



Universidade de Brasília – UnB

IE - Departamento de Estatística

Estágio Supervisionado II

Avaliação do rendimento e evasão de alunos cotistas e não cotistas da Universidade de Brasília

Edmar Ferreira Souto Mourão Bonfim

Relatório Final

Orientadora: Claudete Ruas

Coorientador: André Nunes Maranhão

Brasília

Junho de 2014

Edmar Ferreira Souto Mourão Bonfim

Avaliação do rendimento e evasão de alunos cotistas e não cotistas da Universidade de Brasília

Trabalho apresentado junto ao Curso de Estatística da Universidade de Brasília, para a conclusão do curso de Bacharelado em Estatística.

Orientadora: Claudete Ruas

Coorientador: André Nunes Maranhão

Brasília

Junho de 2014

Agradecimentos

Agradeço à minha família, por todo o suporte prestado durante a graduação, principalmente aos meus pais, que sempre me apoiaram nos desafios enfrentados. Agradeço também às minhas irmãs, Sammille e Camille, por todo o incentivo e compreensão na reta final dessa jornada.

À minha companheira e amiga, Sara, pelo apoio, incentivo e amor em tudo.

Aos professores Claudete e André, pela orientação neste trabalho. Obrigado pelo apoio, pelos conhecimentos compartilhados e por acreditarem em meu potencial.

A todos os professores do curso que de alguma forma contribuíram para a minha formação, especialmente aos professores George, Maria Teresa, Lúcio e James, que deixaram sua marca na minha trajetória.

Aos meus colegas de trabalho, na Codeplan e na Odds&Actions, pela compreensão, apoio e amizade.

Ao Decanato de Graduação da Universidade de Brasília, pelo fornecimento das bases de dados utilizada no presente trabalho.

Sumário

Resumo	3
1. Introdução	4
2. Objetivos	5
2.1 Objetivo Geral	5
2.2 Objetivos Específicos	5
3. Revisão Bibliográfica	6
3.1 Políticas de ações afirmativas	6
3.2 Definições Importantes	10
3.2.1 O Índice de Rendimento Acadêmico	10
3.2.2 Evasão	11
3.3 Métodos utilizados na pesquisa	13
3.3.1 Regressão Linear	13
3.3.2 Análise de dados em painel	17
3.3.3 Regressão Logística	20
4. Metodologia	23
5. Resultados	24
5.1 Instrumentalização das variáveis	24
5.2 Tratamento da base de Dados	25
5.3 Descrição da População	25
5.3.1 Perfil dos ingressantes do 2º Vestibular de 2004	26
5.3.2 Perfil dos ingressantes no 2º Vestibular de 2009	28
5.4 Características dos alunos segundo desempenho e evasão	31
5.4.1 Rendimento acadêmico dos alunos ingressantes no 2º Vestibular de 2004	31
5.4.1.1 Descrição do IRA do semestre	31
5.4.1.2 Descrição do IRA final dos alunos formados entre 2/2008 e 2/2010	33
5.4.2 Rendimento acadêmico dos alunos ingressantes no 2º Vestibular de 2009	36
5.4.2.1 Descrição do IRA do semestre	36
5.4.2.2 Descrição do IRA final dos alunos formados ou ativos	38
5.4.3 Evasão dos alunos ingressantes no 2º Vestibular de 2004	42
5.4.4 Evasão dos alunos ingressantes no 2º Vestibular de 2009	45
5.5 Modelagem	48
5.5.1 Análise do IRA	48

5.5.1.1	Análise do IRA final dos ingressantes no 2º Vestibular de 2004 utilizando regressão linear	48
5.5.1.2	Análise do IRA final dos ingressantes no 2º Vestibular de 2009 utilizando regressão linear	50
5.5.2.1	Análise do IRA do semestre dos ingressantes no 2º Vestibular de 2004 utilizando dados em painel	52
5.5.2.2	Análise do IRA do semestre dos ingressantes no 2º Vestibular de 2009 utilizando dados em painel	56
5.5.3	Análise da Evasão utilizando Regressão Logística	59
5.5.3.1	Análise da Evasão – Ingressantes no 2º Vestibular de 2004	59
5.5.3.2	Análise da Evasão – Ingressantes no 2º Vestibular de 2009	61
6.	Conclusão	66
7.	Bibliografia	69

Lista de Tabelas

Tabela 1: Distribuição de ingressantes por sistema de ingresso e sexo segundo grandes áreas - 2004	26
Tabela 2: Percentual de ingressantes por sistema de ingresso, segundo faixa etária - 2004	27
Tabela 3: Distribuição de ingressantes por sistema de ingresso e tipo de escola, segundo grandes áreas - 2004.....	27
Tabela 4: Razão entre os totais de alunos de escolas privadas e públicas por sistema de ingresso, segundo grandes áreas - 2004.....	28
Tabela 5: Distribuição dos ingressantes por sistema de ingresso e sexo, segundo grandes áreas - 2009	29
Tabela 6: Distribuição de ingressantes por sistema de ingresso, segundo a faixa etária- 2009	29
Tabela 7: Distribuição de alunos por tipo de escola e sistema de ingresso, segundo grandes áreas - 2009	30
Tabela 8: Razão entre os totais de alunos de escolas privadas e públicas por sistema de ingresso, segundo grandes áreas - 2009.....	30
Tabela 9: IRA do semestre médio por Sistema de Ingresso segundo o período - 2004	33
Tabela 10: Distribuição do IRA final médio dos formados entre 2/2008 e 2/2010 por sistema de ingresso, segundo grande área - 2004.....	34
Tabela 11: Distribuição do IRA final médio dos formados entre 2/2008 e 2/2010 por sistema de ingresso, segundo tipo de escola e curso médio - 2004.....	35
Tabela 12: IRA final médio dos formados entre 2/2008 e 2/2010 por sistema de ingresso, segundo maior nível de escolaridade dos pais - 2004.....	35
Tabela 13: IRA final médio dos formados entre 2/2008 e 2/2010 de cotistas e não cotistas por sistema de ingresso, segundo atividade remunerada e curso anterior - 2004.....	36
Tabela 14: IRA do semestre médio por sistema de ingresso, segundo semestre - 2009	38
Tabela 15: IRA final médio de alunos formados e ativos por sistema de ingresso segundo, grande área - 2009	39
Tabela 16: IRA final médio de alunos formados e ativos por sistema de ingresso, segundo tipo de escola e curso médio - 2009	40
Tabela 17: IRA final médio de alunos formados e ativos por sistema de ingresso segundo, maior nível de escolaridade dos pais - 2009	40
Tabela 18: IRA final médio de alunos formados e ativos por sistema de ingresso, segundo atividade remunerada e curso anterior - 2009.....	41
Tabela 19: Percentual de evadidos por sistema de ingresso, segundo ano - 2004	42
Tabela 20: Percentual de evadidos por sistema de ingresso, segundo faixa de idade - 2004	43

Tabela 21: Percentual de evadidos por sistema de ingresso, segundo grandes áreas - 2004	43
Tabela 22: Percentual de evadidos por sistema de ingresso, segundo tipo de escola e curso médio - 2004.....	44
Tabela 23: Percentual de evadidos por sistema de ingresso, segundo atividade remunerada e curso universitário- 2004.....	44
Tabela 24: Percentual de evadidos por sistema de ingresso, segundo ano - 2009	45
Tabela 25: Percentual de evadidos por sistema de ingresso, segundo faixa de idade	45
Tabela 26: Percentual de evadidos por sistema de ingresso, segundo grandes áreas - 2009	46
Tabela 27: Percentual de evadidos por sistema de ingresso, segundo tipo de escola e curso médio.....	46
Tabela 28: Percentual de evadidos por sistema de ingresso, segundo atividade remunerada e curso superior anterior	47
Tabela 29: Estimativas, Desvio Padrão, Estatísticas t do modelo de regressão - 2004.....	49
Tabela 30: Estimativas, Desvio Padrão, Estatísticas t do modelo de regressão - 2009.....	51
Tabela 31: Estimativas dos modelos em painel Pooled, Within e Efeitos aleatórios - 2004..	54
Tabela 32: Estimativas dos modelos em painel Pooled, Within e Efeitos aleatórios - 2009..	57
Tabela 33: Estimativa, Desvio Padrão e Estatísticas de Wald para regressão logística - 2004	59
Tabela 34: Estimativas das Razões de Chances para as variáveis significativas do modelo logístico - 2004	60
Tabela 35: Estimativas, Desvio Padrão, Estatísticas de Wald para o modelo logístico - 2009	62
Tabela 36: Estimativas das Razões de Chances para as variáveis significativas do modelo logístico - 2009	64

Lista de Figuras

Figura 1: IRA médio obtido no semestre - 2004	32
Figura 2: Boxplot do IRA do semestre por período, segundo sistema de ingresso - 2004....	32
Figura 3: Distribuição do IRA final dos formados entre 2/2008 e 2/2010 por sistema de ingresso, segundo Grande Área - 2004	34
Figura 4: IRA médio obtido no semestre - 2009	37
Figura 5: Boxplot do IRA do semestre por período, segundo sistema de ingresso - 2009....	37
Figura 6: Distribuição do IRA final dos alunos formados ou ativos por sistema de ingresso, segundo grande área - 2009.....	39
Figura 7: Resíduos Deviance do modelo logístico - 2004	60
Figura 8: Resíduos Deviance do modelo logístico - 2009	63

Resumo

Este trabalho é um estudo observacional com o objetivo de compreender os efeitos do sistema de cotas da universidade através da avaliação do rendimento acadêmico e da evasão de estudantes que ingressaram na UnB no 2º vestibular de 2004, logo após a implementação da política de cotas na Universidade de Brasília e no 2º Vestibular de 2009, cinco anos após a implementação da política. No estudo são utilizadas análise descritiva e modelagem estatística para identificar os principais fatores que interferem no rendimento acadêmico e na evasão dos alunos. Na realização da pesquisa, foram utilizadas bases de dados do pelo Centro de Seleção e Promoção de Eventos da UnB CESPE-UnB, com as informações relativas ao ingresso pelo vestibular e dados fornecidos pelo Decanato de Graduação da UnB, com informações do histórico do aluno na universidade.

Palavras-chave: Regressão linear, Regressão Logística, Dados em Painel, Sistema de Cotas, UnB, Rendimento Acadêmico, Evasão.

1. Introdução

A pauta de discussões sobre as desigualdades na educação superior sobre o prisma social e racial se ampliou após a adoção de políticas de cotas por grande parte das universidades públicas e a recente aprovação da Lei das Cotas, que já está sendo implementada gradualmente pelas universidades públicas federais e institutos técnicos federais, e devem reservar, no mínimo, 50% das vagas para estudantes que tenham cursado todo o ensino médio em escolas da rede pública. O sistema de cotas da Universidade de Brasília (UnB) instituído em 2003, e vigente entre 2004 e 2013, foi estabelecido através do Plano de Metas para a Integração Social, Ética e Racial da Universidade de Brasília, com duração de dez anos e reserva 20% das vagas do vestibular para negros.

O presente estudo estará centrado na análise de como se comportam o rendimento e a evasão de duas coortes de alunos no decorrer de seis anos. A análise levará em consideração a variabilidade contida nos dados relativos ao rendimento e evasão, os aspectos socioeconômicos dos estudantes e como esses aspectos impactam no comportamento dessas variáveis.

As coortes escolhidas se referem aos ingressantes no segundo vestibular de 2004 e aos ingressantes no segundo vestibular de 2009. Os dois grupos foram escolhidos com o intuito de captar dois momentos diferentes da política. O primeiro, com os alunos que entraram na universidade, assim que o sistema foi implementado e era pouco conhecido no Distrito Federal, e a segunda, com os alunos que foram aceitos na universidade após cinco anos de sua implementação. Outra razão que influenciou essa escolha foi a disponibilidade, em proporção razoável, de respostas ao Questionário Socioeconômico (QSE), feita no momento da inscrição no vestibular. Essas informações permitirão uma análise mais detalhada do desempenho e evasão dos estudantes cotistas e não cotistas.

Tendo como base as informações relativas ao rendimento e evasão e as variáveis socioeconômicas, pretende-se embasar as respostas a alguns questionamentos sobre quais são os efeitos da política de cotas na UnB:

1. Existem diferenças sistemáticas entre o desempenho acadêmico de alunos cotistas e não cotistas? Essa diferença, caso exista, se altera no decorrer do curso?
2. Existem diferenças no desempenho acadêmico dos alunos tomando em conta aspectos socioeconômicos?
3. Como se caracteriza a evasão dos alunos cotistas e não cotistas?
4. Características socioeconômicas influem nas chances de evasão?
5. Existem diferenças nos padrões observados de desempenho e evasão para alunos cotistas e não cotistas no início e no final da política de cotas?

De modo geral, este trabalho buscará complementar, atualizar e embasar as discussões em torno da política de cotas, que em 2013, completou dez anos e está sendo rediscutida por diversos segmentos dentro da Universidade de Brasília.

2. Objetivos

2.1 Objetivo Geral

Avaliar o rendimento e a evasão dos alunos cotistas e não cotistas de dois grupos de estudantes durante todo o período universitário. O primeiro é grupo será formado por alunos que ingressaram em 2004, logo após a implementação da política de cotas na Universidade de Brasília e o segundo, formado por aluno que entraram na universidade em 2009.

2.2 Objetivos Específicos

- Realizar revisão de literatura e aprofundamento sobre o que já foi produzido sobre avaliação das políticas de cotas nas universidades públicas brasileiras.
- Realizar tratamento dos dados fornecidos pelo Centro de Seleção e Promoção de Eventos da UnB CESPE e pelo Decanato de Graduação com o intuito de possibilitar as análises estatísticas adequadas.
- Estudar e compreender o comportamento do rendimento acadêmico dos estudantes cotistas e não cotistas ingressantes no 2º Vestibular de 2004 nos seis

anos seguintes por meio de análises descritivas e modelagem estatística, tomando em conta aspectos sociodemográficos, aspectos relacionados a trajetória pré-universitária e trajetória universitária.

- Estudar e compreender o comportamento do rendimento acadêmico dos estudantes cotistas e não cotistas ingressantes no 2º Vestibular de 2009, por meio de análises descritivas e modelagem estatística, tomando em conta aspectos sociodemográficos, aspectos relacionados a trajetória pré-universitária e trajetória universitária.
- Estudar e compreender o comportamento da evasão dos estudantes cotistas e não cotistas ingressantes no 2º Vestibular de 2004 nos seis anos seguintes, por meio de análises descritivas e modelagem estatística, tomando em conta aspectos sociodemográficos, aspectos relacionados a trajetória pré-universitária e trajetória universitária.
- Estudar e compreender o comportamento da evasão dos estudantes cotistas e não cotistas ingressantes no 2º Vestibular de 2009, por meio de análises descritivas e modelagem estatística, tomando em conta aspectos sociodemográficos, aspectos relacionados a trajetória pré-universitária e trajetória universitária.

3. Revisão Bibliográfica

Este capítulo está dividido em três tópicos essenciais. O primeiro aborda brevemente o histórico das políticas afirmativas, a política de cotas na Universidade de Brasília e alguns estudos que avaliam esses sistemas em universidades brasileiras. O segundo trata sobre os conceitos rendimento acadêmico e evasão no contexto da UnB, essenciais para o entendimento dos objetos da pesquisa. O terceiro capítulo aborda os três principais métodos estatísticos utilizados na pesquisa: Regressão linear, Análise de Dados em Painel e Regressão Logística.

3.1 Políticas de ações afirmativas

As políticas de ações afirmativas, incluídas as cotas para ingresso nas instituições de ensino superior estão presentes em diversos países a mais de 50 anos, sendo atribuída à Índia a história mais longa de políticas de preferência a grupos marginalizados,

desde 1947 (CARDOSO, 2008). Porém, essas políticas ganharam maior relevância em discussões políticas e acadêmicas nos Estados Unidos, quando, na década de 1960, os norte-americanos viviam um momento de reivindicações democráticas internas, expressas principalmente no movimento pelos direitos civis, cuja bandeira central era a extensão da igualdade de oportunidades a todos. Nesse período, as leis segregacionistas começaram a ser eliminadas no país, e o movimento negro surgia como uma das principais forças atuantes, com lideranças de projeção nacional, apoiado por liberais e progressistas brancos, unidos numa ampla defesa de direitos (MOEHLECKE, 2002). É nesse contexto que se desenvolve a ideia de ações afirmativas exigindo que o Estado, para além de garantir leis antissegregacionistas, viesse também a assumir uma postura ativa para a melhoria das condições da população negra. Algumas experiências semelhantes também ocorreram, de forma mais ou menos abrangente, em países da Europa Ocidental, em Israel, Austrália, Canadá, Nigéria, Paquistão, África do Sul, Argentina, Cuba, dentre outros.

No Brasil, apesar de existirem reserva de vagas para deficientes físicos em concursos públicos e cotas mínimas de candidatas mulheres nas eleições desde a década de 90, as ações afirmativas se manifestam, na maioria das vezes, por meio de reservas de vagas de universidades públicas. A primeira e maior ação afirmativa brasileira numa universidade teve início em 2001, na Universidade do Estado do Rio de Janeiro (UFRJ). Nesse sistema, são reservadas 50% das vagas para os alunos cotistas, divididas entre alunos negros ou indígenas, estudantes oriundos de escolas públicas e deficientes que atendem a uma condição de carência em termos de renda per capita mensal, que a instituição estipula ano a ano (MENDES, 2013).

Na Universidade de Brasília, o sistema de cotas foi instituído na forma de um plano de metas para integração social, étnica e racial da universidade, em junho de 2003. O plano previa três pontos fundamentais: reserva de vagas para negros e indígenas, incentivo à permanência do estudante que ingressa via política de cotas e um programa de apoio ao ensino público do Distrito Federal (CARDOSO, 2008). O sistema teve início no segundo vestibular de 2004 e estabelece reserva de 20% das vagas para estudantes negros, ou seja, estudantes que auto declaram como pretos ou pardos. Até o segundo vestibular de 2007, os candidatos que desejavam concorrer pelo sistema de cotas deveriam declarar-se negros e optar pelo sistema de cotas no momento da inscrição, quando seriam fotografados e assinariam declaração específica relativa aos requisitos exigidos para concorrer pelo sistema. O pedido de inscrição e a foto seriam analisados

por uma comissão que decidiria pela homologação, ou não, do pedido. Todo esse processo acontecia antes da realização das provas (CESPE/UnB, 2004).

Com o intuito de evitar fraudes, esses procedimentos foram modificados no primeiro vestibular de 2008. Atualmente, são feitas entrevistas pessoais, em data anterior à realização das provas de conhecimentos, em que são considerados a auto declaração de cor e os traços fenotípicos que o caracterizam como negro. Vale destacar que, caso as vagas do sistema de cotas não sejam preenchidas, são adicionadas às vagas do sistema universal dos respectivos cursos (CESPE, 2007).

A partir do primeiro semestre de 2013, por força da Lei Federal (Lei nº 7.824 de outubro de 2012) foi iniciada a implantação do sistema de cotas para escola pública. A meta é que até 2016, sejam reservadas 50% das vagas no ensino superior público para candidatos que tenham estudado integralmente em escolas públicas.

No dia 3 de Abril de 2014, o CEPE aprovou a redução da reserva de vagas pelo sistema de cotas raciais de 20% para 5% do total de alunos, decisão que já passa a valer para o 2º Vestibular de 2014, que ocorre em junho. Com isso, a Universidade de Brasília, passará a ter em 2016, 55% das vagas reservadas para alunos que se encaixam em uma das ações afirmativas.

Dentre os trabalhos que avaliam empiricamente o sistema de cotas na UnB, destaca-se CARDOSO (2008), dissertação de mestrado que usa os registros de admissão e acadêmicos da universidade para estudar o sistema de cotas. Com base em tabulações e comparações de médias, seus resultados indicam que as cotas aumentaram a proporção de estudantes negros na universidade. Também foi verificado que os alunos cotistas têm taxas de evasão menores do que as dos não cotistas, independente das razões da evasão e, que as notas dos alunos cotistas são similares as dos não cotistas, com exceção de alguns cursos, nos quais as notas são inferiores. Nesse trabalho, foi utilizada a metodologia de análise multivariada de classificação em árvore pela técnica CHAID (Chi-Squares automatic interaction detection), técnica que consiste em identificar as variáveis independentes associadas a uma variável dependente por meio da estatística qui-quadrado.

Também destaca-se TANNURI-PIANTO e FRANCIS (2012), artigo que aborda o acesso e a progressão dos alunos na Universidade de Brasília. No trabalho, utilizam estatísticas descritivas e comparações binárias para caracterizar a relação entre raça e condições socioeconômicas nas amostras de alunos da UnB, candidatos a UnB e

residentes do Distrito Federal. Para estudar o desempenho dos alunos cotistas da UnB em termos de notas e tempo dedicado ao estudo, utiliza-se técnicas de regressão. Dentre as evidências encontradas, também há o indicativo de que as cotas aumentaram a proporção de alunos não brancos e que a população estudantil não se tornou significativamente mais pobre do que seria sem a intervenção da política. Além disso, foram encontradas evidências de que, apesar de terem em geral notas de corte mais baixas no vestibular, os alunos admitidos pelo sistema de cotas tiveram desempenho acadêmico apenas marginalmente inferior ao de não cotistas.

Outro método utilizado em TANNURI-PIANTO e FRANCIS (2012) foi o *diferenças em diferenças* (DID). Essa técnica de avaliação de impacto de políticas públicas possibilita verificar a eficiência de um programa, ou seja, verificar se ele atingiu ou não, o impacto pretendido. Esse método requer dados em painel das unidades tratadas e não tratadas, com períodos de observação antes e depois da intervenção. Ele recebe esse nome porque duas diferenças são calculadas: diferença entre os períodos de tempo para cada unidade observada e a diferença entre as unidades para cada período de tempo. A principal hipótese é que a trajetória da variável resposta de resultado do grupo de controle reflita a do grupo de tratamento na ausência da intervenção (BARROS e FOGUEL, 2012). Neste caso, qualquer diferença entre os grupos que apareça após o programa pode ser interpretada como o impacto da intervenção. Essa metodologia foi utilizada em seu estudo para estimar o efeito do sistema de cotas no esforço pré-universitário e na identidade racial dos estudantes avaliados.

Em outras universidades, destacam-se algumas publicações. Dentre elas, MENDES (2013) avalia o sistema de cotas na Universidade Estadual do Rio de Janeiro (UERJ). Este estudo está dividido em três partes. Na primeira, o autor aborda questões relativas ao acesso à universidade e como as ações afirmativas afetam a composição do corpo estudantil (MENDES, 2013). Dentre os resultados obtidos, a partir da análise dos candidatos afetados pela política, foi possível perceber que o nível de renda dos beneficiados pela reserva de vagas é consideravelmente menor do que o dos prejudicados pelo sistema. Percebe-se também que a política afeta o ingresso de candidatos brancos de classe pobre e média advindos de colégios privados. Na segunda parte, MENDES (2013) analisa a progressão dos alunos cotistas e não cotistas com o intuito de responder como tem se mostrado o desempenho destes grupos. Os resultados mostraram que apesar de obterem rendimento inferior em termos de notas médias, os cotistas estão se formando a taxas maiores que o grupo não beneficiado pela política. Por

fim, o estudo analisa como outras políticas podem alterar as probabilidades de admissão de diferentes indivíduos. Para isto foram feitas simulações de outras políticas de reservas de vagas. Dentre os resultados encontrados, destaca o fato de que uma política voltada para alunos oriundos do ensino médio público pode aumentar a proporção de estudantes pardos e pretos em dois pontos percentuais (MENDES, 2013).

PEDROSA ET.AL (2007), contrariando os resultados citados anteriormente, mostra que os cotistas apresentaram um melhor desempenho relativo na Universidade Estadual de Campinas (UNICAMP) do que os não cotistas quando comparados os *rankings* de classificação no vestibular e ao final da graduação. Utilizando os dados do Exame Nacional de Desempenho de Estudantes (ENADE), WALTENBERG e CARVALHO (2012) revelam que há um diferencial de desempenho entre não cotistas e cotistas para as universidades públicas, porém não encontraram diferenças significativas entre os estudantes de IES privadas.

3.2 Definições Importantes

3.2.1 O Índice de Rendimento Acadêmico

O índice de rendimento acadêmico - IRA é utilizado para comparação dos resultados obtidos pelos alunos no decorrer da vida universitária e é calculado cumulativamente ao final de cada período letivo. O índice varia entre 0 e 5 e é construído a partir das menções obtidas nas disciplinas cursadas, considerando-se o número de créditos de cada disciplina bem como os trancamentos efetuados em disciplinas obrigatórias e optativas. O IRA é calculado de acordo com a seguinte fórmula:

$$IRA = \left[1 - \frac{(0,6 \times DT_b + 0,4 \times DT_p)}{DC} \right] \times \frac{\sum_i P_i \times CR_i \times Pe_i}{\sum_i CR_i \times Pe_i}$$

DT_b : Número de disciplinas obrigatórias trancadas

DT_p : Número de disciplinas optativas trancadas

DC : Número de disciplinas matriculadas

P_i : Peso da menção ($SS = 5$, $MS = 4$, $MM = 3$, $MI = 2$, $II = 1$, $SR = 0$)

Pe_i : Período em que uma dada disciplina foi cursada, obedecendo a seguinte limitação: $\min|6, período|$

CR_i : Número de créditos de uma dada disciplina

O IRA é utilizado pela Universidade como um dos critérios para matrícula em disciplinas, ou seja, quanto maior, maior a prioridade do aluno na obtenção de vaga em disciplinas. Outra utilização do índice é em seleções para certos programas oferecidos pela Universidade como, por exemplo, programas de intercâmbio, bolsas para projetos de iniciação científica e Programas de Educação Tutorial – PET.

De acordo com ARAÚJO (2013):

Esse índice tem grande relevância para o acompanhamento da vida acadêmica dos estudantes e, nesse sentido, é a variável que melhor retrata o desempenho dos alunos nos seus respectivos cursos. Entretanto, deve-se levar em conta que na sua construção o elemento fundamental é a menção obtida pelo estudante em cada disciplina cursada. As diversidades de disciplinas e de instrumentos de avaliação utilizadas pelos diferentes docentes indicam uma grande variabilidade de critérios de avaliação e de atribuição de menções finais, mesmo dentro de uma mesma disciplina onde são ofertadas diferentes turmas ou mesmo quando são cursadas em períodos distintos. Conseqüentemente, essas limitações devem ser consideradas em qualquer estudo sobre o desempenho de alunos com base no IRA.

No estudo, além do índice cumulativo, será usada a variável IRA do semestre, que é igual a média de todas as menções obtidas em determinado semestre.

$$IRA_{Sem_t} = \frac{\sum P_{it}}{DC_t}$$

P_{it} : Peso da menção obtida na matéria i , no semestre t .

DC : Número de disciplinas matriculadas no período t .

3.2.2 Evasão

A avaliação do fenômeno da evasão é essencial para acompanhar a efetividade da política de cotas na Universidade de Brasília que, de acordo com o Plano de Metas

para a Integração Social, Étnica e Racial, tem como um de seus pilares a permanência do estudante que ingressa na universidade via política de ações afirmativas.

A evasão estudantil é um fenômeno educacional complexo e não existe consenso entre os autores em relação ao seu conceito, o que dificulta comparações entre estudos que avaliam a questão no ensino superior e pode induzir a interpretações enganosas.

De acordo com RISTOFF (1995), o significado de evasão não é necessariamente ruim. Considera-se, por exemplo, que nem toda evasão significaria, obrigatoriamente, uma situação de exclusão ou fracasso. O autor procura diferenciar a evasão, onde haveria o abandono dos estudos, da mobilidade, que significa a migração do aluno para outro curso:

Parcela significativa do que chamamos de evasão (...) não é exclusão mas mobilidade, não é fuga, mas busca, não é desperdício mas investimento, não é fracasso – nem do aluno nem do professor, nem do curso ou da instituição, - mas tentativa de buscar o sucesso ou a felicidade, aproveitando as revelações que o processo natural de crescimento dos indivíduos faz sobre suas reais potencialidades. (1995: 56)

Como citado pela Comissão Especial do Ministério da Educação (BRASIL,1995), a proposição de Ristoff alerta para distinções de base, ou seja, quando se fala em evasão escolar no ensino superior importa referi-la a seus diferentes níveis dentro do sistema, como a seguir caracterizados:

- Evasão de curso: quando o estudante desliga-se do curso superior em situações diversas tais como: abandono (deixa de matricular-se), desistência (oficial), transferência ou reopção (mudança de curso), exclusão por norma institucional;
- Evasão da instituição: quando o estudante desliga-se da instituição na qual está matriculado;
- Evasão do sistema: quanto o estudante abandona de forma definitiva ou temporária o ensino superior.

Tendo como base o conceito de evasão definido por Dilvo Ristoff (1995), este trabalho determinou três diferentes formas de saída da universidade: formatura, evasão e mobilidade.

Para fazer parte do grupo de formados, o estudante satisfaz as seguintes condições: ser aprovado em todas as disciplinas obrigatórias do curso e integralizar o número de créditos exigidos.

No grupo de mobilidade estudantil estão as seguintes categorias de alunos: os que se desligaram voluntariamente, os que solicitaram transferência para outra IES, os que solicitaram mudança de curso ou mudança de habilitação e os que fizeram um novo vestibular.

E, finalmente, no grupo dos evadidos, variável a ser estudada neste trabalho, estão duas categorias de ex-alunos: os que foram desligados porque não cumpriram as condições exigidas pela UnB para continuidade dos estudos e os que foram desligados porque abandonaram o curso.

3.3 Métodos utilizados na pesquisa

3.3.1 Regressão Linear

A regressão linear, ou método dos mínimos quadrados ordinários (MQO), será utilizada para analisar como as variáveis independentes relacionadas a aspectos sociodemográficos, trajetória pré universitária e inserção universitária explicam a variável rendimento escolar, medida através do IRA final.

A regressão linear é um modelo estatístico de média condicional, no qual covariáveis condicionais se relacionam linearmente com a variável aleatória de interesse. O modelo é chamado de simples quando envolve uma relação funcional entre duas variáveis. O modelo é múltiplo quando admitimos que o valor da variável dependente é função linear de duas ou mais variáveis independentes. A regressão linear múltipla com k variáveis independentes é dada por:

$$Y_j = \alpha + \beta_1 X_{1j} + \beta_2 X_{2j} + \beta_3 X_{3j} + \dots + \beta_k X_{kj} + e_j, j = 1, \dots, n$$

$$Y_j = \alpha + \sum_{i=1}^k \beta_i X_{ij} + e_j$$

Onde Y_j é a variável resposta, X_{ij} são as variáveis explicativas independentes e e_j o erro. A equação também é composta pelos coeficientes de regressão β_j e pelo termo constante da equação β_0 .

Utilizando a notação matricial para representar a equação acima, tem-se:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

Onde,

$$\mathbf{y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & X_{11} & \dots & X_{k1} \\ 1 & X_{12} & \dots & X_{k2} \\ \dots & \dots & \dots & \dots \\ 1 & X_{1n} & \dots & X_{kn} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ \dots \\ b \end{bmatrix}, \mathbf{e} = \begin{bmatrix} e_1 \\ \dots \\ e_n \end{bmatrix}$$

As pressuposições do modelo são:

- I) A variável dependente Y_j é função linear das variáveis independentes X_{ij} .
- II) Os valores das variáveis independentes são fixos
- III) Os erros são não correlacionados entre si e possuem distribuição Normal com média $E(e_i) = 0$ e variância $E(e_i^2) = \sigma^2$

As estimativas dos parâmetros do modelo ($\beta_0, \beta_1, \dots, \beta_k$) são obtidas utilizando o método dos mínimos quadrados e são dadas por:

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

Onde,

$$\mathbf{b} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \beta_k \end{bmatrix}$$

Para avaliar a significância do modelo, é realizada a análise de variância (ANOVA) com o intuito de verificar se os parâmetros do modelo são significativos. A análise de variância é baseada na decomposição da soma de quadrados e nos graus de liberdade associados a variável resposta Y . O desvio de uma observação em relação à média pode

ser decomposto como o desvio da observação em relação ao valor ajustado pela regressão mais o desvio do valor ajustado em relação à média.

A expressão dessa decomposição é dada por

$$SQ_{Total} = SQ_{Erro} + SQ_{Regressão}$$

Logo, segue-se que a soma dos quadrados dos desvios é:

$$SQE = \mathbf{e}'\mathbf{e} = \mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y}$$

E, a soma de quadrados de regressão é dada por:

$$SQR = \mathbf{y}' - \frac{\sum(Y_j)^2}{n}$$

Com os valores das somas dos quadrados, podemos obter a tabela ANOVA.

Fonte	Graus de Liberdade	Soma de Quadrados	Quadrado Médio
<i>Regressão</i>	p	SQR	$QMR = \frac{SQR}{n - p - 1}$
<i>Erro</i>	$n - p - 1$	SQE	$QME = \frac{SQE}{p}$
<i>Total</i>	$n - 1$	SQT	

Uma medida de avaliação da capacidade explicativa geral do modelo é o coeficiente de determinação, que mede o percentual explicado da variação de interesse (variável resposta) em relação a variação das covariáveis. Nesse estudo, essa medida será importante para determinar se as variáveis socioeconômicas contribuem de alguma maneira para explicar a variação no rendimento do aluno.

O coeficiente de determinação da regressão múltiplo é dado por:

$$R^2 = \frac{SQR}{\mathbf{Y}'\mathbf{Y} - n\bar{Y}} = \frac{\sum(\hat{Y}_i - \bar{Y})}{\sum(Y_i - \bar{Y})}$$

Este valor retorna a razão do quadrado da correlação entre Y , o valor observado da variável resposta e \hat{Y} , seu valor predito. O valor de R^2 varia entre *zero* e *um* e indica a proporção da variação da variável Y explicada pela equação de regressão. Esse valor é normalmente usado como uma medida de ajuste da regressão em explicar a variação nos dados.

Contudo, um grande valor de R^2 não implica necessariamente que o modelo de regressão tenha um bom ajustamento, uma vez que a adição de uma variável aumenta sempre o valor deste coeficiente, sem ter em conta se variável que se adiciona é, ou não, estatisticamente significativa. Assim, modelos com um elevado valor de R^2 podem produzir estimativas pouco confiáveis do valor esperado de Y . Por este motivo, utiliza-se o coeficiente de determinação ajustado dado por:

$$R_a^2 = 1 - \left(\frac{n-1}{n-p} \right) (1 - R^2)$$

Este coeficiente fornece uma melhor ideia da proporção de variação de Y explicada pelo modelo de regressão considerando apenas variáveis independentes estatisticamente significantes. Ao contrário do que acontecia com o coeficiente de determinação múltiplo, o R_a^2 só aumenta se a adição da variável produz uma redução suficientemente grande na soma de quadrados do erro de maneira a que compense a perda de um grau de liberdade, resultante da adição da variável. Quando a diferença entre R^2 e R_a^2 é acentuada, há indícios de que tenham sido incluídos no modelo variáveis estatisticamente não significativas. Apesar disso, o valor dessa estatística deve ser visto apenas como uma indicação inicial de ajuste, ou não, da regressão (DRAPER e SMITH, 1981).

Para avaliar a significância global do modelo, ou seja, testar a existência do modelo, utiliza-se o Teste F-Snedecor:

$$F = \frac{QME}{QMR}$$

O valor F tem distribuição F de Fisher-Snedecor com $v_1 = p$ e $v_2 = n - p$ graus de liberdade e é obtido para testar a hipótese

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$$

A rejeição da hipótese nula implica que pelo menos uma das variáveis explicativas contribui significativamente para explicar a variação de Y .

Para testar a significância individual dos parâmetros de cada covariável, utiliza-se o teste t :

$$t = \frac{b_i}{s(b_i)}$$

Onde $s(b_i)$ é a variância da distribuição de b_i e t possui distribuição de t de Student com $n - p$ graus de liberdade.

3.3.2 Análise de dados em painel

A análise de Dados em Painel ou dados longitudinais é caracterizada por possuir observações em duas dimensões, tempo e o espaço. Ou seja, observações para um mesmo indivíduo, em diferentes momentos do tempo. A análise de dados em painel é utilizada para capturar efeitos de heterogeneidade individual no tempo.

Este tipo de dado contém informações que possibilitam uma melhor investigação sobre a dinâmica das mudanças nas variáveis, tornando possível considerar o efeito das variáveis não-observadas. Além disso, de acordo com BALTAGI (2008), existe a melhoria na inferência dos parâmetros estudados, pois propiciam mais graus de liberdade e maior variabilidade na amostra em comparação com dados em *cross-section* ou em séries temporais, o que refina algumas propriedades desejáveis dos estimadores.

Modelos de dados em painel examinam efeitos individuais com o intuito de lidar com a heterogeneidade que pode, ou não, ser observada. Esses efeitos podem ser fixos ou aleatórios. O modelo de efeitos fixos, examinam se o intercepto do modelo varia entre os indivíduos, enquanto o modelo de efeitos aleatórios explora as diferenças nos componentes da variância do erro entre os indivíduos (PARK, 2011).

A especificação de um modelo de dados em painel consiste em:

$$Y_{it} = \alpha_i + X_{it}\beta + \varepsilon_{it} \quad (1)$$

Em que α_i é um componente fixo ou aleatório que capta a heterogeneidade entre os indivíduos da análise, onde o subscrito i sugere que os interceptos podem ser diferentes em cada indivíduo. X_{it} representa o conjunto de variáveis explicativas e ε_{it} é o termo do erro.

De acordo com LOUREIRO (2009), no caso em que $Cov(\alpha_i, \varepsilon_{it}) \neq 0$, ou seja, há correlação entre o termo de cada indivíduo e a componente dos erros, para que se possa estimar essa equação consistentemente, a abordagem mais usual no contexto de dados longitudinais é a de Efeitos Fixos.

Neste método de estimação, mesmo permitindo que $Cov(\alpha_i, \varepsilon_{it}) \neq 0$, a ideia é eliminar o efeito não-observado α , baseado na seguinte suposição: $E(\varepsilon_{it}|x_i, \alpha_i) = 0$, onde $x_i = (x_{i1}, x_{i2}, \dots, x_{iT})$, conhecida como condição de exogeneidade estrita. A transformação de efeitos fixos (ou transformação *Within*) é obtida em dois passos.

Tirando-se a média da equação anterior no tempo obtêm-se:

$$\bar{y}_i = \bar{x}_i\beta + \alpha_i + \bar{\varepsilon}_i \quad (2)$$

E subtraindo a primeira da segunda para cada t , obtêm-se a equação transformada de efeitos fixos:

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)\beta + \varepsilon_{it} - \bar{\varepsilon}_i \quad (3)$$

removendo assim a heterogeneidade não-observada α_i .

O estimador de Efeitos Fixos é obtido ao se aplicar mínimos quadrados ordinários agrupados na equação acima. Sob a hipótese de exogeneidade estrita, esse estimador é consistente. Este estimador também é conhecido como estimador *Within*, por usar a variação do tempo dentro de cada unidade observacional. O estimador *Between*, obtido a partir das transformações anteriores é obtido ao se aplicar mínimos quadrados ordinários agrupados na equação (2), e leva em consideração somente a variação entre as unidades observacionais.

O modelo de efeitos fixos, também pode ser reescrito usando variáveis *dummies* para indicar os indivíduos. Nesse caso, o modelo é dado por:

$$y_{it} = \sum_{j=1}^N \alpha_j d_{ij} + x_{ij} \beta + \varepsilon_{it} \quad (4)$$

Com as *dummies* d_{ij} , onde $d_{ij} = 1$, se $i = j$. Os parâmetros são estimados via mínimos quadrados ordinários. O estimador β chamado de estimador *Least Square Dummy Variable*.

Outro método de estimação utilizado com dados em painel é o de Efeitos Aleatórios. Assim como nos MQO (Mínimos Quadrados Ordinários) agrupados, em uma análise de efeitos aleatórios, o efeito não-observado α_i é colocado junto com ao termo aleatório ε_i . Entretanto, impõe três suposições adicionais: $E(\varepsilon_{it}|x_i, \alpha_i) = 0$, $E(\alpha_i|x_i) = E(\alpha_i) = 0$, $Var(\alpha_i^2|x_i) = \sigma_i^2$. A primeira diz respeito a exogeneidade estrita. A segunda diz respeito a ortogonalidade entre α_i e cada x_i ser nula, ou seja, o termo do indivíduo não está correlacionado a nenhum regressor. A terceira se refere a homocedasticidade de α_i .

O modelo de efeitos fixos permite a existência de correlação entre os efeitos individuais não-observados com as variáveis incluídas. Entretanto, se esses efeitos forem estritamente não-correlacionados com as variáveis explicativas, pode ser mais apropriado modelar esses efeitos como aleatoriamente distribuídos entre as unidades observacionais, utilizando o modelo de efeitos aleatórios (LOUREIRO, 2009).

Assim, o ponto crucial na decisão de que modelo deve ser utilizado, se efeitos fixos ou aleatórios, reside na questão se α_i e x_i são correlacionados ou não. Esse questionamento deve ser feito de acordo com os dados que se está trabalhando, examinando suas especificidades. Um teste mais formal pode ser realizado, o Teste de Hausman, baseado nas diferenças das estimativas de efeitos fixos e aleatórios. (LOUREIRO, 2009). O teste examina se o efeito individual é não correlacionado com nenhum regressor do modelo.

De acordo com o teste, seja β_{EF} o vetor de estimativas de efeitos fixos e β_{EA} o vetor de estimativas de efeitos aleatórios, sob a hipótese nula de: $H_0 = \hat{\beta}_{EF} - \hat{\beta}_{EA} = 0$, com a seguintes estatística do teste:

$$H = [\hat{\beta}_{EF} - \hat{\beta}_{EA}]' [V(\hat{\beta}_{EF}) - V(\hat{\beta}_{EA})]^{-1} [\hat{\beta}_{EF} - \hat{\beta}_{EA}]$$

Possui distribuição χ^2 com $K - 1$ graus de liberdade. Se a estatística for significativa, há indícios de que o modelo de efeitos fixos é mais indicado para a análise dos dados. Nesse caso, os efeitos individuais são correlacionados com algum regressor e o modelo aleatório viola o pressuposto de Gauss-Markov e não é o melhor estimador linear não viesado. Isso ocorre porque, no modelo aleatório, o efeito individual faz parte do erro. No caso em que não se rejeita a hipótese nula, é preferível a utilização do modelo de efeitos aleatórios.

3.3.3 Regressão Logística

A regressão logística será utilizada para analisar como as variáveis independentes relacionadas a aspectos sociodemográficos, trajetória pré universitária e inserção universitária influenciam nas chances de evasão dos alunos da Universidade de Brasília.

O modelo de regressão logística descreve a relação entre uma variável resposta e uma ou mais variáveis explicativas independentes, também chamadas de covariáveis. Assim como no método de regressão linear, o objetivo é encontrar o melhor modelo que explique a relação entre as variáveis. Nesse modelo, a variável resposta do modelo é binária ou dicotômica (HOSMER e LEMESHOW, 2000).

A primeira diferença significativa entre a regressão logística e linear são os modelos paramétricos adotados, ou seja, a natureza da relação entre a variável resposta e as variáveis independentes. Nos problemas de regressão linear, no caso univariado, a esperança condicional de Y dado X é dada por $E(Y|x) = \beta_0 + \beta_1 x$ e pode tomar qualquer valor entre $-\infty$ e ∞ . Na regressão logística, onde a variável resposta é dicotômica, a esperança condicional, $E(Y|x)$, está entre zero e 1.

Com o intuito de simplificar a notação, será usada a quantidade $\pi(x) = E(Y|x)$ para representar a média condicional de Y dado o vetor de covariáveis x , quando a regressão logística for utilizada. A forma específica do modelo de regressão logística que será utilizada é igual a:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x + \dots + \beta_p x}}{1 + e^{\beta_0 + \beta_1 x + \dots + \beta_p x}}$$

A transformação Logito de $\pi(x)$, essencial no estudo de regressão logística é dada por:

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x + \dots + \beta_p x$$

A importância dessa transformação está no fato de $g(x)$ possuir diversas propriedades desejáveis de uma regressão linear. Ou seja, é linear nos parâmetros, pode ser contínuo e seu intervalo varia entre $-\infty$ e ∞ , dependendo do intervalo de x .

Como visto, o método de estimação dos parâmetros na regressão linear é o de mínimos quadrados. Para a estimação dos parâmetros na regressão logística, utiliza-se o método da máxima verossimilhança. De maneira geral, esse método retorna valores que maximizam a probabilidade de obter o conjunto de dados analisados.

As estimativas de $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)$, são dadas pelas seguintes equações de máxima verossimilhança:

$$\sum_{i=1}^n [y_i - \pi(x_i)] = 0$$

$$\sum_{i=1}^n x_{ij} [y_i - \pi(x_i)] = 0, \text{ para } j = 1, 2, \dots, p.$$

Como as expressões são não lineares em $\boldsymbol{\beta}$, necessitam métodos computacionais para obter os estimadores de máxima verossimilhança (EMV), $\hat{\boldsymbol{\beta}}$.

Na regressão linear, o interesse está focado no tamanho de SQR . Um valor alto sugere que as variáveis independentes são importantes e um valor baixo sugere que as variáveis não são essenciais ao modelo. A mesma lógica é aplicada aos modelos de regressão logística: comparar os valores observados da variável resposta para prever os valores obtidos dos modelos com e sem as variáveis em questão (HOSMER e LEMESHOW, 2000).

A comparação dos valores observados com os valores preditos é baseada na seguinte expressão:

$$D = -2 \ln \left[\frac{(EMV \text{ do modelo})}{(EMV \text{ do modelo saturado})} \right]$$

$$= -2 \sum_{i=1}^n \left[y_i \ln \left(\frac{\hat{\pi}_i}{y_i} \right) + (1 - y_i) \ln \left(\frac{1 - \hat{\pi}_i}{1 - y_i} \right) \right]$$

O modelo saturado é aquele que possui tantos parâmetros quanto dados observados.

A estatística D , faz o mesmo papel que SQE na regressão linear. Para verificar se os parâmetros são significativos, comparamos os valores de D com e sem as variáveis independentes. Logo, essa diferença faz papel semelhante ao teste F na regressão linear.

$$G = D(\text{modelosemvariável}) - D(\text{modelocomvariável})$$

Como o EMV do modelo saturado é igual para ambos valores de D , temos que

$$G = -2 \ln \left[\frac{(EMV \text{ sem as variáveis indep.})}{(EMV \text{ com as variáveis indep.})} \right]$$

Logo, sob a hipótese nula de que os valores do vetor $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)$ são iguais à zero, a estatística G segue distribuição Qui-Quadrado com p graus de liberdade.

Para testar hipóteses acerca de cada coeficiente e determinar o potencial de cada variável no modelo, é utilizado o Teste de Wald. A estatística do teste é dada por:

$$W = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$$

Onde $SE(\hat{\beta}_j)$ é igual ao desvio padrão dos coeficientes estimados e W segue distribuição normal padrão.

4. Metodologia

As fontes de informações do trabalho são os dados relativos ao processo seletivo da Universidade de Brasília, fornecidos pelo Centro de Seleção e Promoção de Eventos da UnB (CESPE/UnB) e os bancos de dados referentes ao rendimento acadêmico dos alunos da UnB, fornecidos pela Centro de Processamento de Dados da Universidade de Brasília (CPD/UnB) com autorização do Decanato de Graduação

Na base relativa ao vestibular, foram identificadas as variáveis curso, sexo, data de nascimento, sistema de entrada e argumento final no vestibular e para os candidatos que responderam, informações socioeconômicas contidas no Questionário Socioeconômico (QSE). As variáveis socioeconômicas utilizadas nas análises do trabalho são: tipo de estabelecimento de ensino - se é público ou privado -, faixa de renda familiar, escolaridade do pai e da mãe, se exercia atividade remunerada, qual curso médio concluiu ou estava concluindo (curso regular, supletivo, ou outros), se já fez - e onde fez - curso superior.

Nos dados fornecidos pelo CPD, foram identificadas informações referentes ao o Índice de Rendimento Acadêmico (IRA) dos alunos em cada semestre até o 1º semestre de 2013, o IRA acumulado no 1º semestre de 2013 ou no semestre de saída, o curso e a forma de saída da universidade.

Para uma melhor avaliação dos fatores importantes para explicar o rendimento acadêmico e a evasão, as variáveis selecionadas foram divididas em: sociodemográficas - idade, sexo, faixa de renda familiar e escolaridade dos pais -, trajetória pré-universitária - tipo de instituição em que fez ensino médio, se exercia alguma atividade remunerada, qual curso médio cursou e se já havia universitário anteriormente -, e inserção universitária – argumento final no vestibular, curso e forma de saída.

Inicialmente, houve o tratamento e cruzamento das bases utilizadas no trabalho com o intuito de verificar e corrigir possíveis inconsistências. Em seguida, foi realizada uma caracterização da distribuição dos cotistas e não cotistas de acordo com algumas variáveis sociodemográficas, de trajetória pré-universitária e de inserção universitária. Na análise descritiva foram estudados os comportamentos do IRA médio do semestre, IRA médio final e evasão de cotistas e não cotistas no tempo e de acordo com os fatores já citados. Posteriormente, a regressão linear e a análise de dados em painel foram utilizadas para verificar quais dessas características estão relacionadas ao IRA final e ao

IRA do semestre dos alunos. Além disso, a regressão logística foi utilizada para verificar quais fatores estão associados a evasão dos alunos na Universidade de Brasília.

O software utilizado para auxiliar na criação de tabelas, gráficos, testes estatísticos e modelagem foi o R Project, versão 3.0.1.

5. Resultados

5.1 Instrumentalização das variáveis

Nesta pesquisa, algumas variáveis foram operacionalizadas de modo a permitir uma análise mais adequada e acurada dos dados.

Tendo como base a data de nascimento, foi calculada a idade de todos os estudantes no início de cada semestre, tornando-se uma variável não constante para cada indivíduo no tempo.

Com o intuito de compreender o rendimento de acordo com a área do conhecimento e agrupar a grande quantidade de cursos de graduação existentes na universidade, a variável curso foi agrupada em oito grandes áreas, classificadas de acordo com a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes): Ciências Exatas e da Terra; Ciências Biológicas; Engenharias; Ciências da Saúde; Ciências Agrárias; Ciências Sociais Aplicadas; Ciências Humanas; Linguística, Letras e Artes e; Multidisciplinar (cursos presentes apenas para a coorte de alunos ingressantes no 2º Vestibular de 2009).

Nas análises iniciais, foi verificado que há correlação entre as variáveis nível de escolaridade do pai e da mãe. Com o intuito de evitar multicolinearidade nos modelos e redundância na descrição dos dados, as duas variáveis foram transformadas em apenas uma, que indica qual o maior nível de escolaridade verificado entre os pais.

Para o grupo de alunos ingressantes no 2º Vestibular de 2004, a variável tipo de escola é derivada da pergunta: “Em que tipo de estabelecimento de ensino você fez o curso médio?”, presente no QSE. Os alunos que responderam que estudaram todo ou maior parte em escola pública, dentro ou fora do Distrito Federal, foram considerados alunos de escolas públicas. Já os alunos que responderam que estudaram todo ou maior parte em escola particular, dentro ou fora do DF, foram considerados alunos de escolas

privadas. Em relação aos alunos da coorte de 2009, essa informação foi fornecida pelo CPD.

A variável forma de saída, disponibilizada pelo CPD, permite identificar as diferentes situações dos estudantes que deixaram o curso de graduação. No grupo de mobilidade estudantil estão as seguintes categorias de alunos: os que se desligaram voluntariamente, os que solicitaram transferência para outra IES, os que solicitaram mudança de curso ou mudança de habilitação e os que fizeram um novo vestibular. No grupo dos evadidos estão duas categorias de ex-alunos: os que foram desligados porque não cumpriram as condições exigidas pela UnB para continuidade dos estudos e os que foram desligados porque abandonaram o curso.

5.2 Tratamento da base de Dados

Após o cruzamento das bases relativas aos ingressantes no 2º Vestibular de 2004, dentre os 2121 aprovados, considerando todas as chamadas, 268 nomes não constavam na base de dados fornecida pelo CPD, o que corresponde a 12,6% dos registros. Dentre os 1853 estudantes restantes, a população estudada nesta coorte, foi verificado que 321 já estavam na UnB anteriormente, sendo que 62 entraram como cotistas e 259 pelo sistema universal.

Em relação aos dados referentes ao 2º Vestibular de 2009, foi verificado que dentre os 3578 aprovados constantes na base de dados, levando em conta todas as chamadas, 397 inscrições não constavam na base de dados fornecida pelo CPD, o que corresponde a 11% dos registros. Dentre os 3181 estudantes restantes, a população estudada nesta coorte, foi verificado que 654 já estavam na Universidade anteriormente, sendo que 122 entraram como cotistas e 528 ingressaram pelo sistema universal.

Para manter o balanceamento dos dados, foram desconsiderados da análise os IRAs obtidos nos períodos de Verão. Para a coorte iniciada em 2004, os alunos desse grupo cursaram matérias no Verão entre os anos 2007 e 2010 e na coorte dos ingressantes em 2009, o verão ocorreu nos anos de 2010 e 2012.

5.3 Descrição da População

A descrição dos ingressantes dos vestibulares do 2º semestre de 2004 e do 2º semestre de 2009 tem como objetivo identificar características que diferenciam e aproximam alunos que entraram pelo sistema universal, não cotistas, e alunos que entraram pelo sistema de cotas, cotistas, com base em variáveis sociodemográficas (idade, sexo, faixa de renda familiar, etc), trajetória pré-universitária (tipo de instituição em que cursou ensino médio) e inserção universitária (grande área ao qual o curso pertence).

5.3.1 Perfil dos ingressantes do 2º Vestibular de 2004

O grupo de estudantes ingressantes no 2º Vestibular de 2004 é composto por 1853 alunos. Dentre eles, 719 são do sexo feminino e 1134 do sexo masculino. Também observa-se que 353 (19%) dos estudantes são cotistas e 1500 (81%) são não cotistas.

A tabela 1 mostra a quantidade e o percentual dos alunos cotistas e não cotistas por sexo de acordo com a grande área a qual pertence seu curso. A tabela indica que a distribuição de cotistas e não cotistas por sexo entre as grandes áreas são semelhantes. Destaca-se que, entre os ingressantes dos cursos de ciências exatas e engenharias, 82,6% e 87,3%, respectivamente, dos estudantes são do sexo masculino.

Tabela 1: Distribuição de ingressantes por sistema de ingresso e sexo segundo grandes áreas - 2004

Área Grande	Cotista		Não Cotista		Total	
	Feminino	Masculino	Feminino	Masculino	Feminino	Masculino
Ciências Agrárias	7 (31,8%)	15 (68,2%)	39 (44,3%)	49 (55,7%)	46 (41,8%)	64 (58,2%)
Ciências Biológicas	5 (71,4%)	2 (28,6%)	22 (53,7%)	19 (46,3%)	27 (56,2%)	21 (43,8%)
Ciências da Saúde	18 (50%)	18 (50%)	65 (48,9%)	68 (51,1%)	83 (49,1%)	86 (50,9%)
Ciências Exatas e da Terra	12 (21,4%)	44 (78,6%)	38 (16,4%)	194 (83,6%)	50 (17,4%)	238 (82,6%)
Ciências Humanas	33 (47,1%)	37 (52,9%)	148 (53,4%)	129 (46,6%)	181 (52,2%)	166 (47,8%)
Ciências Sociais Aplicadas	40 (42,1%)	55 (57,9%)	164 (42,6%)	221 (57,4%)	204 (42,5%)	276 (57,5%)
Engenharias	3 (10,3%)	26 (89,7%)	19 (13,2%)	125 (86,8%)	22 (12,7%)	151 (87,3%)
Letras e Artes, Linguística	19 (50%)	19 (50%)	87 (43,5%)	113 (56,5%)	106 (44,5%)	132 (55,5%)
Total	137 (38,8%)	216 (61,2%)	582 (38,8%)	918 (61,2%)	709 (38,8%)	1134 (61,2%)

A tabela 2 mostra o percentual de ingressantes por sistema de ingresso, segundo faixa etária. Nela, é possível constatar que existem diferenças entre a idade dos ingressantes cotistas e não cotistas. Dentre os cotistas, 62,2% tinham até 21 anos e 27,8% tinham entre 21 e 27 anos. Para os alunos que ingressaram pelo sistema universal essas proporções são iguais a 70,1% e 18,9%, respectivamente. Esse fato indica que

alunos cotistas entraram mais velhos no ensino superior nesse primeiro vestibular com o sistema de cotas.

Tabela 2: Percentual de ingressantes por sistema de ingresso, segundo faixa etária - 2004

Faixa etária	Cotista	Não Cotista	Total
Até quatorze anos	0,0%	0,1%	0,1%
Entre 15 e 17 anos	11,6%	17,7%	16,5%
Entre 18 e 21 anos	51,0%	52,3%	52,1%
Entre 22 e 24 anos	18,7%	13,1%	14,1%
Entre 25 e 27 anos	9,1%	5,9%	6,5%
Entre 28 e 30 anos	3,4%	3,6%	3,6%
Entre 31 e 40 anos	4,8%	5,1%	5,1%
Mais de 40 anos	1,4%	2,2%	2,1%
Total	100,0%	100,0%	100,0%

A tabela 3 mostra a distribuição do tipo de escola cursada pelo aluno durante maior parte do ensino médio, por sistema de ingresso, segundo grande área ao qual o curso pertence. A variável tipo de escola possui uma quantidade considerável de não resposta – ingressantes que não responderam ao questionário -, principalmente para os alunos que ingressaram pelo sistema universal. Para facilitar a interpretação e visualização dos dados, a tabela 4 fornecerá a razão entre os totais de alunos de escolas privadas e públicas por sistema de ingresso, segundo grande área.

Tabela 3: Distribuição de ingressantes por sistema de ingresso e tipo de escola, segundo grandes áreas - 2004

Variáveis	Cotista			Não Cotista		
	Particular	Pública	Não respondeu	Particular	Pública	Não respondeu
Ciências Agrárias	14 (63,6%)	8 (36,4%)	0 (0%)	41 (46,6%)	13 (14,8%)	34 (38,6%)
Ciências Biológicas	3 (42,8%)	4 (57,1%)	0 (0%)	20 (48,8%)	6 (14,6%)	15 (36,6%)
Ciências da Saúde	23 (63,9%)	10 (27,8%)	3 (8,3%)	72 (54,1%)	18 (13,5%)	43 (32,3%)
Ciências Exatas e da Terra	21 (37,5%)	34 (60,7%)	1 (1,8%)	67 (28,9%)	44 (18,9%)	121 (52,2%)
Ciências Humanas	17 (24,3%)	46 (65,7%)	7 (10%)	77 (27,8%)	62 (22,4%)	138 (49,8%)
Ciências Sociais						
Aplicadas	42 (44,2%)	46 (48,4%)	7 (7,4%)	136 (35,3%)	70 (18,2%)	179 (46,5%)
Engenharias	16 (55,2%)	10 (34,5%)	3 (10,3%)	83 (57,6%)	16 (11,1%)	45 (31,3%)
Letras e Artes, Linguística	13 (34,2%)	25 (65,8%)	0 (0%)	48 (24%)	53 (26,5%)	99 (49,5%)
TOTAL	149 (42,2%)	183 (51,8%)	21 (5,9%)	544 (36,3%)	282 (18,8%)	674 (44,9%)

De acordo com a tabela abaixo, é possível verificar que essa razão é maior entre os ingressantes do sistema universal em todas as grandes áreas. Também é observado que, com exceção das grandes áreas Ciências da Saúde, Ciências Agrárias e

Engenharias, todas as outras possuem mais alunos cotistas oriundos de escolas públicas do que de escolas privadas. Em relação aos não cotistas, esse fato é observado apenas na grande área Letras e Artes.

Tabela 4: Razão entre os totais de alunos de escolas privadas e públicas por sistema de ingresso, segundo grandes áreas - 2004

Variáveis	Cotista	Não Cotista	Total
Ciências Agrárias	1,75	3,15	2,62
Ciências Biológicas	0,75	3,33	2,30
Ciências da Saúde	2,30	4,00	3,39
Ciências Exatas e da Terra	0,62	1,52	1,13
Ciências Humanas	0,37	1,24	0,87
Ciências Sociais Aplicadas	0,91	1,94	1,53
Engenharias	1,60	5,19	3,81
Letras e Artes, Linguística	0,52	0,91	0,78
TOTAL	0,81	1,93	1,49

5.3.2 Perfil dos ingressantes no 2º Vestibular de 2009

O grupo de alunos ingressantes no 2º Vestibular de 2009 é composto por 3181 estudantes. Dentre eles, 1479 são do sexo feminino e 1762 do sexo masculino. Também observa-se que 597 (18,7%) dos estudantes são cotistas e 2584 (81,3%) são não cotistas.

A tabela 5 mostra a distribuição de alunos por sistema de ingresso e sexo, segundo grande área. Assim como em 2004, é possível verificar que a distribuição de cotistas e não cotistas por sexo entre as grandes áreas são bastante semelhantes. Além do destaque para as engenharias e os cursos da área de ciências exatas e da terra com, respectivamente 82,3% e 80,4% de alunos do sexo masculino, é possível verificar que em Ciências da Saúde e Letras e Artes e cursos multidisciplinares possuem mais de 60% do corpo discente formado por mulheres.

Tabela 5: Distribuição dos ingressantes por sistema de ingresso e sexo, segundo grandes áreas - 2009

Variáveis	Cotista		Não Cotista		Total	
	Feminino	Masculino	Feminino	Masculino	Feminino	Masculino
Ciências Agrárias	14 (45,2%)	17 (54,8%)	67 (43,8%)	86 (56,2%)	81 (44,0%)	103 (56%)
Ciências Biológicas	12 (54,5%)	10 (45,5%)	44 (52,4%)	40 (47,6%)	56 (52,8%)	50 (47,2%)
Ciências da Saúde	54 (63,5%)	31 (36,5%)	234 (67,0%)	115 (33%)	288 (66,4%)	146 (33,6%)
Ciências Exatas e da Terra	15 (21,1%)	56 (78,9%)	58 (19,2%)	244 (80,8%)	73 (19,6%)	300 (80,4%)
Ciências Humanas	61 (54,4%)	47 (45,6%)	211 (51,6%)	198 (48,4%)	267(52,1%)	245 (47,8%)
Ciências Sociais Aplicadas	61 (51,3%)	58 (48,7%)	220 (46,6%)	252 (53,4%)	281 (47,5%)	310 (52,5%)
Engenharias	21 (21%)	79 (79%)	70 (16,9%)	345 (83,1%)	91 (17,7%)	424 (82,3%)
Letras e Artes, Linguística	35 (62,5%)	21 (37,5%)	189 (59,6%)	128 (40,4%)	224 (60,1%)	148 (39,9%)
Multidisciplinar	7 (70%)	3 (30%)	51 (61,4%)	32 (38,6%)	58 (62,4%)	35 (37,6%)
Total	137 (38,8%)	216 (61,2%)	582 (38,8%)	918 (61,2%)	709 (38,8%)	1134 (61,2%)

A tabela 6 mostra a distribuição de ingressantes por sistema de ingresso, segundo a faixa etária. Nela, é possível constatar que existem diferenças entre a idade dos ingressantes cotistas e não cotistas até 24 anos, porém, de maneira menos acentuada do que ocorreu em 2004. Dentre os cotistas, 68% tinham até 21 anos e 13,7% tinham entre 21 e 24 anos. Já entre os alunos do sistema universal, 73,1% tinham até 21 anos e 9,6% possuíam entre idade entre 21 e 24. De maneira geral, é possível perceber maior homogeneidade entre as distribuições das idades dos cotistas e não cotistas quando comparado com 2004 e, além disso, maior percentual de ingressantes mais novos, passando de 16,5% estudantes com até 18 em 2004, para 26,5% em 2009.

Tabela 6: Distribuição de ingressantes por sistema de ingresso, segundo a faixa etária- 2009

Faixa Etária	Cotista	Não cotista	Total
Entre 15 e 17 anos	27,0%	26,4%	26,5%
Entre 18 e 21 anos	41,0%	47,1%	45,9%
Entre 22 e 24 anos	13,7%	9,6%	10,3%
Entre 25 e 27 anos	6,0%	5,8%	5,8%
Entre 28 e 30 anos	5,9%	4,0%	4,3%
Entre 31 e 40 anos	5,4%	4,6%	4,7%
Mais de 40 anos	1,0%	2,6%	2,3%

A tabela 7 mostra a distribuição de ingressantes por tipo de escola onde fez maior parte do ensino médio e sistema de ingresso, segundo grande área. Foi verificado que a

não resposta é inferior ao que ocorreu em 2009 e aparece de forma mais equilibrada entre cotistas e não cotistas – 12,6% e 9,6% respectivamente. Para facilitar a interpretação e visualização dos dados, a tabela 8 fornecerá a razão entre os totais de alunos de escolas privadas e públicas por sistema de ingresso, segundo grande área.

Tabela 7: Distribuição de alunos por tipo de escola e sistema de ingresso, segundo grandes áreas - 2009

Variáveis	Cotista			Não Cotista		
	Particular	Pública	Sem resposta	Particular	Pública	Sem resposta
Ciências Agrárias	18 (58,1%)	13 (41,9%)	0 (0%)	107 (69,9%)	46 (30,1%)	0 (0%)
Ciências Biológicas	8 (36,4%)	14 (63,6%)	0 (0%)	66 (78,6%)	18 (21,4%)	0 (0%)
Ciências da Saúde	43 (50,6%)	39 (45,9%)	3 (3,5%)	245 (70,2%)	91 (26,1%)	13 (3,7%)
Ciências Exatas e da Terra	24 (33,8%)	30 (42,3%)	17 (23,9%)	180 (59,6%)	73 (24,2%)	49 (16,2%)
Ciências Humanas	42 (40,8%)	47 (45,6%)	14 (13,6%)	245 (59,9%)	123 (30,1%)	41 (10,%)
Ciências Sociais Aplicadas	57 (47,9%)	39 (32,8%)	23 (19,3%)	281 (59,5%)	115 (24,4%)	76 (16,1%)
Engenharias	62 (62%)	32 (32%)	6 (6%)	324 (78,1%)	80 (19,3%)	11 (2,7%)
Letras e Artes, Linguística	13 (23,2%)	33 (58,9%)	10 (17,9%)	163 (51,4%)	109 (34,4%)	45 (14,2%)
Multidisciplinar	5 (50%)	3 (30%)	2 (20%)	22 (26,5%)	48 (57,8%)	13 (15,7%)
Total	272 (45,6%)	250 (41,9%)	75 (12,6%)	1633 (63,2%)	703 (27,2%)	248 (9,6%)

De acordo com a tabela 8, a única grande área que não possui razão maior que um, ou seja, possui mais alunos de escolas públicas do que privadas, entre os ingressantes não cotistas é a Multidisciplinar. Além disso, para os cotistas, dentre as nove grandes áreas avaliadas, quatro possuem razão menor do um, ou seja, possuem mais alunos de escolas públicas do que de escolas particulares.

Tabela 8: Razão entre os totais de alunos de escolas privadas e públicas por sistema de ingresso, segundo grandes áreas - 2009

Curso	Cotista	Não Cotista	Total
Ciências Agrárias	1,38	2,33	2,12
Ciências Biológicas	0,57	3,67	2,31
Ciências da Saúde	1,10	2,69	2,22
Ciências Exatas e da Terra	0,80	2,47	1,98
Ciências Humanas	0,89	1,99	1,69
Ciências Sociais Aplicadas	1,46	2,44	2,19
Engenharias	1,94	4,05	3,45
Letras e Artes, Linguística	0,39	1,50	1,24
Multidisciplinar	1,67	0,46	0,53
Total	1,09	2,32	2,00

Na avaliação dos dois grupos de ingressantes, a proporção de alunos por sexo se manteve semelhante. Porém, em 2009 os cursos Ciências da Saúde e Letras e Artes e

cursos multidisciplinares se destacaram por possuírem mais de 60% dos alunos do sexo feminino.

Destaca-se também que em 2004, os alunos ingressavam com mais idade na universidade, principalmente no grupo de alunos cotistas. Logo, uma das hipóteses levantadas é que alunos que estavam na UnB anteriormente, refizeram o vestibular devido a nova oportunidade dada pelo sistema de cotas.

Também é possível verificar que, acompanhando o crescimento no número de vagas disponíveis, houve um aumento no percentual de alunos oriundos de escolas particulares entre cotistas e não cotistas. Entre os cotistas, a razão entre a proporção de alunos de escola pública sobre a proporção de alunos de escolas particulares era igual a 0,81 em 2004 e passou a ser 1,09 em 2009. Entre os alunos não cotistas, essa razão passou de 1,93 em 2004 para 2,32 em 2009.

5.4 Características dos alunos segundo desempenho e evasão

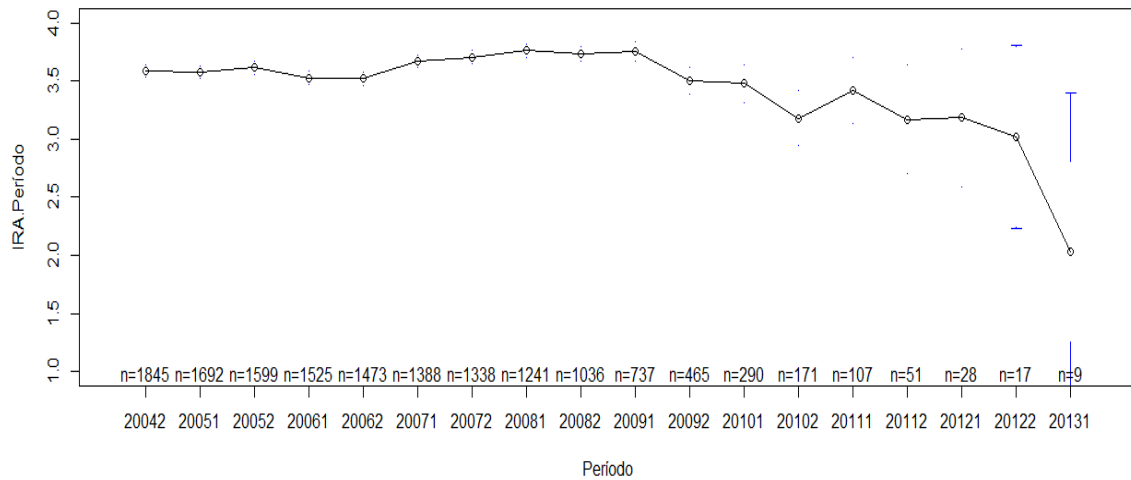
Para melhor compreensão do desempenho e da evasão das populações estudadas e identificação de eventuais semelhanças e diferenças entre cotistas e não cotistas, serão relacionadas as variáveis estudadas com o IRA do semestre, o IRA final e com a evasão dos ingressantes no 2º Vestibular de 2004 e no 2º Vestibular de 2009.

5.4.1 Rendimento acadêmico dos alunos ingressantes no 2º Vestibular de 2004

5.4.1.1 Descrição do IRA do semestre

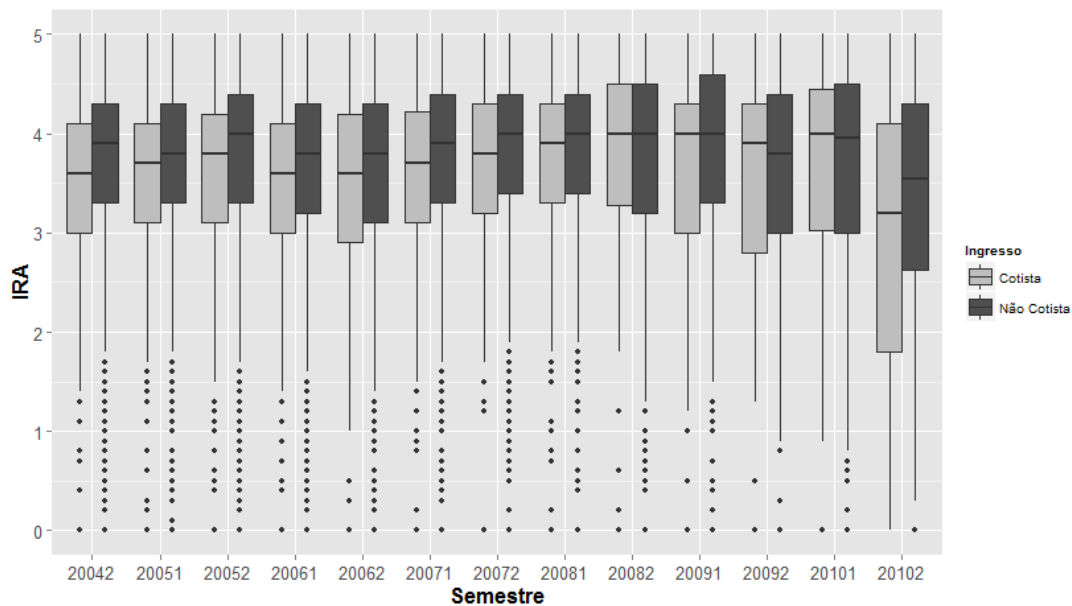
Ao avaliar o IRA médio semestral de todos os alunos, é possível verificar uma tendência crescente até o 1º semestre de 2009, 9 semestres após o ingresso na universidade, para um contínuo decréscimo no decorrer dos anos, até o 1º semestre de 2013.

Figura 1: IRA médio obtido no semestre - 2004



Para analisar o IRA do semestre entre cotistas e não cotistas, foi utilizado o gráfico Boxplot. De acordo com a figura 2, é possível verificar que a distribuição dos IRAs dos cotistas é inferior à dos não cotistas no decorrer dos semestres, porém, a partir do 2º semestre de 2007, essa diferença passa a ser menor.

Figura 2: Boxplot do IRA do semestre por período, segundo sistema de ingresso - 2004



Corroborando com a visualização gráfica, de acordo com a tabela 9, verifica-se que a diferença percentual entre o IRA médio de alunos não cotistas e cotistas diminui no decorrer dos semestres a partir de 2007.

Tabela 9: IRA do semestre médio por Sistema de Ingresso segundo o período - 2004

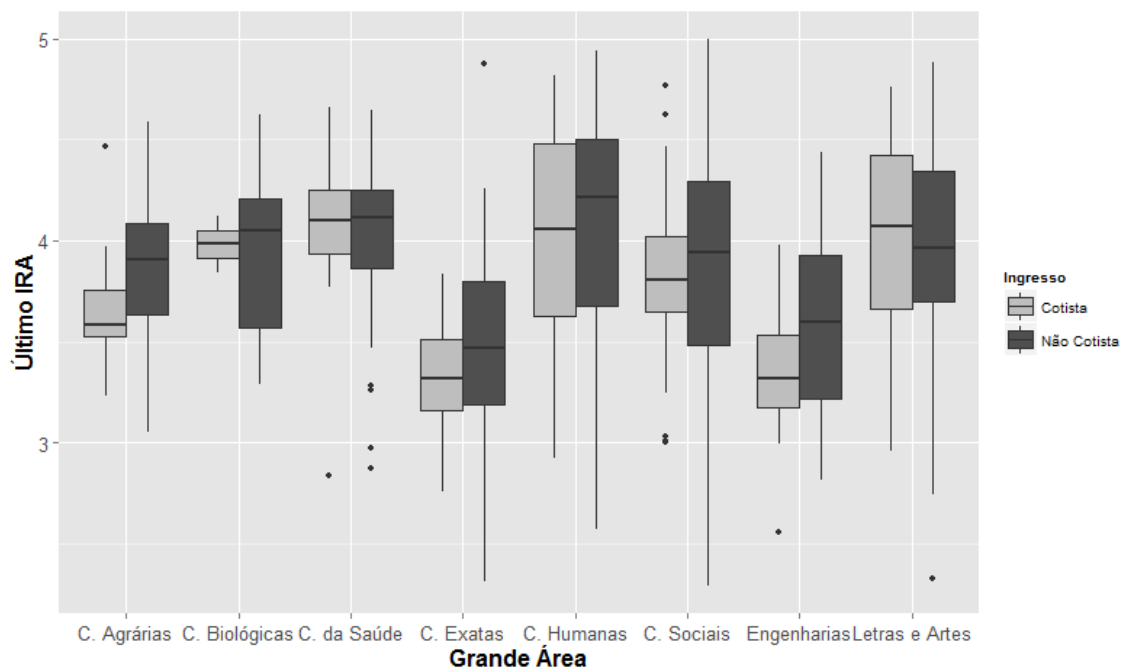
Período	Cotista	Não Cotista	Diferença
2004.2	3,46	3,61	4,3%
2005.1	3,48	3,60	3,5%
2005.2	3,50	3,64	3,9%
2006.1	3,37	3,57	5,7%
2006.2	3,38	3,56	5,4%
2007.1	3,52	3,71	5,4%
2007.2	3,64	3,72	2,1%
2008.1	3,67	3,78	3,0%
2008.2	3,74	3,73	-0,1%
2009.1	3,70	3,77	1,9%
2009.2	3,46	3,51	1,7%
2010.1	3,60	3,44	-4,3%
2010.2	3,02	3,23	6,8%

5.4.1.2 Descrição do IRA final dos alunos formados entre 2/2008 e 2/2010

Para a construção dos próximos gráficos e tabelas será utilizado o IRA final médio dos 851 alunos que se formaram entre o 2º semestre de 2008, quatro anos após esses indivíduos entrarem na universidade e o 2º semestre de 2010, após seis anos no curso superior. Dentre eles, 167 (19,8%) estudantes eram cotistas e 682 (80,1%) eram não cotistas. A figura abaixo mostra a distribuição do último IRA para alunos cotistas e não cotistas de acordo com a grande área a qual pertence seus cursos.

De acordo com o gráfico, é possível identificar que, em média, os cursos de Ciências Exatas e da Terra e de Engenharias possuem formados com IRAs finais menores e os cursos de Ciências Humanas e Ciências da Saúde, possuem IRAs finais maiores. Também é possível verificar maior discrepância entre a distribuição dos IRAs finais de cotistas e não cotistas nos cursos de Ciências Agrárias e Engenharias. Fato inverso ocorre para os cursos de Letras e Artes, que possuem mediana maior no grupo dos cotistas e para os cursos de Ciências da Saúde, que possuem distribuição dos IRAs semelhantes entre cotistas e não cotistas.

Figura 3: Distribuição do IRA final dos formados entre 2/2008 e 2/2010 por sistema de ingresso, segundo Grande Área - 2004



A tabela 10 corrobora a visualização mostrada acima. O IRA final médio de cotistas é maior nos cursos de Ciências Biológicas, Ciências da Saúde e Letras, Artes e Linguística. Para os demais cursos, não cotistas possuem IRA final médio maior, diferença que chega a 7,8% nos cursos de Engenharia.

Tabela 10: Distribuição do IRA final médio dos formados entre 2/2008 e 2/2010 por sistema de ingresso, segundo grande área - 2004

Grande Área	Cotista	Não Cotista	Diferença
Ciências Agrárias	3,65	3,84	5,1%
Ciências Biológicas	3,98	3,92	-1,5%
Ciências da Saúde	4,06	4,04	-0,5%
Ciências Exatas e da Terra	3,34	3,49	4,8%
Ciências Humanas	4,02	4,09	1,7%
Ciências Sociais Aplicadas	3,80	3,86	1,6%
Engenharias	3,33	3,59	7,8%
Letras e Artes, Linguística	3,99	3,97	-0,3%

A tabela 11 mostra o IRA final médio dos formados, por tipo de escola onde o aluno fez maior parte do ensino médio e qual tipo de curso médio foi feito. De acordo com os dados, alunos de escolas privadas possuem IRA final médio semelhante aos dos alunos

de escolas públicas. Em relação ao tipo de curso médio, alunos que fizeram outro tipo de curso, como profissionalizante ou técnico, possuem rendimento maior do que alunos que fizeram curso regular ou supletivo. Por outro lado, os menores IRAs médios foram apresentados pelos formados que fizeram supletivo.

Tabela 11: Distribuição do IRA final médio dos formados entre 2/2008 e 2/2010 por sistema de ingresso, segundo tipo de escola e curso médio - 2004

Tipo de Escola	Cotista	Não cotista	Diferença
Privada	3,83	3,90	1,57%
Pública	3,80	3,90	2,63%
Sem Resposta	3,71	3,84	3,50%
Curso Médio			
Curso Regular	3,72	3,80	2,15%
Supletivo	3,29	3,60	9,42%
Outro	3,83	3,89	1,57%
Sem Resposta	3,81	3,73	-2,10%

A tabela 12 mostra o IRA final médio dos formados por sistema de ingresso, segundo maior nível de escolaridade entre os pais. Foi verificado que existem apenas três formados – dois cotistas e um não cotista – com pais analfabetos. Portanto, a análise descritiva levará em consideração apenas os demais níveis de escolaridade. O grupo de cotistas formados com maior IRA final médio são os que possuem pais com ensino superior. Dentre os não cotistas o maior valor se deu para alunos com pais com ensino fundamental incompleto. Os menores valores foram registrados para alunos, tanto cotistas como não cotistas, com pais com ensino fundamental completo.

Tabela 12: IRA final médio dos formados entre 2/2008 e 2/2010 por sistema de ingresso, segundo maior nível de escolaridade dos pais - 2004

Maior escolaridade dos pais	Cotista	Não cotista	Diferença
Ensino fundamental incompleto	3,64	3,87	6,5%
Ensino fundamental completo	3,52	3,61	2,6%
Ensino médio	3,71	3,82	3,1%
Superior	3,76	3,81	1,2%
Sem Resposta	3,72	3,74	0,7%

A tabela 13 fornece o IRA médio final por atividade remunerada no período de inscrição no vestibular e pelo fato de cursar universidade antes de entrar na UnB. De

acordo com os dados, o grupo de alunos que realizava trabalho em tempo integral e o grupo que cursava universidade fora da UnB, obtiveram maior IRA final médio. Já os grupos de alunos que realizam trabalhos eventuais e cursavam universidade na UnB obtiveram IRA final médio menor.

Tabela 13: IRA final médio dos formados entre 2/2008 e 2/2010 de cotistas e não cotistas por sistema de ingresso, segundo atividade remunerada e curso anterior - 2004

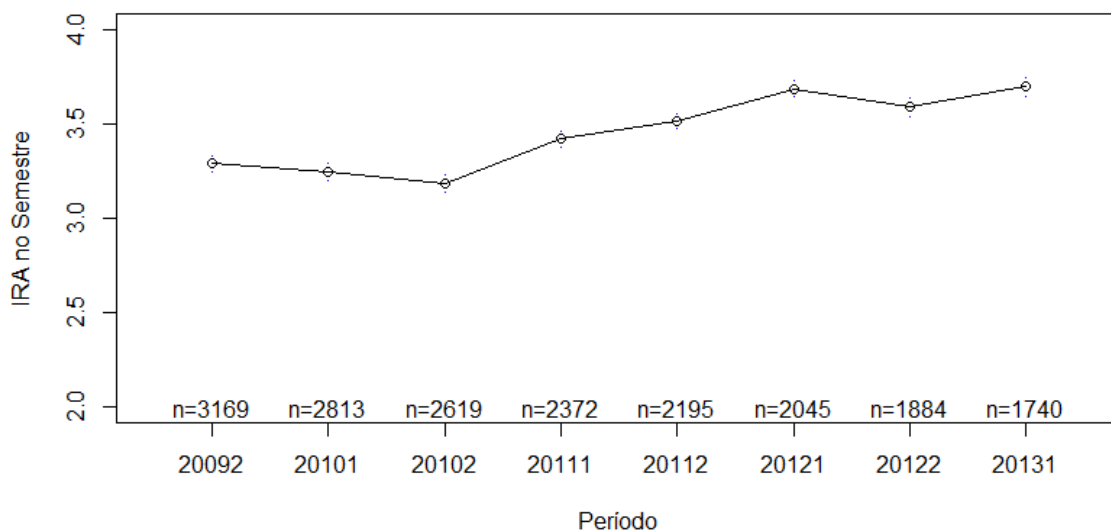
Atividade remunerada	Cotista	Não cotista	Diferença
Não	3,80	3,89	2,42%
Sim, em tempo parcial	3,76	3,90	3,74%
Sim, em tempo integral	4,00	3,99	-0,30%
Sim, mas trata-se de trabalho eventual	3,81	3,65	-4,24%
Sem Resposta	3,75	3,85	2,46%
Cursava universidade			
Não	3,82	3,89	1,94%
Sim, fora da UnB	4,10	3,84	-5,68%
Sim, na UnB	3,64	3,79	3,90%
Sem resposta	3,75	3,84	2,39%

5.4.2 Rendimento acadêmico dos alunos ingressantes no 2º Vestibular de 2009

5.4.2.1 Descrição do IRA do semestre

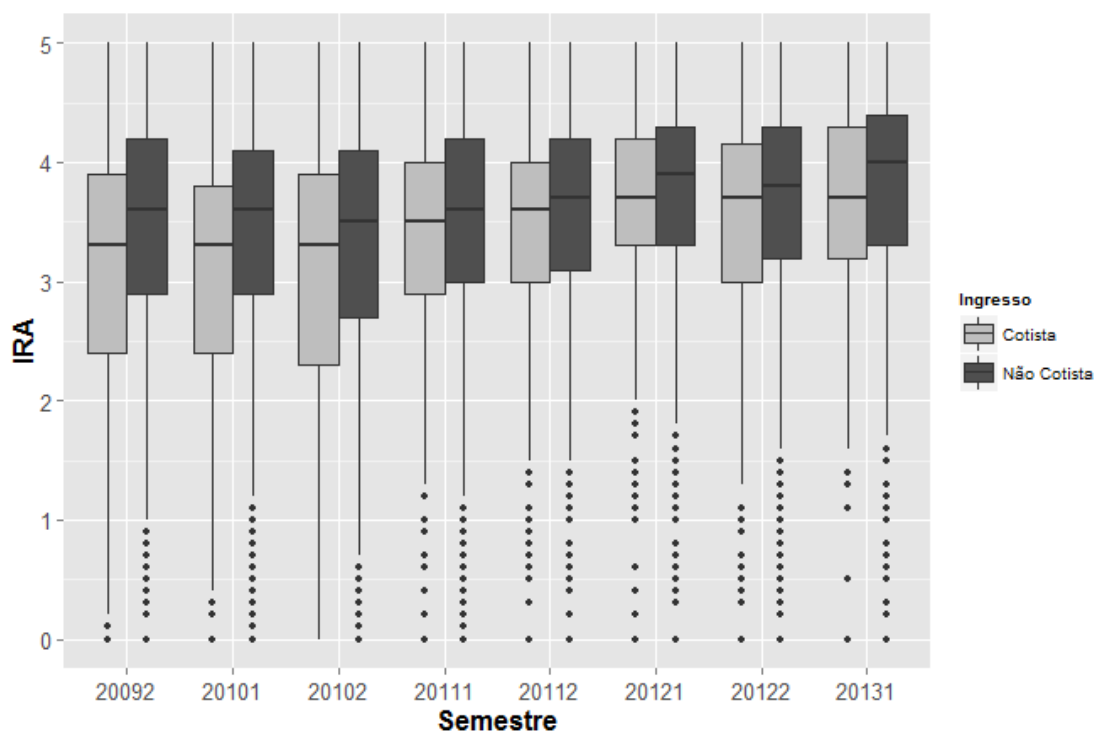
O IRA do semestre médio obtido em cada semestre de todos os alunos que ingressaram no 2º Vestibular de 2009 é dado pela figura 9. Ao analisarmos o gráfico, é possível verificar uma tendência de crescimento a partir do 1º semestre de 2011, fato que se mantém até o 1º semestre de 2013.

Figura 4: IRA médio obtido no semestre - 2009



A Figura 5 destaca a distribuição do IRA em cada semestre para alunos cotistas e não cotistas. Foi identificado que a distribuição dos IRAs dos cotistas é inferior à dos não cotistas no decorrer dos semestres, porém, essa diferença tende a diminuir no decorrer dos anos. Também é perceptível a redução da variabilidade da distribuição do IRA do semestre com o passar dos períodos letivos.

Figura 5: Boxplot do IRA do semestre por período, segundo sistema de ingresso - 2009



Corroborando com a visualização gráfica, de acordo com a tabela 14, verifica-se que a diferença entre o IRA médio de alunos cotistas e não cotistas diminui no decorrer dos semestres, principalmente a partir do 1º semestre de 2011, passando de 10,07% no semestre anterior para 4,02%.

Tabela 14: IRA do semestre médio por sistema de ingresso, segundo semestre - 2009

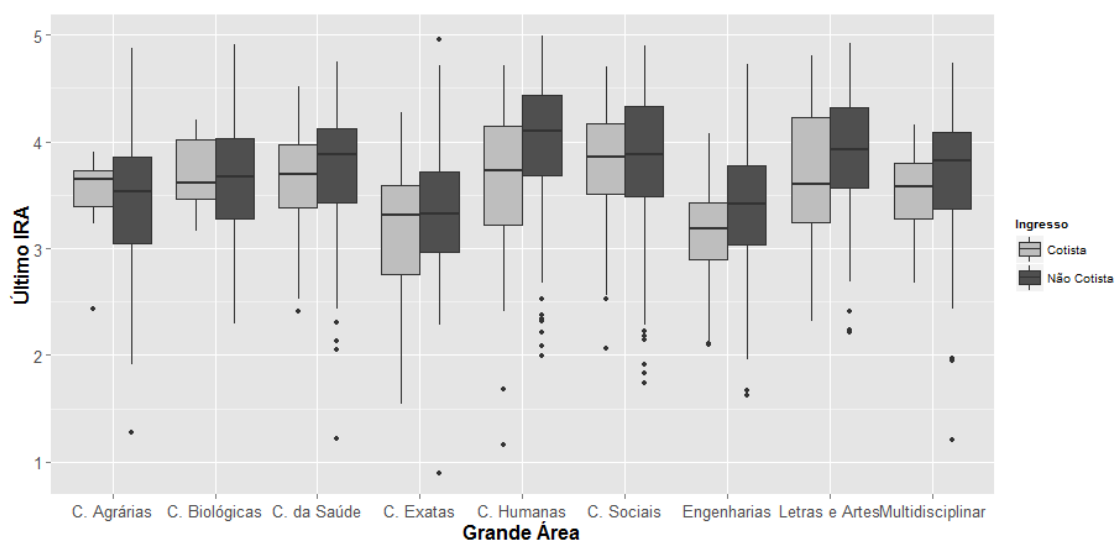
Período	Cotista	Não cotista	Diferença
2009.2	3,00	3,36	12,02%
2010.1	2,98	3,30	10,59%
2010.2	2,94	3,24	10,07%
2011.1	3,31	3,44	4,02%
2011.2	3,40	3,54	3,94%
2012.1	3,57	3,71	3,88%
2012.2	3,46	3,62	4,63%
2013.1	3,61	3,71	3,02%

5.4.2.2 Descrição do IRA final dos alunos formados ou ativos

Para a construção dos próximos gráficos e tabelas foi utilizado o IRA final médio dos 1918 alunos que se formaram até o 1º semestre de 2013 ou estão ativos. Dentre eles, 342 (17,8%) estudantes são cotistas e 1576 (82,2%) são não cotistas. A figura abaixo mostra a distribuição do último IRA para alunos cotistas e não cotistas de acordo com a grande área a qual pertence seus cursos.

De acordo com o gráfico, é possível identificar que, em média, os cursos de Ciências Exatas e da Terra e de Engenharias possuem IRAs mais baixos e os cursos de Ciências Humanas e Ciências Sociais, possuem IRAs mais altos. Também é possível verificar maior discrepância entre a distribuição dos IRAs finais de cotistas e não cotistas nos cursos de Engenharias e Letras e Artes a favor dos não cotistas. Fato inverso ocorre para os cursos de Ciências Agrárias, que possuem mediana maior no grupo dos cotistas e para os cursos de Ciências Biológicas, que possui distribuição dos IRAs semelhantes entre cotistas e não cotistas.

Figura 6: Distribuição do IRA final dos alunos formados ou ativos por sistema de ingresso, segundo grande área - 2009



A tabela 15, que fornece o IRA final médio de alunos cotistas e não cotistas por grande área, corrobora a visualização mostrada na figura 6. O IRA final médio de cotistas é maior nos cursos de Ciências Biológicas e Ciências Agrárias. Para os demais cursos, não cotistas possuem IRA final médio maior, diferença que chega a 9,56% nos cursos de Ciências Humanas e 8,75% para as Engenharias, resultados próximos ao observado em 2004.

Tabela 15: IRA final médio de alunos formados e ativos por sistema de ingresso segundo, grande área - 2009

Período	Cotista	Não cotista	Diferença
C. Agrárias	3,53	3,44	-2,74%
C. Biológicas	3,70	3,66	-1,07%
C. da Saúde	3,64	3,75	3,02%
C. Exatas	3,22	3,34	3,74%
C. Humanas	3,65	4,00	9,56%
C. Sociais	3,80	3,85	1,33%
Engenharias	3,12	3,40	8,75%
Letras e Artes	3,71	3,91	5,43%
Multidisciplinar	3,50	3,61	3,08%

A tabela 16 mostra o IRA médio por tipo de escola onde o formado fez maior parte do ensino médio e curso médio. De acordo com os dados, alunos cotistas de escolas

pública possuem IRA final médio maior do que alunos de escolas privadas. Esse fato, embora em menor grau, se inverte para os não cotistas. Em relação ao tipo de curso médio, alunos não cotistas que fizeram outro tipo de curso, como profissionalizante ou técnico, possuem rendimento maior do que alunos que fizeram curso regular ou supletivo. Já entre os cotistas, alunos que fizeram curso regular possuem maior IRA final médio

Tabela 16: IRA final médio de alunos formados e ativos por sistema de ingresso, segundo tipo de escola e curso médio - 2009

Tipo de Escola	Cotista	Não cotista	Diferença
Privada	3,59	3,71	3,2%
Pública	3,64	3,70	1,9%
Sem Resposta	3,34	3,71	11,2%
Curso médio			
Curso regular.	3,60	3,70	2,8%
Supletivo (Educação de Jovens e Adultos).	3,48	3,32	-4,6%
Outro.	3,40	3,83	12,8%
Sem Resposta	3,60	3,75	3,8%

A tabela 17 mostra o IRA final médio pelo maior nível de escolaridade observado nos pais. Dentre os cotistas, o maior IRA final médio são observados para formados com pelo menos um dos pais com ensino fundamental incompleto, e dentre os não cotistas o maior valor é observado para o grau de escolaridade ensino fundamental completo.

Tabela 17: IRA final médio de alunos formados e ativos por sistema de ingresso segundo, maior nível de escolaridade dos pais - 2009

Maior escolaridade	Cotista	Não cotista	Diferença
Ensino fundamental incompleto	3,72	3,79	1,7%
Ensino fundamental completo	3,39	3,85	13,5%
Ensino médio	3,66	3,71	1,3%
Superior	3,55	3,68	3,8%
Sem Resposta	3,60	3,75	4,3%

A tabela 18 mostra o IRA médio final por atividade remunerada no período de inscrição no vestibular e pelo fato de cursar, ou não, universidade na época. De acordo com os dados, entre os cotistas, o grupo de alunos que realizava trabalho em tempo integral e o grupo que não estava na universidade na época do vestibular, obtiveram maior IRA final médio. Já entre os não cotistas, alunos que trabalhavam em tempo integral ou

que não trabalhavam e os que faziam faculdade fora da UnB obtiveram maior IRA final médio.

Tabela 18: IRA final médio de alunos formados e ativos por sistema de ingresso, segundo atividade remunerada e curso anterior - 2009

Atividade remunerada	Cotista	Não cotista	Diferença
Não.	3,62	3,71	2,5%
Sim, mas se tratava de trabalho eventual.	3,40	3,61	6,1%
Sim, em tempo parcial	3,49	3,52	1,0%
Sim, em tempo integral	3,66	3,71	1,3%
Sem Resposta	3,59	3,76	5,2%
Cursava universidade			
Não	3,63	3,70	2,0%
Sim, fora da UnB	3,47	3,74	7,8%
Sim, na UnB.	3,57	3,67	2,7%
Sem Resposta	3,60	3,74	4,4%

Na avaliação dos IRA médios dos semestres dos alunos para os anos de 2004 e 2009, é possível verificar que existe uma tendência de aumento no IRA semestral médio no decorrer dos anos, além de uma aproximação entre as distribuições dos IRAs dos alunos cotistas e não cotistas.

Quanto a avaliação das grandes áreas, é notável que formados pertencentes aos cursos de engenharia e ciências exatas possuem IRAs Finais menores do que os estudantes de outros cursos. Por lado, os alunos de Ciências Humanas possuem IRAs maiores em ambos grupos.

Quanto ao tipo de escola em que os estudantes fizeram nível médio, em 2004, não foi possível verificar diferenças significativas e, em 2009, foi percebido que formados cotistas de escolas públicas possuem IRA médio maior do que os formados cotistas de escolas particulares. Esse fato se inverte para os não cotistas, ou seja, alunos de escolas particulares que possuem IRA final médio maior. Em relação ao tipo de curso, é notável que alunos que fizeram supletivo saíram da universidade com IRA médio menor nas duas coortes grupos, tanto para cotistas quanto para não cotistas.

Em relação a escolaridade dos pais, os resultados divergiram entre os ingressantes dos dois grupos analisados e entre cotistas e não cotistas.

Quanto a atividade remunerada exercida no período do vestibular, foi verificado que em 2004, os alunos que trabalhavam em tempo integral possuíam maiores IRAs médios. Já em 2009, as maiores médias foram verificadas para alunos que não trabalhavam ou trabalhavam em tempo integral. Também foi constatado que alunos que não faziam curso superior quando ingressaram na Universidade saíram com maiores IRAs Finais em relação aos demais alunos.

5.4.3 Evasão dos alunos ingressantes no 2º Vestibular de 2004

Para mostrar como ocorre a evasão entre cotistas e não cotistas ingressantes no 2º Vestibular de 2004 a tabela 19, mostra o percentual de evadidos entre 2005 e 2009 e a diferença proporcional entre dos dois grupos de alunos. De acordo com a tabela, é possível verificar que alunos cotistas evadem menos do que os alunos não cotistas nos primeiros três anos, diferença que chega a ser de -2,38 pontos percentuais em 2005, e, no quarto e quinto ano o percentual de evadidos é menor entre os não cotistas. A diferença entre cotistas e não cotistas em todos os anos é de 2,3 pontos percentuais, sendo 10,2% maior entre não cotistas.

Tabela 19: Percentual de evadidos por sistema de ingresso, segundo ano - 2004

Ano	Cotista	Não cotista	Diferença
2005	4,8%	7,2%	-2,38
2006	5,1%	5,5%	-0,43
2007	5,4%	5,5%	-0,08
2008	2,8%	1,9%	0,97
2009	2,3%	1,7%	0,53
Total	22,4%	24,7%	-2,29

A tabela 20 mostra o percentual de alunos evadidos por sistema de ingresso, segundo faixa etária. De acordo com a tabela, existe uma tendência de crescimento no percentual de evasão a medida que os alunos têm mais idade. Para os alunos com até 21 anos no momento da realização do vestibular, o percentual de alunos evadidos não chega ser maior que 20%. Porém, entre os alunos com mais de 24 anos, as proporções são maiores que 40% em todas as faixas, tanto para cotistas quanto para não cotistas.

Tabela 20: Percentual de evadidos por sistema de ingresso, segundo faixa de idade - 2004

Faixa Etária	Cotista	Não cotista	Diferença
Entre 15 e 18 anos	14,6%	17,4%	-2,72
Entre 18 e 21 anos	12,2%	16,8%	-4,59
Entre 21 e 24 anos	31,8%	35,2%	-3,39
Entre 24 e 27 anos	43,8%	48,9%	-5,11
Entre 27 e 30 anos	41,7%	46,3%	-4,63
Entre 30 e 40 anos	52,9%	46,8%	6,19
Mais de 40 anos	40,0%	57,6%	-17,58

Em relação às grandes áreas, é possível perceber que nos cursos de Ciências Exatas e Letras e Artes, as proporções de evadidos chegam a ser maior do que 30% tanto entre cotistas e não cotistas. É verificado também que a diferença entre evadidos cotistas e não cotistas chega a 13,64 e 11,59 pontos percentuais nos cursos de ciências agrárias e engenharias, respectivamente, a favor dos não cotistas, e a 9,42 nos cursos de Artes, Letras e Linguística a favor dos cotistas

Tabela 21: Percentual de evadidos por sistema de ingresso, segundo grandes áreas - 2004

Grandes Áreas	Cotista	Não cotista	Diferença
C. Agrárias	27,3%	13,6%	13,64
C. Biológicas	14,3%	12,2%	2,09
C. da Saúde	8,3%	11,3%	-2,94
C. Exatas	39,3%	38,4%	0,92
C. Humanas	22,9%	26,4%	-3,50
C. Sociais	10,5%	17,1%	-6,62
Engenharias	31,0%	19,4%	11,59
Letras e Artes	31,6%	41,0%	-9,42

De acordo com a tabela 22, é possível verificar que alunos de escolas privadas, tanto cotistas quanto não cotistas, evadiram menos do que os alunos de escolas públicas. Além disso, cotistas de escolas privadas evadiram menos do que os não cotistas, fato que se inverte entre alunos de escolas públicas. Quando a análise é feita com o tipo de curso feito antes do ingresso na universidade, verifica-se que alunos que fizeram ensino regular evadem menos do que os alunos que fizeram supletivo ou outros. Nessa tabela se destaca o fato de 60% dos alunos cotistas que fizeram supletivo terem evadido da universidade.

Tabela 22: Percentual de evadidos por sistema de ingresso, segundo tipo de escola e curso médio - 2004

Tipo de Escola	Cotista	Não cotista	Diferença
Privada	15,2%	17,1%	-1,93
Pública	25,1%	20,6%	4,57
Sem Resposta	44,0%	32,4%	11,60
Curso Médio			
Curso regular.	17,1%	17,8%	-0,71
Supletivo (Educação de Jovens e Adultos).	60,0%	16,7%	43,33
Outro.	30,6%	22,6%	7,97
Sem Resposta	52,4%	32,6%	19,79

A tabela 23 mostra o percentual de evadidos por sistema de ingresso, segundo atividade remunerada e ao fato de cursar ensino superior anteriormente. Em relação ao tipo de atividade remunerada, é possível verificar que estudantes que trabalhavam em tempo parcial evadiram mais entre os cotistas e, estudantes que trabalhavam em tempo integral evadiram mais entre os não cotistas. Por outro lado, alunos que não realizavam atividade remunerada evadiram menos do que os estudantes que trabalhavam. Quando se compara o percentual de evadidos de acordo com o fato de cursar ensino superior antes do vestibular, verifica-se que alunos que fizeram faculdade fora da UnB evadiram mais do que os demais alunos. Entretanto, quem teve a UnB como primeiro curso superior evadiu menos do que os demais.

Tabela 23: Percentual de evadidos por sistema de ingresso, segundo atividade remunerada e curso universitário- 2004

Atividade remunerada	Cotista	Não cotista	Diferença
Não	15,2%	13,5%	1,60
Sim, em tempo parcial	40,6%	43,1%	-2,51
Sim, em tempo integral	29,5%	44,9%	-15,38
Sim, mas trata-se de trabalho eventual	30,8%	21,4%	9,34
Sem Resposta	15,2%	13,5%	1,60
Cursava universidade			
Não	15,3%	14,0%	1,29
Sim, fora da UnB	39,3%	34,9%	4,40
Sim, na UnB.	37,3%	34,2%	3,08
Sem Resposta	52,6%	32,6%	20,04

5.4.4 Evasão dos alunos ingressantes no 2º Vestibular de 2009

Para mostrar como se dá a evasão entre cotistas e não cotistas ingressantes no 2º Vestibular de 2009, a tabela 24 mostra o percentual de evadidos entre os anos 2010 e 2013. De acordo com a tabela, é possível verificar que houve um aumento no percentual de alunos evadidos nos primeiros quatro anos em relação a 2004, passando de 20,1% e 23% - cotistas e não cotistas, respectivamente, - para 27,9% e 23,5% em 2009.

Também foi verificado que alunos cotistas, passaram a evadir em maior proporção em 2009, passando de 22,4% para 27,9%. A diferença entre as proporções de alunos evadidos passa de 3,09 a favor dos não cotistas em 2010, para menos de um ponto percentual nos anos seguintes.

Tabela 24: Percentual de evadidos por sistema de ingresso, segundo ano - 2009

Ano	Cotista	Não cotista	Diferença
2010	12,9%	9,8%	3,09
2011	6,4%	6,3%	0,07
2012	5,2%	4,2%	0,98
2013	3,4%	3,1%	0,22
Total	27,9%	23,5%	4,36

A tabela 25 mostra o percentual de evadidos por sistema de ingresso segundo faixa etária. De acordo com os dados, a proporção de evadidos aumenta de acordo com a faixa de idade, chegando a 51% entre cotistas entre 27 e 30 anos e a 65% entre não cotistas de 27 a 30 anos.

Tabela 25: Percentual de evadidos por sistema de ingresso, segundo faixa de idade

Faixa Etária	Cotista	Não cotista	Diferença
Entre 15 e 18 anos	23%	14%	8,73
Entre 18 e 21 anos	21%	17%	4,18
Entre 21 e 24 anos	33%	30%	2,85
Entre 24 e 27 anos	42%	44%	-2,63
Entre 27 e 30 anos	51%	65%	-13,62
Entre 30 e 40 anos	47%	52%	-4,82
Mais de 40 anos	33%	49%	-15,20

A Tabela 26 mostra as proporções de evadidos por sistema de ingresso de acordo com a grande área a qual o curso pertence. Os cursos com maior percentual de evadidos entre os cotistas são os multidisciplinares (60%) e cursos das áreas de ciências exatas (45,1%), engenharias (39%). Entre os não cotistas, os cursos com maior número de evadidos são das Letras e Artes (36,3%), ciências exatas (33,9%) e cursos multidisciplinares (32,5%). Por outro lado, os cursos de ciências biológicas contam com apenas 9,5% de evasão entre cotistas e os cursos de ciências sociais com o percentual de 15,3% de evadidos para cotistas e não cotistas.

Tabela 26: Percentual de evadidos por sistema de ingresso, segundo grandes áreas - 2009

Grandes Áreas	Cotista	Não cotista	Diferença
C. Agrárias	19,4%	19,1%	0,28
C. Biológicas	9,5%	19,3%	-9,75
C. da Saúde	25,9%	18,3%	7,54
C. Exatas	45,1%	33,9%	11,18
C. Humanas	26,2%	27,1%	-0,93
C. Sociais	15,3%	15,3%	0,00
Engenharias	39,0%	17,2%	21,81
Letras e Artes	25,0%	36,3%	-11,28
Multidisciplinar	60,0%	32,5%	27,47

A tabela 27 mostra o percentual de alunos evadidos por sistema de ingresso, segundo tipo de escola onde fez o ensino médio e tipo de curso médio. É possível verificar que os alunos de escolas públicas (26% e 24%) evadiram em maior proporção do que os alunos de escolas privadas (21 e 19%). Já os alunos cotistas evadiram mais para ambos tipos de escola. Quando se analisa o tipo de curso médio realizado, verifica-se percentual bastante alto para alunos cotistas que fizeram supletivo (71,4%) ou outro tipo de curso no ensino médio. Entre os não cotistas a maior proporção de evadidos é observada para aluno que fizeram supletivo.

Tabela 27: Percentual de evadidos por sistema de ingresso, segundo tipo de escola e curso médio

Tipo de Escola	Cotista	Não cotista	Diferença
Privada	21%	19%	2,29
Pública	26%	24%	2,71
Sem Resposta	59%	56%	3,02
Curso médio			
Curso regular.	27,3%	22,7%	4,54
Supletivo (Educação de Jovens e Adultos).	71,4%	40,5%	30,89
Outro.	40,9%	26,3%	14,66
Sem Resposta	23,7%	23,6%	0,12

A tabela 28 mostra o percentual de evadidos por sistema de ingresso segundo atividade remunerada exercida e curso superior antes do vestibular. Os dados mostram que alunos que exerciam algum tipo de atividade remunerada evadem em maior proporção quando comparados com os alunos que não trabalhavam. Em relação ao fato de cursar alguma faculdade antes de ingressar na UnB, os alunos que já haviam feito outro curso superior na Universidade de Brasília evadiram em maior proporção do que os demais.

Tabela 28: Percentual de evadidos por sistema de ingresso, segundo atividade remunerada e curso superior anterior

Atividade remunerada	Cotista	Não cotista	Diferença
Não.	23,2%	16,6%	6,55
Sim, mas se tratava de trabalho eventual.	47,4%	32,6%	14,76
Sim, em tempo parcial	40,5%	38,4%	2,14
Sim, em tempo integral	44,1%	51,5%	-7,46
Sem Resposta	23,3%	23,5%	-0,24
Cursava universidade			
Não	25,6%	17,0%	8,61
Sim, fora da UnB	23,5%	35,4%	-11,86
Sim, na UnB.	48,8%	40,8%	7,98
Sem Resposta	22,9%	22,6%	0,36

A partir da descrição realizada para as duas coortes, verifica-se que houve um aumento na proporção de alunos evadidos nos primeiros quatro anos, passando de 20,1% e 23% - cotistas e não cotistas - em 2004 para 27,9% e 23,5% em 2009.

Também observa-se que os cotistas evadiram em menor proporção, 22,4%, em relação ao não cotistas entre alunos ingressantes em 2004. Porém, em 2009, passaram a evadir em maior proporção, 27,9%, quando comparado aos não cotistas.

Além disso, para as duas coortes, o percentual de estudantes evadidos aumenta de acordo com a faixa de idade.

Outro ponto em comum é que, para ambos grupos de ingressantes, cotistas de cursos de ciências exatas e engenharias possuem as maiores proporções de evadidos. Já entre os não cotistas, as maiores proporções são dos cursos de ciências exatas e Artes e Letras.

Quanto ao tipo de escola onde foi cursado ensino médio, o grupo de alunos que estudou em escola pública possui maior proporção de evadidos tanto em 2004, quanto em 2009. Em relação ao tipo de curso feito, destaca-se que alunos cotistas de ambas coortes que fizeram supletivo possuem taxas altas de evasão, chegando 71,4% em 2009.

Em relação as atividades realizadas antes do ingresso na UnB, tanto em 2004 quanto em 2009, os grupos de estudantes que trabalhavam ou que cursavam nível superior antes do ingresso na universidade possuem maiores proporções de evadidos do que o grupo de alunos que não trabalhavam ou não haviam ingressado no ensino superior.

5.5 Modelagem

Com base na análise descritiva realizada com o IRA e evasão, foi investigado como os determinantes socioeconômicos podem influenciar o comportamento dessas variáveis através da modelagem estatística. Nos modelos, foram testadas relacionadas já as variáveis estudadas.

Para estudar o comportamento do IRA, primeiro foi utilizada a regressão linear, método de mínimos quadrados ordinários (MQO), onde a variável dependente considerada foi o IRA final do aluno. O segundo modelo utiliza dados em painel para explicar a variável IRA por semestre no decorrer do tempo. Para estudar quais fatores estão associadas a evasão dos alunos, foi utilizada a regressão logística.

5.5.1 Análise do IRA

5.5.1.1 Análise do IRA final dos ingressantes no 2º Vestibular de 2004 utilizando regressão linear

Para o primeiro modelo, foram eleitas as 12 variáveis explicativas já estudadas no trabalho. Como foram considerados na modelagem apenas os indivíduos com todas as variáveis disponíveis, a análise contou com 1154 estudantes. A tabela 29 apresenta os resultados encontrados para as seis variáveis selecionadas. As variáveis do modelo foram selecionadas através do teste t para testar a significância de cada preditor individualmente.

Tabela 29: Estimativas, Desvio Padrão, Estatísticas t do modelo de regressão - 2004

	Variável	Estimativa	D. Padrão	Valor t	P- Valor
	Intercepto	2,47	0,64	3,86	0,000***
	Argumento final	0,0007	0,00	3,66	0,000***
Forma de Saída	Evasão	-0,38	0,64	-0,60	0,551
	Formatura	1,34	0,63	2,12	0,034*
	Mobilidade	0,42	0,64	0,66	0,509
Grande Área	C. Biológicas	-0,03	0,11	-0,25	0,804
	C. da Saúde	0,13	0,09	1,46	0,144
	C. Exatas	-0,29	0,09	-3,10	0,0012*
	C. Humanas	0,28	0,09	3,09	0,0014*
	C. Sociais	0,13	0,08	1,59	0,111
	Engenharias	-0,17	0,10	-1,83	0,068.
	Linguística L. e Artes	0,27	0,11	2,52	0,012*
Sistema de Ingresso	Universal	0,09	0,05	1,89	0,059.
Sexo	Masculino	-0,26	0,04	-5,71	0,000***
Fez ou está fazendo curso superior	Sim, fora da UnB	-0,14	0,09	-1,52	0,129
	Sim, na UnB	-0,35	0,09	-4,05	0,000***

Códigos de Significância: <0,0001 “***”, <0,001 “**”, <0,05 “*”, <0,10 “.”

Primeiramente, foram abordadas questões relativas a pressuposições comuns em regressões lineares, como heterocedasticidade, multicolinearidade, viés por variável omitida e qualidade dos ajustes.

De acordo com o teste de Breusch-Pagan (BREUSCH e PAGAN, 1979), rejeitou-se a hipótese de variância constante. Logo, com o intuito de corrigir o efeito da heterocedasticidade e obter melhores estimativas, foram utilizados erros-padrões robustos nos parâmetros do modelo.

De acordo com o Fator de Inflação da Variância (LONGNECKER e OTT, 2004), não foi identificado multicolinearidade entre os as variáveis do modelo.

Para testar a hipótese de que não existem variáveis omitidas no modelo, ou seja, variáveis que são essenciais para explicar o IRA final, foi utilizado o teste Reset (teste de erro de especificação em regressão) de Ramsey (RAMSEY, 1969). Como o p-valor do teste é igual a 0,01, rejeita-se a hipótese nula, ou seja, existem fatores que influenciam o rendimento que não foram captados pelo modelo. Portanto, para compreender melhor os efeitos não observados e capturar a heterogeneidade individual no tempo, o IRA dos alunos será analisado utilizando dados em painel.

Também foi verificada a existência de outliers, testando a hipótese de que não havia resíduos estudentizados com p-valor menor do que 0,05. De acordo com o teste de Bonferroni (Cook e Weisberg, 1984), apenas uma observação foi considerada outlier.

O modelo apresentou p-valor tendendo a zero para o teste F de ajustamento do modelo, ou seja, rejeita-se a hipótese de que as variáveis independentes não exercem influência sobre a nota média de um candidato. O R^2 ajustado do modelo é igual a 0,54.

Em relação aos resultados apresentados pelo modelo, é possível afirmar que entre as variáveis sociodemográfica a única estatisticamente significativa é sexo do estudante, em que mulheres tem notas significativamente mais alta do que os homens.

Em relação a trajetória pré universitária, foi verificado que o fato dele ter feito universidade anteriormente influencia negativamente o IRA final do estudante, quando comparado com os estudantes não que ainda não estavam cursando ensino superior antes de entrar na universidade. Também foi verificado que o Argumento final do Vestibular está correlacionado de maneira significativa com o IRA final do estudante, ou seja, a cada 1 ponto no argumento final no vestibular, há um aumento de 0.0007 pontos no IRA final do aluno, dadas as demais variáveis constantes. O valor baixo desse coeficiente se deve a diferenças entre as magnitudes do IRA e do argumento final no vestibular.

Quanto a inserção universitária, é possível constatar que os alunos não cotistas possuem desempenho marginalmente superior ao dos alunos cotistas. Além disso, é possível verificar que alunos das áreas de Letras e Artes e Ciências Humanas possuem desempenho superior, quando comparado com cursos de Ciências Agrárias. Fato inverso ocorre para os alunos dos cursos de Ciências Exatas e Engenharias.

Também foi constatado que os alunos que se formam, saem com o IRA superior do que os alunos optaram pela mobilidade.

5.5.1.2 Análise do IRA final dos ingressantes no 2º Vestibular de 2009 utilizando regressão linear

A tabela 31 apresenta os resultados encontrados para as oito variáveis selecionadas. Como foram considerados na modelagem apenas os indivíduos com todas as variáveis disponíveis, a análise contou com 2391 estudantes. As variáveis do modelo foram selecionadas através do teste t para testar a significância de cada preditor individualmente.

Tabela 30: Estimativas, Desvio Padrão, Estatísticas t do modelo de regressão - 2009

	Variável	Estimativa	D. Padrão	Valor t	P- Valor
	Intercepto	3,41	0,08	45,38	0,000***
	Argumento final	0,0002	0,00	4,84	0,000***
Forma de Saída	Evasão	-1,65	0,06	-29,94	0,000***
	Formatura	0,45	0,03	14,66	0,000***
	Mobilidade	-0,76	0,06	-11,69	0,000***
Grande Área	C. Biológicas	0,12	0,10	1,22	0,223
	C. da Saúde	0,10	0,08	1,30	0,192
	C. Exatas	-0,15	0,08	-1,88	0,061.
	C. Humanas	0,32	0,08	4,13	0,000***
	C. Sociais	0,15	0,07	2,06	0,040*
	Engenharias	-0,01	0,07	-0,15	0,880
	Linguística L. e Artes	0,36	0,08	4,24	0,000***
	Multidisciplinar	-0,01	0,11	-0,12	0,902
Sistema de Ingresso	Universal	0,11	0,05	2,26	0,024*
Sexo	Masculino	-0,31	0,04	-8,11	0,000***
Fez ou está fazendo curso superior	Sim, fora da UnB	0,04	0,06	0,58	0,559
	Sim, na UnB	-0,16	0,06	-2,79	0,005**
Exerce alguma atividade remunerada	Sim, mas se tratava de trabalho eventual.	-0,13	0,07	-1,93	0,054.
	Sim, em tempo parcial	-0,25	0,09	-2,81	0,005*
	Sim, em tempo integral	-0,14	0,07	-2,03	0,043*
Qual curso médio você concluiu ou estava está concluindo	Supletivo	-0,34	0,10	-3,37	0,001**
	Outros (profissionalizantes, técnico, etc)	-0,02	0,10	-0,19	0,847
Códigos de Significância: <0,0001 “***”, <0,001 “**”, <0,05 “*”, <0,10 “.”					

De acordo com o teste de Breusch-Pagan, rejeitou-se a hipótese de variância constante. Logo, com o intuito de corrigir o efeito da heterocedasticidade e obter melhores estimativas, foram utilizados erros-padrões robustos nos parâmetros do modelo.

De acordo com o fator de inflação da variância (VIF), não foi identificado multicolinearidade entre os as variáveis do modelo.

Para testar a hipótese de que não existem variáveis omitidas no modelo, ou seja, variáveis que são essenciais para explicar o IRA final, foi utilizado o teste Reset (teste de erro de especificação em regressão) de Ramsey. Como o valor do teste foi significativo, rejeita-se a hipótese nula, ou seja, existem fatores que influenciam o rendimento que não foram captados pelo modelo. Portanto, assim como em 2004, reforça-se a necessidade de se compreender os efeitos não observados e capturar a heterogeneidade individual no tempo, o que será feito utilizando dados em painel.

Também foi verificada a existência de outliers, testando a hipótese de que não havia resíduos estudentizados com p-valor menor do que 0,05. De acordo com o teste de Bonferroni, apenas uma observação foi considerada *outlier*.

O modelo apresentou p-valor tendendo a zero para o teste F de ajustamento do modelo, ou seja, rejeita-se a hipótese de que as variáveis independentes não exercem influência sobre a nota média de um candidato. O R^2 ajustado do modelo é igual a 0,53.

Em relação aos resultados apresentados pelo modelo, é possível afirmar que a única variável sociodemográfica significativa é o sexo do estudante, em que, assim como na coorte de alunos ingressantes em 2004, mulheres tem notas estatisticamente mais alta do que os homens.

Em relação a trajetória pré universitária, foi verificado que o fato de ter cursado UnB anteriormente influencia negativamente o IRA final do estudante. Os estudantes que realizavam alguma atividade remunerada possuem IRA final menor do que os alunos que não trabalham. Além disso, estudantes que fizeram supletivo possuem desempenho inferior aos alunos que fizeram ensino regular.

Também foi verificado que o Argumento final do Vestibular está correlacionado de maneira significativa com o IRA final do estudante, ou seja, a cada ponto a mais no Argumento final acrescenta em média 0,002 pontos no IRA final, dadas as demais variáveis constantes.

Quanto a inserção universitária, é possível constatar que os alunos não cotistas possuem desempenho superior ao dos alunos cotistas. Também é possível verificar que alunos das áreas de Letras e Artes e Ciências Humanas possuem desempenho superior, quando comparado com alunos de Ciências Agrárias. Fato inverso ocorre para os alunos dos cursos de Ciências Exatas

Também foi constatado que os alunos que se formam, saem com o IRA superior do que os alunos que optaram pela mobilidade.

5.5.2.1 Análise do IRA do semestre dos ingressantes no 2º Vestibular de 2004 utilizando dados em painel

Para a análise dos fatores associados ao desempenho dos alunos no decorrer dos semestres, foi utilizado o método de dados em painel, tendo como variável resposta o IRA do semestre e, como variáveis explicativas, fatores sociodemográficos, trajetória e inserção universitária. Para tornar o painel mais balanceado foram considerados os alunos que saíram da universidade até o 2º Semestre de 2010. Com a restrição citada e

a o fato de levar em consideração apenas indivíduos com todas as variáveis disponíveis, a análise contou com 5176 observações de 733 estudantes

Inicialmente, foram verificadas quais variáveis são significativas no modelo utilizando o teste t. Dentre as variáveis testadas, foram incluídas duas variáveis quantitativas que variam no tempo e que indicam, para cotistas e não cotistas, em qual semestre determinado IRA foi obtido. O objetivo dessa inclusão é verificar se o IRA do semestre dos cotistas e não cotistas aumentam, ou não, no decorrer do tempo. Para evitar problemas relacionados a multicolinearidade entre idade e as variáveis adicionadas, foi retirada a variável idade da análise.

Em seguida, foi feito o modelo de efeitos fixos e o modelo de efeitos aleatórios. Para a coorte dos alunos que ingressaram no 2º Vestibular de 2004, o teste de Hausman indica que o painel com efeitos fixos é mais adequado do que o painel com efeitos aleatórios, ou seja os efeitos individuais fazem parte do intercepto.

De acordo com o resultado do teste de Hausman, existe evidência estatisticamente significativa da importância da heterocedasticidade individual, ou seja, é possível supor que existe um fator individual de empenho que influencia de maneira significativa o comportamento do IRA do aluno no decorrer do semestre.

A tabela abaixo apresenta as estimativas dos coeficientes, o grau de significância de acordo com o teste t e o valor do R^2 e do Rho , que representa a razão entre variância devido ao indivíduo sobre a variância total para três modelos: Within, o modelo de efeitos aleatórios e o Pooled.

Tabela 31: Estimativas dos modelos em painel Pooled, Within e Efeitos aleatórios - 2004

Variáveis		Pooled	Within	Efeitos Aleatórios
	Intercepto	2,527***	3,82***	2,49***
	Argumento final	0,001***	-	0,0004*
	Ano*Cotista	0,007	0,013*	0,012*
Forma de saída	Formatura	1,229***	-	1,39***
	Mobilidade	0,565***	-	0,65***
Grande Área	C. Biológicas	-0,087	-	-0,08
	C. da Saúde	0,169	-	0,17*
	C. Exatas	-0,245***	-	-0,25**
	C. Humanas	0,363***	-	0,40***
	C. Sociais	0,171***	-	0,20**
	Engenharias	-0,255***	-	-0,18*
	Linguística L. e Artes	0,369***	-	0,40***
Sistema de Ingresso	Não Cotista	0,091***	-	0,10**
Sexo	Masculino	-0,207***	-	-0,22***
Fez ou está fazendo curso superior	Sim, fora da UnB	-0,096**	-	-0,07
	Sim, na UnB	-0,245***	-	-0,28***
Instituição de Ensino	Pública	0,086**	-	0,10*
Exercia alguma atividade remunerada	Sim, em tempo parcial	0,059	-	0,08
	Sim, em tempo integral	0,173***	-	0,14*
	Sim, mas trata-se de trabalho eventual	-0,085*	-	-0,09
Ajuste do modelo	R ²	0,258	0,518	0,22
	Rho	-	0,46	0,27

Códigos de Significância: <0,0001 “***”, <0,001 “**”, <0,05 “*”, <0,10 “.”

De acordo com o teste de Breusch-Pagan, rejeitou-se a hipótese de variância constante para todos os modelos. Portanto, para corrigir o efeito da heterocedasticidade, foram utilizados erros-padrões robustos (HC3) nos parâmetros do modelo.

O R² do modelo Pooled de 0,22 indica a presença de efeitos (fixos ou aleatórios) que contribuem para explicar a variância do IRA semestre a semestre.

O R² do modelo Within é obtido por meio da regressão LSDV (mínimos quadrados com variáveis dummies para cada indivíduo) e é dado por 0,518. Para o modelo aleatório o R² é igual a 0,25.

O valor Rho representa o total da variância dos dados que ocorrem devido ao efeito individual. No modelo fixo, esse valor é igual a 0,46 e no modelo aleatório, igual a 0,22,

o que ressalta a importância da heterogeneidade individual na explicação da variabilidade do IRA no tempo.

Em relação aos resultados, é possível afirmar que as variáveis sociodemográficas que estão relacionadas ao desempenho dos alunos de maneira significativa são o sexo do estudante, indicando que mulheres tem notas mais alta do que os homens.

Em relação a trajetória pré universitária, foi verificado que ter feito universidade anteriormente, principalmente tendo sido na UnB, influencia negativamente o IRA do semestre do estudante. Em relação a atividade remunerada exercida antes do vestibular, é possível perceber que alunos que realizavam trabalhos eventuais possuem IRA do semestre mais baixo do que alunos que não trabalhavam ou exerciam atividades em tempo integral ou parcial. Assim como nas outras análises, também foi verificado que o Argumento final do Vestibular está relacionado ao IRA do semestre do estudante.

Em relação ao tipo de escola em que o aluno cursou o ensino médio, de acordo com o modelo Pooled e o modelo de efeitos aleatórios, essa diferença é significativa, indicando que alunos de escolas públicas tem desempenho superior aos alunos de escolas particulares, dadas as demais variáveis constantes.

Quanto a inserção universitária, os resultados são semelhantes à análise *cross-section*. Em relação as essas variáveis, é possível constatar que os alunos não cotistas possuem desempenho superior ao dos alunos cotistas. Nos modelos Pooled e de efeitos aleatórios, também é possível verificar que alunos das áreas de Letras e Artes e Ciências Humanas possuem desempenho superior aos alunos de Ciências Agrárias. Fato inverso ocorre para os alunos dos cursos de Ciências Exatas e Engenharias, que possuem rendimento inferior aos alunos dos cursos de Ciências Agrárias.

Também foi constatado que os alunos que se formam, possuem IRA do semestre superior do que os alunos que optam pela mobilidade.

Em relação as variáveis indicadoras do semestre para cotistas e não cotistas, é verificado que ela é significativa apenas para os cotistas. De acordo com o parâmetro estimado, a cada semestre o IRA semestral dos cotistas aumenta em média 0,013 pontos, considerando as demais variáveis constantes, ou seja, existem evidências estatisticamente significantes de o desempenho dos cotistas estar melhorando no período considerado.

5.5.2.2 Análise do IRA do semestre dos ingressantes no 2º Vestibular de 2009 utilizando dados em painel

Para a análise dos fatores associados ao desempenho dos alunos ingressantes em 2009 no decorrer dos semestres, foi utilizado o método de dados em painel, tendo como variável resposta o IRA do semestre e, como variáveis explicativas, fatores sociodemográficos, trajetória e inserção universitária. Como foram considerados na modelagem apenas os indivíduos com todas as variáveis disponíveis, a análise contou com 10255 observações de 1685 estudantes.

Inicialmente, foram verificadas quais variáveis são significativas no modelo utilizando o teste t. Assim como em 2004, dentre as variáveis testadas, foram incluídas duas variáveis quantitativas que variam no tempo e indicam, para cotistas e não cotistas, em qual semestre determinado IRA foi obtido. O objetivo dessa inclusão é verificar se o IRA do semestre dos cotistas e não cotistas aumentam, ou não, no decorrer do tempo. Para evitar problemas relacionados a multicolinearidade entre idade e as variáveis adicionadas, foi retirada a variável idade da análise. Como nenhuma das duas variáveis mostrou ter efeitos significativos no modelo, elas foram desconsideradas e foi utilizada a variável Idade, que varia no tempo

Em seguida, foi feito o modelo de efeitos fixos e o modelo de efeitos aleatórios. Com os modelos feitos foi realizado teste de Hausman, que testa hipótese de que os efeitos individuais são não correlacionados com algum regressor. De acordo com a estatística do teste, o painel com efeitos fixos é mais adequado do que o painel com efeitos aleatórios, ou seja os efeitos individuais fazem parte do intercepto.

De acordo com o resultado do teste de Hausman, assim como em 2004, a heterogeneidade individual tem relevância estatística, ou seja, é possível supor que existe um fator individual de empenho que influencia de maneira significativa o comportamento do IRA do aluno no decorrer do tempo.

A tabela abaixo apresenta as estimativas dos coeficientes, o grau de significância de acordo com o teste t e o valor do R^2 e do Rho para três modelos: Within, o modelo de efeitos aleatórios e o Pooled.

Tabela 32: Estimativas dos modelos em painel Pooled, Within e Efeitos aleatórios - 2009

		Pooled	Within	Modelo Aleatório
	Intercepto	3,27***	3,76***	3,26***
	Argumento final	0,0012***	-	0,0011***
	Idade	0,01**	-0,014*	0,005
Forma de Saída	Evadiu	-1,37***	-	-1,45***
	Formatura	0,37***	-	0,38***
	Mobilidade	-0,44**	-	-0,58***
Grande Área	C. Biológicas	0,07	-	0,09
	C. da Saúde	0,13**	-	0,14*
	C. Exatas	-0,11**	-	-0,11
	C. Humanas	0,35**	-	0,37***
	C. Sociais	0,21**	-	0,21**
	Engenharias	-0,02	-	0,01
	Linguística L. e Artes	0,47**	-	0,46***
	Multidisciplinar	0,01	-	0,00
Sistema de Ingresso	Universal	0,08**	-	0,11**
Sexo	Masculino	-0,30**	-	-0,32**
Exercia alguma atividade remunerada	Sim, mas se tratava de trabalho eventual.	-0,11*	-	-0,14*
	Sim, em tempo parcial	-0,20***	-	-0,23**
	Sim, em tempo integral	-0,06*	-	-0,09
Fez ou está fazendo curso superior	Sim, fora da UnB	0,04	-	0,04
	Sim, na UnB	-0,14**	-	-0,17***
Qual curso médio concluiu ou estava está concluindo	Supletivo	-0,33**	-	-0,34***
	Outros (profissionalizantes, técnico, etc)	0,00	-	-0,01
Ajuste do modelo	R ²	0,31	0,57	0,51
	Rho	-	0,70	0,46

Códigos de Significância: <0,0001 “***”, <0,001 “**”, <0,05 “*”, <0,10 “.”

De acordo com o teste de Breusch-Pagan, rejeitou-se a hipótese de variância constante para todos os modelos. Portanto, para corrigir o efeito da heterocedasticidade, foram utilizados erros-padrões robustos (HC3) nos parâmetros do modelo.

O R² do modelo Pooled de 0,31 indica a presença de efeitos (fixos ou aleatórios) que contribuem para explicar a variância do IRA semestre a semestre.

O R² do modelo Within é obtido por meio da regressão LSDV (mínimos quadrados com variáveis dummies para cada indivíduo) e é dado por 0,57. Para o modelo aleatório o R² é igual a 0,51. Portanto, para essa coorte, o modelo de efeitos fixos e aleatórios

apresentam capacidades equiparáveis de explicação da variância do IRA no tempo, diferentes dos resultados observados para 2004.

O valor *Rho* representa a razão entre variância devido ao indivíduo sobre a variância total, ou seja, o total da variância dos dados devido ao efeito individual. No modelo fixo, esse valor é igual a 0,70 e no modelo aleatório, igual a 0,46, reforçando os resultados do teste de Hausman, que indicam forte presença de heterogeneidade individual.

De acordo com o resultado para as variáveis relacionadas aos aspectos sociodemográficos em 2009, nos dois modelos, as mulheres possuem rendimento acadêmico estatisticamente superior ao dos homens. Por outro lado, ao se analisar o modelo de efeitos fixos, a cada ano a mais de idade do aluno, o IRA do semestre se reduz em média 0,014 pontos, considerando as demais variáveis constantes.

Em relação a trajetória pré-universitária, nos modelos Pooled e Aleatório, foi verificado que ter feito universidade na UnB anteriormente influencia negativamente o IRA do semestre do estudante. Em relação a atividade remunerada exercida antes do vestibular, é possível perceber que alunos que trabalhavam possuem IRA do semestre mais baixo do que alunos que não trabalhavam, principalmente se o trabalho era em tempo parcial ou eventual. Assim como nas outras análises, também foi verificado que o Argumento final do Vestibular está relacionado ao IRA do semestre do estudante.

Quanto a inserção universitária, é possível constatar que os alunos não cotistas possuem desempenho superior ao dos alunos cotistas, de acordo com os modelos Pooled e Aleatório. Também é possível verificar que alunos das áreas de Letras e Artes Ciências Sociais e Ciências Humanas possuem desempenho superior aos alunos dos cursos de Ciências Agrárias. Também foi constatado que os alunos que já se formaram, possuem IRA do semestre superior aos alunos que ainda estão ativo. Além disso, é constatado que os alunos que optaram pela mobilidade possuem IRA do semestre menor do que os alunos que ainda estão ativos.

5.5.3 Análise da Evasão utilizando Regressão Logística

5.5.3.1 Análise da Evasão – Ingressantes no 2º Vestibular de 2004

Para a modelagem logística, que tem como variável resposta a evasão, foram inicialmente utilizadas as doze variáveis já mencionadas no estudo. Após testar o efeito de cada variável utilizando o teste Qui-Quadrado individualmente, foram incluídas no modelo apenas cinco. Como foram considerados na modelagem apenas os indivíduos com todas as variáveis disponíveis, a análise contou com 1153 estudantes.

A Tabela 39 mostra as variáveis selecionadas, os valores das estatísticas dos testes de Wald, o desvio padrão e seus respectivos p-valores.

Tabela 33: Estimativa, Desvio Padrão e Estatísticas de Wald para regressão logística - 2004

	Variável	Estimativa	D. Padrão	Valor Z	P- Valor
	Intercepto	-2,02	0,46	-4,42	0,000***
	Argumento final	-0,002	0,00	-2,72	0,006***
	Idade	0,01	0,02	0,68	0,497
Grande Área	C. Biológicas	-1,78	1,09	-1,64	0,102
	C. da Saúde	-0,67	0,50	-1,34	0,179
	C. Exatas	0,79	0,38	2,07	0,039*
	C. Humanas	0,23	0,39	0,60	0,548
	C. Sociais	-0,50	0,39	-1,29	0,198
	Engenharias	0,72	0,43	1,68	0,094.
	Linguística L. e Artes	0,13	0,41	0,33	0,740
Exercia alguma atividade remunerada	Sim, em tempo parcial	0,91	0,28	3,27	0,001***
	Sim, em tempo integral	0,94	0,26	3,63	0,000***
	Sim, mas trata-se de trabalho eventual	0,33	0,35	0,95	0,343
Fez ou está fazendo curso superior	Sim, fora da UnB	1,07	0,24	4,39	0,000***
	Sim, na UnB	0,86	0,23	3,72	0,000***

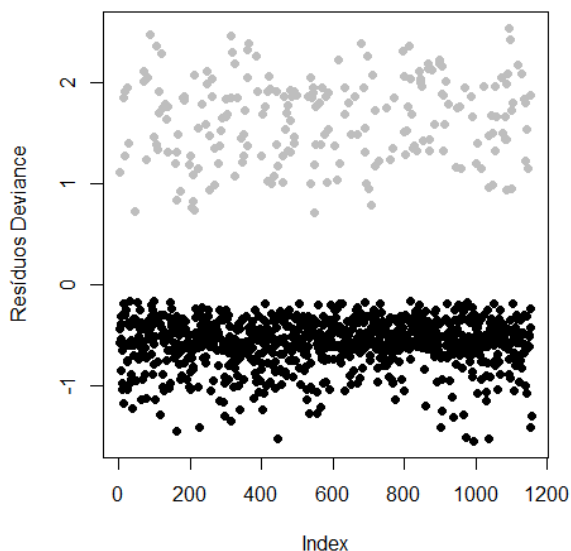
Códigos de Significância: <0,0001 “***”, <0,001 “**”, <0,05 “*”, <0,10 “.”

O teste da razão de verossimilhança com a hipótese nula de que todos os parâmetros do modelo são iguais à zero foi rejeitado com p-valor menor que 0,001, ou seja, o modelo é estatisticamente significativo.

No diagnóstico do modelo foram utilizados os gráficos com resíduos Deviance para identificar observações discrepantes ou que não foram bem explicadas pelo modelo. No

gráfico, as observações estão dentro do intervalo desejado e nenhuma observação se sobressai.

Figura 7: Resíduos Deviance do modelo logístico - 2004



Na análise dos DfBetas, apenas uma observação representou grande influência na estimação do parâmetro Grande Área ao nível Ciências Biológicas, para um aluno que evadiu, portanto, pode-se afirmar que o modelo é adequado aos dados e partir para a análise das razões de chance.

Tabela 34: Estimativas das Razões de Chances para as variáveis significativas do modelo logístico - 2004

	Variável	Razão de chances	Intervalo de confiança (95%)	
	Argumento final	0,997	0,996	0,999
	Idade	1,010	0,977	1,038
Grande Área	C. Biológicas Vs C. Agrárias	0,169	0,009	0,978
	C. da Saúde Vs C. Agrárias	0,514	0,190	1,361
	C. Exatas Vs C. Agrárias	2,218	1,075	4,909
	C. Humanas Vs C. Agrárias	1,260	0,608	2,794
	C. Sociais Vs C. Agrárias	0,609	0,289	1,364
	Engenharias Vs C. Agrárias	2,076	0,913	4,970
	Linguística L. Artes Vs C. Agrárias	1,143	0,528	2,617
Exercia alguma atividade remunerada	Sim, em tempo parcial Vs. Não trabalhava	2,468	1,424	4,242
	Sim, em tempo integral Vs. Não trabalhava	2,542	1,530	4,228
	Sim, mas tratava-se de trabalho eventual Vs. Não trabalhava	1,386	0,678	2,693
Fez ou estava fazendo curso superior	Sim, fora da UnB Vs. Não estava	2,962	1,823	4,780
	Sim, na UnB Vs. Não estava	2,374	1,499	3,730

Em relação aos resultados apresentados pelo modelo, é possível afirmar que a única variável sociodemográfica significativa é a idade do estudante, em que, a cada ano a mais de vida a chance de evadir aumenta em média por um fator de 1,01.

Em relação a trajetória pré universitária, foi verificado que o fato de ter feito algum curso superior anteriormente influencia a chance de evasão dos alunos. Nesse caso, alunos que estavam cursando ensino superior fora da UnB e na Universidade de Brasília possuem 2,96 e 2,34, respectivamente, mais chances de evadir do que os alunos que não faziam curso superior.

A chance do estudante da UnB evadir tendo exercido uma atividade remunerada em tempo integral é 154% maior do que os estudantes que não exerciam. Fenômeno semelhante ocorre para os alunos que exerciam atividades em tempo parcial ou de modo eventual que possuem 146% e 38%, respectivamente, mais chances de evadir do que o aluno que não trabalhava, mantendo as demais variáveis constantes.

Também foi verificado que o aumento no Argumento final do Vestibular reduz a chance de evasão dos estudantes.

Quanto a inserção universitária, não é possível constatar diferenças significativas na chance de evasão entre alunos cotistas e não cotistas. Porém, é possível verificar que alunos das áreas de Ciências da Saúde, Ciências Biológicas e Ciências Sociais possuem menores chances de evasão em relação ao grupo de alunos dos cursos de Ciências Agrárias. Por outro lado, alunos dos cursos de Ciências Exatas e Engenharias possuem maior chance de evasão, com 2,21 e 2,07 mais chances de evadir do que alunos dos cursos de Ciências Agrárias.

5.5.3.2 Análise da Evasão – Ingressantes no 2º Vestibular de 2009

Em relação a coorte de alunos ingressantes no 2º Vestibular de 2009, após testar o efeito de cada variável individualmente utilizando o teste Qui-quadrado, foram incluídas sete variáveis no modelo. Como foram considerados na modelagem apenas os indivíduos com todas as variáveis disponíveis, a análise contou com 1911 estudantes.

A tabela 37 mostra as variáveis selecionadas, os valores das estatísticas dos testes de Wald, o desvio padrão e seus respectivos p-valores.

Tabela 35: Estimativas, Desvio Padrão, Estatísticas de Wald para o modelo logístico - 2009

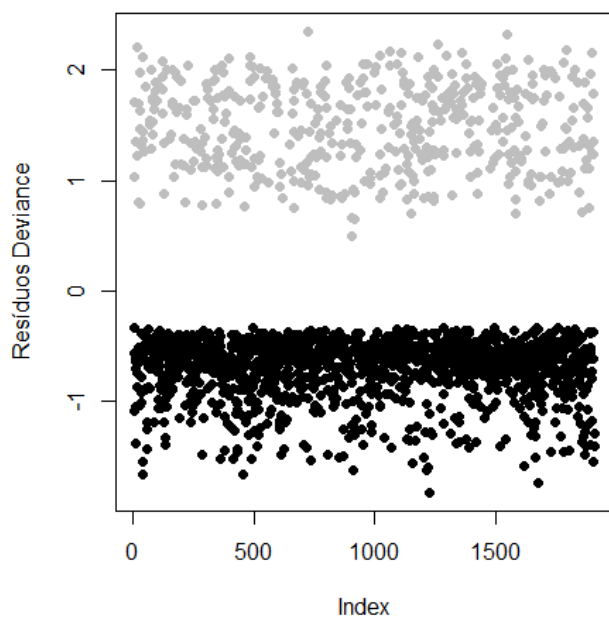
	Variável	Estimativa	D. Padrão	Teste de Wald	P- Valor
	Intercepto	-1,76	0,52	-3,40	0,001***
	Argumento final	0,00	0,00	-4,39	0,000***
	Idade	0,02	0,01	2,30	0,021*
Grande Área	C. Biológicas	0,57	0,45	1,27	0,205
	C. da Saúde	0,55	0,34	1,64	0,102
	C. Exatas	1,15	0,32	3,57	0,000***
	C. Humanas	0,72	0,32	2,26	0,024*
	C. Sociais	0,02	0,33	0,06	0,950
	Engenharias	0,86	0,33	2,61	0,009***
	Linguística L. e Artes	0,47	0,32	1,45	0,148
	Multidisciplinar	1,22	0,40	3,08	0,002***
	Exercia alguma atividade remunerada	Sim, mas se tratava de trabalho eventual.	0,47	0,21	2,24
Sim, em tempo parcial		0,53	0,22	2,46	0,014**
Sim, em tempo integral		0,87	0,18	4,86	0,000***
Fez ou está fazendo curso superior	Sim, fora da UnB Vs. Não estava	0,24	0,18	1,35	0,176
	Sim, na UnB Vs. Não estava	0,83	0,16	5,34	0,000***
Qual curso médio concluiu ou estava está concluindo	Supletivo Vs. Regular	0,71	0,27	2,65	0,008**
	Outros Vs. Regular	0,02	0,28	0,08	0,939
Renda Familiar	1 a 2 SM	-0,83	0,47	-1,79	0,073.
	2 A 3 SM	-0,80	0,43	-1,89	0,059.
	3 A 6 SM	-0,81	0,40	-2,01	0,044*
	6 A 10 SM	-0,78	0,41	-1,92	0,054.
	10 A 14 SM	-0,75	0,41	-1,82	0,069.
	14 A 20 SM	-0,62	0,42	-1,47	0,141
	Mais de 14 SM	-0,56	0,41	-1,36	0,174

Códigos de Significância: <0,0001 “***”, <0,001 “**”, <0,05 “*”, <0,10 “.”

O teste da razão de verossimilhança com a hipótese nula de que todos os parâmetros do modelo são iguais à zero foi rejeitado com p-valor menor que 0,001, ou seja, o modelo é estatisticamente significativo.

No diagnóstico do modelo foram utilizados os gráficos com resíduos Deviance para identificar observações discrepantes ou que não foram bem explicadas pelo modelo. No gráfico, as observações estão dentro do intervalo desejado e nenhuma observação se sobressai.

Figura 8: Resíduos Deviance do modelo logístico - 2009



Na análise dos DfBetas, apenas uma observação representa grande influência na estimação do parâmetro idade, para um estudante que não evadiu. Nos outros gráficos este indivíduo está dentro do intervalo esperado, portanto, no geral, o modelo é adequado aos dados.

Tabela 36: Estimativas das Razões de Chances para as variáveis significativas do modelo logístico - 2009

	Variável	Razão de chances	Intervalo de confiança (95%)	
	Intercepto	0,17	0,06	0,47
	Argumento final	1,00	1,00	1,00
	Idade	1,02	1,00	1,05
Grande Área	C. Biológicas Vs. C. Agrárias	1,77	0,72	4,24
	C. da Saúde Vs. C. Agrárias	1,73	0,91	3,42
	C. Exatas Vs. C. Agrárias	3,16	1,72	6,10
	C. Humanas Vs. C. Agrárias	2,05	1,12	3,93
	C. Sociais Vs. C. Agrárias	1,02	0,55	1,99
	Engenharias Vs. C. Agrárias	2,35	1,26	4,59
	Linguística L. e Artes Vs. C. Agrárias	1,60	0,86	3,10
	Multidisciplinar Vs. C. Agrárias	3,40	1,58	7,51
	Exercia alguma atividade remunerada	Sim, mas se trata de trabalho eventual. Vs. Não trabalhava	1,61	1,05
Sim, em tempo parcial Vs. Não trabalhava		1,70	1,11	2,58
Sim, em tempo integral Vs. Não trabalhava		2,40	1,68	3,41
Fez ou estava fazendo curso superior	Sim, fora da UnB Vs. Não fazia	1,27	0,90	1,78
	Sim, na UnB Vs. Não fazia	2,29	1,69	3,10
Qual curso médio concluiu ou estava concluindo	Supletivo Vs. Regular	2,04	1,20	3,46
	Outros (profissionalizantes, técnico, etc) Vs. Regular	1,02	0,58	1,75
Renda Familiar	1 a 2 SM Vs. Até um SM	0,43	0,17	1,08
	2 A 3 SM Vs. Até um SM	0,45	0,19	1,04
	3 A 6 SM Vs. Até um SM	0,44	0,20	0,99
	6 A 10 SM Vs. Até um SM	0,46	0,21	1,02
	10 A 14 SM Vs. Até um SM	0,47	0,21	1,07
	14 A 20 SM Vs. Até um SM	0,54	0,23	1,24
	Mais de 14 SM Vs. Até um SM	0,57	0,26	1,29

Em relação aos resultados apresentados pelo modelo, as variáveis sociodemográficas significativas no modelo são idade do estudante e renda familiar. Em relação a idade, a cada ano a mais de vida no momento do vestibular a chance de evadir aumenta por um fator médio de 1,02. Quanto a renda familiar, é possível verificar a tendência de crescimento na chance de evasão de acordo com o aumento da faixa de renda. Nesse caso, os estudantes de famílias que ganham menos de 1 salário mínimo possuem chances de evasão maiores do que todas as outras faixas, porém essa chance aumenta à medida que aumenta a faixa de renda dos alunos.

Em relação a trajetória pré universitária foi verificado que o fato de ter feito algum curso superior anteriormente aumenta a chance de evasão dos alunos. Nesse caso, alunos que estavam cursando ensino superior fora da UnB e na Universidade de Brasília possuem 1,27 e 2,29, respectivamente, mais chances de evadir do que os alunos que não faziam curso superior. Resultado diferente do verificado em 2004, quando alunos que estudavam fora da Universidade de Brasília tinham maior chance de evasão do que alunos que estudaram na UnB.

Os resultados para evasão de acordo com o tipo de atividade remunerada exercida no período do vestibular são semelhantes aos encontrados em 2004. No caso de 2009, a chance do estudante da UnB evadir tendo exercido uma atividade remunerada em tempo integral é 140% maior do que os estudantes que não exerciam. Fenômeno semelhante ocorre para os alunos que exerciam atividades em tempo parcial ou de modo eventual que possuem 70% e 61%, respectivamente, mais chances de evadir do que os alunos que não trabalhavam, mantendo as demais variáveis constantes.

Também foi verificado que o aumento no Argumento final do Vestibular reduz a chance de evasão dos estudantes.

Quanto a inserção universitária, também não foi possível constatar diferenças significativas na chance de evasão entre alunos cotistas e não cotistas. Porém, é possível verificar que alunos das áreas de Ciências Agrárias e Ciências Sociais possuem menores chances de evasão. Por outro lado, alunos dos cursos Multidisciplinares, Ciências Exatas possuem maior chance de evasão, com 3,40 e 3,16 mais chances de evadir do que alunos dos cursos de Ciências Agrárias.

6. Conclusão

A pesquisa desenvolvida na presente monografia avaliou as coortes de estudantes ingressantes no 2º vestibular de 2004 e no 2º Vestibular de 2009 da Universidade de Brasília. Abordou aspectos de caracterização de cotistas e não cotistas, assim como do rendimento acadêmico e da evasão desses estudantes, levando em consideração variáveis sociodemográficas, de trajetória pré-universitária e universitária.

Na avaliação do rendimento acadêmico, foi verificado que alunos não cotistas tem rendimento acadêmico superior aos alunos não cotistas. Esse resultado é semelhante com o de pesquisas anteriores como Mendes (2013) na UERJ, e Tannuri e Pianto (2012) na UnB. Porém, apesar das diferenças significativas, ao se avaliar os IRAs médios dos semestres em 2004 e 2009, é possível verificar que existe uma tendência de aumento no IRA semestral médio no decorrer dos anos, além da aproximação entre as distribuições dos IRAs dos alunos cotistas e não cotistas. Essa aproximação indica que a diferença inicial se reduz a partir da convivência e do amadurecimento dos estudantes. Por outro lado, o ponto de destaque no modelo em painel é a presença de efeitos fixos tanto em 2004 quanto em 2009, indicando a relevância da heterogeneidade individual, ou seja, a existência de um fator de esforço e empenho individual significativo ao explicar o rendimento acadêmico, mesmo considerando todas as demais variáveis do estudo.

Dentre as variáveis demográficas estudadas, destaca-se que as mulheres têm rendimento superior aos homens, considerando as demais variáveis constantes. Esse melhor rendimento das mulheres já havia sido detectado em estudos de Tannuri e Pianto (2012).

Indo de encontro com pesquisas que apontam um melhor rendimento acadêmico para alunos com pais com melhor formação e melhores salários, não foi verificado que essas variáveis são significativas ao explicar o IRA final dos alunos. Dentre as suposições para a não detecção da significância dessas variáveis está a possível pouca qualidade da informação fornecida, já que se trata de informações relativas a terceiros, no caso, aos pais. Também é possível supor que o alto índice não resposta possa ter interferido na detecção do efeito desses fatores sobre o IRA final.

As variáveis que levam em consideração a trajetória pré-universitária dos alunos, na maioria dos modelos, se mostraram significativas. Dentre elas se destaca o fato de cursar ensino superior, na UnB ou fora, antes de ingressar na universidade. Nesse caso,

alunos que já faziam curso superior possuem rendimento inferior ao dos alunos que ainda não estavam na faculdade.

Quanto a avaliação das grandes áreas, é notável que alunos pertencentes aos cursos de engenharia e ciências exatas possuem rendimento acadêmico inferior, comparado aos estudantes de outros cursos. Nessas grandes áreas, também há maior discrepância entre as distribuições dos IRA Finais de cotistas e não cotistas para os ingressantes em 2004 e em 2009. Essa constatação vai de acordo com o que foi verificado em Cardoso (2008), que mostrou que nos cursos de Ciências com maior prestígio social, em mais da metade dos cursos há diferenças consideráveis favoráveis aos não cotistas quanto ao rendimento acadêmico dos alunos.

Dentre as limitações do estudo do rendimento acadêmico, verifica-se que os modelos de regressão possuem limitações técnicas ao usarem variável resposta com limites definidos entre 0 e 5. Além disso, o percentual alto de não respostas, que variam entre 60% e 25% nos modelos, não permite resultados tão fieis a realidade dos estudantes da universidade.

Na análise da evasão, foi constatado que os cotistas evadiram em menor proporção em 2004, porém passaram a evadir em maior proporção em 2009. Entretanto, não é possível afirmar que essas diferenças entre cotistas e não cotistas são estatisticamente significativas.

Os resultados corroboraram com diversas pesquisas que indicam que o pico de evasão no ensino superior ocorre nos dois primeiros anos de curso. Além disso, acompanhando o aumento do número de vagas no vestibular, houve um aumento no índice de evasão até o quarto ano em 2009, quando comparado com 2004.

Em relação as variáveis sociodemográficas, se destaca o fator idade do estudante, em que alunos mais velhos possuem maior chance de evadir do que os alunos mais novos.

Quando analisa-se a trajetória pré-universitária, é verificado que o fato de ter feito algum curso superior anteriormente aumenta a chance de evasão dos alunos. Seguindo a mesma lógica, os resultados para a evasão de acordo com o tipo de atividade remunerada exercida no período do vestibular são semelhantes aos resultados encontrados por Cardoso (2008), que mostrou que a situação de trabalho influencia na decisão de abandonar ou permanecer no curso. Estudantes que trabalhavam em tempo integral tem maiores chances de evadir do que os demais. Outro ponto de destaque da

análise é a alta chance de evasão de alunos que fizeram supletivo, quando comparado com os demais alunos.

Destaca-se que alunos de Engenharia, Ciências Exatas e cursos multidisciplinares (em 2009), possuem chances de evasão significativamente maiores do que os demais, como observado em Mendes (2013), que mostra que a evasão entre cotista e não cotistas é grande em cursos considerados de alta dificuldade.

Assim como nos modelos que explicam o rendimento acadêmico, a limitação da análise da evasão está no percentual elevado de não respostas, não permitindo acessar a dimensão completa dos dados dos alunos ingressantes na Universidade nas duas coortes.

Como sugestão para trabalhos futuros, recomenda-se estudar as coortes mais recentes de estudantes, quando o preenchimento do QSE passou a ser obrigatório e o índice de resposta se tornou próximo a 100%. Também é importante, uma análise que utilize técnicas de avaliação de políticas públicas para a comparação entre o desempenho acadêmico dos dois grupos, como por exemplo, escores de propensão e regressão descontínua.

7. Bibliografia

ALA-HARJA, Marjukka; HELGASON. S. 2000. Em direção às melhores práticas de avaliação. *Revista do Serviço Público*, Brasília.

ARAÚJO I. E., 2013. Análise socioeconômica das qualidades de cotas para negros na Universidade de Brasília. Dissertação de Mestrado, Programa de Pós Graduação da Faculdade de Administração, Ciências Contábeis e Economia, Universidade de Brasília.

BARROS, R. P; FOGUEL M. N., 2012. Avaliação econômica de projetos sociais; organizador Naercio Menezes Filho. – 1. Ed - São Paulo: Dinâmica Gráfica e Editora.

BALTAGI B. H.. *Econometric Analysis of panel data*, 4th edition, John Wiley & Son, 2008

BRASIL. Lei nº 12.711, de 29 de Agosto de 2012. Torna obrigatório que instituições federais de educação superior vinculadas ao Ministério da Educação reservem, no mínimo 50% de suas vagas para estudantes que tenham cursado integralmente o ensino médio em escolas públicas. Diário oficial [da] República do Brasil, Brasília, DF, 29 ago. 2012. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12711.htm>. Acesso em: 05 de mai. 2014.

_____, Ministério da Educação e Cultura. Secretaria de Ensino Superior. Comissão Especial de Estudos sobre Evasão nas Universidades Públicas Brasileiras. ANDIFES/ABRUEM, SESu. Brasília: MEC, 1995.

BREUSCH, T.S.; PAGAN, A.R.. "Simple test for heteroscedasticity and random coefficient variation". *Econometrica* (The Econometric Society), 1979.

CAPES. Tabela de Áreas de conhecimento. Disponível em <http://www.ufrb.edu.br/prograd/legislacao-e-normas/doc_download/1097-tabela-de-areas-de-conhecimento-cnpq> Acesso em: 3 de março de 2014

CARDOSO, C. B., 2008. Efeitos da Política de Cotas na Universidade de Brasília: Uma Análise do Rendimento e da Evasão. Tese de Mestrado, Faculdade de Educação, Universidade de Brasília.

CESPE/UnB. Edital n 3/2004, de 18 de março de 2004. Referente ao 2º Vestibular de 2004. Brasília: CESPE/UnB, 2004

_____. Edital n 4/2007, de 1º de outubro de 2007. Referente ao 1º Vestibular de 2008. Brasília: CESPE/UnB, 2007

COOK, R. D. AND WEISBERG, S.. *Residuals and Influence in Regression*. Wiley, 1984.

DRAPER, N. R. & SMITH, H. *Applied Regression Analysis*. Second Edition. New York: John Wiley, 1981.

FRANCIS, A.; TANNURI-PIANTO, M. E.. *Using Brazil's Racial Continuum to examine tem Short-Term Effects of Affirmative Action in Higher Education*. *Journal of Human Resources*, 2012.

HOSMER, D. W.; LEMESHOW, S.. *Applied logistic regression*. New York: John Wiley, 2000.

LONGNECKER, M.T & OTT, R.L. *A First Course in Statistical Methods*. Thomson Brooks/Cole, 2004.

LOUREIRO A. O. F.; SULIANO D. C.. Uma breve discussão sobre os modelos com dados em painel. Governo do Estado do Ceará: SEPLAG/IPECE, Nota Técnica, 2009.

MENDES, A. A. F.. Três ensaios sobre ações afirmativas no ensino superior brasileiro: Acesso, progressão e simulações de diferentes políticas de cotas para a Universidade do Estado do Rio de Janeiro. Tese de Mestrado, Faculdade de Economia, Universidade do Estado do Rio de Janeiro, 2013.

MOEHLECKE, Sabrina. Ação Afirmativa: História e Debates no Brasil. *Cadernos de Pesquisa* (Fundação Carlos Chagas), São Paulo, v. 117, p. 197-218, 2003.

PARK, H. M.. *Regression Models for Panel Data Using Stata*. *Public Management & Policy Analysis Program*, International University of Japan, 2011.

PEDROSA, Renato et al.. *Academic Performance, Students*. Higher Education Management and Policy, Paris, 2009.

RAMSEY, J.B. Tests for Specification Errors in Classical Linear Least Squares Regression Analysis, *Journal of the Royal Statistical Society, Series B.*, 1969.

RISTOFF, Dilvo - *Evasão: Exclusão ou Mobilidade*. (MIMEO) Santa Catarina, UFSC, 1995.

TANNURI-PIANTO, M, E., TORRES, I. O. Estudo sobre o perfil dos candidatos e alunos da Universidade de Brasília no período pós-cotas raciais: 2004-2011, 2011. In: apresentado ao 9º Congresso de Iniciação Científica do DF e 18º Congresso de Iniciação Científica da UnB, Brasília, 2012. Anais do 18º Congresso de Iniciação Científica, 2012

UNIVERSIDADE DE BRASÍLIA. Informações sobre a história e números da Universidade de Brasília. Disponível em: <<http://www.unb.br/sobre>> Acesso em: 4 de abril de 2014

_____. Guia do Calouro 1º semestre de 2013. Disponível em: <http://www.unb.br/administracao/decanatos/deg/downloads/index/guia_calouro_1_2013.pdf> Acesso em 20 de novembro de 2013.

WALTENBERG, Fábio; CARVALHO, Márcia de. Cotas aumentam a diversidade de estudantes sem comprometer o desempenho? *Sinais Sociais*, Rio de Janeiro, 2012.