter os ganhos dos candidatos a um curso na UnB caso eles não consigam completar o curso escolhido $e_{i0}(z)$.

A fim de analisar mais a fundo a nova variável introduzida no modelo, probabilidade de passar no curso, outro modelo foi também considerado. O modelo *Random Parameters Logit* (RPL), que é, na verdade, uma variação do modelo anterior *Mixed Logit Model*, onde se relaxa uma das hipóteses. Neste modelo se permite que os parâmetros sejam aleatórios com distribuições normal ou log-normal. Esse modelo também é conhecido com *Mixed Logit Model*, mas para não confundir com o *Mixed Logit Model* apresentado acima nesse estudo quando se refere ao *Mixed Logit Model*, o modelo será o descrito anteriormente e se utilizará o nome *Random Parameters Logit Model* para se referir ao modelo onde os parâmetros são aleatórios.

O *Random Parameters Logit Model* especifica a função utilidade do indivíduo i da seguinte maneira.

$$\begin{aligned} U_{ij} &= x'_{ij}\beta_i + \varepsilon_{ij}, \text{ para cada } j = 1, 2 \dots J \\ \beta_i &\sim \mathcal{N}[\beta, \Sigma_\beta] \end{aligned} \quad (31)$$

Onde ε_{ij} é iid com distribuição de valor extremo, exatamente como nos modelos anteriores, porém agora adicionalmente tem-se o parâmetro β_i aleatório com uma distribuição normal (distribuição também pode ser log-normal). O modelo pode ser reescrito da seguinte maneira:

$$\begin{aligned} U_{ij} &= x'_{ij}\beta_i + v_{ij} \\ v_{ij} &= x'_{ij}u_i + \varepsilon_{ij} \end{aligned} \quad (32)$$

$$u_i \sim \mathcal{N}[0, \Sigma_\beta]$$

$$\text{cov}[v_{ij}, v_{ik}] = x'_{ij} \Sigma_\beta x_{ik}, \text{ para } j \neq k \quad (33)$$

Dessa forma, observa-se que com a introdução do parâmetro aleatório o modelo passa a ter a atrativa propriedade de permitir a correlação entre as alternativas.

3.2.2 Dados

Para esse estudo são utilizados dados da base de dados do Cespe/UnB para todos os alunos que fizeram vestibular no período de 2004:1 a 2005:2. Nos dados, têm-se informações sobre os dados pessoais dos alunos, a sua escolha de curso, as respostas do questionário sociocultural, as notas dos alunos no vestibular e se o aluno, anos após o ingresso, completou o curso ou não.

No período 2004-1 têm-se 25.645 candidatos, em 2004-2 têm-se 27.386, em 2005-1 têm-se 23.389 candidatos e finalmente na base de dados de 2005-2 têm-se 24.563 candidatos. Após a junção dos quatro semestres e de eliminar alguns candidatos que não fizeram as provas do vestibular, a base de dados final ficou com 86.674 candidatos.

No questionário sociocultural (QSC) os candidatos são submetidos a 18 questões na inscrição para o vestibular, isto é, antes de fazer o exame. QSC cobre temas como raça, *status* socioeconômico e preparação para o vestibular. No entanto, apesar do QSC ser disponibilizado para todos os candidatos, este não é obrigatório responder para o ingresso na universidade. Cerca de 50% dos candidatos responderam o QSC (43.468) para esse período. É importante informar que as taxas de resposta do QSC caíram de cerca de 60% em 2004-1 e 2004-2 para 45% e 56% em 2005-1 e 2005-2, respectivamente. Isso representa uma fonte potencial de seletividade amostral. Para tentar sanar esse problema é estimada uma probabilidade de responder o QSC, através de um modelo probit e utiliza-se algumas variáveis conhecidas para todos os candidatos. Tais variáveis são obtidas da ficha de inscrição e incluem gênero, se cotista, curso para o qual está pleiteando uma vaga na universidade e bairro onde mora. O inverso desta probabilidade é utilizado como peso nas estimações que utilizaram dados do QSC.

Para os retornos esperados futuros dos estudantes é utilizada a base de dados do capítulo 1 para estimar dois modelos, uma para o retorno financeiro caso o candidato termine o curso universitário, $e_{ij}(z)$ e um outro, $e_{i0}(z)$, para caso o mesmo candidato não termine o curso universitário. Utiliza-se os coeficientes estimados dos modelos para os cursos da Universidade de Brasília para inferir os ganhos esperados salariais futuros dos candidatos ao curso universitário pretendido.

3.2.3 Variáveis

As variáveis são retiradas do questionário sociocultural apresentado no anexo e são divididas em 4 grupos, características individuais, características familiares, habilidade, preferências e preparação e por último características demográficas. Segue abaixo a classificação adotada:

Características individuais:

- 1 Gênero - mulher
- 2 Raça – PPI
- 3 Opção de ingresso - cota
- 4 Estado civil - casado

Características Familiares:

- 1 Escolaridade da mãe - mae_edu_NS e mae_edu_NM
- 2 Escolaridade do pai - pai_edu_NS e pai_edu_NM
- 3 Renda familiar – renda_alta, renda_media e renda_baixa

Características demográficas:

- 1 Local onde reside – Bairro (qsc18_n)

Habilidades, preferencias e preparação:

- 1 Nota no vestibular - nota_port e nota_mat
- 2 Estudo em escola pública no ensino médio - publica_EM
- 3 Estudo em escola particular - particular_EM
- 4 Cursinho preparatório para o vestibular – cursinho
- 5 Tentativas anteriores para UnB - tentativas_UnB
- 6 Recém concluinte do ensino médio – concluinte
- 7 Curso regular no ensino médio - curso_regular

A Tabela 9 apresenta as médias, desvios padrão, assim como uma breve descrição das variáveis. Segundo as variáveis de características individuais 53% da amostra é composta por mulheres, 53% também é o percentual de candidatos que se autodeclararam pretos, pardos ou indígenas, PPI, porém apenas 13% optou por tentar o ingresso na UnB pelas vagas de cotas raciais. É possível que muito se considerem pardos acreditem que não seriam pardos o suficiente pelos critérios rigorosos raciais da UnB. Por fim, a amostra apresenta apenas 5% de candidatos casados.

Para as variáveis de características familiares tem-se aproximadamente uma distribuição uniforme entre as variáveis que medem o nível de escolaridade das mães e dos pais dos candidatos a um curso superior, sendo um terço para as mães dos candidatos com nível superior, um terço para as mães com nível médio e um terço para as demais com nível inferior ao nível médio. De forma análoga para os pais dos candidatos, com um ligeiro aumento na escolaridades dos pais com nível inferior ao nível médio em relação aos pais com nível médio completo. Com relação à renda familiar, 54% dos candidatos são de famílias com renda familiar média, ou seja, de 3 a 20 salários mínimos, 24% dos para candidatos são de famílias com renda familiar alta, acima de 20 salários mínimos e 13% dos candidatos são de famílias com renda familiar abaixo de 3 salários mínimos.

Com relação às habilidades, preferências e preparação dos candidatos observa-se que 52% estudou a maior parte do período em escola pública no ensino médio e 48% em escola particular, com cerca de 83% cursando o ensino médio regular e 17% cursando outro tipo, como: supletivo, profissionalizante, técnico e outros. Aproximadamente metade dos candidatos fez algum tipo de cursinho preparatório para o vestibular e aproximadamente 60% já havia feito o vestibular em anos anteriores. A variável *concluinte* indica que cerca de 60% dos candidatos terminaram o ensino médio recentemente, há menos de 1 ano e aponta para uma amostra de candidatos jovens. As notas do caderno de provas de português e ciências sociais e do caderno de provas de matemática e ciências naturais também estão na tabela 9, onde se observa a média de 42,86 para o caderno de português e ciências sociais e uma média um pouco menor para o caderno de matemática e ciências naturais.

Tabela 9 - Estatísticas descritivas

Variáveis	Obs	Média	DP	Min	Max	Descrição
Características Individuais						
Mulher	693392	0,53	0,50	0	1	Variável <i>dummy</i> , mulher = 1 e homem = 0
PPI	345520	0,53	0,50	0	1	Variável <i>dummy</i> , PPI = 1 e brancos ou asiáticos = 0
Cota	693392	0,13	0,34	0	1	Variável <i>dummy</i> , cotista = 1 e não cotista = 0
Casado	348672	0,05	0,21	0	1	Variável <i>dummy</i> , casado = 1 e não casado = 0
Características Familiares						
Mae_educ_NS	347296	0,33	0,47	0	1	Variável <i>dummy</i> , mãe com nível superior = 1 e mãe sem nível superior = 0
Mae_educ_NM	347296	0,34	0,47	0	1	Variável <i>dummy</i> , mãe com nível médio = 1 e mãe com nível superior ou menos de nível médio = 0
Pai_educ_NS	346568	0,33	0,47	0	1	Variável <i>dummy</i> , pai com nível superior = 1 e pai sem nível superior = 0
Pai_educ_NM	346568	0,30	0,46	0	1	Variável <i>dummy</i> , pai com nível médio = 1 e pai com nível superior ou menos de nível médio = 0
Renda Baixa	346504	0,13	0,33	0	1	Variável <i>dummy</i> , renda baixa = 1 e renda média ou alta = 0
Renda Média	346504	0,54	0,50	0	1	Variável <i>dummy</i> , renda média = 1 e renda baixa ou alta = 0
Renda Alta	346504	0,24	0,43	0	1	Variável <i>dummy</i> , renda alta = 1 e renda baixa ou média = 0
Habilidade, preferências e preparação						
nota_port	554848	42,86	20,93	-77	110	Nota no caderno de português e ciências sociais
nota_mat	554848	31,56	24,00	-71	144	Nota no caderno de matemática e ciências naturais
publica	346128	0,52	0,50	0	1	Variável <i>dummy</i> , mais da metade do ensino médio em escola pública = 1 e caso contrário = 0
particular	346128	0,48	0,50	0	1	Variável <i>dummy</i> , mais da metade do ensino médio em escola particular = 1 e caso contrário = 0
cursinho	346672	0,51	0,50	0	1	Variável <i>dummy</i> , estudou em cursinho = 1 e caso contrário = 0
tentativas_UnB	346912	0,59	0,49	0	1	Variável <i>dummy</i> , fez vestibular para UnB em anos anteriores = 1 e primeira vez = 0
concluinte	346856	0,59	0,49	0	1	Variável <i>dummy</i> , concluiu o ensino médio há pelo menos 1 ano = 1 e há mais de um ano = 0
curso_regular	345440	0,83	0,37	0	1	Variável <i>dummy</i> , fez o ensino médio em curso regular = 1 e caso contrário = 0

3.3 Resultados

Os resultados são apresentados em três etapas. Primeiramente os ganhos futuros esperados para cada indivíduo, tanto para o curso que ele escolheu quanto para as demais opções de cursos, já que a escolha será sempre por um curso de um determinado grupo em detrimento dos demais cursos dos outros grupos. Em seguida se apresenta os resultados para cada candidato da probabilidade deste

passar em um curso de um determinado grupo, faz-se isto para todas as possibilidades de grupos de cursos de cada candidato. Por fim, apresenta-se o modelo de escolha de carreiras em função dos ganhos esperados, das probabilidades de passar e das características individuais, sociais, familiares, habilidades, preferências e preparação de cada candidato.

3.3.1 Ganhos futuro esperado

Para o cálculo dos ganhos futuro esperado, ou remunerações futura esperada, é preciso estimar a probabilidade de o candidato concluir o curso, os ganhos no mercado de trabalho se o candidato se formar e os ganhos caso este não consiga completar o curso.

$$y_{ij}^* = pc_{ij}(x)e_{ij}(z) + (1 - pc_{ij}(x))e_{i0}(z) \quad (34)$$

Onde: y_{ij}^* é a remuneração futura esperada

$pc_{ij}(x)$ é probabilidade de concluir o curso escolhido j do indivíduo i (sucesso)

$e_{ij}(z)$ é a remuneração futura esperada caso o indivíduo i conclua o curso j

$e_{i0}(z)$ é a remuneração futura esperada caso o indivíduo i não conclua o curso j

3.3.1.1 Probabilidade de conclusão de curso

Os modelos de escolha de carreiras tradicionais assumem uma probabilidade de conclusão no curso universitário constante, igual a 1 para todos os candidatos, como também uma remuneração constante para cada curso universitário. Flexibilizar essas hipóteses trás um ganho em relação aos modelos tradicionais, já que os candidatos são diferentes, possuem formação, habilidades e características socioeconômicas distintas. Nesse estudo assume-se que os candidatos a um curso universitário verificam suas reais chances de concluir o curso para cada opção de grupos de cursos e consideram a sua remuneração esperada para estes na sua escolha de carreira.

A probabilidade de conclusão de curso, nesse estudo, é estimada pelo modelo logit e apresentada na tabela 10 para cada grupo de cursos (engenharias, ciências, saúde, medicina/direito, negócios, educação e outros). As variáveis explicativas são de características individuais, de habilidades e familiares. Para as variáveis de características individuais, escolheu-se gênero, raça e estado civil, para de habilidades escolheu-se as notas nos cadernos de provas de português e ciências sociais e matemática e ciências naturais. Porém as notas no caderno de português e ciências sociais não se mostram significativas para nenhum grupo de curso, enquanto as notas do caderno de matemática e ciências naturais estão entre as variáveis que são significativas para a maior quantidade de grupos. Para finalizar, as variáveis escolhidas de características familiares são as relacionadas a escolaridade da mãe dos candidatos.

Tabela 10 - Coeficientes do modelo Logit para a probabilidade de conclusão do curso

Modelo - probabilidade de concluir	Engenharias	Ciências	Saúde	Medicina /Direito	Ciências Sociais	Negócios	Educação	Outros
(coluna 1)	(coluna 2)	(coluna 3)	(coluna 4)	(coluna 5)	(coluna 6)	(coluna 7)	(coluna 8)	(coluna 9)
Numero de alunos	315	131	173	127	372	226	504	313
Wald chi2(7)	9,89	15,97	17,60	26,14	14,81	11,40	28,74	14,57
prob > Chi2	19,5%	1,4%	0,7%	0,0%	3,9%	12,2%	0,0%	4,2%
Log pseudolikelihood	-347,92	-132,39	-82,17	-23,82	-309,98	-160,48	-508,22	238,13
<i>Variáveis</i>	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
Mulher	1,601 ** (0,626)	-0,139 (0,526)	0,845 (0,548)	0,658 * (0,357)	0,658 * (0,357)	0,062 (0,453)	1,127 *** (0,228)	0,520 (0,339)
PPI	0,286 (0,545)	-0,023 (0,505)	-0,541 (0,596)	0,781 ** (0,376)	0,781 ** (0,376)	-0,145 (0,543)	0,008 (0,224)	0,137 (0,349)
Casado	-1,429 (1,201)	(omitted)	(omitted)	-1,229 (0,980)	-1,229 (0,980)	-1,641 ** (0,826)	0,021 (0,438)	-1,206 ** (0,642)
Nota Port.	0,028 (0,023)	0,025 (0,022)	0,015 (0,023)	0,011 (0,022)	0,011 (0,022)	-0,011 (0,020)	-0,006 (0,010)	-0,027 (0,017)
Nota Mat.	0,009 (0,009)	0,031 *** (0,011)	0,001 (0,016)	0,018 * (0,010)	0,018 * (0,010)	0,026 (0,020)	0,001 (0,007)	0,019 ** (0,010)
Mãe_edu_NS	-0,530 (0,549)	1,259 ** (0,609)	1,323 * (0,741)	0,390 (0,470)	0,390 (0,470)	0,195 (0,662)	-0,038 (0,292)	-1,243 ** (0,540)
Mãe_edu_NM	-0,838 (0,732)	1,184 * (0,651)	1,171 (0,783)	0,094 (0,495)	0,094 (0,495)	-0,098 (0,644)	0,309 (0,266)	-1,072 ** (0,532)
constante	-0,867 (1,926)	-3,649 ** (1,778)	0,068 (2,338)	-0,972 (1,564)	-0,972 (1,564)	1,427 (1,987)	0,797 (0,639)	3,344 *** (1,198)

OBS : *** significante a menos de 1%, ** significante a 5% e * significante a 10% e erros-padrão entre parênteses. Nessa tabela se apresentam os coeficientes dos modelos *logit* para a probabilidade de conclusão do curso para todas as alternativas de grupos de cursos

Os resultados mostram que pelo menos uma das variáveis é significativa para um dado grupo. Vale ressaltar que ser mulher gera impactos positivos para a conclusão dos cursos nos grupos de engenharia, medicina/direito, ciências sociais e educação, enquanto que ser do grupo PPI é apenas relevante com impacto positivo para a conclusão dos grupos de medicina/direito e ciências sociais. Já ser casado prejudica a conclusão dos cursos no grupo de negócios. A nota nos cadernos de matemática e ciências naturais foi relevante para os cursos dos grupos de ciência, medicina/direito e ciências sociais, porém não é significativa para o grupo de engenharia. Talvez a amostra pequena não tenha conseguido capturar o impacto nesse grupo ou a variável não varia muito nesse grupo. O nível de escolaridade da mãe dos candidatos é relevante e com impacto positivo sobre a probabilidade de conclusão nos grupos de ciências e saúde.

Após a estimação dos coeficientes do modelo *logit*, para cada candidato e para cada grupo de cursos, é projetado uma probabilidade de conclusão individual para cada opção de grupo de cursos. As projeções são individuais e de acordo com o modelo variam dependendo do gênero, raça, estado civil, notas no vestibular, nível de escolaridade da mãe e para cada grupo de cursos, as tabelas 11 e 12 apresentam as probabilidades de conclusão dos candidatos para cada grupo. Na tabela 11 se apresenta a probabilidade de conclusão observada para cada grupo de cursos, ou seja,

de todos os alunos daquele grupo quantos realmente alcançaram o diploma, e a média das probabilidades projetadas para estes alunos pelo modelo. Porém vale ressaltar que aproximadamente metade dos alunos não respondeu o QSC e, portanto, não se tem informação sobre todas as características individuais e familiares destes para projetar a probabilidade de concluir.

Tabela-11 - Probabilidade de conclusão do curso por grupos						
		Obs	Mean	Std. Dev.	Min	Max
Engenharias	Prob. de conclusão observada	547	82,8%	37,8%	0	1
	Prob. de conclusão projetada	315	80,8%	9,5%	0,49	0,98
Ciências	Prob. de conclusão observada	256	65,2%	47,7%	0	1
	Prob. de conclusão projetada	131	74,4%	16,8%	0,12	0,98
Saúde	Prob. de conclusão observada	291	90,7%	29,1%	0	1
	Prob. de conclusão projetada	173	91,7%	6,2%	0,63	0,97
Medicina /Direito	Prob. de conclusão observada	219	94,5%	22,8%	0	1
	Prob. de conclusão projetada	127	97,0%	6,1%	0,71	1,00
Ciências Sociais	Prob. de conclusão observada	706	82,6%	38,0%	0	1
	Prob. de conclusão projetada	372	85,7%	9,1%	0,26	0,98
Negócios	Prob. de conclusão observada	448	83,9%	36,8%	0	1
	Prob. de conclusão projetada	226	88,5%	8,4%	0,46	0,98
Educação	Prob. de conclusão observada	917	70,3%	45,7%	0	1
	Prob. de conclusão projetada	504	74,2%	11,1%	0,56	0,89
Outros	Prob. de conclusão observada	582	81,4%	38,9%	0	1
	Prob. de conclusão projetada	313	86,2%	7,9%	0,37	0,97

OBS :. Nessa tabela se apresenta a probabilidade de conclusão observada para cada grupo de cursos, ou seja, de todos os alunos daquele grupo quantos realmente alcançaram o diploma, e a média das probabilidades projetadas para estes alunos pelo modelo.

Na tabela 11 se observa que 94,5% dos alunos que cursaram os cursos de medicina e direito concluíram o curso, enquanto apenas 65,2% dos alunos que cursaram um curso do grupo de ciências concluíram seus cursos. Talvez isto esteja relacionado com a dificuldade de ingressar no curso e a dificuldade do curso em si.

A tabela 12 apresenta um resumo das probabilidades projetadas de conclusão de curso de um candidato. Para cada candidato foi projetada uma probabilidade de conclusão para todas as alternativas de grupos de cursos, ou seja, como se tem oito grupos (engenharias, ciências, saúde, medicina/direito, ciências sociais, negócios, educação e outros) para cada candidato tem-se 8 probabilidades de conclusão de curso, uma para cada grupo. Porém o candidato opta por apenas um

dos cursos de um dado grupo para concorrer à vaga na universidade. Dessa forma a tabela 12 agrupa nas linhas todos os candidatos que estão concorrendo a uma vaga em um determinado grupo de cursos e nas colunas apresentam as médias das probabilidades de conclusão desses candidatos para cada opção de grupo de cursos que estes poderiam ter escolhidos, independente da sua escolha.

Tabela-12 - Probabilidade de conclusão do curso por grupos (projetada)

Média das probabilidades dos candidatos que escolheram um dos grupos indicados nas linhas concluir um curso do grupo da coluna

Candidatos ao Grupo (coluna 1)	Engenharias (coluna 2)	Ciências (coluna 3)	Saúde (coluna 4)	Medicina /Direito (coluna 5)	Ciências Sociais (coluna 6)	Negócios (coluna 7)	Educação (coluna 8)
Engenharias	64,37%	45,44%	82,38%	23,68%	69,67%	86,31%	69,65%
Ciências	66,61%	37,81%	81,29%	16,58%	68,78%	84,77%	73,29%
Saúde	74,91%	33,12%	84,70%	10,93%	70,26%	83,75%	81,23%
Medicina /Direito	76,39%	51,70%	87,02%	33,73%	75,29%	85,89%	75,42%
Ciências Sociais	74,11%	35,77%	84,28%	19,96%	70,31%	82,66%	77,47%
Negócios	66,91%	32,72%	80,17%	16,44%	66,42%	81,99%	74,22%
Educação	66,81%	21,66%	76,65%	5,64%	63,06%	79,60%	77,36%

Obs: nas linhas estão os candidatos agrupados nos grupos que eles realmente optaram por fazer o vestibular e nas colunas estão os grupos para qual se calcula as médias das probabilidades de conclusão deste

Observa-se que para todos os candidatos que estão concorrendo para o grupo de engenharia a média de sua probabilidade de concluir o curso de engenharia é de 64,37%, bem menor que a apresentada na tabela 11 de 80,8%. Isto ocorre porque nessa amostra da tabela 12 estão todos os candidatos que optaram por um dos cursos do grupo de engenharia e na amostra da tabela 11 estão apenas os alunos que passaram para um dos cursos do grupo de engenharia. A tabela 12, na coluna 3, mostra que a media das probabilidades de conclusão de cursos do grupo de ciências é de 37,81% se os candidatos são os que escolheram fazer vestibular para um dos cursos do grupo de ciências, entretanto se os candidatos são os que escolheram um dos cursos do grupo de engenharia a média da probabilidade de concluir um dos cursos do grupo de ciência destes é de 45,44%. Já para os candidatos que escolheram um dos cursos do grupo de educação tem em média apenas 21,66% de chances de concluir um dos cursos do grupo de ciências. Observe que a média de probabilidade de conclusão de um dos cursos do grupo de ciências (coluna3) é maior para os candidatos que escolheram cursos do grupo de medicina/direito (51,70%) e engenharias (45,44%) seguindo dos próprios candidatos que escolheram essa opção (37,81%).

Em geral os candidatos que optam pelo grupo medicina/direito têm as maiores médias de probabilidade de conclusão, ou médias entre as maiores médias, para concluir cursos de qualquer um dos grupos comparando com os outros candidatos. Para a conclusão de um dos cursos dos grupos de saúde (coluna 4) e ciências sociais (coluna 6) a média das probabilidades de conclusão dos candidatos que optaram por fazer vestibular para um curso desses grupos só é inferior que a média dos candidatos que escolheram os cursos de medicina ou direito.

Os candidatos a um curso do grupo de negócios possuem a maior das médias das probabilidades de conclusão para exatamente concluir o grupo de negócios, 81,99%, grupo este escolhido por esses candidatos. Isto também ocorre para os candidatos que escolheram um curso do grupo de saúde, 84,70%. Já entre os candidatos que optaram por um dos cursos do grupo de educação a menor das medias das probabilidades de conclusão do curso é para um curso no grupo de medicina/direito, ou seja, se estes candidatos fizerem o curso de medicina ou de direito a média das suas probabilidades para concluir um desses cursos é de 5,64% e por outro lado para concluir um dos cursos de sua própria escolha, grupo educação é de 77,36%.

3.3.1.2 Remuneração futura, caso o candidato conclua o curso(e_{ij})

Para calcular a remuneração futura de cada candidato para cada grupo de cursos, opta-se por simplificar a trajetória de todos os ganhos futuros devido a graduação pela remuneração logo após a formatura. Utiliza-se o estudo do capítulo 1 para remuneração dos recém-formados e se estima novamente para a mesma base de dados do capítulo 1 o modelo HLM abaixo:

$$\begin{aligned} \ln(Y_{ij}) &= \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + \beta_{4j}X_{4ij} + \varepsilon_{ij} \\ \beta_{0j} &= \gamma_{00} + \gamma_{01}Z_{1j} + \gamma_{02}Z_{2j} + \gamma_{03}Z_{3j} + \gamma_{04}Z_{4j} + \gamma_{05}Z_{5j} + \gamma_{06}Z_{6j} + u_{0j} \\ \beta_{1j} &= \gamma_{10} \\ \beta_{2j} &= \gamma_{20} \\ \beta_{3j} &= \gamma_{30} \\ \beta_{4j} &= \gamma_{40} \end{aligned} \tag{35}$$

Onde para o grupo de variáveis explicativas em nível do aluno, utiliza-se como características individuais, X_{1ij} , Mulher, PPI, Casado e Filhos, utiliza-se como características familiares, X_{2ij} , renda familiar (renda alta, ou renda familiar acima de 20 salários mínimos, e renda baixa, ou renda familiar inferior a 3 salários mínimos), utiliza-se como medida de desempenho do aluno, X_{3ij} , a nota do Enade e utiliza-se como variáveis de mercado de trabalho, X_{4ij} , tipo de jornada (trabalho parcial, ou inferior a 40 horas semanais e trabalho acima de 44 horas semanais) e tipo de emprego como estatutário e diploma não necessário. Para o grupo de variáveis explicativas em nível da instituição de ensino superior Z_{kj} , se escolhem a indicadora da instituição de ensino superior ser particular, o percentual de alunos PPI na instituição, o percentual de pais com nível superior e médio em uma dada instituição e as variáveis de avaliação da instituição (média_ingressante, conceito e CPC).

O modelo e a base de dados são os mesmos do capítulo 1, porém devido a ausência de informação da idade dos candidatos na base de dados para o modelo de escolhas de carreiras, opta-se por re-estimar o modelo do capítulo 1 o mais próximo possível sem a variável idade. O resultado da estimação é apresentado na tabela 13 e é muito parecido com o apresentado no capítulo 1. O termo u_{0j} é o termo que diferencia o intercepto pelas instituições de ensino superior e no modelo estimado seu valor para a UnB e para cada grupo da universidade é apresentado na tabela 14.

Tabela 13 - Coeficientes do modelo HLM para e_{ij} dos recém-formados (continua)

Modelo HLM para a Remuneração futura e_{ij} (coluna 1)	Engenharias (coluna 2)	Ciências (coluna 3)	Saúde (coluna 4)	Medicina /Direito (coluna 5)	Ciências Sociais (coluna 6)	Negócios (coluna 7)	Educação (coluna 8)	Outros (coluna 9)
Numero de alunos	8655	3986	9497	8719	21123	42036	29918	18464
Numero de instituições	106	60	144	126	226	532	278	225
Wald chi2(19)	1777,21	943,19	3502,53	4981,78	4736,44	13600,64	7471,33	4173
Prob > Chi2	0	0	0	0	0	0	0	0
Log likelihood	-7464,08	-3665,25	-8640,56	-8834,41	-18753,52	-37103,43	-24794,40	-15614,42
Efeito fixo - Variáveis do nível Aluno								
Mulher	-0,268 *** (0,016)	-0,249 *** (0,021)	-0,097 *** (0,013)	-0,201 *** (0,015)	-0,138 *** (0,009)	-0,194 *** (0,006)	-0,242 *** (0,008)	-0,146 *** (0,010)
PPI	-0,026 (0,016)	-0,057 * (0,025)	0,009 (0,018)	-0,050 ** (0,018)	-0,032 *** (0,010)	-0,043 *** (0,007)	0,008 (0,007)	-0,030 ** (0,010)
Casado	0,113 *** (0,022)	0,193 *** (0,034)	-0,048 * (0,024)	0,142 *** (0,019)	0,117 *** (0,012)	0,166 *** (0,008)	0,029 *** (0,008)	0,055 *** (0,011)
Filhos	0,006 (0,023)	0,079 * (0,035)	-0,006 (0,023)	0,087 *** (0,020)	0,059 *** (0,012)	0,095 *** (0,009)	0,069 *** (0,009)	0,075 *** (0,012)
Renda Baixa	-0,180 *** (0,026)	-0,311 *** (0,037)	0,000 (0,018)	-0,331 *** (0,023)	-0,233 *** (0,010)	-0,288 *** (0,008)	-0,171 *** (0,007)	-0,213 *** (0,010)
Renda Alta	0,103 *** (0,017)	0,177 *** (0,031)	0,019 (0,018)	0,170 *** (0,021)	0,150 *** (0,017)	0,221 *** (0,013)	0,126 *** (0,022)	0,127 *** (0,022)
Nota no Enade	0,002 *** (0,0005)	0,005 *** (0,0006)	0,001 (0,0005)	0,006 *** (0,0005)	0,003 *** (0,0003)	0,005 *** (0,0002)	0,001 *** (0,0002)	0,002 *** (0,0003)
Trabalho Parcial	-0,174 *** (0,023)	-0,245 *** (0,029)	-0,072 *** (0,014)	-0,080 *** (0,018)	-0,072 *** (0,010)	0,039 *** (0,009)	-0,068 *** (0,007)	0,001 (0,010)
Trabalho +44h/sem	0,272 *** (0,039)	0,372 *** (0,050)	0,616 *** (0,023)	0,358 *** (0,040)	0,404 *** (0,017)	0,507 *** (0,025)	0,489 *** (0,011)	0,481 *** (0,015)
Diploma ã Necessario	-0,382 *** (0,015)	-0,198 *** (0,020)	-0,529 *** (0,013)	-0,317 *** (0,015)	-0,171 *** (0,008)	-0,129 *** (0,006)	-0,145 *** (0,007)	-0,130 *** (0,009)
Estatutario	-0,187 *** (0,024)	-0,016 (0,032)	0,275 *** (0,015)	0,444 *** (0,016)	0,200 *** (0,010)	0,101 *** (0,011)	0,195 *** (0,007)	0,204 *** (0,010)

Tabela 13 - Coeficientes do modelo HLM para e_{ij} dos recém-formados (conclusão)

Modelo HLM para a Remuneração futura e_{ij} (coluna 1)	Engenharias (coluna 2)	Ciências (coluna 3)	Saúde (coluna 4)	Medicina /Direito (coluna 5)	Ciências Sociais (coluna 6)	Negócios (coluna 7)	Educação (coluna 8)	Outros (coluna 9)
Efeito fixo - Variáveis do nível Instituição								
Particular	-0,352 *** (0,097)	-0,526 *** (0,106)	-0,094 (0,071)	-0,794 *** (0,079)	-0,512 *** (0,043)	-0,662 *** (0,035)	-0,341 *** (0,037)	-0,443 *** (0,043)
Particular * idade	0,024 *** (0,0019)	0,023 *** (0,0027)	0,003 (0,0016)	0,027 *** (0,0012)	0,019 *** (0,0008)	0,028 *** (0,0006)	0,011 *** (0,0006)	0,017 *** (0,0008)
%_PPI_na_inst	0,149 (0,191)	0,339 (0,234)	0,028 (0,1440)	0,683 *** (0,160)	0,069 (0,094)	-0,071 (0,064)	-0,006 (0,080)	-0,020 (0,087)
%_pai_educ_NS	0,060 (0,225)	0,884 *** (0,276)	0,362 * (0,176)	0,758 *** (0,197)	0,609 *** (0,095)	0,709 *** (0,099)	0,839 *** (0,209)	0,267 * (0,126)
%_pai_educ_NM	-1,016 * (0,431)	-1,042 (0,591)	0,146 (0,343)	0,116 (0,461)	0,345 (0,184)	-0,356 * (0,144)	0,063 (0,223)	0,200 (0,215)
Media_ingressantes	0,016 (0,009)	-0,013 (0,010)	0,004 (0,005)	0,013 * (0,006)	0,008 * (0,004)	0,009 ** (0,003)	-0,001 (0,003)	-0,001 (0,004)
Conceito	0,136 * (0,065)	0,161 * (0,068)	0,002 (0,046)	-0,064 (0,047)	-0,078 (0,040)	0,069 ** (0,023)	-0,035 (0,034)	-0,017 (0,033)
CPC	-0,087 (0,064)	0,008 (0,077)	-0,060 (0,041)	-0,029 (0,047)	0,023 (0,033)	-0,020 (0,024)	0,073 * (0,033)	0,054 (0,032)
constante	7,337 *** (0,295)	7,235 *** (0,333)	6,974 *** (0,216)	6,827 *** (0,228)	6,728 *** (0,119)	6,646 *** (0,097)	6,885 *** (0,125)	6,819 *** (0,127)
Efeito Aleatório								
var(constante)	0,065 (0,010)	0,038 (0,008)	0,032 (0,005)	0,052 (0,008)	0,037 (0,004)	0,044 (0,003)	0,042 (0,004)	0,035 (0,004)
var(Residual)	0,318 (0,005)	0,357 (0,008)	0,351 (0,005)	0,431 (0,007)	0,338 (0,003)	0,333 (0,002)	0,300 (0,002)	0,310 (0,003)
teste LR - chibar2(01)	1010,26	243,41	507,7	544,31	1498,58	4010,94	2815,09	1343,22

OBS : *** significante a menos de 0,1%, ** significante a 1% e * significante a 5% e erros-padrão entre parênteses. Nessa tabela se apresentam os coeficientes dos modelos HLM, para todos os grupos de cursos, para a remuneração caso o candidato conclua o curso.

Tabela 14 - Modelo HLM para a Remuneração futura e_{ij} - u_{0j} para a UnB

Engenharias	0,432
Ciências	0,036
Saúde	0,000
Medicina /Direito	0,525
Ciências Sociais	0,427
Negócios	0,313
Educação	0,543
Outros	0,633

OBS: Nessa tabela se apresentam termo u_{0j} , o termo que diferencia o intercepto pelas instituições de ensino superior e no modelo, para a UnB.

Após a estimação dos coeficientes do modelo projeta-se a remuneração futura para cada candidato e para cada opção de escolha de grupos de cursos. Para esta projeção utiliza-se os dados da UnB da base de dados do capítulo 1 para as variáveis em nível da instituição de ensino superior, ou seja, para as variáveis %_PPI_na_inst, %_pai_educ_NS, %_pai_educ_NM, Média_ingressante, Conceito e CPC e para as variáveis Particular e Particular_idade têm-se 0, uma vez que a UnB é pública. Para as variáveis de características individuais, familiares, de desempenho no curso universitário e do mercado de trabalho utilizam-se as respostas do questionário sociocultural dos candidatos. Entretanto algumas variáveis não estão disponíveis nessa base, como é o caso do desempenho no curso universitário, as variáveis de mercado de trabalho e se o candidato possui filhos ou não.

No modelo todas essas variáveis são centradas na média, então supor zero para todas significa que o candidato está na média para estas variáveis, dessa forma adota-se o valor zero para as variáveis nota Enade, Filhos, Trabalho Parcial, Trabalho +44h/sem, Diploma não Necessário e Estatutário para todos os candidatos.

A tabela 15 apresenta um resumo das médias das projeções das remunerações dos candidatos (e_{ij}) para cada opção de grupo. Na linha tem-se a real escolha do grupo do candidato e nas colunas todas as opções de grupos de curso que este poderia ter optado. Observa-se que se a escolha dependesse apenas da remuneração futura todos os candidatos escolheriam o curso de Medicina ou Direito e nunca escolheriam os cursos do grupo de saúde.

Tabela-15- Média das remunerações (e_{ij}) projetadas por grupos de cursos

Média das remunerações (e_{ij}) dos candidatos que escolheram um dos grupos indicados nas linhas para um curso do grupo da coluna

Candidatos ao Grupo (coluna 1)	Engenharias (coluna 2)	Ciências (coluna 3)	Saúde (coluna 4)	Medicina /Direito (coluna 5)	Ciências Sociais (coluna 6)	Negócios (coluna 7)	Educação (coluna 8)
Engenharias	R\$ 3.853,19	R\$ 3.518,44	R\$ 1.313,01	R\$ 4.593,36	R\$ 2.370,13	R\$ 2.946,12	R\$ 2.983,42
Ciências	R\$ 3.638,07	R\$ 3.290,66	R\$ 1.292,17	R\$ 4.323,76	R\$ 2.268,63	R\$ 2.776,10	R\$ 2.828,89
Saúde	R\$ 3.284,90	R\$ 2.999,28	R\$ 1.243,62	R\$ 4.013,56	R\$ 2.156,29	R\$ 2.582,42	R\$ 2.576,53
Medicina /Direito	R\$ 3.569,26	R\$ 3.312,30	R\$ 1.269,39	R\$ 4.386,62	R\$ 2.302,32	R\$ 2.827,04	R\$ 2.783,01
Ciências Sociais	R\$ 3.435,66	R\$ 3.139,77	R\$ 1.262,22	R\$ 4.170,69	R\$ 2.216,29	R\$ 2.686,01	R\$ 2.687,00
Negócios	R\$ 3.595,24	R\$ 3.262,62	R\$ 1.284,81	R\$ 4.286,27	R\$ 2.257,36	R\$ 2.763,75	R\$ 2.796,60
Educação	R\$ 3.364,93	R\$ 2.982,84	R\$ 1.266,50	R\$ 3.928,83	R\$ 2.116,51	R\$ 2.534,33	R\$ 2.618,98

Fonte Rais 2007 - Valores nominais em Reais no ano de 2007.

Obs: nas linhas estão os candidatos agrupados nos grupos que eles realmente optaram por fazer o vestibular e nas colunas estão as alternativas de grupos de cursos para qual se calcula as médias das remunerações caso conclua esse curso

3.3.1.3 Remuneração futura, caso o candidato não conclua o curso(e_{i0})

Para obter a remuneração dos candidatos caso estes não concluam o curso opta-se por investigar quanto um aluno universitário recebe após ingressar em um curso universitário e antes de concluir seu curso no mercado de trabalho. A remuneração média dos trabalhadores com apenas o nível médio é preterida, já que só o fato do trabalhador estar cursando um curso universitário já os diferencia dos demais com apenas o nível médio completo.

Para obter a remuneração desses alunos universitários que ainda não concluíram o curso utiliza-se os dados do Enade de 2004, 2005 e 2006, junto com a Rais 2007, da mesma forma que foi feito para o capítulo 1, porém agora com os alunos ingressante no curso universitário. Dessa forma, os alunos que ingressaram em um curso universitário em 2004, 2005 ou 2006 ainda não terminaram o curso em 2007, ano o qual se observa se estes alunos estão no mercado de trabalho (Rais 2007). Como o estudo é para os candidatos a uma vaga na Universidade de Brasília, opta-se por escolher os alunos que fizeram as provas do Enade apenas de instituições no mesmo município da UnB, já que a não se encontrou dados suficientes para estimar um modelo apenas com alunos da UnB. Opta-se por instituições de ensino superior no mesmo município da UnB para assegurar que o emprego é na região da Universidade de Brasília e, portanto, evitar uma distorção no salário devido a diferentes custos de vida de regiões distintas.

A metodologia utilizada é a mesma do capítulo 1. Primeiro estima-se os ganhos salariais individuais usando regressão por mínimos quadrados ordinários (MQO), em seguida utiliza-se a metodologia de Hox(2010) para verificar a necessidade de utilizar o modelo Hierárquico Linear (HLM).

Primeiramente, para o modelo MQO, se estima uma função de ganho relacionando de forma linear a remuneração do recém-formado com variáveis de características individuais, de perfil familiar, de característica da educação e do mercado de trabalho. O modelo básico é escrito da seguinte maneira:

$$\ln(Y_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon_1 \quad (36)$$

Onde Y_i é a remuneração do recém-formado; X_1 é um vetor de características individuais, que para este estudo são gênero (mulher) e estado civil (casado); X_2 é um vetor de variáveis da base familiar, que é representado pela variável Renda Baixa, ou renda familiar até 3 salários mínimos, neste estudo; X_3 é um vetor de características da educação, representado aqui pelos cursos graduados (se do grupo de engenharias, ciências, saúde, medicina/direito, ciências sociais, negócios, educação ou outros), pela variável de desempenho escolar do aluno medida pela nota do Enade (nota na prova de conhecimento específico do Enade) e pelas medidas da qualidade da instituição de ensino superior (se cursou em uma universidade particular, e o percentual de membros PPI no curso, %_PPI_na_inst); por fim, X_4 um vetor de variáveis do mercado de trabalho, que neste estudo são as variáveis trabalho além das 44h semanais e se o emprego é no setor público, estatutário.

Os resultados da estimação são apresentados na tabela 16. Todas as variáveis foram significativas, exceto as indicadoras dos grupos de saúde, ciências sociais e educação, o que significa que o coeficiente não é diferente de zero e, portanto, não existe diferença entre esses

grupos e o grupo de referência outros. O sinal do impacto de todas as variáveis são na mesma direção do estudo do capítulo 1 e as demais variáveis utilizadas no capítulo anterior que não estão presente nesta estimação são suprimidas por não serem significativas e a amostra não ser tão grande quanto a amostra do capítulo 1.

Tabela 16 - Modelo MQO para a Remuneração futura e_{10}

Numero de alunos	2.628	
F(16,2611)	143,49	
prob > F	-	
R ²	0,47	
R ² ajustado	0,46	
<i>Variáveis</i>	<i>Coef.</i>	<i>Erro Padrão</i>
Mulher	-0,127 ***	0,024
Casado	0,181 ***	0,029
Renda Baixa	-0,236 ***	0,027
Nota no Enade	0,004 ***	0,001
Engenharias	0,197 **	0,079
Ciências	0,115 **	0,053
Saúde	0,124	0,128
Medicina /Direito	0,078 *	0,044
Ciências Sociais	0,009	0,043
Negócios	0,077 **	0,037
Educação	-0,017	0,043
Particular	-0,900 ***	0,069
Particular * idade	0,029 ***	0,002
%_PPI_na_inst	-0,438 ***	0,113
Trabalho +44h/sem	0,537 ***	0,112
Estatutario	0,735 ***	0,030
Constante	7,096 ***	0,013

OBS : *** significante a menos de 1%, **significante a 5% e * significante a 10%. Nessa tabela se apresentam os coeficientes do modelo MQO para a remuneração caso o candidato não conclua o curso.

Para a análise de regressão multinível é utilizado o modelo HLM e o método proposto por Hox (2010), o qual implica em 5 passos de análise: o passo 1 analisa o modelo sem nenhuma variável explicativa, também chamado de modelo nulo ou modelo 0. Essa estimação serve de base de comparação para todos os demais modelos subsequentes e também para o cálculo da correlação intraclasse (ICC), a estimativa de *deviance* e análise da necessidade da técnica.

Os resultados do modelo em três níveis, grupo, instituição e aluno, são apresentados na tabela 17. Verifica-se que o ICC do grupo e o ICC conjunto do grupo e instituição são 1,48% e 9,86% respectivamente. Segundo Lee(2008) quando o ICC for maior que 10% o método multinível deve ser

considerado, uma vez existem diferenças significativas entre os grupos hierárquicos. Entretanto para remuneração futura, caso o candidato não conclua o curso (e_{i0}) não se observa um percentual maior que 10% mesmo no ICC dos dois grupos de maior nível hierárquico conjuntamente. Desta forma optou-se pelo modelo MQO apresentado acima.

Tabela 17 - Modelo - 3 níveis para a Remuneração futura e_{i0}	
Modelo 0	
Numero de alunos	5061
Numero de instituições	154
Numero de grupos	8
Intercepto	R\$ 1.272,39
Log likelihood	-5787,90
Efeito Aleatório	
Variância do Nivel 1 - Aluno	0,5551 (0,011)
Variância do Nivel 2 - Instituição	0,0516 (0,009)
Variância do Nivel 3 - Grupo	0,0091 (0,007)
ICC -grupo	1,48% (0,0118)
ICC instituição grupo	9,86% (0,017)

Obs: erros-padrão entre parênteses. Nesta tabela se apresentam os resultados do modelo HLM em três níveis para a remuneração dos candidatos caso eles não concluem o curso universitário

Após a estimação dos coeficientes do modelo MQO projeta-se a remuneração para o caso do candidato não concluir o curso para cada candidato e para cada opção de escolha de grupos de cursos deste. Para esta projeção observe que as variáveis são todas centradas na média e, portanto, como, no estudo, todos os candidatos concorrem a uma vaga para uma única universidade, UnB, todas as variáveis de nível da instituição possuem valor igual a zero. Dessa forma as variáveis %_PPI_na_inst e Particular são zero. Para as variáveis que não se têm informação, como é o caso do desempenho no curso universitário, nota no Enade e as variáveis de mercado de trabalho, estatutário e trabalhar +44h/sem, assume-se o valor zero o que significa para a projeção o candidato não ser diferente da média dos demais nessas variáveis.

A tabela 18 apresenta um resumo das médias das projeções das remunerações dos candidatos (e_{i0}) para cada opção de grupo. Nas linhas têm-se a real escolha do grupo do candidato e nas colunas todas as opções de grupos de cursos que este poderia ter optado. Observa-se que os

valores apresentados são bem menos dispersos que os apresentados na tabela 15, o que é natural, uma vez que todos os candidatos não possuem o diploma universitário, e, portanto, todos são potenciais trabalhadores com nível médio completo. Entretanto uma análise mais cuidadosa mostra uma leve valorização do mercado para os alunos que optem pelo grupo de engenharia (coluna 2), mesmo que estes não terminem os cursos e uma leve desvalorização dos alunos que optam pelo do grupo de educação (coluna 8). Em outras palavras, fazer um dos cursos do grupo de engenharia, mesmo que não se conclua o curso agrega mais valor ao mercado que fazer um dos cursos do grupo de educação e não conclui-lo.

Tabela 18 - Média das remunerações (e_{i0}) projetadas por grupos de cursos

Média das remunerações (e_{i0}) dos candidatos que escolheram um dos grupos indicados nas linhas para um curso do grupo da coluna

Candidatos ao Grupo (coluna 1)	Engenharias (coluna 2)	Ciências (coluna 3)	Saúde (coluna 4)	Medicina /Direito (coluna 5)	Ciências Sociais (coluna 6)	Negócios (coluna 7)	Educação (coluna 8)
Engenharias	R\$ 1.450,00	R\$ 1.336,60	R\$ 1.348,90	R\$ 1.287,22	R\$ 1.201,43	R\$ 1.287,02	R\$ 1.171,28
Ciências	R\$ 1.402,68	R\$ 1.292,98	R\$ 1.304,88	R\$ 1.245,21	R\$ 1.162,22	R\$ 1.245,02	R\$ 1.133,06
Saúde	R\$ 1.339,61	R\$ 1.234,84	R\$ 1.246,21	R\$ 1.189,22	R\$ 1.109,96	R\$ 1.189,03	R\$ 1.082,11
Medicina /Direito	R\$ 1.404,91	R\$ 1.295,03	R\$ 1.306,96	R\$ 1.247,19	R\$ 1.164,06	R\$ 1.246,99	R\$ 1.134,86
Ciências Sociais	R\$ 1.367,34	R\$ 1.260,40	R\$ 1.272,00	R\$ 1.213,84	R\$ 1.132,93	R\$ 1.213,65	R\$ 1.104,51
Negócios	R\$ 1.391,57	R\$ 1.282,73	R\$ 1.294,54	R\$ 1.235,35	R\$ 1.153,01	R\$ 1.235,15	R\$ 1.124,08
Educação	R\$ 1.334,42	R\$ 1.230,05	R\$ 1.241,38	R\$ 1.184,61	R\$ 1.105,65	R\$ 1.184,43	R\$ 1.077,92

Fonte Rais 2007 - Valores nominais em Reais no ano de 2007.

Obs: nas linhas estão os candidatos agrupados nos grupos que eles realmente optaram por fazer o vestibular e nas colunas estão as alternativas de grupos de grupos de cursos para qual se calcula as médias das remunerações caso não conclua esse curso

3.3.1.4 Remuneração futura esperada

Nesse estudo, se projeta uma remuneração esperada para todos os candidatados a um curso universitário e para todas as opções de grupos de cursos que estes podem escolher. As projeções são individuais e são funções das probabilidades de concluir os cursos e das remunerações caso o candidato conclua o curso e caso ele não conclua o curso. Um resumo das médias das remunerações esperadas individuais para cada opção de grupo de cursos é apresentado na tabela 19.

Tabela 19 - Média das remunerações esperadas (Y) por grupos de cursos

Média das remunerações esperadas (Y) dos candidatos que escolheram um dos grupos indicados nas linhas para um curso do grupo da coluna

Candidatos ao Grupo (coluna 1)	Engenharias (coluna 2)	Ciências (coluna 3)	Saúde (coluna 4)	Medicina /Direito (coluna 5)	Ciências Sociais (coluna 6)	Negócios (coluna 7)	Educação (coluna 8)
Engenharias	R\$ 2.967,64	R\$ 2.381,35	R\$ 1.314,96	R\$ 2.136,31	R\$ 2.020,19	R\$ 2.732,73	R\$ 2.411,57
Ciências	R\$ 2.849,99	R\$ 2.099,59	R\$ 1.289,46	R\$ 1.820,89	R\$ 1.927,28	R\$ 2.557,08	R\$ 2.350,26
Saúde	R\$ 2.761,63	R\$ 1.856,97	R\$ 1.239,67	R\$ 1.530,16	R\$ 1.847,96	R\$ 2.366,49	R\$ 2.275,20
Medicina /Direito	R\$ 3.026,27	R\$ 2.375,70	R\$ 1.271,16	R\$ 2.377,68	R\$ 2.027,37	R\$ 2.617,36	R\$ 2.354,70
Ciências Sociais	R\$ 2.858,35	R\$ 1.974,53	R\$ 1.259,32	R\$ 1.870,05	R\$ 1.897,62	R\$ 2.442,14	R\$ 2.303,84
Negócios	R\$ 2.812,51	R\$ 1.973,70	R\$ 1.279,35	R\$ 1.802,76	R\$ 1.887,00	R\$ 2.495,22	R\$ 2.334,24
Educação	R\$ 2.625,54	R\$ 1.615,41	R\$ 1.252,09	R\$ 1.344,22	R\$ 1.734,17	R\$ 2.251,11	R\$ 2.238,79

Fonte Rais 2007 - Valores nominais em Reais no ano de 2007.

Obs: nas linhas estão os candidatos agrupados nos grupos que eles realmente optaram por fazer o vestibular e nas colunas estão as alternativas de grupos de cursos para qual se calcula as médias das remunerações esperadas para esse grupo

Na linha tem-se a real escolha do grupo do candidato e nas colunas todas as opções de grupos de curso que este poderia ter optado. Comparando com a tabela 15, nota-se que os valores apresentados são menores e menos dispersos. Observa-se também que para os candidatos que optaram por cursos do grupo de engenharias a remuneração esperada para esse grupo passa a ser a maior entre todas as opções. Nota-se também que para os grupos dos candidatos que optaram por ciências sociais, negócios e educação a remuneração do seu próprio grupo é mais atrativa em relação às demais opções de grupos que na tabela 15. No caso dos candidatos que optaram pelo grupo de negócios a média das remunerações esperadas para essa opção de grupo é a segunda maior, apenas atrás da hipótese destes optarem pelo grupo de engenharias e na tabela 15 a opção pelo grupo de negócios era apenas a quinta maior, atrás das opções dos grupos de medicina/direito, engenharias, ciências e educação.

3.3.2 Probabilidade de Passar

Sabe-se das dificuldades de ingressar, devido à forte concorrência, em um curso universitário em instituições de ensino superior públicas no Brasil, dessa forma, surge à ideia que a probabilidade de passar no curso seja relevante para a escolha de carreiras. Um dos objetivos desse estudo é verificar se a probabilidade de passar no vestibular é relevante para a escolha das carreiras dos candidatos.

A probabilidade de passar, nesse estudo, é estimada também por um modelo logit. As variáveis independentes escolhidas estão relacionadas com preparação dos candidatos, a escolaridade dos pais, o gênero e, em especial, a opção do candidato por tentar o ingresso no curso via cota racial. Como variáveis relacionadas à preparação dos candidatos utilizou-se tentativas_UnB, que indica se o candidato já prestou vestibular para a UnB em anos anteriores e portanto já conhece o processo seletivo e o nível de dificuldade da prova, a variável Cursinho, que indica se o candidato cursou cursinho preparatório para o vestibular, as variáveis Publica-EM e Curso_regular que indicam se o candidato estudou a maior parte do ensino médio em escola pública e fez um curso regular, respectivamente e a variável concluinte, que indica se o candidato terminou o ensino médio dentro de um prazo de um ano. Esta variável também pode ser utilizada com uma proxy para a idade do candidato.

O resultado da estimação encontra-se na tabela 20. Observa-se que o modelo é significativo para todos os grupos e a grande maioria das variáveis é significativa para todos os grupos. Vale ressaltar que tentativas_UnB, Cursinho e Curso_regular são significativas para todos os grupos com impacto positivo. Também apresentaram impacto positivo para todos os grupos, as quais são significativas, as variáveis cota e de nível de escolaridade dos pais, enquanto impactam de forma negativa as variáveis que indicam se o candidato é do sexo feminino, se estudou a maior parte do ensino médio em escola pública e se concluiu o ensino médio recentemente dentro de um prazo de um ano.

Após a estimação dos coeficientes do modelo logit, para cada candidato e para cada grupo de cursos, foi projetado uma probabilidade de passar individual para cada opção de grupo de cursos. As projeções são individuais e de acordo com o modelo variam dependendo das variáveis independentes descritas acima. De forma análoga aos resultados das probabilidades projetadas de conclusão de curso, as probabilidades projetadas de passar nos cursos são apresentadas nas tabelas 21 e 22 para cada opção de grupos de cursos. Na tabela 21 apresentam-se por grupo a probabilidade de passar observada, ou seja, de todos os alunos daquele grupo quanto realmente alcançaram a vaga no curso, e a média das probabilidades projetadas para estes alunos pelo o modelo. Porém, exatamente como ocorreu para as probabilidades de conclusão de curso, vale ressaltar que a amostra das probabilidades projetadas é menor, uma vez que nem todos os alunos responderam o Questionário Sociocultural (QSC) e, portanto, não se tem informação sobre todas as características individuais de preparação e familiares destes.

Para essa amostra, observa-se que a probabilidade de passar é em torno de 10% para quase todos os cursos, com destaque para os cursos de medicina e direito, grupo medicina/direito, que possui uma probabilidade de passar bem inferior aos demais grupos, cerca de 2%.

Tabela 20 - Coeficientes do modelo logit para a probabilidade de passar

	Engenharias	Ciências	Saúde	Medicina /Direito	Ciências Sociais	Negócios	Educação	Outros
(coluna 1)	(coluna 2)	(coluna 3)	(coluna 4)	(coluan 5)	(coluna 6)	(coluna 7)	(coluna 8)	(coluna 9)
Numero de alunos	4119	3270	4398	6615	8025	3552	7816	4513
Wald chi2(9)	211,02	151,21	202,23	107,87	248,20	133,34	278,90	200,48
prob > Chi2	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Log pseudolikelihood	-3016,30	1817,67	1705,11	-1395,35	3954,35	-2258,77	-4208,66	2809,51
<i>Variaveis</i>	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
Mulher	-0,415 *** (0,157)	-0,750 *** (0,157)	-0,350 ** (0,162)	-0,745 *** (0,170)	-0,524 *** (0,107)	-0,119 (0,158)	-0,268 *** (0,089)	-0,353 *** (0,105)
Cota	0,403 *** (0,135)	0,357 ** (0,151)	0,218 (0,160)	0,291 (0,202)	0,334 *** (0,109)	0,270 * (0,142)	0,057 (0,092)	0,016 (0,121)
Pai_edu_NS	0,676 *** (0,175)	0,138 (0,159)	0,292 * (0,162)	0,297 (0,195)	0,380 *** (0,122)	0,502 *** (0,159)	0,293 ** (0,127)	0,295 ** (0,129)
Mãe_edu_Ns	-0,128 (0,194)	0,577 *** (0,155)	0,447 *** (0,160)	0,557 *** (0,207)	0,626 *** (0,121)	0,528 *** (0,157)	0,415 *** (0,128)	0,383 *** (0,135)
Tentativas_UnB	1,022 *** (0,125)	0,969 *** (0,185)	1,124 *** (0,191)	1,646 *** (0,289)	1,062 *** (0,133)	1,257 *** (0,177)	1,168 *** (0,126)	0,947 *** (0,147)
Cursinho	0,541 *** (0,124)	0,397 ** (0,161)	0,975 *** (0,178)	0,555 *** (0,202)	0,561 *** (0,111)	0,429 *** (0,137)	0,230 ** (0,091)	0,354 *** (0,106)
Publica_EM	-0,848 *** (0,175)	-0,675 *** (0,156)	-0,787 *** (0,174)	-0,116 (0,241)	-0,190 (0,121)	-0,291 * (0,154)	-0,372 *** (0,101)	-0,344 *** (0,124)
Curso_Regular	0,649 *** (0,208)	0,480 * (0,246)	1,070 *** (0,314)	1,259 *** (0,419)	0,678 *** (0,170)	0,411 ** (0,197)	0,599 *** (0,111)	0,559 *** (0,163)
Concluintes	-0,178 (0,201)	-0,402 ** (0,157)	-0,557 *** (0,145)	-0,319 * (0,182)	-0,491 *** (0,125)	-0,200 (0,162)	-0,494 *** (0,107)	-0,426 *** (0,112)
constante	-3,360 *** (0,264)	-3,224 *** (0,316)	-4,678 *** (0,400)	-6,441 *** (0,583)	-4,177 *** (0,215)	-3,897 *** (0,295)	-3,280 *** (0,181)	-3,134 *** (0,230)

OBS : *** significante a menos de 1%, ** significante a 5% e * significante a 10% e erros-padrão entre parenteses

Tabela 21 - Probabilidade de passar no curso por grupos						
		Obs	Mean	Std. Dev.	Min	Max
Engenharias	Probabilidade de passar observada	8768	10,0%	30,0%	0	1
	Probabilidade de passar projetada	4119	12,8%	9,7%	0,01	0,48
Ciências	Probabilidade de passar observada	6960	8,4%	27,7%	0	1
	Probabilidade de passar projetada	3270	8,9%	7,5%	0,01	0,42
Saúde	Probabilidade de passar observada	8431	5,4%	22,5%	0	1
	Probabilidade de passar projetada	4398	6,2%	6,5%	0,00	0,37
Medicina /Direito	Probabilidade de passar observada	13613	2,0%	14,1%	0	1
	Probabilidade de passar projetada	6615	2,4%	2,6%	0,00	0,14
Ciências Sociais	Probabilidade de passar observada	17628	6,5%	24,6%	0	1
	Probabilidade de passar projetada	8025	7,0%	6,2%	0,00	0,37
Negócios	Probabilidade de passar observada	7735	9,3%	29,0%	0	1
	Probabilidade de passar projetada	3552	10,0%	8,0%	0,01	0,38
Educação	Probabilidade de passar observada	14661	10,1%	30,1%	0	1
	Probabilidade de passar projetada	7816	9,7%	6,8%	0,01	0,37
Outros	Probabilidade de passar observada	8878	10,3%	30,4%	0	1
	Probabilidade de passar projetada	4513	10,8%	6,8%	0,01	0,36

OBS.: Nessa tabela se apresenta a probabilidade de passar observada para cada grupo de cursos, ou seja, de todos os alunos daquele grupo quantos realmente alcançaram o ingresso no curso, e a média das probabilidades projetadas para estes alunos pelo modelo.

A tabela 22 apresenta um resumo das probabilidades individuais projetadas de passar em um grupo de cursos para os candidatos. Para cada candidato é projetada uma probabilidade de passar para todas as alternativas de grupos de cursos, ou seja, como se tem oito grupos (engenharias, ciências, saúde, medicina/direito, ciências sociais, negócios, educação e outros) para cada candidato tem-se oito probabilidades de passar em grupo de cursos. Porém o candidato opta por apenas um dos cursos de um dado grupo para concorrer à vaga na universidade. Dessa forma a tabela 22 agrupa, de maneira análoga a tabela 12, na linha todos os candidatos que estão concorrendo a uma vaga em um determinado grupo de cursos e na coluna apresenta a média das probabilidades de passar desses candidatos para cada opção de grupo de cursos que estes poderiam ter escolhidos, independente da sua escolha.

Tabela 22 - Probabilidade de passar no curso por grupos (projetada)

Média das probabilidades dos candidatos que escolheram os grupos indicados nas linhas para cada grupo alternativo na coluna

Candidatos ao Grupo (coluna 1)	Engenharias (coluna 2)	Ciências (coluna 3)	Saúde (coluna 4)	Medicina /Direito (coluna 5)	Ciências Sociais (coluna 6)	Negócios (coluna 7)	Educação (coluna 8)
Engenharias	12,82%	10,71%	7,11%	2,65%	8,06%	10,71%	12,04%
Ciências	11,13%	8,95%	6,09%	2,23%	6,99%	9,60%	10,82%
Saúde	10,44%	7,35%	6,17%	1,79%	6,24%	9,64%	10,31%
Medicina /Direito	13,21%	10,07%	8,03%	2,44%	8,15%	11,81%	12,73%
Ciências Sociais	11,01%	8,37%	6,34%	2,10%	6,96%	10,30%	11,12%
Negócios	11,23%	9,03%	6,29%	2,28%	7,24%	10,00%	11,32%
Educação	8,44%	6,90%	4,74%	1,78%	5,92%	8,08%	9,68%

Obs: nas linhas estão os candidatos agrupados nos grupos que eles realmente optaram por fazer o vestibular e nas colunas estão os grupos para qual se calcula as médias das probabilidades de passar nele

A tabela 22 mostra que os candidatos a um curso dos grupos de medicina/direito e engenharia são em média os mais preparados para passar em um curso de qualquer grupo, as médias das probabilidades de passar em um curso de um determinado grupo deles são as maiores medias para todas as oito opções de grupos. No caso dos candidatos a um curso de engenharia estes possuem a segunda maior média para passar em um dos cursos do grupo de engenharias (12,82%), seu próprio grupo, a maior média para passar nos grupos de ciências (10,71%) e medicina/direito (2,65%), e a segunda maior media para os demais grupos: saúde (7,11%), ciências sociais (8,06%), negócios (10,71%) e educação (12,04%). Já os candidatos a um dos cursos do grupo medicina/direito possuem a segunda maior média para passar em seu próprio grupo (2,44%) e para o grupo ciências (10,07%) e as maiores médias de probabilidade de passar para os demais grupos: engenharias (13,21%), saúde (8,03%), ciências sociais (8,15%), negócios (11,81%) e educação (12,73%).

Por outro lado, os candidatos a um curso do grupo de educação são os menos preparados e possuem as menores média de probabilidade de passar para qualquer grupo de cursos: 8,44% para passar em um dos curso do grupo de engenharias, 6,90% para o grupo de ciências, 4,74% para o grupo de saúde, 1,78% para o grupo de medicina/direito, 5,92% para o grupo de ciências sociais, 8,08% para o grupo de negócios e 9,68% para seu próprio grupo, educação. Entretanto observe que estes candidatos possuem em média a maior probabilidade de passar para seu próprio grupo, educação (9,68%). Isto também ocorre com os candidatos a um dos cursos de engenharias, dentre todas as opções de cursos possíveis os candidatos a um curso do grupo de engenharias possuem em média a maior chance de passar para o próprio grupo de cursos, engenharias.

Outro resultado importante é que os candidatos a um dos cursos do grupo de medicina/direito têm em média a menor probabilidade de passar exatamente no seu próprio grupo medicina/direito, ou seja, apesar de serem os mais preparados seus cursos são os mais concorridos e com menores chances de passar comparando com as outras possibilidades de grupos de cursos que estes candidatos poderiam optar.

3.3.3 Escolha de carreiras

O resultado principal desse estudo é apresentado na tabela 23, nessa tabela apresentam-se os resultados do *Mixed Logit Model* para a escolha de carreiras. Neste modelo os candidatos optam por um grupo de cursos entre os oito grupos de cursos disponíveis.

As variáveis explicativas que variam entre as alternativas são a remuneração esperada e a probabilidade de passar e as variáveis que não variam com as alternativas são as características individuais (mulher e PPI) as características familiares (pai com nível superior completo) e para representar habilidade, preferências e preparação se escolheu a variável *Particular_EM*, que indica se o candidato estudou a maior parte do ensino médio em escolas particulares. Opta-se por escolher ao menos uma variável de cada grupo, para evitar que as mesmas variáveis sejam utilizadas no modelo e nas construções das variáveis remuneração esperada e probabilidade de passar. Esta opção de não escolher todas as variáveis explicativas disponíveis para o modelo principal, assim como foi a opção dos autores Montmarquette(2002), Casari(2006) e Costa da Silva e Silveira Neto(2015), também é motivada pelo desejo de comparar este modelo com o *Random Parameters Logit Model* e este último, por sua vez, exige para cada variável explicativa que não varia com as alternativas do *Mixed Logit Model* sete variáveis no modelo *Random Parameters Logit Model*. Entretanto, no apêndice, se apresenta uma tabela do *Mixed Logit Model* com todas as variáveis explicativas (tabela A10 dos apêndices).

Observa-se na tabela 23 que o modelo é considerado globalmente válido, uma vez que o p-valor do teste de validade global do modelo é altamente significativo ($\text{Prob} > \chi^2 = 0$). As variáveis explicativas remuneração esperada e probabilidade de passar são positivas e extremamente significativas com p-valor igual a zero e as variáveis de características do candidato estão todas em relação ao grupo de cursos educação que foi utilizado como referência. Nota-se que todas as variáveis são significativas, exceto a variável PPI nos grupos ciências e negócios, o que significa que entre esses grupos e o grupo de educação não se observa diferenças para essa variável. Observa-se que para grupo de engenharias ser do gênero feminino (mulher) e do grupo PPI diminui as chances de escolha desse grupo em relação ao grupo de educação. Enquanto o pai possui o nível superior completo e o candidato ter estudado em uma escola particular no ensino médio aumenta a probabilidade de escolha do grupo. O mesmo se observa para o grupo de negócios e ciências. Entretanto para os grupos de saúde, medicina/direito, ciências sociais e outros ser do sexo feminino aumenta a probabilidade de escolher esses grupos em relação ao grupo de educação.

Tabela 23 - Resultado <i>Mixed Logit Model</i>				
Número de obsevações		260.880		
Número de casos		32.610		
Wald chi2(30)		4720,68		
Prob > chi2		-		
Log pseudolikelihood		-127714		
		Coeficiente	Erro Padrão	P - valor
	Remuneração esperada	0,00045	0,00002	-
	Probabilidade de Passar	3,186	0,361	-
Grupo	Variável	Coeficiente	Erro Padrão	P - valor
Engenharias	mulher	-1,458	0,056	-
	PPI	-0,238	0,057	-
	particular_EM	1,104	0,065	-
	pai_edu_NS	0,486	0,066	-
	_cons	-0,707	0,052	-
Ciências	mulher	-0,369	0,064	-
	PPI	-0,071	0,058	0,221
	particular_EM	0,735	0,067	-
	pai_edu_NS	0,396	0,073	-
	_cons	-0,705	0,054	-
Saúde	mulher	0,938	0,064	-
	PPI	-0,151	0,057	0,008
	particular_EM	1,131	0,062	-
	pai_edu_NS	0,538	0,070	-
	_cons	-1,109	0,071	-
Medicina /Direito	mulher	0,375	0,050	-
	PPI	-0,249	0,050	-
	particular_EM	1,471	0,058	-
	pai_edu_NS	0,756	0,058	-
	_cons	-0,681	0,058	-
Ciências Sociais	mulher	0,323	0,047	-
	PPI	-0,144	0,048	0,003
	particular_EM	0,958	0,056	-
	pai_edu_NS	0,696	0,058	-
	_cons	-0,194	0,048	-
Negócios	mulher	-0,346	0,056	-
	PPI	-0,034	0,055	0,541
	particular_EM	0,667	0,063	-
	pai_edu_NS	0,499	0,067	-
	_cons	-0,885	0,054	-
Educação		base alternativa		
Outros	mulher	0,353	0,049	-
	PPI	-0,098	0,049	0,048
	particular_EM	0,308	0,060	-
	pai_edu_NS	0,476	0,063	-
	_cons	-0,791	0,050	-

Obs: Nessa tabela apresentam-se os resultados do Mixed Logit Model para a escolha de carreiras

Na tabela 24 se apresentam os efeitos marginais sobre a probabilidade de escolha. No primeiro bloco os efeitos marginais de um aumento na remuneração esperada, no segundo bloco de um aumento na probabilidade de passar no curso e no terceiro e ultimo bloco os efeitos de um aumento nas variáveis de características individuais, familiares e de preparação, habilidade e preferências.

Tabela 24 - Efeitos marginais sobre a probabilidade de escolha (continua)

Grupo	Engenharias	Ciências	Saúde	Medicina /Direito	Ciências Sociais	Negócios	Educação	Outros
Probabilidade de escolha do grupo	8,81%	8,61%	9,45%	15,34%	22,47%	9,52%	15,36%	10,44%
Efeitos marginais de um aumento no Redimento Esperado (R\$100,00) de um dado grupo nos demais								
Engenharias	0,0036	-0,0003	-0,0004	-0,0006	-0,0009	-0,0004	-0,0006	0,0000
Ciências	-0,0003	0,0035	-0,0004	-0,0006	-0,0009	-0,0004	-0,0006	0,0000
Saúde	-0,0004	-0,0004	0,0038	-0,0007	-0,0010	-0,0004	-0,0007	0,0000
Medicina /Direito	-0,0006	-0,0006	-0,0007	0,0058	-0,0015	-0,0007	-0,0010	0,0000
Ciências Sociais	-0,0009	-0,0009	-0,0010	-0,0015	0,0078	-0,0010	-0,0015	0,0000
Negócios	-0,0004	-0,0004	-0,0004	-0,0007	-0,0010	0,0038	-0,0007	0,0000
Educação	-0,0006	-0,0006	-0,0007	-0,0010	-0,0015	-0,0007	0,0058	0,0000
Outros	-0,0004	-0,0004	-0,0004	-0,0007	-0,0010	-0,0004	-0,0007	0,0002
Efeitos marginais de um aumento na Probabilidade de Passar (1 ponto percentual) de um dado grupo nos demais								
Engenharias	0,256	-0,024	-0,027	-0,043	-0,063	-0,027	-0,043	0,003
Ciências	-0,024	0,251	-0,026	-0,042	-0,062	-0,026	-0,042	0,003
Saúde	-0,027	-0,026	0,273	-0,046	-0,068	-0,029	-0,046	0,004
Medicina /Direito	-0,043	-0,042	-0,046	0,414	-0,110	-0,047	-0,075	0,006
Ciências Sociais	-0,063	-0,062	-0,068	-0,110	0,555	-0,068	-0,110	0,009
Negócios	-0,027	-0,026	-0,029	-0,047	-0,068	0,274	-0,047	0,004
Educação	-0,043	-0,042	-0,046	-0,075	-0,110	-0,047	0,414	0,006
Outros	-0,029	-0,029	-0,031	-0,051	-0,075	-0,032	-0,051	0,034

Tabela 24 - Efeitos marginais sobre a probabilidade de escolha (conclusão)								
Grupo	Engenharias	Ciências	Saúde	Medicina /Direito	Ciências Sociais	Negócios	Educação	Outros
Probabilidade de escolha do grupo	8,81%	8,61%	9,45%	15,34%	22,47%	9,52%	15,36%	10,44%
Efeitos marginais de um aumento nas variáveis de características dos candidatos								
Mulher	-0,14	-0,03	0,08	0,05	0,06	-0,04	-0,01	0,03
PPI	-0,01	0,00	0,00	-0,02	0,00	0,01	0,02	0,00
Ensino médio em escola particular	0,03	0,00	0,03	0,10	0,04	-0,01	-0,13	-0,05
Pai com nível superior completo	0,00	-0,01	0,00	0,04	0,04	0,00	-0,07	0,00

Obs: Nessa tabela se apresentam os efeitos marginais sobre a probabilidade de escolha. No primeiro bloco os efeitos marginais de um aumento de R\$ 100,00 na remuneração esperada de um dado grupo na probabilidade de escolha de todos os grupos, no segundo bloco de um aumento de 1 ponto percentual na probabilidade de passar de um dado grupo na probabilidade de escolha de todos os grupos e no terceiro e último bloco os efeitos de um aumento de 1 ponto nas variáveis *dummies* de características individuais, familiares e de preparação, habilidade e preferências.

No primeiro bloco da tabela 24 têm-se os impactos de um aumento de R\$100,00 na remuneração esperada de um dado grupo na probabilidade de escolha de todos os grupos. Na diagonal principal em negrito têm-se o impacto do aumento de R\$100,00 na remuneração de um grupo na probabilidade de escolha deste e fora da diagonal principal os impactos nas probabilidades de escolha dos demais grupos. Desta forma tem-se na primeira linha o impacto da elevação de R\$100,00 (cem reais) na remuneração esperada do grupo de engenharias gera um aumento de 0,0036 ponto percentual ou de 4% ($100 \times 0,000036 / 0,0881$) na probabilidade de escolha desse grupo e uma diminuição de 0,0009 ponto percentual ou 0,4% ($100 \times 0,000009 / 0,2247$) na probabilidade de escolha do grupo ciências sociais. O grupo mais sensível a um aumento na sua remuneração esperada em pontos percentuais é o grupo de ciências sociais, uma elevação de R\$100,00 (cem reais) na remuneração do grupo de Ciências Sociais gera um aumento de 0,0078 ponto percentual na probabilidade de escolha desse grupo, entretanto, em termos percentuais de impacto na probabilidade de escolha, todos os grupos possuem praticamente o mesmo impacto, cerca de 3,5% a 4% de aumento na probabilidade de escolha do grupo devido a um aumento de R\$100,00 na remuneração deste grupo. No caso do grupo de ciências sociais tem-se o aumento de 3,5% na probabilidade escolha do grupo para um aumento de R\$100,00 na sua remuneração esperada. Este resultado está de acordo com os alcançados por Montmarquette (2002), Bartalotti e Menezes(2007) e Costa de Silva e Silveira Neto(2015). Tais autores também obtiveram em seus estudos impactos positivos e significativos do retorno esperado nas probabilidades de escolha dos cursos.

No segundo bloco da mesma tabela têm-se os impactos de um aumento de 1 ponto percentual na probabilidade de passar de um dado grupo na probabilidade de escolha de todos os grupos. Na diagonal principal em negrito têm-se o impacto do aumento de 1 ponto percentual na probabilidade de passar em um grupo na probabilidade de escolha deste e fora da diagonal principal

os impactos nas probabilidades de escolha dos demais grupos. Desta forma na primeira linha observa-se o impacto da elevação de 1 ponto percentual na probabilidade de passar para um dos curso do grupo de engenharias gera um aumento de 0,256 ponto percentual na probabilidade de escolha deste grupo e uma diminuição de 0,063 ponto percentual a probabilidade de escolha do grupo ciências sociais. O grupo mais sensível a um aumento na probabilidade passar, em termos percentuais, é o grupo de ciências sociais, uma elevação de 1 ponto percentual na probabilidade de passar em um dos cursos do grupo de ciências sociais gera um aumento de 0,555 ponto percentual na probabilidade de escolha desse grupo. Em termos percentuais de impacto na probabilidade de escolha, todos os grupos possuem praticamente o mesmo impacto, cerca de 3% de aumento na probabilidade de escolha do grupo devido a um aumento de 1 ponto percentual na probabilidade de passar deste grupo.

No terceiro bloco se apresentam os efeitos marginais para impactos das características individuais, familiares e preparação e habilidades dos candidatos sobre a probabilidade de escolha dos grupos de cursos. Nas colunas estão os grupos de cursos e nas linhas as variáveis dummies explicativas. Observa-se que o candidato ser do sexo feminino reduz em 0,14 ponto percentual a probabilidade de este escolher o grupo de engenharias e aumenta em 0,08 ponto percentual a probabilidade do candidato escolher um dos cursos do grupo de saúde. O impacto do candidato ser do sexo feminino na probabilidade de escolha é positivo para os grupos de saúde, medicina/direito e ciências sociais e negativo para os grupos de cursos mais quantitativos engenharias, ciências e negócios. Para educação o impacto foi nulo. O impacto racial, do candidato ser PPI, foi diferente de zero para alguns grupos. O grupo de medicina direito é preterido pelos que se autodeclararam preto, pardo ou indígena. Ser PPI reduz a probabilidade de escolher o grupo medicina/direito em 0,02 ponto percentual. Por outro lado o candidato ser PPI aumenta em 0,02 ponto percentual a probabilidade de escolha de um dos cursos do grupo de educação. Os candidatos que cursaram o ensino médio em escolas particulares aumentam em 0,10 ponto percentual a probabilidade de escolha do grupo de medicina/direito enquanto reduz em 0,13 pontos percentual a probabilidade de escolha do grupo de educação. Por fim os candidatos com pai com nível superior completo aumenta a probabilidade de escolher o grupo de medicina/direito e ciências sociais em 0,04 ponto percentual, enquanto reduz a probabilidade de escolha do grupo de educação em 0,07 ponto percentual.

Apesar dos grupos cursos serem divididos um pouco diferentes o resultado para o impacto do gênero na probabilidade de escolha de carreiras foi muito semelhante do apresentado por Costa de Silva e Silveira Neto(2015) para os candidatos a uma vaga na Universidade Federal de Pernambuco. Este verificou que os cursos quantitativos também são preteridos pelos candidatos, diminui a probabilidade de escolha do curso, se estes são do sexo feminino, tais como engenharias e ciências, que a raça, diferente do resultado obtido nesse estudo, só foi relevante para o grupo de artes e comunicações e, de maneira semelhante aos candidatos a uma vaga na UnB, os candidatos a uma vaga na UFPE preferem os cursos de saúde se cursaram o ensino médio em escolhas particulares e preterem os cursos do grupo de educação se cursaram o ensino médio em escolas particulares e se o pai possui curso superior completo.

Adicionalmente ao modelo apresentado acima, também foi estimado o *Random Parameters Logit Model*, cujos resultados são apresentados na tabela 25.

Tabela 25 - Resultado <i>Random Parameters Logit Model</i>				
Número de obsevações	266.665			
Wald chi2(37)	7.819			
Prob > chi2	-			
Log pseudolikelihood	-130848			
		Coeficiente	Erro Padrão	P - valor
Remuneração esperada		0,00045	0,00002	-
Probabilidade de Passar (PP)		3,297	0,399	-
Desvio Padrão da PP		1,407	1,928	0,465
Grupo	Variável	Coeficiente	Erro Padrão	P - valor
Engenharias	mulher	-1,451	0,055	-
	PPI	-0,253	0,055	-
	particular_EM	1,092	0,063	-
	pai_edu_NS	0,499	0,065	-
	_cons	-0,734	0,051	-
Ciências	mulher	-0,358	0,064	-
	PPI	-0,071	0,057	0,21
	particular_EM	0,721	0,066	-
	pai_edu_NS	0,404	0,072	-
	_cons	-0,714	0,054	-
Saúde	mulher	0,942	0,064	-
	PPI	-0,152	0,056	0,01
	particular_EM	1,122	0,061	-
	pai_edu_NS	0,548	0,069	-
	_cons	-1,112	0,071	-
Medicina /Direito	mulher	0,386	0,050	-
	PPI	-0,252	0,049	-
	particular_EM	1,460	0,057	-
	pai_edu_NS	0,761	0,058	-
	_cons	-0,688	0,057	-
Ciências Sociais	mulher	0,335	0,046	-
	PPI	-0,146	0,046	0,002
	particular_EM	0,951	0,055	-
	pai_edu_NS	0,708	0,056	-
	_cons	-0,215	0,047	-
Negócios	mulher	-0,365	0,054	-
	PPI	-0,023	0,053	0,66
	particular_EM	0,660	0,061	-
	pai_edu_NS	0,489	0,066	-
	_cons	-0,881	0,052	-
Educação		base alternativa		
Outros	mulher	0,361	0,047	-
	PPI	-0,103	0,048	0,03
	particular_EM	0,296	0,058	-
	pai_edu_NS	0,492	0,061	-
	_cons	-0,803	0,049	-

Obs: Nessa tabela apresentam-se os resultados do *Random Parameters Logit Model* para a escolha de carreiras.

Nesse modelo se considerou o parâmetro da probabilidade de passar aleatório com distribuição normal a fim de induzir correlação através das alternativas. Observa-se que o modelo considerado globalmente válido, uma vez que o p-valor do teste de validade global do modelo é altamente significativo ($\text{Prob} > \chi^2 = 0$). Entretanto o resultado da probabilidade de passar apesar de ser significativo possui o parâmetro do seu desvio padrão não significativo (p-valor = 0,465). Nota-se também que os demais resultados são praticamente idênticos aos apresentados na tabela 23, onde o parâmetro da probabilidade de passar é fixo. Desta forma se conclui que a aleatoriedade do parâmetro da probabilidade de passar não é relevante neste estudo.

3.4 Sumário e conclusões

Este estudo se propõe a avaliar como os estudantes escolhem suas carreiras, quais os fatores relevantes: retorno esperado? Probabilidade de ingressar no curso? Habilidades, preferências, preparação, influencia familiar? Existem diferenças na opção dos estudantes devido ao sexo e raça dos estudantes?

Primeiramente, reconhece-se a limitação do estudo para o universo restrito dos candidatos a cursos da Universidade de Brasília, uma vez que a base de dados é composta apenas por estudantes que se candidataram a uma vaga na UnB. Entretanto, as evidências encontradas no estudo estão de acordo com grande parte da literatura internacional, tais como Berger(1988), Montmarquette et al(2002), Arcidiacomo(2004) e Arcidiacomo(2010) e com a literatura nacional, como Bartalotti e Menezes-filho(2007) e Costa da Silva e Silveira Neto(2015).

Os resultados obtidos confirmam que o retorno esperado é importante na decisão de escolha do curso dos estudantes, uma aumento de cerca de R\$100,00 na remuneração esperada de um grupo gera um aumento na probabilidade de escolha desse grupo de aproximadamente 3,5% a 4% para todos os grupos. A probabilidade de passar no curso se apresentou igualmente importante na escolha do curso, a variável se apresentou extremamente significativa e com um impacto de aproximadamente de 3% no aumento da probabilidade de escolha do grupo para um aumento de 1 ponto percentual na probabilidade de passar desse grupo.

As variáveis de habilidade, preferencias e preparação foram significativas para a construção da probabilidade de passar, como também o nível de escolaridade dos pais e o gênero do candidato, entretanto a variável cota, que foi significativa par o cálculo da probabilidade de passar, merece um destaque. Esta variável apresentou o coeficiente positivo para todos os grupos sendo significativa para quatro deles (engenharias, ciências, ciências sociais e negócios) e, portanto, como altera a probabilidade de passar dos candidatos pode ser de fato uma ferramenta importante para políticas de inclusão na UnB.

O estudo também apresentou que os candidatos pretos, pardos ou indígenas preferem o grupo de educação, quando se compara esse grupo com os demais. Quanto às mulheres, o estudo apresentou que elas preferem os grupos de saúde, medicina/direito e ciências sociais ao grupo de educação e este é preferido aos grupos de engenharias, ciências e negócios. Acredito que esse estudo possa abrir caminho para se estudar políticas de inclusão na UnB, ter uma ideia de como os

estudantes escolhem seus cursos pela suas características socioeconômicas, habilidades, preferências, preparação, gênero, família, probabilidade de passar e remuneração esperada pode auxiliar em descobrir o melhor mecanismo para influenciar a escolha dos candidatos a fim de minimizar as distorções existentes na sociedade.

4 CONCLUSÃO

O presente estudo, dividido em duas partes, investigou: a) três fontes de influências qualitativas sobre os ganhos iniciais dos recém-formados universitários, a saber: 1) o curso universitário; 2) a qualidade da Instituição de Ensino Superior (IES); e, 3) o desempenho do formado na IES.

Abrangendo todas as regiões do Brasil, os resultados demonstram que as três fontes influenciam as remunerações dos recém-graduados. Os cursos universitários possuem maior efeito na remuneração e, quando os grupos de cursos são analisados separadamente, é possível observar que os impactos do desempenho acadêmico e da qualidade da instituição diferem substancialmente. Por exemplo, quando se observa o impacto na remuneração dos recém-formados da variável indicadora de uma instituição de nível superior ser particular para cada grupo de cursos separadamente, nota-se que o impacto, no caso negativo, é maior para o grupo de cursos de engenharias; e quando se observa o impacto na remuneração dos recém-formados da variável conceito da instituição, o impacto, no caso positivo, é maior para o grupo de cursos de ciências. Tal fato aponta que os impactos das variáveis na remuneração dos graduados são diferentes entre os grupos de cursos. No caso do impacto da variável particular na remuneração, o mercado valoriza mais os formados de engenharias de instituições públicas em relação às instituições particulares que os formados de outro grupo de cursos de instituições públicas em relação às instituições particulares. Os resultados apresentados também indicam que o mercado valoriza de forma diferente um recém-formado do sexo feminino e um recém-formado do sexo masculino, mesmo que ambos possuam diploma do mesmo curso e da mesma IES. Ser do sexo feminino implica em um impacto negativo na remuneração da recém-graduada, independente do diploma alcançado, em relação aos seus pares do sexo masculino. O mesmo ocorre quando se compara os recém-formados que se autodeclararam Preto, Pardo ou Indígena (PPI) com as demais (brancos e asiáticos). Assim, é possível observar que o mercado valoriza de forma diferente os profissionais por gênero e raça, o que justificaria uma possível implementação de políticas afirmativas, a fim de minimizar tais distorções.

A pesquisa também investiga: b) a escolha de carreiras para os candidatos a uma vaga na Universidade de Brasília (UnB). Os resultados obtidos confirmam que o retorno esperado é importante na decisão de escolha do curso dos estudantes. Verifica-se que um aumento de cerca de R\$100,00 (cem reais) na remuneração esperada de um grupo gera um aumento na probabilidade de escolha desse grupo de aproximadamente 3,5% a 4% – fato de ocorrência em todos os grupos. Igualmente importante para a decisão de escolha, os resultados apresentam a probabilidade de aprovação para o ingresso no curso. Observa-se que para cada aumento de 1 ponto percentual na probabilidade de aprovação para o ingresso no curso de um dado grupo tem-se um impacto de

aproximadamente 3% no aumento da probabilidade de escolha do grupo a qual pertence o curso. Isto ocorre para todos os grupos de cursos.

O estudo também indica que as variáveis de habilidade, preferências e preparação foram significativas para a construção da probabilidade de aprovação para o ingresso no curso superior, bem como o nível de escolaridade dos pais e o gênero do candidato, além da cota racial. Esta última merece destaque, uma vez que é a ferramenta utilizada com mais frequência nas políticas de inclusão do governo. A cota racial apresenta o coeficiente positivo para todos os grupos, sendo significativa para quatro deles (engenharias, ciências, ciências sociais e negócios).

In conclusio, esta parte da pesquisa aponta que as políticas de inclusão na UnB são mais eficientes nos quatro grupos supramencionados que nos demais.

5 REFERÊNCIAS

- [1] Albright, Jeremy J. and Marinova, Dani M. (2010). "Estimating Multilevel Models Using SPSS, STATA, SAS and R" [PDF document].
Retrieved from <http://www.indiana.edu/~statmath/stat/all/hlm/hlm.pdf>.
- [2] Arcidiacono P. 2004. Ability sorting and the returns to college major. *J. Econom.* 121(1-2):343_75
- [3] Arcidiacono P, Hotz VJ, Kang S. 2010. Modeling college major choices using elicited measures of expectations and counterfactuals. *IZA Disc. Pap.* 4738
- [4] Altonji JG, Blom E, Meghir C. 2012. Heterogeneity in human capital investments: High school curriculum, college major, and careers. *NBER Working Paper No. 17985*, april 2012.
- [5] Bartalotti, O.; Menezes-Filho, N. A relação entre o desempenho da carreira no mercado de trabalho e a escola profissional dos jovens. *Revista Economia Aplicada*, v. 11 n. 4, 2007.
- [6] Berger MC. 1988. Predicted future earnings and choice of college major. *Ind. & Labor Relat. Rev.* 41(3):418_29
- [7] Betts JR. 1996. What do students know about wages? Evidence from a survey of undergraduates. *J. Hum. Resour.* 31(1):27_56
- [8] Cameron, A. C.; Trivedi, P. K. *Microeconometrics: Methods and applications*. Cambridge University Press, New York, 2005
- [9] Casari P 2006. Retorno esperado e escolha profissional: Fatores associados à escolha da carreira dos alunos da Universidade de São Paulo. *Dissertação de mestrado – USP – Ribeirão Preto*
- [10] Costa da Silva, D. F.; Silveira Neto, R. M. Escolhas de carreiras universitárias e mercado de trabalho: Um análise da influência dos incentivos econômicos. *Nova Economia Belo Horizonte* 25[3] 519-552, setembro-dezembro 2015.
- [11] Eide E, Waehrer G. 1998. The role of the option value of college attendance in college major choice. *Econ. Educ. Rev.* 17(1):73_82
- [12] Francis, Andrew M. e Maria Tannuri-Pianto. 2010. Uma Avaliação do Sistema de Cotas da Unb. In: Sampaio de Sousa, Coelho, Versiani e Tannuri-Pianto *Economia Pública Brasileira*, ESAF, Brasília-DF, p. 329-366.

- [13] Francis, Andrew M. e Maria Tannuri-Pianto. 2012a. "Using Brazil's Racial Continuum to Examine the Short-Term Effects of Affirmative Action in Higher Education." *Journal of Human Resources*, 47(3): 754-784
- [14] Francis, Andrew M. e Maria Tannuri-Pianto. 2012b. "The Redistributive Equity of Affirmative Action: Exploring the Role of Race, Socioeconomic Status, and Gender in College Admissions." *Economics of Education Review*, 31(1): 45-55.
- [15] GWARTNEY, J.D. and LONG, J.E. (1978) The relative earnings of blacks and other minorities. *Znd.Labor Ret. Rev.* 31,336-346.
- [16] Hox, Joop J. (2010) *Multilevel Analysis: Techniques and applications* (2nd ed.) New York: Routledge.
- [17] James,E., ALSALAM, N., CONATY, J.C. and To, D.-L. (1989) College quality and future earnings: where should YOU send your child to college? *Am. Econ. Rev.* 79, 247-252.
- [18] KARABEL, J. and ASTIN, A.W. (1975) Social class, academic ability, and college "quality". *Soc. Forces* 53, 381-398.
- [19] KARABEL, J. and MCCLELLAND, K. (1987) Occupational advantage and the impact of college rank on labor market outcomes. *Soc. Inquiry* 57, 323-347.
- [20] Lee, V. L. (2008). Utilização de modelos lineares hierárquicos para estudar contextos sociais: o caso dos efeitos da escola. In N. Brooke e J. F. Soares (Eds), *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 273-296). Belo Horizonte: Editora UFMG.
- [21] MEISENHEIMERJ., R. (1990) Black college graduates in the labor market, 1979 and 1989. *Monthly LaborRev.*, November 1990, 13-21.
- [22] Montmarquette C, Cannings K, Mahseredjian S. 2002. How do young people choose college majors? *Econ. Educ. Rev.* 21(6):543_56
- [23] MUELLER, R.O. (1988) The impact of college selectivity on income for men and women. *Res. Higher Educ.* 29, 175-191.
- [24] MURPHY, K. and WELCH, F. (1989) Wage premiums for college graduates: recent growth and possible explanations. *Educ. Res.* 18, 17-26.

- [25] RAUDENBUSH, S.W. (1988) Applications of hierarchical linear models in educational research: a review. *J. Educ. Stat.* 13, 85-116.
- [26] Rumberger, R. W., & Thomas, S. L. (1993). The economic returns to college major, quality and performance: A multilevel analysis of recent graduates. *Economics of Education Review*, 12, 1–19.
- [27] Robles V, Krishna K. 2012. Affirmative action in higher education in India: targeting, catch up and mismatch. NBER Working Paper No. 17727, January 2012
- [28] Saks R and Shore S. 2005. Risk and Career Choice. *Advances in economic analysis & policy*. Vol 5, issue 1, article 7
- [29] SMART, J.C. (1988) College influences on graduates' income levels. *Res. Higher Educ.* 29, 41-59.
- [30] SOLMON, L.C. (1973) The definition and impact of college quality. In *Does College Matter?* (Edited by SOLMON, L.C. and TAIJBMAN, PI). New York: Academic Press.
- [31] SOLMON, L.C. (1975) The definition of college quality and its impact on earnings. *Expior. &on. Res.* 2,537-588.
- [32] SOLNICK, L. M. (1990) Black college attendance and job success of black college graduates. *Econ. Educ. Rev.* 9, 135-148.
- [33] Stinebrickner TR, Stinebrickner R. 2011. Math or science? Using longitudinal expectations data to examine the process of choosing a college major. NBER working paper 16869
- [34] TIENDA. M. and La, D.-T. (1987) Minority concentration and earnings inequality: blacks, Hispanics, and Asians compared. *Am. J. Social.* 93, 141-165.
- [35] TRUSHEIM, D. and CROUSE, J. (1981) Effects of college prestige on men's occupational status and income. *Res. Higher Educ.* 14, 283-304.
- [36] Zafar B. 2009. How do college students form expectations? Sta_ Report 378, Federal Reserve Bank, New York, NY
- [37] Zafar B. 2009. College major choice and the gender gap. Sta_ Report 364, Federal Reserve Bank, New York, NY

6 ANEXOS

Anexo 1 - Questionário Socioeconômico respondido pelos alunos que fizeram o Enade.

01. Qual o seu estado civil?

- (A) Solteiro(a).
- (B) Casado(a).
- (C) Separado(a)/desquitado(a)/divorciado(a).
- (D) Viúvo(a).
- (E) Outro.

02. Quantos irmãos você tem?

- (A) Nenhum.
- (B) Um.
- (C) Dois.
- (D) Três.
- (E) Quatro ou mais.

03. Quantos filhos você tem?

- (A) Nenhum.
- (B) Um.
- (C) Dois.
- (D) Três.
- (E) Quatro ou mais.

04. Como você se considera?

- (A) Branco(a).
- (B) Negro(a).
- (C) Pardo(a) / mulato(a).
- (D) Amarelo(a) (de origem oriental).
- (E) Indígena ou de origem indígena.

05. Com quem você mora atualmente?

- (A) Com os pais e/ou outros parentes.
- (B) Com esposo(a) e/ou filho(s).
- (C) Com amigos (compartilhando despesas ou de favor).
- (D) Com colegas em alojamento universitário.
- (E) Sozinho(a).

06. Qual a faixa de renda mensal da sua família?

- (A) Até 3 salários-mínimos.
- (B) De 3 a 10 salários-mínimos.
- (C) De 11 a 20 salários-mínimos.
- (D) De 21 a 30 salários-mínimos.
- (E) Mais de 30 salários-mínimos.

07. Quantos membros de sua família moram com você?

- (A) Nenhum.
- (B) Um ou dois.
- (C) Três ou quatro.
- (D) Cinco ou seis.
- (E) Mais do que seis.

08. Assinale a situação abaixo que melhor descreve seu caso.

- (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.

9. Se você trabalha ou trabalhou, qual é ou foi a carga horária aproximada de sua atividade remunerada (não contar estágio e bolsas de pesquisa)?

- (A) Não exerço / não exerci atividade remunerada.
- (B) Trabalho / trabalhei eventualmente.
- (C) Trabalho / trabalhei até 20 horas semanais.
- (D) Trabalho / trabalhei mais de 20 horas semanais e menos de 40 horas semanais.
- (E) Trabalho / trabalhei em tempo integral – 40 horas semanais ou mais.

10. Que tipo de bolsa de estudos ou financiamento você recebe ou recebeu para custeio das despesas do curso?

- (A) Financiamento Estudantil – FIES.
- (B) Bolsa integral ou parcial (inclusive descontos em mensalidades) oferecida pela própria instituição.

- (C) Bolsa integral ou parcial oferecida por entidades externas.
- (D) Outro(s).
- (E) Nenhum.

11. Qual o grau de escolaridade do seu pai?

- (A) Nenhuma escolaridade.
- (B) Ensino Fundamental: de 1ª a 4ª série.
- (C) Ensino Fundamental: de 5ª a 8ª série.
- (D) Ensino Médio.
- (E) Superior.

12. Qual o grau de escolaridade de sua mãe?

- (A) Nenhuma escolaridade.
- (B) Ensino Fundamental: de 1ª a 4ª série.
- (C) Ensino Fundamental: de 5ª a 8ª série.
- (D) Ensino Médio.
- (E) Superior.

13. Em que tipo de escola você cursou o ensino médio?

- (A) Todo em escola pública.
- (B) Todo em escola privada (particular).
- (C) A maior parte do tempo em escola pública.
- (D) A maior parte do tempo em escola privada (particular).
- (E) Metade em escola pública e metade em escola privada (particular).

14. Que tipo de curso de ensino médio você concluiu?

- (A) Comum ou de educação geral, no ensino regular.
- (B) Profissionalizante técnico (eletrônica, contabilidade, agrícola, etc.), no ensino regular.
- (C) Profissionalizante de magistério de 1ª a 4ª série (Curso Normal), no ensino regular.
- (D) Supletivo.
- (E) Outro cursos

Anexo 2 - Questionário Sociocultural respondido pelos candidatos a uma vaga na UnB

01 - Qual é o seu estado civil? (TAMBÉM PAS)

- Solteiro.....1
- Casado.....2
- Outros.....3

02- Qual é o tipo de residência de sua família? (TAMBÉM PAS)

- Própia, não-quitada.....1
- Própia, quitada.....2
- Alugada.....3
- Funcional.....4
- Outra situação..... 5

03 - Qual é a renda mensal de sua família? (TAMBÉM PAS)

- Até R\$ 250,00.....1
- De R\$ 250,01 até R\$ 500,00.....2
- De R\$ 500,01 até R\$ 750,00.....3
- De R\$ 750,01 até R\$ 1.500,00...4
- De R\$ 1.500,01 até R\$ 2.500,00..5
- De R\$ 2500,01 até R\$ 3.500,00...6
- De R\$ 3.500,01 até R\$ 5.000,00... 7
- Mais de R\$ 5.000,00.....8
- Não sei.....9

04 - Qual é o nível de escolaridade de seu pai? (TAMBÉM PAS)

- Analfabeto.....1
- Primeiro grau incompleto.....2
- Primeiro grau completo.....3
- Segundo grau.....4
- Superior.....5
- Não sei.....6

05 - Qual é o nível de escolaridade de sua mãe? (TAMBÉM PAS)

- Analfabeto.....1
- Primeiro grau incompleto..... 2
- Primeiro grau completo..... 3
- Segundo grau..... 4
- Superior..... 5
- Não sei..... 6

06 - Qual a sua raça/cor? (TAMBÉM PAS)

- Branca..... 1
- Preta.....2
- Amarela..... 3
- Parda..... 4
- Indígena..... 5
- Sem declaração... 6

07- Você se declara negro? (TAMBÉM PAS)

- Sim..... 1
- Não..... 2

08 - Você exerce alguma atividade remunerada? (TAMBÉM PAS)

- Não.....1
- Sim, em tempo parcial (até 30 horas semanais).....2
- Sim, em tempo integral (mais de 30 horas semanais)3
- Sim, mas trata-se de trabalho eventual.....4

09 - Em que tipo de estabelecimento de ensino você fez o curso médio (2º grau)?

- Em escola do exterior..... 1
- Todo em escola pública, fora do Distrito Federal..... 2
- Maior parte em escola pública, fora do Distrito Federal... 3
- Todo em escola particular, fora do Distrito Federal..... 4
- Maior parte em escola particular, fora do Distrito Federal.. 5
- Todo em escola pública do Distrito Federal.....6
- Maior parte em escola pública do Distrito Federal..... 7
- Todo em escola particular do Distrito Federal..... 8
- Maior parte em escola particular do Distrito Federal..... 9

10 Que curso médio você concluiu ou está concluindo?

- Curso regular..... 1
- Supletivo.....2
- Profissionalizante... 3
- Técnico..... 4
- Outro..... 5

11- Há quanto tempo você concluiu o curso médio?

- Ainda estou cursando a 1ª série..... 1
- Ainda estou cursando a 2ª série..... 2

Ainda não concluí, mas estou cursando a 3ª série.....	3
Há menos de 6 meses.....	4
Entre 6 meses e 1 ano.....	5
Entre 1 e 2 anos.....	6
Entre 2 e 3 anos.....	7
Entre 3 e 4 anos.....	8
Há mais de 4 anos.....	9

12 - Como você se preparou para o vestibular?

Apenas estudei durante a 3ª série regular....	1
Fiz cursinho.....	2
Estudei sem o auxílio de professores.....	3
Não estudei.....	4

13 - Quantos vestibulares você já fez na UnB?

Nenhum, este é o primeiro....	1
Um.....	2
Dois.....	3
Três.....	4
Mais de três.....	5

14 - Você já fez ou está fazendo curso superior?

Não.....	1
Sim, fora da UnB...	2
Sim, na UnB.....	3

15 - Como você se considera com relação ao curso em que está se inscrevendo? (TAMBÉM PAS)

Absolutamente decidido....	1
Com alguma dúvida.....	2
Totalmente indeciso.....	3

16 - Qual dos fatores seguintes mais influenciou na sua opção de curso?

Área de atuação profissional de meus pais....	1
Professor(es) que tive.....	2
Amigo(s).....	3
A relação candidato/vaga no vestibular.....	4
Perspectiva profissional.....	5
Outro.....	6

17 - Se você já estuda na Unb, porque vai fazer outro vestibular?

Porque não fui selecionado para o curso que queria...	1
Porque me decepcionei com o curso.....	2
Porque mudei minha opção profissional.....	3
Por pressão familiar.....	4
Para melhorar o histórico escolar.....	5
Por outro motivo.....	6

18 - Onde você mora? (TAMBÉM PAS)

Fora do DF.....	1
Zona rural do DF.....	2
Águas Claras.....	3
Areal.....	4
Brazlândia.....	5
Candangolândia.....	6
Ceilândia.....	7
Cruzeiro.....	8
Gama.....	9
Guará.....	10
Lago Norte.....	11
Lago Sul.....	12
Núcleo Bandeirante.....	13
Octogonal.....	14
Paranoá.....	15
Planaltina.....	16
Plano Piloto (Asa N/S).....	17
Recanto das Emas.....	18
Riacho Fundo.....	19
Samambaia.....	20
Santa Maria.....	21
São Sebastião.....	22
Sobradinho.....	23
Sudoeste.....	24
Taguatinga.....	25
Outro.....	26

7 APÊNDICES

Apêndices 1 – Correlação entre as variáveis em nível da Instituição

Correlações	Particular	%_PPI_na_inst	%_pai_educ_NS	%_pai_educ_NM	%_mãe_educ_NS	%_mãe_educ_NM	Media_ingressantes	Conceito	CPC
Particular	1								
%_PPI_na_inst	-0.3414	1							
%_pai_educ_NS	-0.0393	-0.3641	1						
%_pai_educ_NM	0.0983	-0.1115	0.3301	1					
%_mãe_educ_NS	0.0409	-0.0428	0.4139	0.8536	1				
%_mãe_educ_NM	-0.1044	-0.3158	0.9566	0.3292	0.3609	1			
Media_ingressantes	-0.0564	-0.1000	0.2105	-0.0740	-0.0419	0.1882	1		
Conceito	-0.3309	-0.1416	0.3483	0.0017	0.0547	0.3645	0.5212	1	
CPC	-0.3038	-0.2707	0.4152	0.1040	0.1186	0.4273	0.3125	0.6030	1

Apêndices 2 – Coeficientes da regressão MQO com as variáveis de nível de escolaridade dos pais

Modelos MQO	Sem Mercado de Trabalho	Com Mercado de Trabalho	Homens	Mulheres	Branco / Asiáticos	PPI
(coluna 1)	(coluna 2)	(coluna 3)	(coluna 4)	(coluna 5)	(coluna 6)	(coluna 7)
<i>Variáveis</i>	Coef. **	Coef. **	Coef. **	Coef. **	Coef. **	Coef. **
Mulher	-0,2271 (0,004)	-0,2272 (0,004)	-	-	-0,2217 (0,004)	-0,2361 (0,007)
PPI	-0,0415 (0,004)	-0,0416 (0,004)	-0,0562 (0,007)	-0,0359 (0,005)	-	-
Idade	0,0198 (0,0005)	0,0176 (0,0005)	0,0201 (0,0009)	0,0163 (0,0005)	0,0186 (0,0006)	0,0156 (0,0007)
Casado	0,0992 (0,005)	0,0988 (0,005)	0,1890 (0,009)	0,0307 (0,006)	0,1088 (0,006)	0,0780 (0,008)
Filhos	0,0174 (0,005)	0,0092 (0,005)	0,0624 (0,009)	-0,0298 (0,006)	-0,0062 (0,006)	0,0345 (0,009)
Pai com nível médio	0,0200 (0,005)	0,0247 (0,005)	0,0085 (0,008)	0,0350 (0,006)	0,0231 (0,005)	0,0309 (0,009)
Pai com nível superior	0,0372 (0,006)	0,0417 (0,006)	0,0339 (0,009)	0,0534 (0,007)	0,0417 (0,006)	0,0365 (0,012)
Mãe com nível médio	0,0120 (0,005)	0,0098 (0,005)	-0,0048 (0,008)	0,0191 (0,006)	0,0077 (0,005)	0,0168 (0,009)
Mãe com nível superior	0,0211 (0,006)	0,0142 (0,006)	-0,0066 (0,009)	0,0316 (0,007)	0,0127 (0,006)	0,0201 (0,011)
Nota no Enade	0,0039 (0,006)	0,0036 (0,006)	0,0044 (0,009)	0,0032 (0,007)	0,0035 (0,006)	0,0038 (0,011)

	(0,0001)	(0,0001)	(0,0002)	(0,0001)	(0,0001)	(0,0002)
Trabalho Parcial	-	-0,0771 (0,004)	-0,1093 (0,008)	-0,0513 (0,005)	-0,0775 (0,005)	-0,0767 (0,007)
Trabalho +44h/sem	-	0,4474 (0,007)	0,3478 (0,013)	0,4994 (0,009)	0,4438 (0,009)	0,4573 (0,012)
Diploma ã Necessario	-	-0,2160 (0,003)	-0,2154 (0,006)	-0,2133 (0,004)	-0,2346 (0,004)	-0,1704 (0,006)
Estatutario	-	0,1488 (0,004)	0,1281 (0,008)	0,1782 (0,005)	0,1289 (0,005)	0,1855 (0,008)
Particular	-0,1038 (0,016)	-0,0822 (0,016)	-0,1119 (0,028)	-0,0349 (0,018)	-0,0582 (0,020)	-0,1510 (0,027)
Particular * idade	0,0067 (0,0005)	0,0068 (0,0005)	0,0081 (0,0009)	0,0050 (0,0006)	0,0066 (0,0007)	0,0077 (0,0008)
%_PPI_na_inst	-0,1256 (0,012)	-0,1679 (0,012)	-0,1377 (0,021)	-0,1589 (0,014)	-0,1127 (0,014)	-0,3334 (0,022)
%_pai_educ_NS	0,6352 (0,017)	0,6549 (0,016)	0,6447 (0,026)	0,6683 (0,020)	0,6952 (0,019)	0,4708 (0,035)
%_pai_educ_NM	0,3209 (0,026)	0,4513 (0,025)	0,2749 (0,044)	0,4900 (0,030)	0,4894 (0,031)	0,4897 (0,044)
Media_ingressantes	-0,0022 (0,0005)	-0,0009 (0,0004)	0,0017 (0,0009)	-0,0011 (0,0005)	0,0001 (0,0005)	-0,0035 (0,0008)
Conceito	0,0194 (0,004)	0,0072 (0,004)	0,0272 (0,007)	-0,0126 (0,005)	0,0113 (0,005)	-0,0035 (0,007)
CPC	0,0165 (0,004)	0,0116 (0,004)	-0,0156 (0,007)	0,0313 (0,005)	-0,0018 (0,005)	0,0523 (0,008)
Engenharias	0,6719 (0,009)	0,6616 (0,009)	0,7868 (0,015)	0,4829 (0,016)	0,6809 (0,011)	0,6148 (0,018)
Ciências	0,3694 (0,013)	0,3819 (0,012)	0,5334 (0,019)	0,2785 (0,018)	0,4009 (0,014)	0,3340 (0,024)
Saúde	-0,0786 (0,009)	-0,0640 (0,009)	-0,0596 (0,018)	-0,0481 (0,010)	-0,0554 (0,011)	-0,0542 (0,019)
Medicina /Direito	0,1988 (0,009)	0,2189 (0,009)	0,2750 (0,015)	0,1743 (0,011)	0,2156 (0,010)	0,2596 (0,017)
Ciências Sociais	0,0095 (0,007)	0,0211 (0,006)	0,0314 (0,013)	0,0268 (0,007)	0,0334 (0,008)	-0,0037 (0,011)
Negócios	0,1812 (0,006)	0,2570 (0,006)	0,2973 (0,012)	0,2443 (0,007)	0,2751 (0,007)	0,2200 (0,011)
Educação	0,0153 (0,006)	0,0059 (0,006)	0,1392 (0,014)	-0,0126 (0,007)	0,0194 (0,008)	-0,0203 (0,010)
constante	7,2164 (0,002)	7,2158 (0,002)	7,3533 (0,003)	7,1120 (0,002)	7,2287 (0,002)	7,1940 (0,004)

OBS : *** significãnte a menos de 0,1%, ** significãnte a 1% e * significãnte a 5%. Erros-padrão entre parenteses

Apêndices 3 – Modelos HLM em três níveis com as variáveis de nível de escolaridade dos pais

Tabela A3 – Modelos HLM em três níveis com as variáveis de nível de escolaridade dos pais			
Modelos - 3 níveis (coluna 1)	Modelo 0 (coluna 2)	Modelo 1 (coluna3)	Modelo 2 (coluna 4)
Numero de alunos	201363	143286	142253
Numero de instituições	1713	1713	1697
Numero de grupos	8	8	8
	R\$	R\$	R\$
Intercepto	1.472,75	1.464,47	1.019,41
Log likelihood	-203952,31	-129911,42	-128899,78
Deviance	407904,62	259822,84	257799,56
AIC	407912,60	259858,8	257851,60
BIC	407953,50	260036,5	258108,10
Variância do Nível 1 - Aluno	0,4342 (0,001)	0,3482 (0,001)	0,3484 (0,001)
Variância do Nível 2 - Instituição	0,0575 (0,002)	0,0598 (0,002)	0,0512 (0,002)
Variância do Nível 3 - Grupo	0,0816 (0,041)	0,0812 (0,040)	0,0727 (0,037)
ICC -grupo	14,24% (0,061)	16,60% (0,0697)	15,39% (0,066)
ICC instituição grupo	24,27% (0,054)	28,81% (0,060)	26,22% (0,057)
Variaveis do nivel aluno			
Mulher	-	-0,2141 *** (0,004)	-0,2127 *** (0,004)
PPI	-	-0,0377 *** (0,004)	-0,0380 *** (0,004)
Idade	-	0,0217 *** (0,0003)	0,0163 *** (0,0005)
Casado	-	0,0980 *** (0,004)	0,0961 *** (0,004)
Filhos	-	0,0201 *** (0,005)	0,0193 *** (0,005)
Pai com nível médio	-	0,0121 *** (0,004)	0,0131 *** (0,004)
Pai com nível superior	-	0,0198 *** (0,005)	0,0211 *** (0,005)
Mãe com nível médio	-	0,0187 *** (0,004)	0,0196 *** (0,004)
Mãe com nível superior	-	0,0365 *** (0,005)	0,0368 *** (0,005)
Nota no Enade	-	0,0038 *** (0,0001)	0,0037 *** (0,0001)
Trabalho Parcial	-	-0,0552 *** (0,004)	-0,0549 *** (0,004)
Trabalho +44h/sem	-	0,4626 *** (0,007)	0,4633 *** (0,007)
Diploma não Necessario	-	-0,2015 *** (0,003)	-0,2006 *** (0,003)
Estatutario	-	0,1752 *** (0,004)	0,1766 *** (0,004)

Variáveis do nível Instituição			
Particular	-		-0,1401 *** (0,022)
Particular * idade	-		0,0078 *** (0,001)
%_PPI_na_inst	-	-	0,0906 * (0,038)
%_pai_educ_NS	-	-	0,5412 *** (0,052)
%_pai_educ_NM			-0,1142 (0,084)
Media_ingressante	-	-	0,0022 (0,001)
Conceito	-	-	0,0285 * (0,013)
CPC	-	-	0,0071 (0,014)
constante	7,2949	7,2892 *** (0,101)	6,9270 *** (0,109)
Variância explicada			
Nível Aluno	-	24,71%	24,63%
Nível Instituição	-	-3,74%	12,38%
Nível Grupo	-	0,55%	12,32%

OBS : *** significante a menos de 0,1%, ** significante a 1% e * significante a 5%.
Erros-padrão entre parenteses

Apêndices 4 – Modelo nulo em dois níveis com as variáveis de nível de escolaridade dos pais

Tabela A4 – Modelo nulo em dois níveis com as variáveis de nível de escolaridade dos pais

Modelo 0	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8								
	Engenharias	Ciências	Saúde	Medicina /Direito	Ciências Sociais	Negócios	Educação	Outros								
(coluna1)	(coluna2)	(coluna3)	(coluna4)	(coluna5)	(coluna6)	(coluna7)	(coluna8)	(coluna9)								
	R\$															
Intercepto	2.621,93	R\$ 1.876,78	R\$ 1.216,60	R\$ 1.784,50	R\$ 1.183,19	R\$ 1.360,77	R\$ 1.165,27	R\$ 1.108,78								
Variância do Nível 1 - Aluno	0,3760	0,4426	0,4855	0,6786	0,4261	0,4546	0,3838	0,3895								
Variância do Nível 2 - Instituição	0,0999	0,0646	0,0347	0,0826	0,0466	0,0641	0,0505	0,0361								
ICC	21,0%	12,7%	6,7%	10,8%	9,9%	12,4%	11,6%	8,5%								
Numero de alunos	14408	9323	12502	10568	30366	55517	44877	23802								
Numero de instituições	109	60	144	127	227	539	281	226								
Log likelihood	-13.582,23	-9.517,72	-13.356,60	-13.085,43	-30.421,01	-57.578,08	-42.575,44	-22.796,53								
Deviance	27.164,46	19.035,45	26.713,20	26.170,87	60.842,02	115.156,16	85.150,88	45.593,05								
	Var(β_x)	% explicado	Var(β_x)	% explicado	Var(β_x)	% explicado	Var(β_x)	% explicado	Var(β_x)	% explicado	Var(β_x)	% explicado	Var(β_x)	% explicado	Var(β_x)	% explicado
modelo 1	0,1044	-4,48%	0,0685	-5,96%	0,0336	3,14%	0,0849	-2,81%	0,0402	13,88%	0,0668	-4,10%	0,0522	-3,39%	0,0383	-6,23%
modelo 2	0,0646	35,36%	0,0414	35,93%	0,0318	8,27%	0,0565	31,52%	0,0390	16,35%	0,0491	23,41%	0,0444	12,12%	0,0359	0,49%

Apêndices 5 - Modelo 1 em dois níveis com as variáveis de nível de escolaridade dos pais

Tabela A5 - Modelo 1 em dois níveis com as variáveis de nível de escolaridade dos pais

Modelo 1 (coluna1)	Engenharias (coluna2)	Ciências (coluna3)	Saúde (coluna4)	Medicina /Direito (coluna5)	Ciências Sociais (coluna6)	Negócios (coluna7)	Educação (coluna8)	Outros (coluna9)
Numero de alunos	9.180	3.980	9.557	8.760	21.110	42.120	30.122	18.457
Numero de instituições	109	60	144	127	227	539	281	226
Wald chi2(14)	1568,82	816,81	3497,29	4670,00	4240,94	12006,00	7390,78	3913,12
Prob > Chi2	-	-	-	-	-	-	-	-
Log likelihood	-7925,99	-3712,21	-8717,83	-8987,85	-18957,59	37864,16	24988,72	15688,54
Deviance	15851,97	7424,42	17435,65	17975,70	37915,19	75728,31	49977,43	31377,08
Efeito Fixo - Variáveis do nível aluno								
Mulher	-0,2671 *** (0,015)	-0,2550 *** (0,022)	-0,0992 *** (0,013)	-0,1998 *** (0,015)	-0,1424 *** (0,009)	-0,2050 *** (0,006)	-0,2638 *** (0,008)	-0,1453 *** (0,010)
PPI	-0,0367 * (0,016)	-0,0660 ** (0,025)	0,0059 (0,018)	-0,0704 *** (0,018)	-0,0456 *** (0,010)	-0,0602 *** (0,007)	-0,0021 (0,007)	-0,0381 *** (0,010)
Idade	0,0148 *** (0,002)	0,0247 *** (0,003)	0,0027 (0,002)	0,0310 *** (0,001)	0,0209 *** (0,001)	0,0310 *** (0,001)	0,0149 *** (0,001)	0,0199 *** (0,001)
Casado	0,1303 *** (0,022)	0,1851 *** (0,035)	-0,0465 ** (0,024)	0,1355 *** (0,019)	0,1054 *** (0,012)	0,1622 *** (0,008)	0,0254 ** (0,008)	0,0574 *** (0,012)
Filhos	0,0168 (0,023)	0,0306 (0,036)	-0,0040 (0,023)	0,0369 (0,021)	-0,0081 (0,013)	0,0574 *** (0,009)	0,0057 (0,009)	-0,0025 (0,013)
Pai com nível médio	0,0132 (0,017)	-0,0368 (0,025)	-0,0139 (0,017)	0,0089 (0,020)	0,0195 (0,011)	0,0038 (0,008)	0,0435 *** (0,009)	0,0085 (0,012)
Pai com nível superior	0,0172 (0,018)	0,0120 (0,030)	-0,0180 (0,018)	-0,0112 (0,022)	0,0388 ** (0,014)	0,0402 *** (0,010)	0,0295 ** (0,013)	-0,0080 (0,014)
Mãe com nível médio	-0,0038 (0,017)	0,0129 (0,025)	0,0261 (0,017)	0,0161 (0,020)	0,0247 * (0,011)	0,0145 (0,008)	0,0102 (0,009)	0,0307 *** (0,011)
Mãe com nível superior	0,0324 (0,019)	0,0142 (0,030)	0,0298 (0,018)	0,0133 (0,023)	0,0297 * (0,014)	0,0211 * (0,010)	0,0636 *** (0,012)	0,0555 *** (0,014)
Nota no Enade	0,0017 *** (0,0004)	0,0055 *** (0,0006)	0,0008 (0,0005)	0,0078 *** (0,0005)	0,0035 *** (0,0003)	0,0058 *** (0,0002)	0,0020 *** (0,0002)	0,0033 *** (0,0003)
Trabalho Parcial	-0,1831 *** (0,022)	-0,2549 *** (0,030)	-0,0745 *** (0,014)	-0,0779 *** (0,018)	-0,0791 *** (0,010)	0,0438 *** (0,009)	-0,0729 *** (0,007)	0,0024 (0,010)
Trabalho +44h/sem	0,2662 *** (0,039)	0,3285 *** (0,051)	0,6130 *** (0,023)	0,3381 *** (0,040)	0,3960 *** (0,017)	0,5059 *** (0,025)	0,4885 *** (0,011)	0,4838 *** (0,015)
Diploma ã Necessario	-0,3861 *** (0,015)	-0,2026 *** (0,021)	-0,5286 *** (0,013)	-0,3147 *** (0,015)	-0,1702 *** (0,008)	-0,1351 *** (0,006)	-0,1527 *** (0,007)	-0,1353 *** (0,009)
Estatutario	-0,2096 *** (0,023)	-0,0590 (0,033)	0,2768 *** (0,015)	0,4477 *** (0,016)	0,1770 *** (0,010)	0,0815 *** (0,011)	0,1728 *** (0,007)	0,1730 *** (0,010)
constante	7,8678 *** (0,032)	7,5247 *** (0,035)	7,0980 *** (0,017)	7,4834 *** (0,027)	7,0769 *** (0,016)	7,2088 *** (0,012)	7,0564 *** (0,014)	7,0060 *** (0,014)
Efeito Aleatório								
var(constante)	0,1044 (0,015)	0,0685 (0,014)	0,0336 (0,005)	0,0849 (0,012)	0,0402 (0,060)	0,0668 (0,004)	0,0522 (0,005)	0,0383 (0,004)

var(Residual)	0,3170 (0,005)	0,3643 (0,008)	0,3527 (0,005)	0,4399 (0,007)	0,3436 (0,003)	0,3419 (0,002)	0,3001 (0,002)	0,3119 (0,003)
Variância explicada								
Nível aluno	15,69%	17,69%	27,34%	35,17%	19,37%	24,79%	21,81%	19,91%
Nível Instituição	-4,48%	-5,96%	3,14%	-2,81%	13,88%	-4,10%	-3,39%	-6,23%

OBS : *** significante a menos de 0,1%, ** significante a 1% e * significante a 5%. Erros-padrão entre parenteses

Apêndices 6 - Modelo 2 em dois níveis com as variáveis de nível de escolaridade dos pais

Tabela A6 - Modelo 2 em dois níveis com as variáveis de nível de escolaridade dos pais

Modelo 2 (coluna1)	Engenharias (coluna2)	Ciências (coluna3)	Saúde (coluna4)	Medicina /Direito (coluna5)	Ciências Sociais (coluna6)	Negócios (coluna7)	Educação (coluna8)	Outros (coluna9)
Numero de alunos	8644	3980	9557	8714	21061	41857	30020	18420
Numero de instituições	106	60	144	126	226	532	278	225
Wald chi2(22)	1676,67	849,61	3506,99	4684,95	4291,55	12013,47	7393,72	3928,83
Prob > Chi2	-	-	-	-	-	-	-	-
Log likelihood	-7500,49	3699,02	-8713,53	-8921,03	18878,29	-37562,92	-24890,40	-15651,65
Deviance	15000,99	7398,04	17427,06	17842,07	37756,58	75125,85	49780,80	31303,30
Efeito Fixo - Variáveis do nível aluno								
Mulher	-0,2709 *** (0,016)	-0,2555 *** (0,022)	-0,0995 *** (0,013)	-0,1994 *** (0,015)	-0,1426 *** (0,009)	-0,2049 *** (0,006)	-0,2634 *** (0,008)	-0,1448 *** (0,010)
PPI	-0,0301 (0,017)	-0,0662 ** (0,025)	0,0061 (0,018)	-0,0710 *** (0,018)	-0,0456 *** (0,010)	-0,0605 *** (0,007)	-0,0023 (0,007)	-0,0380 *** (0,010)
Idade	-0,0045 (0,003)	0,0238 *** (0,004)	-0,0022 (0,004)	0,0308 *** (0,003)	0,0169 *** (0,001)	0,0262 *** (0,001)	0,0149 *** (0,001)	0,0187 *** (0,001)
Casado	0,1133 *** (0,023)	0,1843 *** (0,035)	-0,0473 * (0,024)	0,1342 *** (0,020)	0,1060 *** (0,012)	0,1621 *** (0,008)	0,0246 ** (0,008)	0,0559 *** (0,012)
Filhos	0,0087 (0,024)	0,0306 (0,036)	-0,0046 (0,023)	0,0363 (0,021)	-0,0075 (0,013)	0,0574 *** (0,009)	0,0068 (0,009)	-0,0027 (0,013)
Pai com nível médio	0,0066 (0,017)	-0,0366 (0,025)	-0,0138 (0,017)	0,0072 (0,020)	0,0208 (0,011)	0,0045 (0,008)	0,0437 *** (0,009)	0,0090 (0,012)
Pai com nível superior	0,0136 (0,019)	0,0120 (0,030)	-0,0183 (0,018)	-0,0122 (0,023)	0,0394 ** (0,014)	0,0406 *** (0,010)	0,0298 * (0,013)	-0,0075 (0,014)
Mãe com nível médio	0,0012 (0,017)	0,0130 (0,025)	0,0263 (0,017)	0,0171 (0,020)	0,0255 * (0,011)	0,0147 (0,008)	0,0110 (0,009)	0,0313 ** (0,011)
Mãe com nível superior	0,0283 (0,0193)	0,0141 (0,030)	0,0298 (0,018)	0,0147 (0,023)	0,0305 * (0,014)	0,0206 * (0,010)	0,0626 *** (0,012)	0,0560 *** (0,014)
Nota no Enade	0,0016 *** (0,0005)	0,0055 *** (0,001)	0,0007 (0,0005)	0,0078 *** (0,0005)	0,0034 *** (0,0003)	0,0057 *** (0,0002)	0,0020 *** (0,0002)	0,0033 *** (0,0003)
Trabalho Parcial	-0,1838 *** (0,023)	-0,2550 *** (0,030)	-0,0743 *** (0,014)	-0,0745 *** (0,018)	-0,0785 *** (0,010)	0,0446 *** (0,009)	-0,0728 *** (0,007)	0,0028 (0,010)
Trabalho +44h/sem	0,2625 *** (0,040)	0,3292 *** (0,051)	0,6131 *** (0,023)	0,3390 *** (0,040)	0,3949 *** (0,017)	0,5017 *** (0,025)	0,4883 *** (0,011)	0,4839 *** (0,015)
Diploma ã Necessario	-0,3869 *** (0,016)	-0,2029 *** (0,021)	-0,5286 *** (0,013)	-0,3154 *** (0,015)	-0,1703 *** (0,008)	-0,1346 *** (0,006)	-0,1527 *** (0,007)	-0,1352 *** (0,009)
Estatutario	-0,1905 *** (0,024)	-0,0590 (0,033)	0,2768 *** (0,015)	0,4488 *** (0,016)	0,1761 *** (0,010)	0,0829 *** (0,011)	0,1727 *** (0,007)	0,1735 *** (0,010)
Efeito Fixo - Variáveis do nível Instituição								
Particular	-0,5032 *** (0,118)	0,0378 (0,150)	-0,1622 (0,127)	0,0499 (0,116)	-0,1300 ** (0,050)	-0,0052 (0,049)	0,0202 (0,042)	-0,0099 (0,050)
Particular * idade	0,0291 *** (0,003)	0,0013 (0,005)	0,0055 (0,004)	0,0003 (0,003)	0,0062 *** (0,001)	0,0056 *** (0,001)	0,0000 (0,001)	0,0021 (0,001)
%_PPI_na_inst	0,1246 (0,190)	0,4061 (0,243)	0,0397 (0,144)	0,8499 *** (0,167)	0,1095 (0,095)	0,0652 (0,068)	0,0502 (0,082)	0,0235 (0,088)
%_pai_educ_NS	0,0900 (0,225)	0,8986 *** (0,286)	0,3841 * (0,177)	0,7592 *** (0,205)	0,5142 *** (0,097)	0,6247 *** (0,105)	0,8356 *** (0,215)	0,1957 (0,128)
%_pai_educ_NM	-0,9748 * (0,429)	-1,0946 (0,613)	0,1788 (0,343)	0,1273 (0,480)	0,2595 (0,187)	-0,4095 ** (0,152)	0,0579 (0,229)	0,1709 (0,219)
Media_ingressantes	0,0153 (0,009)	-0,0126 (0,010)	0,0039 (0,005)	0,0158 * (0,007)	0,0066 (0,004)	0,0096 ** (0,003)	0,0001 (0,003)	-0,0016 (0,004)
Conceito	0,1313 * (0,065)	0,1523 * (0,071)	0,0019 (0,046)	-0,0849 (0,049)	-0,0807 * (0,041)	0,0667 ** (0,024)	-0,0374 (0,035)	-0,0146 (0,033)
CPC	-0,0860 (0,063)	-0,0087 (0,080)	-0,0596 (0,041)	-0,0417 (0,049)	0,0385 (0,034)	-0,0204 (0,025)	0,0744 * (0,034)	0,0554 (0,033)

constante	7,3403 (0,293) ***	7,3156 (0,346) ***	6,9561 (0,216) ***	6,8003 (0,238) ***	6,7598 (0,122) ***	6,6315 (0,103) ***	6,8468 (0,128) ***	6,8473 (0,130) ***
Efeito Aleatório								
var(constante)	0,0646 (0,010)	0,0414 (0,009)	0,0318 (0,005)	0,0565 (0,008)	0,0390 (0,004)	0,0491 (0,003)	0,0444 (0,004)	0,0359 (0,004)
var(Residual)	0,3211 (0,005)	0,3643 (0,008)	0,3526 (0,005)	0,4402 (0,007)	0,3432 (0,003)	0,3421 (0,002)	0,3003 (0,002)	0,3120 (0,003)
Variância explicada								
Nível aluno	14,58%	17,68%	27,36%	35,12%	19,45%	24,74%	21,77%	19,90%
Nível Instituição	35,36%	35,93%	8,27%	31,52%	16,35%	23,41%	12,12%	0,49%

OBS : *** significante a menos de 0,1%, ** significante a 1% e * significante a 5%. Erros-padrão entre parenteses

Apêndices 7 – Modelos com os efeitos randômicos

Tabela A7 - Modelos com os efeitos randômicos

Modelo 3 (coluna 1)	Engenharias (coluna 2)	Ciências (coluna 3)	Saúde (coluna 4)	Medicina /Direito (coluna 5)	Ciências Sociais (coluna 6)	Negócios (coluna 7)	Educação (coluna 8)	Outros (coluna 9)
Numero de alunos	8655	3986	9497	8719	21123	42036	29918	18464
Numero de instituições	106	60	144	126	226	532	278	225
Wald chi2(19)	879,12	618,48	1848,61	2694,02	2748,80	9107,21	3552,11	2883,00
Log likelihood	-7367,64	-3625,86	-8564,97	-8727,16	18467,68	36762,13	-24152,26	15288,35
Deviance	14.735,29	7.251,72	17.129,94	17.454,33	36.935,37	73.524,27	48.304,52	30.576,70
AIC	14.803,29	7.291,72	17.169,94	17.494,33	37.003,37	73.592,27	48.372,52	30.644,70
BIC	15.043,53	7.417,53	17.313,12	17.635,79	37.273,94	73.886,24	48.654,93	30.910,70
Efeito Fixo - Variáveis do nível Aluno								
Mulher	-0,2558 *** 0,019	-0,2419 *** 0,026	-0,1053 *** 0,017	-0,1868 *** 0,016	-0,1263 *** 0,010	-0,1919 *** 0,007	-0,2640 *** 0,012	-0,1313 *** 0,013
PPI	-0,0242 0,016	-0,0610 * 0,025	0,0119 0,017	-0,0442 * 0,018	-0,0332 *** 0,010	-0,0425 *** 0,007	0,0071 0,008	-0,0220 0,011
Idade	-0,0051 0,003	0,0167 ** 0,006	-0,0030 0,004	0,0270 *** 0,003	0,0155 * 0,002	0,0229 *** 0,002	0,0129 *** 0,001	0,0191 *** 0,001
Casado	0,1117 *** 0,026	0,1799 *** 0,034	-0,0427 0,025	0,1231 *** 0,021	0,0961 *** 0,014	0,1487 *** 0,008	0,0230 * 0,009	0,0521 *** 0,013
Filhos	0,0083 0,023	0,0384 0,036	-0,0088 0,025	0,0400 * 0,020	0,0013 0,013	0,0579 *** 0,009	0,0210 * 0,010	0,0116 0,013
Renda Baixa	-0,1623 *** 0,032	-0,3022 *** 0,039	0,0024 0,018	-0,3297 *** 0,022	-0,2186 *** 0,012	-0,2804 *** 0,008	-0,1560 *** 0,009	-0,2008 *** 0,010
Renda Alta	0,0943 *** 0,017	0,1733 *** 0,035	0,0191 0,019	0,1564 *** 0,031	0,1618 *** 0,025	0,2171 *** 0,019	0,1315 *** 0,031	0,1370 *** 0,026
Nota no Enade	0,0018 *** 0,0006	0,0045 *** 0,0007	0,0010 0,0006	0,0072 *** 0,0007	0,0029 *** 0,0003	0,0049 *** 0,0002	0,0019 *** 0,0003	0,0025 *** 0,0003
Trabalho Parcial	-0,1591 *** 0,031	-0,2370 *** 0,039	-0,0842 *** 0,018	-0,0626 * 0,026	-0,0640 *** 0,013	0,0553 *** 0,012	-0,0706 *** 0,011	-0,0028 0,014
Trabalho +44h/sem	0,2939 *** 0,042	0,3478 *** 0,051	0,6112 *** 0,025	0,3679 *** 0,046	0,4033 *** 0,018	0,5050 *** 0,025	0,4701 *** 0,014	0,4717 *** 0,015
Diploma ã Necessario	-0,3922 *** 0,025	-0,2034 *** 0,023	-0,5219 *** 0,020	-0,2933 *** 0,019	-0,1736 *** 0,010	-0,1278 *** 0,006	-0,1613 *** 0,010	-0,1276 *** 0,011
Estatutario	-0,1983 *** 0,037	-0,0473 0,043	0,2737 *** 0,021	0,4386 *** 0,021	0,1951 *** 0,017	0,0695 *** 0,014	0,1880 *** 0,014	0,1996 *** 0,016
Efeito Fixo - Variáveis do nível Instituição								
Particular	-0,4777 *** 0,142	-0,1012 0,198	-0,1748 0,127	-0,0504 0,129	-0,1434 * 0,065	-0,0490 0,058	-0,0068 0,051	0,0189 0,062
Particular * idade	0,0293 *** 0,004	0,0068 0,007	0,0057 0,004	0,0039 0,003	0,0069 *** 0,002	0,0072 *** 0,002	0,0009 0,001	0,0012 0,002

%_PPI_na_inst	-0,0162	0,3383	0,0420	0,8206 ***	0,1368	-0,0192	0,0520	0,0431
	0,185	0,242	0,135	0,163	0,095	0,062	0,081	0,086
%_pai_educ_NS	0,2343	0,9763 ***	0,3634 *	0,7495 ***	0,5438 ***	0,5958 ***	0,8532 ***	0,2151
	0,222	0,281	0,176	0,201	0,097	0,104	0,182	0,125
Media_ingressantes	0,0158	-0,0134	0,0043	0,0150 *	0,0069	0,0095 **	-0,0003	-0,0008
	0,009	0,010	0,005	0,007	0,004	0,003	0,003	0,003
Conceito	0,1445 *	0,1522 *	0,0049	-0,0765	-0,0833 *	0,0746 **	-0,0343	-0,0173
	0,066	0,071	0,046	0,048	0,041	0,024	0,034	0,033
CPC	-0,0965	-0,0075	-0,0596	-0,0395	0,0386	-0,0304	0,0761 *	0,0574
	0,065	0,080	0,041	0,049	0,034	0,025	0,034	0,032
constante	6,9677 ***	6,9412 ***	7,0012 ***	6,8309 ***	6,8085 ***	6,5422 ***	6,8556 ***	6,8558 ***
	0,257	0,285	0,190	0,195	0,117	0,097	0,122	0,128

Efeito Aleatório

var(mulher)	0,0081	0,0104	0,0120 †	0,0026	0,0045 †	0,0038	0,0165	0,0104 †
var(PPI)	0,0000	0,0000	0,0000	0,0010 †	0,0015 †	0,0016 †	0,0011	0,0051
var(idade)	0,0001	0,0002 †	0,0000	0,0001	0,0001 †	0,0001	0,0000	0,0001
var(casado)	0,0112 †	0,0000	0,0057	0,0028	0,0053	0,0016	0,0026 †	0,0039 †
var(filhos)	0,0007	0,0002	0,0069	0,0000	0,0000	0,0014 †	0,0029 †	0,0023
var(renda_baixa)	0,0299 †	0,0053	0,0000	0,0000	0,0035	0,0023	0,0038	0,0015
var(renda_alta)	0,0000	0,0092 †	0,0038	0,0427 †	0,0412 †	0,0434 †	0,0593 †	0,0178 †
var(nota_enade)	0,0000 †	0,0000	0,0000	0,0000 †	0,0000 †	0,0000	0,0000 †	0,0000
var(trabalho_parcial)	0,0358 †	0,0343 †	0,0160 †	0,0294 †	0,0119 †	0,0208 †	0,0156 †	0,0140
var(trabalho_+44h/sem)	0,0132	0,0071	0,0060	0,0286	0,0032	0,0000	0,0118	0,0003
var(dip_nao_nec)	0,0340	0,0050	0,0268	0,0118	0,0055	0,0025	0,0114 †	0,0085 †
var(estatutario)	0,0615 †	0,0357	0,0274 †	0,0174	0,0313 †	0,0250 †	0,0299 †	0,0281 †
var(constante)	0,0690	0,0422	0,0328	0,0562	0,0401	0,0492	0,0440	0,0357
var(Residual)	0,2963	0,3370	0,3312	0,4035	0,3161	0,3182	0,2745	0,2868

OBS : *** significante a menos de 0,1%, ** significante a 1%, * significante a 5%, † desvio padrão maior que o 50% do coeficiente e erros-padrão entre parenteses

Apêndices 8 - Modelos com os efeitos randômicos e interações – parte 1

Tabela A8 - Modelos com os efeitos randômicos e interações – parte 1									
Modelo 4	Engenharias		Ciências		Saúde		Medicina /Direito		
Numero de alunos	8.655		3.986		9.497		8.719		
Numero de instituições	106		60		144		126		
Log likelihood	-7404,84		-3631,48		-8591,55		-8703,53		
Deviance	14809,68		7262,96		17183,11		17407,06		
AIC	14867,68		7316,96		17237,11		17471,06		
BIC	15072,59		7486,80		17430,39		17697,40		
Ganho na modelagem (AIC)	-64,39		-25,24		-67,17		23,27		
Ganho na modelagem (BIC)	-29,06		-69,27		-117,27		-61,61		
Variáveis - nível Aluno	Coef	EP	Coef	EP	Coef	EP	Coef	EP	
Mulher	-0,262 **		-0,243 **		-0,076 *		-0,190 **		

		*	0,01	*	0,02		0,03	*	0,01
			6		1		3		4
PPI	-0,026		0,01	*	0,02		0,01	*	0,13
			6		5		8		1
Idade	-0,006	*	0,00	**	0,00		0,00	**	0,00
			3		6		4		3
Casado	0,114	**	0,02	**	0,03		0,02	**	0,01
		*	7	*	4	*	4	*	9
Filhos	0,010		0,02		0,03		0,02	*	0,02
			3		5		3		0
Renda Baixa	-0,172	**	0,03	**	0,03		0,01	**	0,02
		*	3	*	7		8	*	2
Renda Alta	0,098	**	0,01	**	0,20		0,01	*	0,15
		*	7	**	5		7	*	2
Nota no Enade	0,002	**	0,00	**	0,00		0,00	**	0,00
		*	1	*	1		0	*	1
Trabalho Parcial	-0,167	**	0,03	**	0,04		0,01	**	0,04
		*	2	*	0	*	8	*	9
Trabalho +44h/sem	0,290	**	0,03	**	0,04		0,02	**	0,03
		*	9	*	9	*	3	*	9
Diploma ã Necessario	-0,376	**	0,01	**	0,02		0,01	**	0,01
		*	5	*	0	*	3	*	5
Estatutario	-0,112		0,18		0,03	**	0,02	**	0,01
			9		2	*	1	*	6

Variáveis - nível Instituição	Coef	EP	Coef	EP	Coef	EP	Coef	EP
Particular	-0,500	**	0,12	0,20	-0,158	0,12	0,056	0,11
		*	1	1		6		4
Particular * idade	0,030	**	0,00	0,00	0,005	0,00	0,000	0,00
		*	3	7		4		3
%_PPI_na_inst	-0,020		0,18	0,24	0,043	0,13	0,839	0,16
			5	3		4		3
%_pai_educ_NS	0,239		0,22	**	0,28	0,17	0,761	0,20
			1	*	3	5		1
Media_ingressantes	0,016		0,00	0,01	0,004	0,00	0,015	0,00
			9	0		5		7
Conceito	0,143	*	0,06	0,07	0,005	0,04	-0,082	0,04
			6	2		6		9
CPC	-0,098		0,06	0,08	-0,059	0,04	-0,040	0,04
			5	1		1		9

Variáveis - interação entre níveis - efeito cross-level

mulher_particular	-		-		-0,280	**	0,09	-
						0		
mulher_particular_idade	-		-		0,009	**	0,00	-
						3		
mulher_perc_PPI	-		-		-		-	-
mulher_educ_NS	-		-		-		-	-
mulher_cpccontinuo	-		-		-		-	-
PPI_media_ingr	-		-		-		-0,011	**
								0,00
PPI_perc_PPI	-		-		-		-	4

casado_particular	-	-	-	-	-	-	-	-
casado_particular_idade	-	-	-	-	-	-	-	-
casado_pai_educ_NS	-	-	-	-	-	-	-	-
alta_particular	-	-	-	-	-0,510	*	0,11	1
alta_particular_idade	-	-	-	-	0,019	*	0,00	3
alta_pai_educ_NS	-	0,548	*	0,28	-	-	-	-
alta_cpccontinuo	-	-0,226	**	0,07	-	-0,102	*	0,05
trabalho_parcial_particular	-	-	-	-	0,428	*	0,09	4
trabalho_parcial_part_idade	-	-	-	-	-0,009	*	0,00	2
trabalho_parcial_media_ingr	-	-	-	-	-	-	-	-
trabalho_parcial_cpccontinuo	-	-	-	-	-	-	-	-
dip_nao_nec_particular	-	-	-	-	-	-	-	-
dip_nao_nec_particular_idade	-	-	-	-	-	-	-	-
dip_nao_nec_perc_PPI	-	-	-	-	-	-	-	-
dip_nao_nec_media_ingr	-	-	-	-	-	-	-	-
dip_nao_nec_conceito	-	-	-	-	-	-	-	-
estatutario_particular	-	-	-	-	-	-	-	-
estatutario_particular_idade	-	-	-	-	-	-	-	-
estatutario_pai_educ_NS	-	-	-	-	-	-	-	-
estatutario_media_ingr	-0,023	**	0,00	-	-	-	-	-
estatutario_conceito	-	-	-	-	-	-	-	-
estatutario_cpccontinuo	0,220	**	0,07	-	-	-	-	-
Constante	6,976	**	0,25	6,950	**	0,28	7,010	**
		*	6		*	7		*
								0,18
								9
								5
var(mulher)	-	-	-	1,35%	†	-	-	-
var(PPI)	-	-	-	-	-	0,04%	-	-
var(idade)	-	0,02%	†	-	-	-	-	-
var(casado)	1,85%	†	-	-	-	-	-	-
var(filhos)	-	-	-	-	-	-	-	-
var(renda_baixa)	3,39%	†	-	-	-	-	-	-
var(renda_alta)	-	0,08%	-	-	-	3,20%	†	-
var(nota_enade)	0,00%	†	-	-	-	0,00%	-	-
var(trabalho_parcial)	4,17%	†	3,57%	†	1,60%	†	2,21%	†
var(trabalho_+44h/sem)	-	-	-	-	-	-	-	-
var(dip_nao_nec)	-	-	-	-	-	-	-	-
var(estatutario)	4,74%	-	-	2,80%	†	-	-	-
var(constante)	6,84%	-	4,28%	3,24%	-	5,63%	-	-
var(Residual)	30,44%	-	34,49%	33,97%	-	41,15%	-	-
número de parâmetros (#)	6	-	4	4	-	5	-	-
teste LR - chibar2(#)	1185,79	-	295,06	597,46	-	630,48	-	-

OBS : *** significante a menos de 0,1%, ** significante a 1% , * significante a 5%, † desvio padrão maior que o 50% do coeficiente

Apêndices 9 - Modelos com os efeitos randômicos e interações – parte 2

Tabela A8 - Modelos com os efeitos randômicos e interações – parte 2									
Modelo 4	Ciências Sociais		Negócios		Educação		Outros		
Numero de alunos	21.123		42.036		29.918		18.464		
Numero de instituições	226		532		278		225		
Log likelihood	-18415,33		-36715,45		-24178,00		-15276,06		
Deviance	36830,66		73430,90		48356,00		30552,12		
AIC	36910,66		73498,90		48440,00		30624,12		
BIC	37228,99		73792,87		48788,86		30905,76		
Ganho na modelagem (AIC)	92,71		93,37		-67,48		20,58		
Ganho na modelagem (BIC)	44,95		93,37		-133,93		4,94		
Variáveis - nível Aluno	Coef	EP	Coef	EP	Coef	EP	Coef	EP	
Mulher	-0,364 ***	0,072	-0,188 ***	0,006	-0,240 ***	0,008	-0,160 ***	0,021	
PPI	-0,033 ***	0,010	-0,013	0,017	0,006	0,007	-0,026 **	0,009	
Idade	0,016 ***	0,002	0,024 ***	0,001	0,014 ***	0,001	0,017 ***	0,001	
Casado	0,087 ***	0,012	0,147 ***	0,008	-0,013	0,017	-0,013	0,020	
Filhos	0,002	0,013	0,060 ***	0,009	0,021	0,011	0,011	0,012	
Renda Baixa	-0,221 ***	0,010	-0,283 ***	0,008	-0,165 ***	0,007	-0,204 ***	0,010	
Renda Alta	0,271 ***	0,045	0,204 ***	0,043	0,135 ***	0,030	0,312 ***	0,058	
Nota no Enade	0,003 ***	0,000	0,005 ***	0,000	0,002 ***	0,000	0,003 ***	0,000	
Trabalho Parcial	-0,049 *	0,022	0,420 ***	0,093	0,100	0,060	-0,009	0,010	
Trabalho +44h/sem	0,405 ***	0,017	0,503 ***	0,025	0,480 ***	0,011	0,468 ***	0,015	
Diploma ã Necessario	-0,175 ***	0,008	-0,128 ***	0,006	-0,131	0,076	-0,066 **	0,022	
Estatutario	0,197 ***	0,028	0,022	0,024	0,349 ***	0,097	0,222 ***	0,027	
Particular	-0,149 *	0,065	0,010	0,050	0,015	0,044	-0,030	0,050	
Particular * idade	0,007 ***	0,002	0,005 ***	0,001	0,000	0,001	0,003 *	0,001	
%_PPI_na_inst	0,139	0,094	-0,005	0,062	0,055	0,081	0,042	0,086	
%_pai_educ_NS	0,545 ***	0,097	0,599 ***	0,104	0,849 ***	0,182	0,218	0,124	
Media_ingressantes	0,007	0,004	0,010 **	0,003	0,000	0,003	-0,001	0,003	
Conceito	-0,084 *	0,041	0,074 **	0,024	-0,034	0,034	-0,019	0,032	
CPC	0,039	0,034	-0,030	0,025	0,075 *	0,034	0,058	0,032	
mulher_particular	0,201 ***	0,049	-	-	-	-	0,318 ***	0,060	
mulher_particular_idade	-0,005 ***	0,001	-	-	-	-	-0,009 ***	0,002	
mulher_perc_PPI	0,131 *	0,065	-	-	-	-	-	-	
mulher_educ_NS	0,197 **	0,066	-	-	-	-	-	-	
mulher_cpccontinuo	0,045 *	0,022	-	-	-	-	-	-	
PPI_media_ingr	-	-	-	-	-	-	-	-	
PPI_perc_PPI	-	-	-0,089 *	0,045	-	-	-	-	
casado_particular	-	-	-	-	0,124 **	0,043	-	-	

casado_particular_idade	-	-	-0,004 ***	0,001	-	
casado_pai_educ_NS	-	-	0,354 ***	0,106	0,360 ***	0,089
alta_particular	-0,708 ***	0,090	-0,604 ***	0,074	-	-0,332 *** 0,107
alta_particular_idade	0,021 ***	0,003	0,020 ***	0,002	-	0,010 *** 0,003
alta_pai_educ_NS	-	-	-	-	-0,419 ***	0,139
alta_cpccontinuo	-	-	-	-	-	
trabalho_parcial_particular	0,184 ***	0,054	0,220 ***	0,052	0,204 ***	0,043
trabalho_parcial_part_idade	-0,007 ***	0,002	-0,010 ***	0,002	-0,006 ***	0,001
trabalho_parcial_media_ingr	-	-	-0,010 ***	0,003	-	-
trabalho_parcial_cpccontinuo	-	-	-0,073 ***	0,022	-	-
dip_nao_nec_particular	-	-	-0,112 **	0,041	-	-
dip_nao_nec_particular_idade	-	-	0,003 **	0,001	-	-
dip_nao_nec_perc_PPI	-	-	-	-	-0,185 ***	0,057
dip_nao_nec_media_ingr	-	-	0,006 **	0,002	-	-
dip_nao_nec_conceito	-	-	-0,061 **	0,022	-	-
estatutario_particular	0,135 *	0,062	-	-	0,178 ***	0,048
estatutario_particular_idade	-0,004 **	0,002	-	-	-0,005 ***	0,001
estatutario_pai_educ_NS	-	-	0,322 **	0,124	-	-
estatutario_media_ingr	-	-	-	-	-	-
estatutario_conceito	-	-	-0,056 *	0,028	-	-
estatutario_cpccontinuo	-	-	-	-	-	-
Constante	6,809 ***	0,117	6,539 ***	0,097	6,855 ***	0,122
Efeito Aleatório						
var(mulher)	0,003	-	-	-	0,009 †	
var(PPI)	0,002 †	0,001	-	-	-	
var(idade)	0,000 †	-	-	-	-	
var(casado)	-	-	0,002	-	0,004	
var(filhos)	-	0,004 †	0,006	-	-	
var(renda_baixa)	-	-	-	-	-	
var(renda_alta)	0,025 †	0,032 †	0,058 †	-	0,002	
var(nota_enade)	0,00001 †	-	0,00001 †	-	-	
var(trabalho_parcial)	0,012 †	0,017	0,013	-	-	
var(trabalho_+44h/sem)	-	-	-	-	-	
var(dip_nao_nec)	-	-	0,010	-	0,007 †	
var(estatutario)	0,033 †	0,023	0,031 †	-	0,030 †	
var(constante)	0,040	0,049	0,044	-	0,035	
var(Residual)	0,317	0,322	0,279	-	0,292	
número de parâmetros (#)	8	6	8	-	6	
teste LR - chibar2(#)	1893,14	4388,58	3645,98	-	1621,15	

OBS : *** significante a menos de 0,1%, ** significante a 1% , * significante a 5%, † desvio padrão maior que o 50% do coeficiente

Apêndices 10 – Resultado Mixed Logit Model com todas as variáveis explicativas.

Tabela A10 - Resultado *Mixed Logit Model* com todas as variáveis explicativas

		Modelo 1			Modelo 2			Modelo3		
Número de obsevações		260880			260880			260880		
Prob > chi2		0			0			0		
Log pseudolikelihood		-130848			-128214			-123507		
Variável	Coeficiente	Erros Padrão	P - valor	Coeficiente	Erros Padrão	P - valor	Coeficiente	Erros Padrão	P - valor	
Remuneração esperada	0,00055	(0,00002)	-	0,00051	(0,00002)	-	0,00017	(0,00003)	-	
Probabilidade de Passar (PP)	1,632	(0,344)	-	3,297	(0,359)	-	2,213	(0,846)	0,01	
Grupo	Variável	Coeficiente	Erros Padrão	P - valor	Coeficiente	Erros Padrão	P - valor	Coeficiente	Erros Padrão	P - valor
Engenharias	mulher	-1,525	(0,054)	-	-1,458	(0,056)	-	-1,314	(0,059)	-
	PPI	-0,339	(0,061)	-	-0,167	(0,064)	0,01	-0,054	(0,069)	0,44
	cota	-0,547	(0,077)	-	-0,410	(0,078)	-	-0,093	(0,088)	0,29
	pai_edu_NS				0,597	(0,084)	-	0,229	(0,092)	0,01
	pai_edu_NM				0,356	(0,068)	-	0,138	(0,072)	0,05
	mae_edu_NS				1,198	(0,089)	-	0,486	(0,103)	-
	mae_edu_NM				0,758	(0,069)	-	0,305	(0,073)	-
	nota_port							0,000	(0,002)	0,87
	nota_mat							0,028	(0,002)	-
	particular_EM							0,561	(0,069)	-
	cursinho							-0,104	(0,059)	0,08
	tentativas_UnB							-0,453	(0,061)	-
	concluinte							0,660	(0,081)	-
	curso_regular							0,152	(0,073)	0,04
	cons	0,028	(0,041)	0,49	-1,035	(0,062)	-	-1,792	(0,103)	-
Ciências	mulher	-0,403	(0,063)	-	-0,345	(0,064)	-	-0,427	(0,067)	-
	PPI	-0,132	(0,062)	0,04	-0,042	(0,063)	0,50	-0,016	(0,066)	0,81
	cota	-0,241	(0,072)	0,00	-0,187	(0,073)	0,01	0,057	(0,075)	0,45
	pai_edu_NS				0,560	(0,084)	-	0,226	(0,090)	0,01
	pai_edu_NM				0,218	(0,070)	0,00	0,038	(0,072)	0,60
	mae_edu_NS				0,452	(0,092)	-	0,162	(0,097)	0,09

	mae_edu_NM			0,280	(0,069)	-	0,135	(0,072)	0,06	
	nota_port						0,001	(0,002)	0,60	
	nota_mat						0,020	(0,002)	-	
	particular_EM						0,411	(0,069)	-	
	cursinho						-0,200	(0,060)	0,00	
	tentativas_UnB						-0,376	(0,071)	-	
	concluente						0,682	(0,061)	-	
	curso_regular						0,457	(0,079)	-	
	_cons	-0,256	(0,045)	-	-0,725	(0,063)	-	-1,869	(0,106)	-
	mulher	0,900	(0,065)	-	0,938	(0,064)	-	1,037	(0,063)	-
	PPI	-0,298	(0,061)	-	-0,141	(0,062)	0,02	-0,060	(0,064)	0,35
	cota	-0,387	(0,071)	-	-0,233	(0,070)	0,00	-0,028	(0,071)	0,69
	pai_edu_NS				0,873	(0,097)	-	0,420	(0,096)	-
	pai_edu_NM				0,380	(0,067)	-	0,151	(0,069)	0,03
	mae_edu_NS				0,739	(0,097)	-	0,187	(0,106)	0,08
	mae_edu_NM				0,553	(0,068)	-	0,222	(0,069)	0,00
Saúde	nota_port						-0,005	(0,002)	0,00	
	nota_mat						0,019	(0,002)	-	
	particular_EM						0,806	(0,065)	-	
	cursinho						0,334	(0,059)	-	
	tentativas_UnB						-0,400	(0,066)	-	
	concluente						0,377	(0,061)	-	
	curso_regular						0,271	(0,069)	-	
	_cons	-0,285	(0,069)	-	-1,155	(0,077)	-	-2,300	(0,106)	-
	mulher	0,387	(0,047)	-	0,409	(0,049)	-	0,413	(0,052)	-
	PPI	-0,485	(0,052)	-	-0,313	(0,054)	-	-0,045	(0,059)	0,45
	cota	-0,268	(0,062)	-	-0,074	(0,064)	0,25	0,170	(0,067)	0,01
	pai_edu_NS				1,233	(0,071)	-	0,578	(0,079)	-
	pai_edu_NM				0,464	(0,061)	-	0,133	(0,065)	0,04
	mae_edu_NS				0,736	(0,087)	-	0,194	(0,096)	0,04
	mae_edu_NM				0,645	(0,060)	-	0,165	(0,064)	0,01
	nota_port						0,018	(0,002)	-	
	nota_mat						0,028	(0,002)	-	
	particular_EM						1,013	(0,064)	-	
	cursinho						0,113	(0,053)	0,03	
	tentativas_UnB						-0,753	(0,081)	-	
	concluente						0,130	(0,064)	0,04	

	curso_regular							-0,073	(0,070)	0,30
	_cons	0,342	(0,053)	-	-0,662	(0,064)	-	-2,285	(0,108)	-
Ciências Sociais	mulher	0,263	(0,046)	-	0,325	(0,047)	-	0,364	(0,048)	-
	PPI	-0,298	(0,051)	-	-0,132	(0,054)	0,02	-0,040	(0,056)	0,47
	cota	-0,304	(0,059)	-	-0,169	(0,059)	0,00	-0,112	(0,062)	0,07
	pai_edu_NS				0,836	(0,069)	-	0,424	(0,074)	-
	pai_edu_NM				0,266	(0,059)	-	0,056	(0,060)	0,35
	mae_edu_NS				0,822	(0,075)	-	0,372	(0,077)	-
	mae_edu_NM				0,423	(0,057)	-	0,147	(0,059)	0,01
	nota_port							0,032	(0,002)	-
	nota_mat							-0,008	(0,002)	-
	particular_EM							0,579	(0,060)	-
	cursinho							0,013	(0,048)	0,78
	tentativas_UnB							-0,209	(0,060)	-
	concluinte							0,486	(0,055)	-
	curso_regular							0,102	(0,059)	0,08
		_cons	0,552	(0,043)	-	-0,264	(0,054)	-	-1,658	(0,083)
Negócios	mulher	-0,321	(0,055)	-	-0,332	(0,055)	-	-0,377	(0,057)	-
	PPI	-0,141	(0,060)	0,02	-0,045	(0,062)	0,46	-0,055	(0,063)	0,38
	cota	-0,090	(0,069)	0,19	-0,021	(0,070)	0,76	0,047	(0,071)	0,51
	pai_edu_NS				0,636	(0,079)	-	0,387	(0,085)	-
	pai_edu_NM				0,141	(0,067)	0,04	0,007	(0,068)	0,92
	mae_edu_NS				0,383	(0,086)	-	0,062	(0,089)	0,49
	mae_edu_NM				0,136	(0,065)	0,04	-0,075	(0,068)	0,27
	nota_port							0,014	(0,002)	-
	nota_mat							0,001	(0,002)	0,53
	particular_EM							0,487	(0,064)	-
	cursinho							-0,026	(0,056)	0,64
	tentativas_UnB							-0,105	(0,062)	0,09
	concluinte							0,314	(0,063)	-
	curso_regular							0,205	(0,069)	0,00
		_cons	-0,476	(0,046)	-	-0,858	(0,060)	-	-1,559	(0,103)
Educação		base alternativa								
Outros	mulher	0,329	(0,049)	-	0,361	(0,049)	-	0,351	(0,049)	-
	PPI	-0,172	(0,053)	0,00	-0,093	(0,055)	0,09	-0,100	(0,056)	0,08
	cota	-0,066	(0,063)	0,29	0,015	(0,063)	0,81	-0,029	(0,064)	0,65

pai_edu_NS	0,469	(0,074)	-	0,316	(0,080)	-
pai_edu_NM	0,078	(0,060)	0,19	0,001	(0,060)	0,98
mae_edu_NS	0,371	(0,081)	-	0,178	(0,083)	0,03
mae_edu_NM	0,188	(0,058)	0,00	0,052	(0,059)	0,38
nota_port				0,017	(0,002)	-
nota_mat				-0,009	(0,002)	-
particular_EM				0,153	(0,061)	0,01
cursinho				-0,050	(0,050)	0,31
tentativas_UnB				0,298	(0,057)	-
concluinte				0,198	(0,053)	-
curso_regular				0,166	(0,061)	0,01
_cons	-0,484	(0,044)	-	-0,862	(0,056)	-
