

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO, CONTABILIDADE E CIÊNCIA
DA INFORMAÇÃO E DOCUMENTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA
MESTRADO EM ECONOMIA

**O Programa de Avaliação Seriada da Universidade de Brasília: análise dos
determinantes de acesso ao ensino superior e avaliação do desempenho dos alunos na
universidade**

VANESSA CARVALHO DOS SANTOS

Brasília
2013

VANESSA CARVALHO DOS SANTOS

**O Programa de Avaliação Seriada da Universidade de Brasília: análise dos
determinantes do acesso ao ensino superior e avaliação do desempenho dos alunos na
universidade**

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade de Brasília como parte dos requisitos para obtenção do título de Mestre em Economia.

Orientadora: Professora Doutora Maria Eduarda Tannuri-Pianto.

Brasília

2013

Santos, Vanessa Carvalho dos.

O Programa de Avaliação Seriada da Universidade de Brasília: análise dos determinantes do acesso e avaliação do desempenho dos alunos na universidade / Vanessa Carvalho dos Santos. – 2013.

80 f.: 31 cm

Dissertação (Mestrado) – Universidade de Brasília, Programa de Pós-Graduação em Economia, 2013.

1. Avaliação. 2. Ensino Superior. 3. Modelo Logit 4. Propensity Score Matching. 5. Programa de Avaliação Seriada. I. Tannuri-Pianto, Maria Eduarda. II. Universidade de Brasília. III. Título.

VANESSA CARVALHO DOS SANTOS

**O Programa de Avaliação Seriada da Universidade de Brasília: análise dos
determinantes do acesso e avaliação do desempenho dos alunos na universidade**

A banca examinadora abaixo identificada aprova a dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade de Brasília como parte dos requisitos para obtenção do título de Mestre em Economia.

Prof.^a Dra. Maria Eduarda Tannuri-Pianto (Orientadora)
Departamento de Economia – Universidade de Brasília

Prof. Dr. Rafael Terra de Menezes
Departamento de Economia - Universidade de Brasília

Prof. Dr. Donald Matthew Pianto
Departamento de Estatística – Universidade de Brasília

Brasília, 12 de agosto de 2013

AGRADECIMENTOS

Agradeço primeiramente a Deus, por seu amor infinito e por todas as bênçãos que Ele não cessa de derramar sobre mim e sobre toda a minha família.

Agradeço a minha família, que proporcionou a sólida base sobre a qual caminha todo o meu estudo. Vocês foram, e são, a minha fonte constante de ânimo e perseverança. Obrigada pela paciência e compreensão, e por me ensinarem, desde cedo, que o investimento em capital humano é muito importante, por mais elevado que seja o seu custo de oportunidade no presente, medido em horas de convívio familiar.

Agradeço à minha orientadora, Maria Eduarda, pela confiança, paciência e empenho durante toda a realização do trabalho, pelas reuniões em horários alternativos e pela contribuição fundamental ao resultado final. Obrigada por acreditar em meu potencial e por desatar os nós que pareciam intransponíveis. Nunca me esquecerei de sua atenção e carinho. Agradeço, ainda, aos membros da banca, formada pelos professores Donald Pianto e Rafael Terra, membros da banca, por todas as importantes sugestões e críticas durante a qualificação, as quais possibilitaram que este trabalho evoluísse bastante nos últimos meses.

Agradeço ao André por ser o namorado mais compreensivo que uma estudante de mestrado poderia desejar ter. Obrigada pelo apoio, pelo incentivo incessante, por não se importar em ter que dividir, cada vez mais, o nosso tempo com meus livros e minhas bases de dados.

Agradeço aos meus amigos de longa data, de graduação e aqueles que o mestrado me proporcionou, pelas palavras de incentivo, e por nunca demonstrarem desinteresse ao escutarem, repetidamente, as tensas notícias sobre a construção das bases, os resultados dos testes e das estimações. Em particular, agradeço à Isabel, Liliane, Patrícia e Viviane, pela ajuda essencial para a conclusão deste trabalho. Vocês foram fenomenais. Nunca me esquecerei do apoio tão essencial para a conclusão, à tempo, desta dissertação.

Por fim, agradeço ao Chefe da Pós Graduação em Economia, professor Roberto Ellery, por todo o apoio durante estes dois anos de mestrado, aos funcionários da Secretaria de Pós Graduação, Joana e Carina, por contribuírem sobremaneira para a resolução das questões administrativas atinentes ao programa de mestrado, e à Capes, pelo apoio financeiro concedido durante o primeiro ano de mestrado.

*O Senhor, porém, respondeu: Marta, Marta, preocupas-te e andas agitada
com muitas coisas; porém, uma só coisa é necessária.
Maria escolheu a melhor parte, que lhe não será tirada.*

Lucas 10:41-42

RESUMO

Esta dissertação é formada por dois diferentes capítulos. O primeiro busca analisar os determinantes do acesso à Universidade de Brasília por meio do Programa de Avaliação Seriada (PAS), processo de seleção alternativo ao exame vestibular que ocorre em três etapas. Para tal, estima-se um modelo de regressão logística para os candidatos que desejam ingressar na universidade no primeiro semestre de 2005. O modelo foi estimado para duas diferentes amostras: a primeira composta pelo conjunto de candidatos ao PAS e ao vestibular, e a segunda, pelos candidatos ao PAS e ao vestibular pelo sistema universal e de cotas. Os resultados para ambas as amostras revelam que o PAS possui grande efeito sobre a probabilidade de aprovação. Dentre as variáveis inclusas no modelo, a cidade de residência e o sexo do candidato são as que mais impactam a probabilidade de sucesso no exame. O PAS amplia a chance de aprovação das mulheres na seleção, bem como as chances de aprovação em cursos seletivos.

O segundo capítulo investiga se o PAS exerce efeitos sobre o desempenho do aluno dentro da universidade. Dentre as variáveis de interesse para esta avaliação estão o rendimento acadêmico do aluno, seu tempo de permanência na universidade e a taxa de evasão. Para tanto, utiliza-se a metodologia de escore de propensão (*propensity score*) para se criar um grupo de comparação para os alunos selecionados pelo PAS a partir do grupo de estudantes que ingressaram na universidade por meio do vestibular tradicional, pelo sistema universal. Os resultados indicam que os alunos do PAS possuem melhor desempenho acadêmico medido pelo Índice de Rendimento Acadêmico (IRA) da universidade. Os efeitos do ingresso pelo Programa sobre o tempo de permanência na universidade e sobre a taxa de evasão não foram estatisticamente significativos.

Palavras-chave: Avaliação. Ensino Superior. Modelo Logit. Propensity Score.

ABSTRACT

This thesis is composed by two main chapters. The first one aims to investigate the determinants of access to the University of Brasilia through the Program of Serial Evaluation (PAS), an alternative selection process to the main entrance exam, known as 'vestibular'. First, we estimate a logistic regression model for applicants who wish to attend university in the first semester of 2005. The model was estimated using two different samples. The first one is composed by the set of candidates that will take the PAS exam and also by the candidates that will take the vestibular exam. The second one is also composed by candidates to the racial quota system. The results for both samples reveal that PAS has a great impact on the likelihood of approval. Among the variables included in the model, city of residence and sex are those that most affect the probability of success in the exam. PAS exam increases the likelihood of approval for women as well as the likelihood of approval for selective courses.

The second chapter investigates if PAS affects the student university performance. The variables of interest are the academic performance of the student, the time spent in university until graduation and the dropout rate. We use the propensity score method to create a comparison group to the PAS candidates, which will be selected within the group of students who entered the university through the vestibular. The results indicate that PAS students have better academic performance, measured by the university Academic Performance Index (IRA). The effects of entering the university through PAS on the time spent at the university until graduation and on the dropout rate were not statistically significant.

Key words: Evaluation. University. Logit Model. Propensity Score

LISTA DE ABREVIATURAS E SIGLAS

Cespe	Centro de Seleção e Promoção de Eventos
IRA	Índice de Rendimento Acadêmico
MQO	Mínimo Quadrado Ordinário
PAS	Programa de Avaliação Seriado
QSC	Questionário Socioeconômico
SSA	Secretaria de Assuntos Acadêmicos / UnB
UnB	Universidade de Brasília

SUMÁRIO

CAPITULO 1 DETERMINANTES DO ACESSO AO ENSINO SUPERIOR PELO PROGRAMA DE AVALIAÇÃO SERIADA: O CASO DA UNIVERSIDADE DE BRASÍLIA	13
1.1 Introdução	13
1.2 Revisão da Literatura	14
<u>1.2.1 Os Programas de Avaliação Seriada no Brasil</u>	14
<u>1.2.2 O Programa de Avaliação Seriada da Universidade de Brasília</u>	16
<u>1.2.3 Os Determinantes de Acesso ao Ensino Superior</u>	18
1.3 Metodologia	21
<u>1.3.1 Base de Dados</u>	21
<u>1.3.2 Seleção da Amostra</u>	26
<u>1.3.3 Análise Descritiva</u>	27
<u>1.3.4 O Modelo <i>Logit</i> Binário</u>	30
<u>1.3.5 Estratégia de Estimação</u>	32
1.4 Resultados	35
1.5 Conclusão	40
Anexos	41
CAPITULO 2 AVALIAÇÃO DO PAS DA UNIVERSIDADE DE BRASÍLIA: EFEITOS SOBRE DESEMPENHO, PERMANÊNCIA E EVASÃO DOS ALUNOS	
2.1 Introdução	43
2.2 Revisão da Literatura	45
<u>2.2.1 Desempenho Acadêmico na Universidade</u>	45
<u>2.2.2 Permanência na Universidade e Evasão</u>	46
<u>2.2.3 Avaliação de Programas Seriadados</u>	48
2.3 Metodologia	50
<u>2.3.1 O Modelo de Resultados Potenciais</u>	51
<u>2.3.2 O Método do Pareamento</u>	54
<u>2.3.3 O Uso do Escore de Propensão</u>	54
<u>2.3.4 Implementação do Método de Pareamento</u>	55
<u>2.3.4 O Método de Pareamento e o Viés de Estimação</u>	60
2.4 Investigação Empírica	60
<u>2.4.1 Base de Dados</u>	62
<u>2.4.2 Grupo de Tratamento e de Controle</u>	63

2.4.3	<u>Estatísticas Descritivas e Teste de Médias</u>	66
2.4.4	<u>Covariadas</u>	68
2.4.5	<u>Variáveis de Interesse</u>	70
2.5	Resultados	70
2.4.1	<u>Estimação do Modelo <i>Logit</i></u>	71
2.4.1	<u>Estimação do Efeito do PAS</u>	77
2.6	Conclusão	75
	REFERÊNCIAS BIBLIOGRÁFICAS	77

CAPITULO 1 DETERMINANTES DO ACESSO AO ENSINO SUPERIOR PELO PROGRAMA DE AVALIAÇÃO SERIADA: O CASO DA UNIVERSIDADE DE BRASÍLIA

1.1 Introdução

Existem atualmente diversos mecanismos de seleção para o ingresso nas universidades públicas brasileiras. O vestibular, forma tradicional utilizada pela maioria das instituições de ensino superior, pode ser realizado paralelamente a outras formas de seleção como o Exame Nacional do Ensino Médio (Enem), transferência facultativa, segunda graduação e cotas para determinados grupos com base em critérios de raça, *status* socioeconômico e localidade geográfica¹. Nesse contexto, insere-se também a admissão por meio de avaliação seriada, adotada por universidades estaduais e federais brasileiras, entre elas, a Universidade de Brasília (UnB). Essa avaliação ocorre ao final de cada ano do ensino médio e ao término do terceiro ano, caso o aluno atinja nota suficiente, será selecionado para o curso escolhido.

Estudos relevantes com foco na compreensão dos determinantes de acesso ao ensino superior em diferentes universidades públicas foram realizados ao longo da última década utilizando-se, em sua maioria, dados de todos os candidatos que prestaram o vestibular em determinado ano, tal como se verá na revisão da literatura. Trabalhos importantes sobre ações afirmativas também buscaram investigar que tipos de candidatos tendem a se beneficiar mais do sistema de cotas raciais e quais mudanças ocorreram no perfil dos estudantes que ingressaram em universidades públicas, como a UnB e Universidade Federal do Rio de Janeiro (UFRJ) (FRANCIS e TANNURI-PIANTO, 2012; MENDES JUNIOR, 2013). Contudo, a literatura econômica que busca compreender os programas seriados ainda é escassa.

Tendo em vista que existem diferenças significantes entre os principais processos de seleção adotados pela UnB – Programa de Avaliação Seriada (PAS) e vestibular – levanta-se uma série de questionamentos acerca das chances de sucesso proporcionadas por cada mecanismo. Esse tipo de questionamento busca uma melhor compreensão sobre os fatores que determinam a aprovação, dentre os já apontados pela literatura, como renda, escolaridades dos

¹ Esse último critério é utilizado por algumas universidades, tais como a UnB. Estudantes que moram em determinadas regiões administrativas usufruem de um fator de correção em seus escores no vestibular caso optem por estudar em algum campi situado nestas regiões (Gama, Planaltina e Ceilândia). Detalhes sobre esse fator são encontrados nos editais do vestibular da UnB.

país, raça. Nesse sentido, pode-se questionar qual forma de seleção beneficia mais os estudantes recém-formados e quais chances de um candidato conquistar uma vaga em um curso seletivo na universidade a depender do tipo de ingresso (PAS ou vestibular).

Este capítulo buscará avaliar quais os fatores que mais influenciam a probabilidade de sucesso dos candidatos que desejam entrar na Universidade de Brasília pelos processos seletivos PAS e vestibular por meio de um modelo de regressão logística. A seção 1.2 apresenta o Programa de Avaliação Seriada e a revisão da literatura. Em seguida, a seção 1.3 descreve a metodologia empregada, com a apresentação das bases de dados, seleção da amostra e estatísticas descritivas. A seção 1.4 expõe os resultados encontrados e a seção 1.5 apresenta as conclusões deste capítulo.

1.2 Revisão de Literatura

Nesta seção, apresenta-se, primeiramente, o conceito de Programa de Avaliação Seriada (PAS). O PAS é descrito em seguida, com o detalhamento acerca da origem do Programa, modo de funcionamento e de seleção dos estudantes. Posteriormente, é feita a revisão da literatura acerca dos determinantes de acesso ao ensino superior, os quais servirão como referência para a escolha das características que serão analisadas na parte metodológica.

1.2.1 Os Programas de Avaliação Seriada no Brasil

O Programa de Avaliação Seriada, ou vestibular seriado, consiste, de forma geral, na realização de provas ao final de cada uma das séries do ensino médio, consecutivamente. Assim, o Programa é realizado em três etapas que correspondem a cada uma das séries do ensino médio. Os grupos são identificados pelo ano de inscrição na primeira etapa do processo. O candidato somente efetua a opção por um dos cursos de graduação oferecidos pela universidade no momento de sua inscrição na terceira etapa. Cada Programa determina como a pontuação obtida em cada etapa será ponderada para o cálculo da nota final do candidato. A porcentagem das vagas que será destinada aos candidatos inscritos no Programa também varia entre os diversos programas existentes. Há atualmente 24 vestibulares seriados no país (Quadro 1.1).

Quadro 1.1 – Descrição dos programas de avaliação seriada por região geográfica brasileira e universidade – 2013

Universidade	Nome do Programa	Ano de Implantação	Quantidade de Vagas
<i>Região Norte</i>			
Universidade Estadual do Amazonas (UEAM)	Sistema de Avaliação para Acesso ao Ensino Superior (SAES)*	2005	50%
Universidade Estadual do Pará (UEPA)	Programa de Ingresso Seriado (PRISE)	1995	50%
Universidade Federal do Amazonas (UFAM)	Programa Seletivo Contínuo (PSC)	1998	50%
Universidade Federal do Pará (UFPA)	Programa Seletivo Seriado (PSS)	2004	Lista única
Universidade Federal de Roraima (UFRR)	Processo Seletivo Seriado (PSS)	2007	Lista única
<i>Região Nordeste</i>			
Universidade de Pernambuco (UPE)	Sistema Seriado de Avaliação (SSA)	2008	20%
Universidade Federal de Campina Grande (UFCG)	Vestibular apenas na modalidade seriada	2002	Lista única
Universidade Federal da Paraíba (UFPB)	Processo Seletivo Seriado (PSS)	1999	Lista única
Universidade Federal do Sergipe (UFS)	Processo Seletivo Seriado (PSS)	2002	Lista única
<i>Região Centro-Oeste</i>			
Universidade de Brasília (UnB)	Programa de Avaliação Seriada (PAS)	1996	50%
Universidade Estadual do Goiás (UEG)	Sistema de Avaliação Seriado (SAS)	2005	20%
<i>Região Sudeste</i>			
Universidade de São Paulo (USP)	Programa de Avaliação Seriada da Universidade de São Paulo (Pasusp)**	2006	
Universidade Federal do Vale do Jequitinhonha e Mucuri (UFVJM)	Processo Seletivo por Avaliação Seriada (Sasi)	2005	20%
Universidade Federal de Viçosa (UFV)	Programa de Avaliação Seriada para o Ensino Superior (Pases)	2001	Lista única
Universidade Federal de Uberlândia (UFU)	Programa Alternativo de Ingresso ao Ensino Superior (Paies)***	1997	50% (cursos semestrais) 25% (cursos anuais)
Universidade de Uberaba (Unibe)	Programa de Ingresso por Avaliação Seriada (Pias)		Determinada por edital
Universidade Estadual de Montes Claros (Unimontes)	Programa de Avaliação Seriada para Acesso ao Ensino Superior (PAES)	1998	40%
Universidade Federal de Juiz de Fora (UFJF)	Programa de Ingresso Seletivo Misto (Pism)	2001	30%
Universidade Federal de Lavras (UFLA)	Programa de Avaliação Seriada (PAS)	2000	40%
Fundação Armando Álvares Penteado (Faap – SP)	Processo Seletivo Contínuo	1997	Definidas por edital

Universidade	Nome do Programa	Ano de Implantação	(conclusão)
			Quantidade de Vagas
<i>Região Sul</i>			
Universidade Estadual de Maringá (UEM)	Processo de Avaliação Seriada (PAS)	2009	20%
Universidade Federal de Santa Maria (UFSM)	Processo Seletivo Seriado	1995	-
Universidade Estadual de Ponta Grossa (UEPG)	Processo Seletivo Seriado (PSS)	2001	20%
Universidade Federal de Pelotas (UFPel)	Programa de Avaliação da Vida Escolar (Pave)	2004	Proporcional ao número de inscritos

Fonte: Informações retiradas dos *websites* das universidades. Acesso entre 2 e 27 de agosto de 2012.

Nota: *Apenas para os cursos de Engenharia e Tecnologia. **O Pasusp destina-se aos alunos que cursaram o primeiro e segundo ano do ensino médio em escolas públicas brasileiras e que estejam cursando em 2010 o terceiro ano em escolas públicas estaduais ou municipais paulistas. O resultado da avaliação será considerado na nota do vestibular para ingresso na USP e conferirá ao candidato bônus adicional de até 3%, dependendo de seu desempenho. ***Substituiu o Programa Experimental de Ingresso ao Ensino Superior (Peies).

As primeiras discussões com vistas a propor um novo modelo de avaliação para ingressantes no ensino superior datam de 1980 (SCHLICHTING, SOARES e BIANCHETTI, 2004). As propostas surgiram em seminários organizados pelo Ministério da Educação que também contavam com a participação de instituições de ensino superior. Como apontam os autores da proposta, as primeiras avaliações seriadas ocorreram entre 1992 e 1995 com a implantação do Sistema de Avaliação Progressiva para Ingresso no Ensino Superior (Sapiens) realizadas pela Fundação Cesgranrio por meio de um projeto piloto na cidade do Rio de Janeiro. O sistema era constituído por seis momentos distintos de avaliação da aprendizagem, a qual se dava por meio de provas objetivas, discursivas, redações e testes de aptidão verbal, numérica e abstrata.

O primeiro vestibular seriado surgiu por iniciativa da Universidade de Santa Maria (UFSM) em 1995 por meio do Programa Experimental de Ingresso ao Ensino Superior (Peies). A promulgação da Lei de Diretrizes e Bases da Educação Nacional (LDB) – Lei nº 9.394/96 – forneceu a base jurídica que possibilitou a criação de outras modalidades de seleção, além do vestibular, para o ingresso em universidades públicas.

1.2.2 O Programa de Avaliação Seriada da Universidade de Brasília

O PAS foi criado em 1995 com o objetivo de se tornar uma forma alternativa de ingresso na UnB e de melhorar a qualidade do ensino nas redes públicas e privadas (UNB, 1995). O projeto do Programa foi apresentado à comunidade acadêmica e externa em 1985 por Lauro Morhy, presidente da antiga Comissão Permanente de Concurso Vestibular da UnB

(Copeve), durante o 1º Seminário sobre Vestibular da Universidade de Brasília. A proposta apresentada defendia a aplicação de provas ao longo do ensino médio com o objetivo de proporcionar à universidade uma melhor forma de avaliar os candidatos. A busca pela integração da universidade ao ensino médio almejava, por sua vez, auxiliar na melhoria da qualidade do ensino do segundo grau nas redes públicas e privadas. Em 1986, o projeto consolidado foi apresentado Ministério da Educação. Apenas em dezembro de 1995 foi publicado o primeiro edital de seleção que regulamentou o Subprograma 1996-1998. A primeira prova foi realizada no ano de 1996 por alunos do primeiro ano do ensino médio, com o ingresso dos aprovados no primeiro semestre de 1999. A terceira etapa do Subprograma 1996 teve a participação de 9.676 candidatos, sendo 7.030 do Distrito Federal.

Dentre os objetivos gerais do Programa de Avaliação Seriada destaca-se a implantação de um processo de seleção capaz de integrar a educação básica à superior, visando a melhoria da qualidade de ensino em ambos os níveis e buscando atingir o princípio de que a vida escolar deve ser um *continuum*. Dentre os objetivos específicos está a seleção gradual e sistemática dos estudantes ao ensino universitário por meio de um processo que busque “avaliar a aprendizagem significativa, em que se privilegie a reflexão sobre a memorização, a qualidade sobre a quantidade de informações” (Centro de Seleção e Promoção de Eventos – Cespe, 2012). Por fim, busca-se um modelo de avaliação que privilegie a contextualização e a interdisciplinaridade “com ênfase no desenvolvimento de competências e habilidades”.

O PAS baseia-se em alguns pressupostos. Primeiramente, sabe-se que os sistemas de acesso à universidade influenciam de forma significativa o ensino médio em relação ao conteúdo ensinado e ao enfoque epistemológico, o que pode proporcionar consequências positivas como a “formação da cidadania, a preparação geral para o trabalho e o desenvolvimento de competências e habilidades” (Cespe, 2012). Há também uma participação mais ampla e efetiva dos professores de ensino médio no estabelecimento dos conteúdos programáticos do PAS. Os conteúdos cobrados constam na Matriz de Objetos de Avaliação, disponível no *website* do Cespe, que é elaborada de forma conjunta pelos professores de ensino médio do Distrito Federal e cidades do entorno e por docentes da UnB. De acordo com o Cespe, o PAS objetiva com a Matriz “selecionar o aluno capaz de compreender, raciocinar, analisar e propor questões relevantes para a própria formação como cidadão e capaz de elaborar propostas de intervenção na realidade” (Cespe, 2012).

Para participar do PAS, os candidatos devem nele se inscrever anualmente a partir do 1º ano do ensino médio. Podem se inscrever na primeira etapa de um subprograma² os alunos que estiverem regularmente matriculados em escola pública ou particular na primeira série do ensino médio, na modalidade de ensino regular de três anos completos, ou na segunda série do ensino médio, em escola cuja estrutura curricular seja de quatro anos completos. Os estudantes que não se inscreveram para a primeira etapa poderão ainda ingressar no Programa durante o 2º ano do ensino médio, mas farão apenas a terceira etapa do Programa³.

Como visto anteriormente, os estudantes são avaliados ao final de cada ano do ensino médio. Os candidatos são selecionados por *campus*, curso e turno de acordo com o seu desempenho no PAS. A pontuação obtida em cada etapa é somada, de forma ponderada, para gerar a pontuação final que considerada para determinar se o estudante foi ou não aprovado. As provas da primeira, segunda e terceira etapas possuem pesos iguais a 1, 2 e 3, respectivamente. Cada subprograma visa o preenchimento de 50% das vagas oferecidas no primeiro semestre letivo subsequente ao término do triênio no qual o candidato se inscreveu.

1.2.3 Os determinantes de Acesso ao Ensino Superior

Os primeiros trabalhos sobre os determinantes de acesso à universidade foram realizados na década de 1970. Citam-se como primeiros estudos os de Siano (1977), Freitas (1979) e Oliveira (1980), os quais usam o método dos mínimos quadrados ordinários (MQO) e realizam uma ampla análise descritiva dos dados. Dentre as principais variáveis independentes utilizadas para explicar o número de pontos obtidos no vestibular estão a educação dos pais, sexo e tipo de escola na qual os candidatos estudaram durante o ensino médio.

O estudo de Honorato e Wing (2011) buscou investigar quais os principais fatores que contribuem para sucesso dos candidatos no ingresso na Universidade Federal de Pernambuco (UFPE) no *campus* Recife ou Agreste ou do curso selecionado. O trabalho utiliza um modelo *logit* binário para calcular a probabilidade de aprovação no vestibular em quatro cursos que existem em ambos os *campi* da UFPE. Os autores utilizam dados obtidos pela Comissão de

² Os subprogramas agregam três anos de avaliação. O último subprograma iniciado foi o de 2012, voltado para os alunos que estavam cursando o primeiro ano do ensino médio daquele ano.

³ O aluno que já tenha participado da segunda etapa de um subprograma e venha a concluir o ensino médio (ou equivalente) no exterior poderá se inscrever na terceira etapa desde que apresente declaração de equivalência de estudos expedida pelo Conselho de Educação competente (CESPE, 2012)

Processos Seletivos e Treinamentos (Covest) provenientes de questionário socioeconômico respondido por todos os candidatos que se inscrevem para a seleção naquele ano, mas restringem sua análise para candidatos que optaram pelos cursos de Administração, Economia, Design e Pedagogia. O modelo não inclui como variável explicativa o gênero, pois esta informação não foi disponibilizada. Os principais resultados do estudo demonstram a existência de diferenças significativas entre as características que mais influenciam a probabilidade de sucesso no vestibular para cada *campus*. Dentre as variáveis mais importantes estão a renda, se o candidato fez curso pré-vestibular (variável com efeito ainda maior para os cursos mais concorridos dentre os que foram investigados), número de tentativas de acesso, local de residência e educação dos pais.

Guimarães e Arraes (2010) analisam os determinantes de desempenho de candidatos ao vestibular da Universidade Federal do Ceará (UFC) utilizando uma base de dados com informações oriundas de questionário socioeconômico preenchido pelos candidatos. Utilizando os modelos *probit* e *probit* ordenado, autores verificam que a renda, educação dos pais, qualidade do ensino público no nível médio de ensino são os principais fatores na determinação do sucesso no exame. Da análise do desempenho por área de concentração, concluiu-se que os candidatos com maior chance de sucesso em determinada área não manteria tais chances em outras áreas.

Velloso (2006) realizou agrupamentos dos cursos de graduação na UnB que refletissem o prestígio social dos cursos a fim de testá-lo como substituto aproximado do nível socioeconômico dos candidatos. Verificou-se que a renda média familiar dos aprovados se reduz a medida que se passa de grupos de maior para os de menor prestígio em termos de concorrência, sendo essa queda maior no grupo das Humanidades. Em análise realizada a partir de dados da Universidade Federal da Bahia (UFBA), Queiroz (2008) observou que as características socioeconômicas, bem com a cor da pele do estudante, possuem relação com o prestígio do curso. Cursos de maior prestígio têm maior proporção de pessoas de pele mais clara e *status* econômico mais elevado. A autora também observou que alunos de escolas particulares e pais com nível superior tiveram melhor desempenho no vestibular.

Guimarães e Sampaio (2007) analisam os fatores que afetam o desempenho dos candidatos ao vestibular da UFPE. Os autores obtiveram resultados condizentes com a literatura: a educação dos pais e o ambiente de estudos possuem valor relevante para o desempenho dos estudantes, bem como as variáveis religião, raça e gênero. Os efeitos de cada

variável, no entanto, variam ao longo da distribuição condicional da nota, resultado esse obtido por meio de regressão quantílica. De acordo com os autores, a função quantílica condicional caracteriza melhor a relação entre as características dos estudantes e as notas obtidas no vestibular do que o método de mínimos quadrados ordinários ao fornecer uma reta de regressão para cada quantil τ . Neste sentido, estudantes oriundos de escolas públicas tendem a alcançar piores resultados, e quanto maior o número de anos de ensino cursados no ensino público, pior o desempenho. O coeficiente estimado pelos autores mostra que mulheres tendem a obter piores resultados que os homens ao longo da distribuição condicional: nos primeiros quantis as mulheres possuem vantagem na nota de um ponto percentual em relação à nota dos homens. Nos quantis superiores, as mulheres passam a ter desvantagem de até três pontos percentuais. Após controlar para diferentes departamentos da universidade, observou-se que mulheres tendem, na média, a ter melhores notas em departamentos onde a parcela de estudantes homens é predominantemente maior e no departamento de medicina. Os resultados da análise nos quantis revelam, no entanto, que nos departamentos de ciências humanas e sociais as notas das mulheres nos quantis superiores da distribuição são inferiores as notas dos homens. Os resultados apresentados por Guimarães e Sampaio (2007) são similares aos encontrados por parte da literatura.

Young e Fisher (2000) examinam as notas no exame de admissão à universidades americanas (*Scholastic Aptitude Test* - SAT) de alunos no último ano do ensino médio e encontram que homens tendem a obter notas superiores às obtidas por mulheres. No entanto, os autores verificaram que os homens possuíam origem socioeconômica distinta do grupo de mulheres, o que explicaria parte da diferença encontrada. O estudo de Dayiglu e Turut-Asik (2004) buscou determinar se há diferenças de gênero significantes no desempenho acadêmico de universitários em uma universidade na Turquia. Com base em três indicadores - notas no exame de entrada e na escola de inglês preparatória e curso escolhido pelo aluno, o estudo encontra que um número reduzido de alunas consegue ingressar na universidade e, quando ingressam, obtém em geral notas menores que às obtidas pelos homens. Contudo, após serem admitidas na universidade, o seu desempenho é superior ao dos alunos homens, mesmo após controlar por atributos individuais.

Ainda de acordo com o estudo de Guimarães e Sampaio (2007), os autores observam que renda familiar também tem maior impacto nos quantis superiores. Candidatos que trabalham tendem a obter piores resultados, apesar de não haver consenso na literatura acerca

do efeito dessa variável sobre o desempenho acadêmico corrente e futuro do aluno (STINEBRICKNER e STINEBRICKNER, 2003).

Cesar (2013) realiza um estudo sobre o PAS e o vestibular com dados do Cespe de 2011 com o objetivo de avaliar as diferenças entre estes exames e discutir os efeitos dos mecanismos de seleção de caráter meritocrático sobre a composição do corpo discente da universidade. A autora procura identificar os principais determinantes do sucesso na aprovação e fornecer alternativas ao modelo vigente, utilizando para isso um modelo *logit* com variáveis independentes referentes à renda, escolaridade dos pais e raça. Entre as variáveis de interação adicionadas ao modelo, apenas aquela referente à interação entre PAS e curso preparatório apresentou coeficiente estatisticamente significativo. O estudo conclui que o PAS não se diferencia substantivamente do vestibular em relação ao perfil socioeconômico dos candidatos selecionados.

1.3 Metodologia

Primeiramente, apresenta-se a base de dados formada pelas informações fornecidas pelo Cespe referentes aos candidatos pelo vestibular sistema universal e PAS. Em seguida, apresenta-se o procedimento para seleção da amostra e a análise descritiva dos dados. Apresenta-se o modelo que será utilizado neste capítulo, bem como a estratégia empírica adotada para a sua estimação.

1.3.1 Base de Dados

A base fornecida pelo Cespe possui dois grupos de informação: o primeiro refere-se ao desempenho de todos os candidatos que se inscreveram no exame PAS e vestibular para o ano de 2004 e 2005, com informações sobre o total de inscritos no exame, sistema de seleção – universal ou PAS – e variável indicadora de aprovação. Para os inscritos que disputam vagas pelo vestibular, estão disponíveis as notas brutas nas provas de língua estrangeira, português e ciências sociais, matemática e ciências naturais, escore bruto e argumento final. Para os inscritos que disputam vagas pelo PAS, estão disponíveis as notas obtidas em cada etapa do exame. O segundo grupo de informações refere-se àquelas fornecidas pelos candidatos no questionário socioeconômico, doravante intitulado de QSC. O questionário possui perguntas sobre estado civil, renda familiar, educação dos pais, raça/cor, raça

autodeclarada, trabalho e cidade de residência, que serão posteriormente exploradas nesse capítulo.

A partir da base formada pelas informações sobre desempenho dos inscritos para o PAS e vestibular no ano de 2005, calcula-se a taxa de aprovação nos processos seletivos dada pela divisão do total de aprovados pelo total de inscritos. Essa taxa pode ser utilizada para uma primeira investigação acerca das diferenças que cada processo pode gerar na probabilidade de aprovação. A Tabela 1.2 sumariza tais informações.

Tabela 1.2 - Taxa de aprovação no Programa de Avaliação Seriada e no primeiro vestibular de 2005 da Universidade de Brasília – resultado geral e para cursos selecionados

Curso	Sistema	Não Aprovados	Aprovados	Inscritos	Taxa de Aprovação (%)
Geral	Vestibular	24482	1167	25649	4,55
	PAS	9426	1028	10454	9,83
Medicina	Vestibular	1413	17	1430	1,19
	PAS	315	18	333	5,41
Direito	Vestibular	1640	20	1.660	1,20
	PAS	452	27	479	5,64
Engenharia Mecatrônica	Vestibular	337	13	350	3,71
	PAS	125	13	138	9,42
Pedagogia	Vestibular	702	47	749	6,28
	PAS	441	41	482	8,51

Fonte: Cespe.

A Tabela 1.2 apresenta dados para os cursos mais concorridos no primeiro vestibular de 2005, à saber: Medicina, Direito e Engenharia Mecatrônica. Apresenta, ainda, dados para o curso de pedagogia, que possui uma das menores notas de corte⁴ no vestibular. Observa-se que no PAS a taxa de aprovação média é mais que o dobro daquela registrada no vestibular.

Para cursos mais seletivos, como Medicina e Direito, essa proporção tende a ser ainda mais elevada: a taxa de aprovação no PAS é mais que quatro vezes aquela registrada no vestibular. À medida que o curso torna-se menos seletivo, com notas de corte inferiores, essa diferença tende a cair. Para o curso de Pedagogia, a taxa de aprovação é apenas 35% mais elevada.

⁴ A nota de corte para Letras e Artes são inferiores às de Pedagogia. No entanto, esses cursos tentem a atrair um grupo específico de estudantes e por isso tendem a ser cursos que não apresentam a representatividade que se procura mostrar pela análise da taxa de admissão.

A análise acima é um instrumento importante para compreender o destaque que a avaliação seriada possui entre os estudantes que almejam ingressar na universidade. O PAS possibilita ao estudante disputar vagas com estudantes que ainda cursam ensino médio, evitando a concorrência com aqueles que já possuem maior experiência ou preparo em exames de vestibular. Ao mesmo tempo, ele possibilita ao aluno adequar a escolha do curso a sua nota acumulada até a segunda etapa, o que contribui para reduzir as chances de reprovação, tendo em vista que o aluno pode optar por não concorrer a cursos com nota superior àquela que o aluno acredita poder obter no exame. Por fim, como o exame destina 50% das vagas disponíveis do primeiro semestre do ano para os aprovados pelo processo, a quantidade de vagas para cada curso acaba sendo superior a quantidade destinada aos aprovados pelo vestibular.

Com os dados do Cespe é realizada uma segunda análise para investigar as relações entre as características socioeconômicas e raciais com o desempenho dos candidatos no exame PAS e vestibular sistema universal⁵. A Tabela 1.3 apresenta o perfil dos candidatos por percentil da nota nos exames. Os percentis são definidos utilizando-se os escores brutos obtidos no vestibular formados pela soma algébrica das notas obtidas nas provas objetivas. Ressalta-se que se optou pelo escore bruto e não pelo argumento final - nota utilizada para determinar a aprovação – pois o cômputo do argumento final elimina os estudantes que não atingiram pontuação mínima. Além disso, essa nota assume valores negativos, o que dificulta sua comparação com as notas dos inscritos ao PAS. Para os inscritos ao PAS utiliza-se a soma das notas obtidas em cada uma das três etapas do exame.

Para candidatos a vagas pelo vestibular, a taxa de admissão para candidatos com notas superiores ao décimo percentil é de 32,6%, enquanto no PAS a taxa chega a 54,2%. A Tabela 1.3 aponta ainda que, em termos de gênero, características socioeconômicas e raça, o perfil dos candidatos muda significativamente com o avanço nos percentis do escore, tanto no PAS quanto no vestibular. Em relação aos candidatos com notas superiores ao nono decil, uma maior proporção de candidatos a vagas pelo ao PAS é do sexo feminino – apesar de haver a mesma tendência de aumento na proporção de homens à medida que se move para maiores percentis da nota.

⁵ A comparação entre os percentis dos escores e notas para o grupo de candidatos a vagas pelo PAS e pelo vestibular sistema universal e de cotas está disponível no Anexo A.

Tabela 1.3 - Perfil dos candidatos por percentil de nota obtida no vestibular e PAS

	Percentis do Escore Bruto - Vestibular Tradicional						Percentis da Nota – PAS					
	<10%	10% 30%	30% 50%	50% 70%	70% 90%	90%<	<10%	10% 30%	30% 50%	50% 70%	70% 90%	90%<
Taxa de Aprovação	0,00	0,00	0,24	2,96	11,43	32,55	0,00	0,11	1,14	8,11	22,29	54,2
Masculino	37,0	42,8	45,8	47,8	51,0	59,3	31,0	34,8	39,5	44,1	48,2	51,9
Renda Mensal Familiar (R\$)												
Renda < 500	25,6	18,3	10,4	5,4	3,4	1,5	27,6	17,1	7,1	3,5	1,8	1,5
500 < Renda ≤ 1500	37,3	38,6	30,8	23,1	15,7	10,8	32,8	30,1	24,7	20,3	8,9	6,0
1500 < Renda ≤ 2500	13,5	16,2	16,7	18,7	17,9	13,2	12,8	18,9	17,5	15,3	13,5	8,2
2500 < Renda ≤ 5000	13,1	14,7	24,1	26,2	26,9	27,9	15,5	19,1	24,4	30,8	31,3	28,4
Renda > 5000	5,9	6,2	9,1	15,4	23,5	35,1	7,6	8,8	19,2	22,4	36,0	46,1
Renda < 500	4,5	6,1	8,9	11,2	12,6	11,6	3,8	6,0	7,1	7,7	8,5	9,9
Mora em Casa Própria	66,9	70,7	70,5	74,8	73,4	83,3	75,7	75,7	77,6	78,0	80,5	84,4
Educação do Pai												
Ensino Básico Incompleto	38,4	30,2	19,6	13,9	7,1	3,5	35,4	24,9	13,9	9,0	3,5	2,8
Ensino Básico Completo	11,0	11,2	8,1	7,3	6,9	3,4	11,7	8,4	7,6	5,6	4,3	1,9
Ensino Médio Completo	28,2	32,7	37,8	32,2	28,9	24,0	24,5	35,0	31,9	29,5	26,8	19,4
Ensino Superior	16,7	19,8	30,1	42,5	54,4	66,9	20,6	26,0	42,6	51,8	63,7	75,2
Não sabem	5,7	6,1	4,4	4,1	2,7	2,3	7,8	5,7	4,0	4,1	1,7	0,8
Educação da Mãe												
Ensino Básico Incompleto	37,8	29,1	17,3	11,2	5,9	2,9	38,3	19,9	11,4	5,8	3,6	1,3
Ensino Básico Completo	10,8	10,5	10,2	7,3	5,8	2,6	10,0	10,9	8,7	5,2	2,1	1,1
Ensino Médio Completo	34,3	34,5	36,2	36,7	32,5	27,3	26,8	28,9	37,9	32,4	26,5	23,8
Ensino Superior	15,3	24,2	34,7	43,1	54,3	66,2	21,5	27,9	41,6	55,4	67,1	73,8
Não sabem	1,8	1,6	1,7	1,7	1,5	1,0	3,5	2,5	0,4	1,2	0,8	0,2

(conclusão)

	Percentis do Escore Bruto - Vestibular Tradicional						Percentis da Nota – PAS					
	<10%	10%-30%	30%-50%	50%-70%	70%-90%	90%<	<10%	10%-30%	30%-50%	50%-70%	70%-90%	90%<
Raça												
Branca	42,9	46,4	46,0	46,6	51,5	51,7	36,4	38,4	45,7	51,8	55,9	55,4
Preta	6,4	3,5	3,4	3,4	1,9	2,3	12,3	11,6	6,8	5,0	3,1	1,9
Amarela	6,6	6,2	4,8	3,7	3,3	2,1	3,5	4,7	3,6	2,7	3,1	2,4
Parda	36,9	35,6	36,4	34,5	32,7	30,7	43,7	38,6	39,4	32,3	30,1	28,5
Indígena	1,0	0,6	0,8	0,7	0,4	0,5	0,4	0,9	0,4	0,8	0,4	0,9
Sem declaração	6,2	7,7	8,6	11,1	10,2	12,7	3,8	5,8	4,1	7,4	7,3	10,8
Negro	17,1	17,6	18,8	17,7	18,1	18,4	34,4	31,9	23,7	17,4	13,7	11,5
Não trabalha	71,7	72,0	76,7	82,2	85,8	91,0	80,9	85,0	87,0	89,1	92,6	92,7
Residência												
Demais Regiões Administrativas	72,4	65,7	56,8	51,8	38,6	28,7	75,1	72,8	63,7	57,5	37,4	32,0
Brasília	11,1	16,5	23,8	24,9	31,2	35,7	15,7	14,2	24,0	25,8	38,4	44,9
Fora do DF	16,5	17,8	19,4	23,4	30,2	35,7	9,2	13,1	12,2	16,7	24,3	23,1

Fonte: Cespe (2012b). Nota: *As notas do vestibular referem-se ao escore bruto, enquanto as notas do PAS referem-se à soma das notas obtidas nas três etapas do exame.

1.3.2 Seleção da Amostra

Como apontado anteriormente, a base Questionário Socioeconômico (QSC) é constituída pelas informações fornecidas pelos candidatos que prestaram o vestibular e a terceira etapa do PAS em 2004 e 2005. Como nenhum inscrito no PAS de 2004 respondeu ao QSC, optou-se pela não utilização dos dados desse ano.

Como o objetivo deste trabalho é comparar de forma apropriada os candidatos que ingressam por cada tipo de processo seletivo, entende-se que os estudantes que prestaram o vestibular para ingresso no segundo semestre do ano não apresentam as mesmas características daqueles que desejam ingressar no primeiro semestre. A concorrência por vaga no vestibular do meio do ano é inferior, pois o número de inscritos é menor e não há a alocação de vagas para o PAS, isto é, todas as vagas são destinadas ao vestibular pelo sistema universal e de cotas. Os estudantes que prestam o vestibular no meio do ano, em geral, já concluíram o ensino médio, possuem maior experiência em exames de seleção e tiveram acesso a outros tipos de preparação que não aquela fornecida durante o ensino médio, decorrente do período em que estiveram fora da universidade. Por estes motivos, tais dados também foram retirados da base. A amostra é, portanto, composta apenas pelos estudantes que se candidataram à terceira etapa do PAS e ao primeiro vestibular no ano de 2005.

Uma análise adicional foi realizada com uma segunda amostra que se diferencia da primeira por adicionar ao conjunto de candidatos aqueles que se inscreveram no vestibular pelo sistema de cotas, a fim de se avaliar como a introdução desse sistema modifica as chances de aprovação. Tal como anteriormente, esta amostra é composta apenas pelos inscritos nos exames de seleção no ano de 2005. Assim, a amostra contrapõe o conjunto de inscritos no PAS com o conjunto de inscritos no vestibular pelo sistema universal e sistema de cotas.

Optou-se ainda por manter na amostra os candidatos que não compareceram ao exame vestibular ou PAS pelo motivo descrito a seguir: a maior parte dos candidatos prestes a se formar inscreve-se no vestibular, sendo o questionário socioeconômico respondido no momento da inscrição neste exame, antes de o aluno realizar a prova do PAS. A depender do resultado do exame PAS, o aluno opta por não comparecer à prova de vestibular, o que o eliminaria da amostra de candidatos ao vestibular. Entende-se, no entanto, que as informações

sobre esse candidato são relevantes para esse estudo já que esse aluno tinha a intenção de ser aprovado para determinado curso, ainda que tenha desistido de fazer a prova.

Como o questionário não é de resposta obrigatória, investiga-se a taxa de resposta. Conforme mostrado na Tabela 1.4, enquanto 41,8% dos candidatos ao PAS responderam ao questionário, essa taxa foi de 32,1% para os candidatos ao vestibular. A taxa de resposta pelos cotistas foi bastante superior, conforme apontam Francis e Tannuri-Pianto (2012).

Tabela 1.4 – Taxa de Resposta ao Questionário Socioeconômico (QSC)

	Vestibular Sistema Universal	PAS	Total
Número de Respondentes	6.135	4.374	6.812
Taxa de Resposta (%)	32,14	41,84	26,36

Fonte: Cespe.

A comparação da taxa de resposta em relação às características socioeconômicas e raciais dos estudantes entre os candidatos ao PAS e ao vestibular sistema universal indica que não há diferenças significativas entre esses grupos na maioria das variáveis que serão utilizadas nas estimativas, como renda, educação do pai, educação da mãe e raça.

Os questionários utilizados por cada processo de seleção possuem perguntas similares, referentes às características socioeconômicas, sexo, escolaridade do pai e da mãe, cor, identidade racial, emprego e residência. Essas perguntas darão origem às variáveis que serão incluídas no modelo.

1.3.3 Análise Descritiva

A Tabela 1.5 apresenta as estatísticas descritivas referentes à base de dados QSC. Optou-se pela comparação entre os seguintes grupos: (i) aprovados e não aprovados pelo PAS; (ii) aprovados e não aprovados pelo vestibular sistema universal; (iii) inscritos no PAS e vestibular sistema universal e (iv) aprovados no PAS e no vestibular sistema universal⁶.

A comparação entre os candidatos ao PAS e ao vestibular pelo sistema universal visa compreender em que medida os candidatos aprovados e não aprovados diferem entre si, bem

⁶ Essas mesmas estatísticas, mas para os grupos de candidatos ao PAS e ao vestibular pelo sistema universal e de cotas, estão dispostas no Anexo B.

como apontar se há diferenças significativas entre o conjunto de inscritos e de aprovados em cada um dos processos seletivos. A amostra é composta por 10.482 estudantes – destes, 6.135 estavam inscritos no vestibular, enquanto 4.374 inscreveram-se para o PAS⁷. Para os candidatos ao vestibular que responderam ao questionário, há 410 aprovados e 5.725 reprovados. Por sua vez, entre os candidatos ao PAS há 592 aprovados e 3.782 reprovados.

Como esperado, há diferenças em várias variáveis entre o grupo de aprovados e não aprovados no PAS e no vestibular. O grupo de aprovados no vestibular é formado em sua maioria por homens (57,1%), enquanto as mulheres consistem na maioria dos selecionados pelo PAS (55,1%). A proporção de candidatos ao vestibular com renda inferior a R\$ 500,00, equivalente a 1,7 salários mínimos à época, decresce quando se passa do grupo de não aprovados para o de aprovados – o mesmo movimento também é observado entre os candidatos ao PAS. Da mesma forma, a proporção de candidatos com renda superior a R\$ 5.000,00 eleva-se consideravelmente entre o grupo de aprovados. Nota-se que o conjunto de aprovados pelos dois sistemas possui pais com maior nível de instrução – a proporção de pais com ensino superior eleva-se de 38,5% entre o grupo de não aprovados no vestibular para 51,5% entre os aprovados.

Em relação à cor, para os candidatos ao PAS há diferença na proporção de pretos entre aprovados e não aprovados. Este fato pode estar ligado à existência da seleção por cotas que ocorre concomitantemente ao vestibular. Como o sistema universal destina vagas aos estudantes que se declaram negros, entende-se que a proporção de pretos no vestibular sistema universal seja inferior àquela vista no PAS. Quando se compara o conjunto de aprovados pelo PAS e pelo vestibular não se encontra, todavia, diferenças estatisticamente significantes em relação à cor, embora os grupos difiram em relação à raça autodeclarada. Os grupos de aprovados e não aprovados apresentam diferenças também quanto à cidade de residência para ambos os processos seletivos, embora a proporção de candidatos oriunda de cada uma das regiões descritas não difira entre os aprovados pelos dois sistemas.

⁷ Ressalta-se que essa amostra não corresponde ao total de inscritos no primeiro vestibular de 2005, pois parte deles não respondeu o questionário socioeconômico.

Tabela 1.5 - Estatísticas Descritivas do Vestibular e PAS

Variáveis	Vestibular		PAS		Inscritos		Aprovados PAS e Vestibular	Diferença nas Médias		
	Não aprovados (%)	Aprovados (%)	Não aprovados (%)	Aprovados (%)	Vestibular (%)	PAS (%)				
Masculino	47,52	***	57,09	41,19	**	44,94	47,97	***	41,56	***
Renda Mensal Familiar (R\$)										
Renda < 500	9,85	***	3,92	8,91	***	3,54	9,46	**	8,18	
500 < Renda ≤ 1500	26,10	***	17,16	21,02	***	9,44	25,50	***	19,44	***
1500 < Renda ≤ 2500	16,63	-	17,65	14,82		13,83	16,70	***	14,69	*
2500 < Renda ≤ 5000	22,20	**	26,72	25,94		27,15	22,50	***	26,11	
Renda > 5000	15,74	***	24,02	22,59	***	36,26	16,29	***	24,45	***
Não sabem	9,47	-	10,54	6,71	***	9,78	9,54	***	7,12	
Possui Casa Própria	72,95		73,41	78,48	**	82,35	72,98	***	79,01	***
Educação do Pai										
Ensino Básico Incompleto	18,44	***	8,29	13,79	*	5,76	17,76	***	12,69	
Ensino Básico Completo	8,04		8,05	6,73	**	4,07	8,04	***	6,37	***
Ensino Médio Completo	30,60		29,02	29,00	**	25,60	30,50	**	28,5	
Ensino Superior	38,52	***	51,46	46,72	***	63,22	39,38	***	48,98	***
Não sabem	4,42		3,17	3,77	***	1,36	4,34	**	3,44	**
Educação da Mãe										
Ensino Básico Incompleto	17,07	***	9,29	11,99	***	4,72	16,55	***	11,00	***
Ensino Básico Completo	8,10	**	4,89	6,18	***	3,04	7,89	***	5,75	
Ensino Médio Completo	33,96	**	27,87	32,36	***	26,64	33,55	**	31,58	
Ensino Superior	39,30	***	56,23	48,16	***	65,43	40,43	***	50,52	***
Não sabem	1,58		1,71	1,31	**	0,17	1,58	*	1,15	**
Raça										
Branca	47,87		45,96	48,85		48,98	47,74	-	48,87	
Preta	3,32		1,96	6,97	***	3,22	3,23	***	6,46	
Amarela	4,64		3,18	3,10		3,73	4,54	***	3,18	
Parda	34,06		34,47	34,86		31,69	34,09	-	34,43	
Indígena	0,61		0,49	0,59	***	1,53	0,61	-	0,72	
Sem declaração	9,50	***	13,94	5,64	***	10,85	9,80	***	6,35	
Negro	17,29		19,41	21,26	***	16,27	17,43	***	20,58	
Não trabalha	79,06		77,56	87,75	**	90,85	78,96	***	88,17	***
Residência										
Demais Regiões Administrativas	25,24	***	42,82	25,84	***	47,89	26,41	***	28,85	
Brasília	49,90	***	40,09	52,00	***	37,77	49,30	-	50,00	
Fora do DF	24,83	***	17,08	22,19	***	14,33	24,31	***	21,11	
<i>N</i>	5725		410	3782		592	6135		4374	

Fonte: Cespe. Nota: *Estatisticamente significante a 1%. **Estatisticamente significante a 5%. ***Estatisticamente significante a 10%.

A comparação entre o conjunto de inscritos no PAS e no vestibular objetiva esclarecer questionamentos acerca da existência de diferenças entre tais grupos. Há diferença estatisticamente significativa para a maior parte das características analisadas. No entanto, quando se compara o grupo de aprovados em cada processo seletivo, há diferença em apenas onze das 24 variáveis analisadas, ponto também ressaltado por Cesar (2013) em análise dos dados relativos ao vestibular de 2011.

1.3.4. O Modelo *Logit* Binário

A estimação dos determinantes de acesso à universidade pelo PAS e pelo vestibular será realizada por meio de um modelo *logit* binomial. O modelo de regressão logística é uma opção ao uso do MQO, que ignora o fato de a variável dependente ser binária, o que pode levar a probabilidades preditas com valores fora do intervalo de zero a um.

A variável de interesse desse trabalho, definida como a probabilidade de sucesso em um exame de seleção, pode ser, portanto, definida por meio de um modelo de resposta binária. Este modelo é desenvolvido por meio do uso de variável não observada denominada latente, que substitui uma variável binária, para a qual se assume determinada distribuição de probabilidade (DAVIDSON e MACKINNON, 2004).

Seja y a variável dependente referente à aprovação no exame de seleção, dada por

$$y = \mathbf{x}\boldsymbol{\beta} + \mu \quad (1)$$

Essa variável pode assumir dois valores, sendo:

$$y = \begin{cases} 1 & \text{com probabilidade } p \\ 0 & \text{com probabilidade } 1 - p \end{cases} \quad (2)$$

Tem-se que p é a probabilidade de se obter êxito no exame, isto é, a aprovação. Um modelo de regressão é formado por meio da parametrização da probabilidade p , a qual irá depender de uma função índice $\mathbf{x}\boldsymbol{\beta}$, em que \mathbf{x} é um vetor de regressores de dimensão $K \times 1$ (com o primeiro elemento sendo unidade) e $\boldsymbol{\beta}$ é um vetor de parâmetros desconhecidos $K \times 1$ (CAMERON e TRIVEDI, 2005). Sendo y um indicador de aprovação, \mathbf{x} contém as características individuais dos candidatos, tais como nível de escolaridade dos pais, renda, raça e demais fatores que afetam a aprovação. A probabilidade condicional em um modelo de resultado binário padrão é dada por

$$p_i \equiv \Pr(y_i = 1|x) = G(x\beta) = p(x) \quad (3)$$

Como a estimação do modelo dado por (1) não garante que y estará contida num intervalo entre zero e um, faz-se uma reformulação por meio da criação de uma variável latente y^* . O modelo na equação (3) é denominado modelo índice por restringir a forma na qual a probabilidade de resposta depende de x . $P(x)$ é uma função de x apenas por meio do índice $x\beta = \beta_1 + \beta_2 x_2 + \dots + \beta_k x_k$. Na maioria dos casos, G é uma função de distribuição acumulada - f.d.a (WOOLDRIDGE, 2011). Os modelos índice nos quais G é uma f.d.a podem ser derivados de um modelo de variável latente:

$$y^* = x\beta + e, y = 1[y^* > 0] \quad (4)$$

Em que e é uma variável com distribuição contínua e independente de x , com distribuição simétrica em torno de zero e $1[\cdot]$ é uma função indicadora. Ainda que y^* não seja observada, é possível observar:

$$y = \begin{cases} 1 & \text{se } y^* > 0 \\ 0 & \text{se } y^* \leq 0. \end{cases} \quad (5)$$

O limiar zero é uma normalização que não traz problemas ao modelo caso x inclua um intercepto (CAMERON E TRIVEDI, 2005). Se G é a f.d.a de e , como a f.d.a de e é simétrica em torno de zero, $1 - G(-z) = G(z) \forall z$ real. Assim:

$$P(y = 1|x) = P(y^* > 0|x) = P(e > -x\beta|x) = 1 - G(x\beta) = G(x\beta) \quad (6)$$

Chegando, portanto, a equação (3).

O modelo *probit* é um caso especial da equação (3) em que $G(z) \equiv \Phi(z) \equiv \int_{-\infty}^z \phi(v) dv$, sendo $\phi(v)$ a densidade normal padrão (WOOLDRIDGE, 2011). O modelo *probit* pode ser derivado da formulação da variável latente quando e tem distribuição normal padrão. Por sua vez, o modelo *logit* é um caso especial de (3) com $G(z) = \frac{e^z}{1+e^z} = \Lambda(z)$, o qual é usado quando e tem distribuição logística padrão.

Por meio da transformação exponencial obtém-se:

$$\log \frac{p_i}{(1-p_i)} = X_i' \beta + e \quad (7)$$

Por meio do modelo *logit*, busca-se estimar os efeitos parciais das variáveis independentes sobre a variável dependente. Se x_K é uma variável independente binária, o efeito parcial da mudança de x_K de zero para um, mantendo todas as demais variáveis fixas, é:

$$G(\beta_1 + \beta_2 x_2 + \dots + \beta_{K-1} x_{K-1} + \beta_K) - G(\beta_1 + \beta_2 x_2 + \dots + \beta_{K-1} x_{K-1}) \quad (9)$$

Assim como no caso para variáveis contínuas, o valor do efeito parcial depende dos valores de todas as outras variáveis x_j . Pode-se, todavia, apresentar os resultados na forma da razão de chance entre o grupo em questão, no caso, o PAS, e o grupo de referência, formado pelos candidatos ao vestibular. Se o coeficiente da razão de chance (OR) for maior que um, significa que os candidatos ao PAS tem probabilidade de ser aprovado (OR -1) vezes maior que a dos candidatos ao vestibular, mantendo-se todas as demais variáveis constantes. Uma razão de chance inferior a um significa que os candidatos ao PAS tem probabilidade (1 – OR) vezes menor que os candidatos ao vestibular.

A estimação do modelo *logit* ocorre pelo Método de Máxima Verossimilhança, haja vista que o modelo é não linear com variável dependente dicotômica que segue uma distribuição de Bernoulli com função de probabilidade discreta dada por $P(Y = y_i) = p_i^{y_i} (1 - p_i)^{1-y_i}$

Como $p_i = \Lambda(X' \beta) \equiv \frac{e^{X' \beta}}{1 + e^{X' \beta}}$, o logaritmo natural da função de verossimilhança torna-se:

$$\mathcal{L}_N = \sum_{i=1}^n \ln \Lambda(X' \beta) + (1 - Y_i) * \ln(1 - \Lambda(X' \beta)) \quad (12)$$

O método de Máxima Verossimilhança objetiva maximizar o logaritmo natural da função de verossimilhança, isto é, obter, por meio de um processo iterativo, os parâmetros do modelo que maximizem a probabilidade de se obter os valores de y_i . Os estimadores obtidos são, em geral, consistentes, uma vez que se assegura a ausência de erros na especificação da variável dependente binária.

1.3.5 Estratégia de Estimação

Visando analisar se há diferencial na probabilidade de acesso ao ensino superior da Universidade de Brasília entre os dois processos seletivos em questão, realiza-se a estimação

de dois modelos *logit*, os quais se distinguem em função da amostra empregada. Para o Modelo 1, utiliza-se a amostra formada por candidatos ao PAS e ao vestibular sistema universal, enquanto para o Modelo 2, a amostra é formada por candidatos ao PAS e ao vestibular tanto pelo sistema universal como pelo sistema de cotas.

O Quadro 1.6 descreve as variáveis explicativas dos modelos. Para o Modelo 1, a variável PAS divide a amostra em dois grupos, os que realizaram o exame PAS e os que realizaram o exame vestibular. Como o Modelo 2 é gerado utilizando também os candidatos cotistas, para este modelo a *dummy* assume valor 0 caso o aluno tenha se candidatado ao vestibular pelo sistema universal ou de cotas. Optou-se pela inclusão de apenas uma *dummy* para captar o efeito da renda e uma *dummy* para captar os efeitos da escolaridade do pai em virtude das restrições impostas pelo tamanho da amostra. A informação sobre escolaridade da mãe não foi incluída no modelo para se evitar problemas de multicolinearidade, já que a variável possui alta correlação com escolaridade do pai.

Em relação à residência, existem três possibilidades: morar na Asa Norte, Asa Sul, Lago Sul, Lago Norte ou Sudoeste – regiões que formam a variável “brasilía”⁸ - morar nas demais regiões administrativas do DF e morar fora do DF, sendo “demais regiões administrativas” a base de comparação. A *dummy* de identidade racial “naonegro” assume valor 1 caso o aluno não tenha se declarado negro e 0 caso contrário. Optou-se por esta classificação para gerar uma distinção no modelo entre “negros e não negros”, ao invés da distinção tradicional entre “brancos e não brancos”, pois se entende que esta variável fornece uma informação mais relevante para a avaliação do efeito do componente racial sobre as chances de aprovação.

Foram adicionados aos modelos termos de interação entre a variável PAS e as demais covariadas com o objetivo de se investigar se o efeito de se candidatar ao PAS sobre as chances de aprovação diferem em relação às características sociais, econômicas e raciais incluídas no modelo.

Algumas observações devem ser feitas quanto às variáveis definidas acima. Os questionários respondidos pelos candidatos ao vestibular e ao PAS possuem perguntas em comum e perguntas distintas. Para que o modelo forneça informações comparáveis, optou-se por utilizar apenas as perguntas existentes em ambos os questionários.

⁸ Optou-se pelo uso de uma definição expandida da região de Brasília em função das similaridades entre esta região e as regiões administrativas do Lago Sul, Lago Norte e Sudoeste, que possuem os maiores níveis de renda *per capita* do DF.

Quadro 1.6 - Descrição das variáveis explicativas dos modelos *logit*

Variável	Descrição
pas	1 se prestou o exame PAS, 0 caso contrário
asculino	1 se o candidato é do sexo masculino, 0 caso contrário
rendaalta	1 se possui renda superior a R\$ 5.000,00, 0 caso contrário
paiensinosuperior	1 se pai possui ensino superior, 0 caso contrário
naonegro	1 se não se declarou negro, 0 caso contrário
brasilial	1 se reside em Brasília, zero caso contrário
foradf	1 se mora fora do DF, 0 caso contrário
cursoseletivo	1 se prestou exame para um curso seletivo, 0 caso contrário

Fonte: Elaboração própria.

Salienta-se que as variáveis que designam se o aluno estudou em escola pública ou particular não estavam presentes nos questionários. Ainda, a pergunta sobre curso preparatório apenas estava presente no questionário do vestibular. Por esses motivos, tais variáveis não foram incluídas no modelo econométrico. Entende-se que a omissão destas informações podem enviesar as estimativas, haja vista que são apontadas pela literatura como determinantes importantes do acesso ao ensino superior. Uma vez que não dispomos destas variáveis, as variáveis “renda alta” e “brasilial” podem estar capturando parte desses efeitos.

Investiga-se, ainda, se há diferença nas chances de aprovação no PAS para cursos mais seletivos, isto é, cursos com maior dificuldade de aprovação em virtude de altas notas de corte. Para criação da variável “cursoseletivo”, os cursos foram organizados em ordem crescente de nota de corte. O procedimento descrito a seguir foi aplicado tendo em vista que cada processo seletivo gera uma classificação dos cursos por nota de corte ligeiramente distinto dos demais, sendo necessário, portanto, chegar a uma lista final de cursos seletivos que fosse comum aos três processos. Foram gerados, portanto, três *rankings*, um para cada processo seletivo (PAS, sistema universal e sistema de cotas). O grupo de cursos denominados seletivos é formado pelos sete cursos com maior nota de corte dentre os dez primeiros de cada *ranking*, sendo eles: Medicina, Direito, Relações Internacionais, Engenharia Mecatrônica, Engenharia Elétrica, Engenharia de Redes e Ciência da Computação. É válido ressaltar que o critério da nota de corte é preferido ao critério da

concorrência por vaga, como adotado por Mendes Junior (2013), pois este critério reflete, necessariamente, a alta seletividade do curso⁹.

Para avaliar o ajuste do modelo, são computados o R^2 de MacFadden, o teste de Hosmer e Lemeshow e o percentual de observações corretamente previstas. O teste de Hosmer e Lemeshow é realizado por meio de um agrupamento com base nos valores das probabilidades estimadas (HOSMER e LEMESHOW, 1980). Primeiro, ordena-se as probabilidades previstas, que podem ser divididas em grupos de mesmo tamanho (HOSMER e LEMESHOW, 2000), sendo dez o número de divisões mais utilizado. Em seguida, comparam-se as frequências previstas às frequências observadas por meio de uma estatística qui-quadrado. Têm-se, como hipótese nula, que os valores esperados são iguais aos observados, sendo esta hipótese rejeitada quando p-valor obtido é inferior ao nível de significância $p < 0,05$. A não rejeição da hipótese nula indica um bom ajuste do modelo construído.

O Método de Matrizes de Classificação fora desenvolvido com o objetivo de se avaliar a capacidade preditiva do modelo de regressão com variável dependente binária. Da matriz de classificação, encontra-se a porcentagem de valores corretamente previstos obtida pela comparação entre os valores reais e previstos para a variável y (HOSMER e LEMESHOW, 2000).

1.4 Resultados

Nos modelos *logit*, as estimativas dos coeficientes fornecem de imediato apenas informação sobre o sentido da influência da variável explicativa sobre a probabilidade de sucesso, isto é, se essa influência é positiva ou negativa. A análise da magnitude do efeito requer, portanto, cálculos adicionais. O coeficiente estimado por meio da regressão logística, quando exponenciado fornece a razão de chance, que será o objeto de análise nesta seção. A Tabela 1.7 apresenta os resultados obtidos.

Primeiramente, reportam-se a *baseline odds*, que fornecem as chances de um resultado positivo quando todas as covariadas assumem valor zero. Para o Modelo 1, espera-se encontrar 0,051 aprovados para cada não aprovado (ou, dito de outra forma, 1 aprovado para

⁹Por exemplo, no PAS de 2005 a taxa de aprovação para o curso de Educação Física foi superior à registrada no curso de Medicina.

cada 20 não aprovados) para os candidatos que prestaram o exame vestibular pelo sistema universal, são mulheres, não negras, possuem renda familiar mensal inferior a R\$ 5.000,00, residem no DF, porém fora da região de Brasília, cujos pais tem nível de escolaridade até, no máximo, o ensino médio e que não prestaram o exame para cursos seletivos.

Para o Modelo 1, constata-se que o exame de seleção escolhido (PAS ou vestibular) e a cidade de residência são os principais determinantes do acesso à universidade. O candidato que disputa as vagas pelo PAS tem probabilidade de aprovação 71,2% maior do que aquele que concorre pelo vestibular, o que revela, portanto, a importância deste mecanismo de seleção para a entrada dos estudantes de ensino médio recém-formados.

Os candidatos que moram na região de Brasília tem probabilidade de aprovação 74,4% maior do que os que residem nas demais regiões administrativas. Levando-se em consideração que o modelo não possui variáveis acerca do tipo de escola que o aluno frequentou (pública e particular), ou se o aluno fez curso preparatório, acredita-se que a variável em questão esteja captando parte destes efeitos. A variável renda também possui efeito relevante: a probabilidade de aprovação é 41,7% maior para candidatos com renda superior a R\$ 5.000,00 do que para aqueles que possuem menor renda. Famílias com maior renda são capazes de fornecer melhores condições para o estudo, que tendem a melhorar o desempenho dos estudantes, como o acesso à internet, apoio na preparação para os exames, escola de idiomas (GUIMARÃES e SAMPAIO, 2007), influenciando positivamente as chances de sucesso na seleção. O menor efeito da renda sobre a probabilidade de aprovação, quando comparada ao efeito da cidade de residência, pode estar relacionado com o fato de os estudantes não possuírem, muitas vezes, um conhecimento preciso acerca da renda auferida por seus pais.

As chances de aprovação aumentam 54,1% quando o candidato é do sexo masculino, o que mostra a vantagem que os candidatos homens possuem nos vestibular vis-à-vis o exame PAS – efeito este que será discutido em maiores detalhes durante a análise do termo de interação. Ter pais com ensino superior eleva as chances de aprovação por um fator de 1,38 em relação aos candidatos cujos pais possuem menor grau de instrução, resultado também em consonância com a literatura. Constata-se, deste modo, que a escolaridade do pai possui grande papel no sucesso dos filhos em exames de admissão. As chances de aprovação reduzem-se 66,5% quando o candidato se inscreve para um curso seletivo. As variáveis “não se declarar negro” e “residir fora do DF”, por sua vez, não apresentaram razão de chance estatisticamente significativa.

Tabela 1.7 - Variáveis Determinantes de Aprovação no PAS e no Vestibular

Variáveis Explicativas	Modelo 1 (PAS X Sistema Universal)		Modelo 2 (PAS X Sistema Universal e de Cotas)	
	Coefficientes	OR	Coefficientes	OR
pas	0,538*** (0,198)	1,712*** (0,339)	0,730*** (0,157)	2,074*** (0,327)
masculino	0,432*** (0,106)	1,541*** (0,164)	0,412*** (0,086)	1,510*** (0,130)
rendaalta	0,348** (0,135)	1,417** (0,192)	0,334*** (0,117)	1,396*** (0,163)
paiensinosuperior	0,323*** (0,119)	1,381*** (0,165)	0,393*** (0,097)	1,482*** (0,143)
naonegro	-0,077 (0,134)	0,926 (0,125)	0,046 (0,088)	1,047 (0,092)
brasilias	0,556*** (0,127)	1,744*** (0,222)	0,589*** (0,102)	1,802*** (0,183)
foradodf	-0,017 (0,154)	0,983 (0,152)	0,047 (0,130)	1,048 (0,136)
cursoseletivo	-1,097*** (0,162)	0,334*** (0,054)	-0,989*** (0,130)	0,372*** (0,048)
pasxmasculino	-0,331** (0,141)	0,718** (0,101)	-0,310** (0,126)	0,733** (0,093)
pasxrendaalta	-0,070 (0,173)	0,933 (0,161)	-0,055 (0,159)	0,947 (0,150)
pasxpaiensinosuperior	0,014 (0,160)	1,014 (0,162)	-0,057 (0,144)	0,945 (0,136)
pasxnaonegro	0,246 (0,184)	1,279 (0,235)	0,123 (0,153)	1,131 (0,173)
pasxbrasilias	0,149 (0,168)	1,161 (0,195)	0,117 (0,149)	1,124 (0,168)
pasforadodf	-0,072 (0,209)	0,931 (0,195)	-0,136 (0,192)	0,873 (0,167)
pasxcursoseletivo	0,619*** (0,208)	1,857*** (0,386)	0,511*** (0,184)	1,667*** (0,308)
constante	-2,980*** (0,151)	0,051*** (0,008)	-3,171*** (0,092)	0,042*** (0,004)
χ^2		424,3304		538,4147
Homer e Lemeshow		0,77 (4,88)		0,74 (5,12)
Pseudo R ²		0,0659		0,0668
Corretos (%)		90,34%		91,44%
N		10,147		13,800

Fonte: Elaboração própria. Nota: ***Estatisticamente significantes a 1%. **Estatisticamente significantes a 5%. *Estatisticamente significantes a 10%. Os erros-padrão e o grau de significância do teste de Hosmer e Lemeshow estão entre parênteses.

Para se compreender a magnitude dos efeitos encontrados, é importante compará-los às *baseline odds*. Quando o candidato concorre às vagas pelo PAS, as chances mudam de 0,051 aprovados a cada não aprovado (1 aprovado a cada 20 não aprovados) para 0,087 aprovados para cada não aprovado (1 para 11), o que é uma variação substantiva. Em relação às chances

de aprovação em um curso seletivo, estas se alteram para 1 aprovado a cada 61 não aprovados, mudança relevante em relação as *baseline odds*.

Quando se analisam as variáveis de interação, observa-se significância estatística apenas em relação ao sexo do candidato e a seletividade do curso. Quando o candidato é do sexo masculino, a razão de chance de ser aprovado pelo PAS reduz-se por um fator de 0,72. Revela-se, portanto, que processos seletivos que ocorrem em mais de uma etapa aumentam as chances de mulheres ingressarem no ensino superior, o que se apresenta como um resultado relevante quando se considera que os homens tendem a ter desempenho superior em exames de seleção (GUIMARÃES e SAMPAIO, 2007) e a apresentarem maiores taxas de aprovação nos vestibulares tradicionais (EMILIO, JUNIOR e ALVES, 2004). Estudos da área de psicologia e psiquiatria mostram que as mulheres apresentam níveis de ansiedade significativamente mais elevados do que os candidatos homens em exames vestibulares (RODRIGUES e PELISOLI, 2008). O fato de o PAS ser realizado em três etapas pode contribuir para a redução da ansiedade, já que a pressão do evento vestibular passa a ser diluída em três eventos, reduzindo o impacto da ansiedade no desempenho do aluno.

Se o estudante se candidatou a um curso seletivo, a razão de chance de ser aprovado pelo PAS aumenta por um fator de 1,86, o que ressalta o papel do exame seriado na admissão em cursos com maior nota de corte, os quais possuem, na maioria das vezes, as maiores concorrências por vaga. Observa-se, portanto, que as maiores oportunidades de acesso a cursos como Medicina, Direito e Engenharias são oferecidas pelo PAS. Caso esse aluno dispute tais vagas pelo vestibular, ele enfrentará uma concorrência por vaga mais acirrada e formada por um conjunto de candidatos mais experientes em exames e com maior tempo de preparo, fatores que dificultam o acesso do recém-formado à universidade.

Como descrito anteriormente, o Modelo 2 diferencia-se do Modelo 1 em relação ao conjunto de candidatos analisados: enquanto este compara os que se inscreveram para o PAS com os que se inscreveram para vestibular pelo sistema universal, aquele realiza a comparação entre os candidatos ao PAS e ao vestibular tradicional, que inclui o sistema universal e o sistema de cotas.

Observa-se uma diferença relevante nas chances de aprovação entre os dois modelos para o candidato que se inscreveu para o PAS. No Modelo 2, o candidato que está disputando as vagas por este Programa tem probabilidade de aprovação 107,2% maior do que aquele que

disputa as vagas pelo vestibular. Ser do sexo masculino, possuir renda alta, residir em Brasília e ter pais com ensino superior aumentam a chance de aprovação. A chance de aprovação se reduzem em 62,8% quando o estudante se inscreve para um curso seletivo.

Tal como no primeiro modelo, residir fora do DF e ser negro não apresentaram razões de chances estatisticamente significativa se os únicos termos de interação a apresentarem significância estatística foram “masculino” e “cursos seletivos”. Se o candidato é do sexo masculino, reduzem-se as chances de ser aprovado pelo PAS por um fator de 0,73. Para um candidato inscrito para um curso seletivo, as chances de aprovação pelo PAS aumentam por um fator de 1,67. Comparando-se a razão de chance da interação em ambos os modelos, observa-se, portanto, que a probabilidade de aprovação em um curso seletivo é menor quando se considera o conjunto de estudantes cotistas. Assim, a inclusão dos estudantes cotistas no conjunto de candidatos ao vestibular impacta negativamente a probabilidade de aprovação em cursos seletivos por meio do PAS.

As *baseline odds* para o Modelo 2 são de 0,042 aprovados para cada não aprovado (1 para 24). Nota-se, portanto, que a probabilidade de aprovação para o candidato de referência é menor quando a amostra considera os inscritos em todas as modalidades – PAS, vestibular sistema universal e cotas. Ainda que o efeito do PAS sobre a aprovação seja mais elevado que o visto no Modelo 1, ao se comparar este efeito em relação a *baseline odds* verifica-se que ele está próximo daquele encontrado no Modelo 1: as chances passam de 1 aprovado a cada 24 não aprovados, para 1 aprovado a cada 11 não aprovados. Prestar o exame para um curso seletivo muda a proporção entre aprovados e não aprovados de 0,042 para 0,016 (1 aprovado a cada 64 não aprovados).

Como medidas de ajustamento, utilizam-se, além do R^2 de McFadden, o percentual das observações corretamente preditas e o teste de Hosmer Lemeshow. Embora o R^2 de MacFadden apresente um valor considerado baixo, ele não difere daquele encontrado em artigos que estimam as chances de aprovação em processos seletivos como o vestibular (CESAR, 2013; MENDES JUNIOR, 2013) e concursos (CASTELAR et al, 2010). Em particular, o modelo *logit* utilizado por CESAR (2013) obteve R^2 de MacFadden ligeiramente superior ao do Modelo 1, mesmo tendo sido estimado com uma amostra maior – a taxa de resposta ao questionário que originou a amostra utilizada pela autora foi bastante superior à taxa de resposta ao QSC de 2005 – e empregado mais variáveis que não estavam disponíveis no questionário de 2005.

No teste de Hosmer e Lemeshow realizado, a estatística qui-quadrado foi de 4,88, com 8 graus de liberdade e p-valor de 0,77. Logo, não se rejeita a hipótese nula de que há diferenças entre os valores preditos e os observados, de onde se conclui que o modelo pode ser considerado adequado para prever as chances de aprovação.

O valor ponderado de predições corretas foi de 90,34% para o Modelo 1 e 91,44% para o modelo 2. Valores altos para esta medida de grau de ajuste também foram obtidos em Mendes Junior (2013) e em Cesar (2013). Em Cesar (2013), o modelo prevê os valores da variável dependente de maneira correta em 81,13% das observações.

1.5 Conclusão

Este trabalho buscou mensurar os determinantes do acesso à Universidade de Brasília por meio do PAS, processo de seleção alternativo ao exame vestibular, cuja avaliação é feita em três etapas. O estudo objetiva analisar se há diferenças nos determinantes de aprovação entre as duas principais forma de acesso à UnB, à saber, PAS e Vestibular. Para tal, utilizou-se um modelo de regressão logística. Foram incluídas *dummies* de interação a fim de investigar se há diferenças na probabilidade de acesso entre PAS e vestibular para uma série de variáveis independentes, como sexo, renda, residência, escolaridade do pai e seletividade do curso. O modelo *logit* foi empregado duas vezes: na primeira, considera-se a amostra formada apenas por candidatos ao PAS e ao vestibular pelo sistema universal, isto é, sem considerar os candidatos cotistas. Na segunda, estima-se o modelo para a amostra formada pelos candidatos ao PAS e ao vestibular tanto pelo sistema universal quando pelo sistema de cotas.

Os resultados mostram que o PAS possui um efeito relevante sobre a probabilidade de aprovação. Dentre as variáveis inclusas no modelo, a cidade de residência e o sexo do candidato foram as que mais impactam a probabilidade de sucesso no exame. Ter renda alta e pais com ensino superior também aumentam a chance de aprovação, enquanto a inscrição para um curso seletivo a reduz substancialmente. Os resultados mostram que a probabilidade de aprovação pelo PAS para homens é menor do que para mulheres. Ainda, observa-se que o Programa amplia a probabilidade de aprovação em cursos seletivos.

Anexo A - Distribuição dos Candidatos por Percentil de Nota – Vestibular Sistema Universal e Cotas

	Percentis do Escore Bruto – Vestibular Sistema Universal e Cotas						
	<10%	10% 30%	30% 50%	50% 70%	70% 90%	90% <	
Taxa de Aprovação	0,00	0,00	0,19	2,82	11,60	32,33	
Masculino	37,5	42,1	45,8	47,4	50,8	57,5	
Renda Mensal Familiar (R\$)							
Renda < 500	28,0	20,7	12,9	8,1	4,4	2,2	
500 < Renda ≤ 1500	38,5	38,8	34,9	27,5	19,3	12,0	
1500 < Renda ≤ 2500	12,2	16,8	17,5	19,3	18,0	13,9	
2500 < Renda ≤ 5000	11,8	13,0	19,9	23,5	26,4	28,8	
Renda > 5000	4,9	4,8	7,2	11,8	20,2	31,7	
Não sabem	4,6	5,9	7,7	9,9	11,8	11,5	
Mora em Casa Própria	66,2	68,5	70,3	71,7	73,2	84,5	
Educação do Pai							
Ensino Básico Incompleto	41,8	33,1	25,0	17,0	10,2	4,5	
Ensino Básico Completo	11,6	11,6	9,0	8,2	7,2	3,8	
Ensino Médio Completo	27,2	32,2	35,1	32,8	31,0	25,8	
Ensino Superior	13,7	17,8	25,2	37,6	48,8	63,4	
Não sabem	5,8	5,4	5,8	4,4	2,9	2,6	
Educação da Mãe							
Ensino Básico Incompleto	41,0	33,5	23,9	15,0	8,4	2,8	
Ensino Básico Completo	10,2	11,6	10,7	8,3	6,7	3,4	
Ensino Médio Completo	33,0	33,9	35,2	38,7	33,4	28,4	
Ensino Superior	13,7	19,7	28,5	36,6	50,0	64,7	
Não sabem	2,1	1,4	1,8	1,5	1,5	0,8	
Raça							
Branca	23,9	24,4	27,1	28,5	34,5	37,9	
Preta	22,3	19,8	15,3	12,5	8,2	5,6	
Amarela	4,0	3,4	2,9	2,3	2,4	1,8	
Parda	45,3	47,2	48,3	49,2	47,4	44,8	
Indígena	1,1	0,7	0,7	0,6	0,4	0,7	
Sem declaração	3,5	4,6	5,8	6,8	7,1	9,3	
Negro	52,2	53,8	50,4	48,3	43,2	37,5	
Não trabalha	69,6	69,8	74,0	80,2	84,6	90,4	
Residência							
Demais Regiões Administrativas	73,0	68,7	61,8	55,3	43,5	28,8	
Brasília	12,7	17,0	22,2	25,8	32,4	39,0	
.Fora do DF	14,3	14,3	16,0	18,9	24,1	32,2	

Fonte: Cespe.

Anexo B - Estatísticas Descritivas para o Vestibular Sistema Universal, Cotas e PAS

	Vestibular Universal e Cotas		PAS		Inscritos		Aprovados PAS, Vestibular Cotas		Diferença nas Médias	
	Não aprovados (%)	Aprovados (%)	Não aprovados (%)	Aprovados (%)	Vestibular e Cotas (%)	PAS (%)				
Masculino	47,2	***	56,1	41,2	**	44,9	47,6	***	41,6	***
Renda Mensal Familiar (R\$)										
Renda < 500	12,2	***	6,5	8,9	***	3,5	11,9	***	8,2	**
500 < Renda ≤ 1500	29,3	***	21,43	21	***	9,4	28,8	***	19,4	***
1500 < Renda ≤ 2500	17,0	-	17,37	14,8	-	13,8	17,0	***	14,7	*
2500 < Renda ≤ 5000	20,3	**	24,03	25,9	-	27,2	20,5	***	26,1	-
Renda > 5000	12,6	***	20,45	22,6	***	36,3	13,1	***	24,5	***
Não sabem	8,7	-	10,23	6,7	***	9,8	8,8	***	7,1	-
Possui casa própria	71,2	-	72,77	78,48	**	82,35	71,3	***	79,0	***
Educação do Pai										
Ensino Básico Incompleto	22,56	***	10,86	13,8	***	5,7	21,8	***	12,7	***
Ensino Básico Completo	8,8	-	7,46	6,7	**	4,1	8,7	***	6,4	**
Ensino Médio Completo	30,8	***	30,95	29	***	25,6	30,8	***	28,5	**
Ensino Superior	33,2	***	47,49	46,7	***	63,2	34,1	***	49	***
Não sabem	4,7	*	3,24	3,8	***	1,4	4,6	***	3,4	**
Educação da Mãe										
Ensino Básico Incompleto	21,5	***	11,02	12	***	4,7	20,9	***	11	***
Ensino Básico Completo	9,0	***	5,35	6,2	***	3	8,7	***	5,8	**
Ensino Médio Completo	34,2	-	31,6	32,4	***	26,6	34,0	***	31,6	*
Ensino Superior	33,8	***	50,57	48,2	***	65,4	34,9	***	50,5	***
Ensino Básico Incompleto	1,5	-	1,46	1,3	**	0,2	1,5	*	1,2	**
Raça										
Branca	29,6	-	30,47	48,9	-	48,9	29,7	***	48,9	***
Preta	14,2	***	6,65	6,9	***	3,2	13,7	***	6,4	***
Amarela	3,0	-	2,11	3,1	-	3,7	2,9	-	3,2	*
Parda	46,4	**	50,57	34,9	-	31,7	46,7	***	34,4	***
Indígena	0,6	-	0,49	0,6	**	1,5	0,7	-	0,7	*
Sem declaração	6,2	***	9,72	5,6	***	10,9	6,4	-	6,3	-
Negro	47,1	-	44,26	21,3	***	16,3	46,9	***	20,6	***
Não trabalha	76,1	*	79,13	87,8	**	90,8	76,3	***	88,2	***
Residência										
Demais Regiões Administrativas	56,1	***	43,86	51,97	***	37,77	55,34	***	50,03	**
Brasília	23,9	***	41,24	25,84	***	47,89	24,98	***	28,85	**
Fora do DF	20	***	14,89	22,2	***	14,3	19,68	**	21,1	-
<i>N</i>	<i>617</i>		<i>9.303</i>	<i>592</i>		<i>3782</i>	<i>9920</i>		<i>4374</i>	

Fonte: Cespe.

CAPÍTULO 2 AVALIAÇÃO DO PAS DA UNIVERSIDADE DE BRASÍLIA: EFEITOS SOBRE DESEMPENHO, PERMANÊNCIA E EVASÃO DOS ALUNOS

2.1 Introdução

No capítulo 1 foram apresentados a estrutura e os principais objetivos do Programa de Avaliação Seriada (PAS) da UnB. Conforme discutido, o PAS é uma modalidade alternativa de acesso ao ensino superior que surgiu como uma tentativa de amenizar o impacto da passagem pelo vestibular para o ingresso na universidade. O propósito de integrar o ensino médio à universidade é realizado por meio de fóruns entre os discentes do ensino médio e do ensino superior para decisão do conteúdo programático do exame. Diferentemente do vestibular, as provas do PAS possuem a característica da contextualização e interdisciplinaridade, com ênfase no desenvolvimento de competências e habilidades.

Entre os objetivos específicos do Programa está a busca por um exame que avalie a aprendizagem significativa, “em que se privilegie a reflexão sobre a memorização, a qualidade sobre a quantidade de informações, o ensino sobre o adestramento, o processo sobre o produto” (UNB, 1996). A seleção de candidatos via PAS é entendida, deste modo, como um processo que almeja selecionar candidatos com maiores habilidades de reflexão e capazes de aplicar de forma mais eficaz os conteúdos apreendidos em sala de aula. O Programa, portanto, não se define apenas como um exame de seleção em três episódios, mas como um instrumento por meio do qual se busca estimular os estudantes a investirem continuamente no preparo para a prova e no desenvolvimento de capacidades e habilidades que lhes serão úteis no futuro, tanto acadêmico, na universidade, quanto profissional, após a sua formação.

Destarte, a avaliação de um programa de seleção como o PAS permeia duas questões. A primeira, analisada no capítulo 1, surge quando se busca entender os fatores que determinam o acesso dos alunos à universidade pelo Programa, de forma a avaliar se há diferenças nas chances de aprovação a depender do exame de seleção escolhido. A segunda questão, que será desenvolvida no presente capítulo, remete-se à avaliação do desempenho dos alunos dentro da universidade como forma de compreender em que medida as diferenças existentes em cada processo seletivo refletem no desempenho do aluno que está ingressando na UnB.

A literatura sobre avaliação do ensino superior aponta diferentes indicadores para o desempenho do aluno universitário. Primeiramente, pode-se investigar a existência de diferenças no rendimento acadêmico do estudante avaliado por meio de indicadores de

desempenho gerados a partir das notas ou menções obtidas nas disciplinas cursadas durante a graduação. Em segundo lugar, a avaliação dos alunos pode ser feita em relação ao tempo de permanência na universidade, que reflete o aproveitamento geral dos estudantes nas disciplinas tendo em vista que um bom aproveitamento nas disciplinas resulta em uma aderência ao fluxo escolar recomendado para cada curso. Pode-se, ainda, avaliar a proporção de alunos que efetivamente concluíram a graduação, bem como a proporção daqueles que saíram da universidade antes de concluí-la, como um indicador acerca da efetividade do PAS e do vestibular em selecionar alunos que irão, de fato, atingir o objetivo maior da universidade, qual seja: formar profissionais em diversas especialidades.

A metodologia de avaliação de programas descrita neste capítulo visa encontrar um grupo de controle para os alunos que ingressaram na universidade por meio do PAS para que se possa avaliar, de forma adequada, os efeitos do ingresso pelo Programa sobre o desempenhos dos universitários. Este grupo será formado a partir do conjunto de alunos que não participaram do PAS, isto é, que ingressaram na universidade por meio de outro processo seletivo, a saber, o vestibular. O método do pareamento por escore de propensão irá, então, selecionar no grupo de alunos do vestibular aqueles que formarão o grupo de comparação e cujos resultados serão utilizados para se estimar o efeito do PAS sobre os ingressantes por esse sistema em relação às variáveis de desempenho acadêmico, permanência na universidade e evasão nos cursos de graduação.

Este capítulo está dividido da seguinte forma: a seção 2.2 é formada pela revisão da literatura de avaliação de estudantes universitários em três dimensões: desempenho, permanência na universidade e evasão. Nesta mesma seção, apresentam-se os trabalhos que focam na avaliação dos programas seriados. A seção 2.3, por sua vez, descreve a metodologia de avaliação que será adotada, com a exposição das hipóteses e detalhamento sobre o procedimento de estimação. Em seguida, na seção 2.4, descreve-se a base de dados bem como a escolha do grupo de tratamento e controle, das covariadas e das variáveis de interesse. A seção 2.5 expõe os resultados obtidos e, finalmente, a seção 2.6 explicita as principais conclusões e questões atinentes ao trabalho.

2.2. Revisão da Literatura

Nessa seção serão apresentados os artigos que motivaram a seleção das variáveis de interesse deste capítulo, bem como os artigos que contribuíram para a escolha das covariadas a serem utilizadas para o cálculo do escore de propensão. Primeiramente, são citados os artigos sobre desempenho acadêmico na universidade. Em seguida, apontam-se os autores que estudaram a questão da permanência e da evasão no ensino superior. Por fim, realiza-se a revisão da literatura sobre avaliação de programas seriados, em que é dado destaque aos trabalhos sobre o PAS da UnB.

2.2.1 Desempenho Acadêmico na Universidade

A literatura sobre avaliação do desempenho acadêmico universitário encontra-se bastante consolidada. Os estudos que serão apresentados analisam a relevância de variáveis como renda, tipo de escola, educação dos pais e raça sobre indicadores de rendimento acadêmico, que variam de acordo com o país e a universidade avaliada. Betts e Morell (1999) estimaram os determinantes do desempenho acadêmico de estudantes da Universidade de San Diego utilizando um modelo com efeito fixo. Na amostra usada, encontram-se todos os estudantes que se matricularam na universidade entre os anos de 1991 e 1993 e frequentaram escolas públicas durante o ensino médio. Corroborando a literatura, os autores encontram que o *background* dos alunos influencia significativamente o GPA (*Grade Point Average*), principal medida de desempenho acadêmico em universidades americanas. A escola em que o aluno concluiu o ensino médio também afeta o desempenho do estudante, mesmo depois de se controlar por diversas variáveis. Outro efeito encontrado pelos autores refere-se à experiência de trabalho dos professores do ensino médio, a qual, apesar de positiva e significante, é de pequena dimensão. Ademais, cabe salientar que não foram encontrados efeitos da escolaridade dos professores sobre o desempenho acadêmico dos alunos.

Cohn et al (2004) estimam como algumas variáveis que medem o desempenho dos alunos durante o ensino médio afetam o desempenho acadêmico na universidade medido pelo GPA nos Estados Unidos. Entre tais variáveis pode-se destacar: a pontuação obtida no SAT – exame para admissão à universidades americanas, o GPA durante o ensino médio e a classificação dos alunos em relação aos demais estudantes de sua sala (*class rank*) durante o ensino médio. A amostra foi construída com dados de estudantes da Universidade da Carolina

do Sul entre 2000 e 2001. Por meio de um MQO, os autores encontraram que as três variáveis independentes investigadas possuem efeito positivo e significativo ao nível de 1% sobre o GPA dos universitários. No entanto, a variável SAT tem um maior poder de explicar o desempenho acadêmico do que o GPA realizado durante o ensino médio e a classificação dos alunos.

Bai e Chi (2011) intentam, por sua vez, analisar o papel do teste nacional de seleção para as universidades chinesas sobre o desempenho dos alunos na universidade. O National College Entrance Examination (CEE) é o principal fator considerado na admissão dos estudantes nas universidades do país. Apesar da importância do teste, há pouca evidência empírica sobre a relação causal entre seu resultado e o desempenho acadêmico dos estudantes uma vez admitidos nas faculdades. Os autores encontraram que o CEE é um bom preditor das notas obtidas pelos alunos durante os quatro anos da graduação. Os resultados do artigo mostram que o *background* escolar dos alunos (desempenho durante o ensino médio, por exemplo) também possui forte influência sobre a performance acadêmica.

Com o objetivo de encontrar os determinantes do rendimento dos estudantes, Lima (2011) utiliza dados inéditos de 240 questionários aplicados entre alunos de Ciências Econômicas da UnB, no período de agosto a setembro de 2011. Ao empregar o método dos Mínimos Quadrados Ordinários, os modelos mostraram que a quantidade de horas de estudo e a frequência às aulas são fundamentais para o desempenho acadêmico medido pelo Índice de Rendimento Acadêmico (IRA) da UnB. Corroborando pesquisas anteriores, entre os resultados encontrados pelo autor podemos destacar que mulheres apresentaram rendimento superior aos homens e que cotistas obtiveram desempenho inferior à média dos outros alunos. O autor também analisa os efeitos da vida social dentro do *campus* sobre o rendimento acadêmico, embora não tenham sido encontrados efeitos significativos sobre esse aspecto.

2.2.2 Permanência na Universidade e Evasão

As análises que buscam compreender o comportamento do aluno dentro da universidade em relação ao seu tempo de permanência e sobre a evasão são relevantes não apenas para o estudante – que será futuramente recompensado por seu investimento em capital humano – como também para a universidade, tendo em vista que um dos maiores objetivos da universidade é formar alunos aptos de atuar profissionalmente naquilo em que se especializaram. A análise sobre a permanência na universidade foi realizada, em sua maioria,

por meio de trabalhos com modelos de análise de sobrevivência, os quais também abordam a questão do abandono (MURTAUGH, BURNS e SCHUSTER, 1999; SILVA, FREITAS e CAVALCANTI, 2012; JUNIOR, SILVEIRA e OSTERMAN, 2012).

Guimarães, Sampaio e Sampaio (2010) realizaram um estudo relevante para a avaliação dos motivos que levam alunos a desistirem de seus cursos. Os autores utilizam um modelo *probit* para estimar os determinantes das matrículas e das desistências na Universidade Federal de Pernambuco (UFPE). Observou-se que as desistências são positivamente correlacionadas com a renda e a nota do vestibular dos alunos. Além disso, as mulheres têm menor probabilidade de entrar na universidade, assim como de desistir do curso. Os alunos que estudaram em escolas públicas têm menor probabilidade de entrar na universidade, mas não há efeito da escola sobre a probabilidade de desistência.

Com dados da mesma universidade, Sampaio et al. (2011) analisam a decisão de evasão observando fatores como renda e desempenho dos alunos no exame de seleção e na universidade. Os autores observaram que a evasão é maior nos cursos cuja nota de entrada é menor. Ao mesmo tempo, em relação aos alunos de um mesmo curso, a taxa de evasão é maior entre aqueles de melhor desempenho. A explicação dada pelos autores é que tais indivíduos têm maior renda e maior oportunidade de tentar uma nova carreira.

Também foram realizados estudos referentes à evasão na UnB para determinados cursos. Cunha, Tunes e Silva (2001) utilizam dados da trajetória acadêmica e informações obtidas em entrevistas dos alunos do curso de Química na universidade em questão para analisar as razões que levam à evasão. O estudo conclui que a decisão de sair da universidade está relacionada a aspectos anteriores à entrada na universidade, expectativas não atendidas e problemas encontrados durante o curso.

Silva, Freitas e Cavalcanti (2012) analisam dados do curso de Estatística da Universidade Federal da Paraíba buscando identificar os fatores que influenciam a desistência no curso. Para tal, utilizam modelos de regressão em análise de sobrevivência. Esses autores observaram que mais de 50% dos evasores são do sexo masculino e que 72,9% dos alunos desistentes ingressaram na universidade via processo seletivo seriado. Não foi encontrado efeito do ensino básico em escola pública ou privada.

Também com um modelo de análise de sobrevivência, Murtaugh, Burns e Schuster (1999) estudaram a retenção e evasão na Universidade do Estado de Oregon, Estados Unidos. Os autores concluíram que a taxa de evasão aumenta com a idade e decresce com a média de

notas no ensino médio do aluno, e que os estudantes que tiveram orientação de início de curso apresentaram menor risco de desistência.

2.2.3 Avaliação de Programas Seriadados

Manzano (2011) discute o surgimento e a consolidação dos vestibulares seriados no Brasil comparando vinte e dois programas. Em uma análise de questionários respondidos por professores de diferentes universidades, a autora conclui que há uma preferência pelo sistema seriado ao vestibular tradicional, com exceção da Região Nordeste, onde 75% dos professores discordam da superioridade do vestibular seriado em relação ao vestibular tradicional. Grande proporção dos professores afirma que o programa seriado oferece mais oportunidades que o vestibular tradicional, apesar de essa concordância ser superior entre os professores de escolas públicas (90% e 80%, respectivamente).

Maciel e Lopes (2001) compararam o desempenho de alunos do curso de Engenharia Civil da Universidade Federal de Santa Maria (UFSM) que ingressaram via vestibular com aqueles que ingressaram via o Programa de Ingresso ao Ensino Superior – Peies. Para isso, analisaram o número de reprovações dos estudantes de cinco turmas, entre 1998 e 2000, e o desempenho médio desses alunos na universidade. A variável de desempenho acadêmico investigada no trabalho foi a retenção, definida como a reprovação em mais de duas disciplinas no semestre cursado ou a reprovação repetida na mesma disciplina. O aluno retido fica impedido de avançar no curso até que logre aprovação nas disciplinas. Os alunos que ingressaram na UFSM via Peies apresentaram diferença no desempenho médio em relação aos alunos que ingressaram via vestibular a partir do 5º semestre: apesar de apresentarem desempenho médio de 60%, os alunos do Peis recuperaram-se, tendo o desempenho chegado a 100%. Em contrapartida, os que ingressaram via vestibular mantiveram o desempenho de 70%, valor próximo ao auferido nos semestres anteriores.

Em uma análise qualitativa, Schlichting, Soares e Biancheti (2004) discutem a experiência de vestibular seriado do Sistema de Avaliação do Ensino Médio (Saem) entre 1998 e 2002 da Universidade Federal de Santa Catarina (UFSC). Utilizando como base aquela formada a partir dos questionários enviados a professores e orientadores educacionais, assim como redações de alunos sobre o Programa, os autores concluíram que o Saem contribuiu para que os alunos da escola pública tivessem maior probabilidade de sucesso no ingresso à universidade.

Reis (2006) avaliou o desempenho de alunos da Faculdade de Odontologia da Universidade Federal de Uberlândia, admitidos pelo vestibular e pelo Programa Alternativo de Ingresso ao Ensino Superior (Paies). Foram analisados os registros de notas e avaliações subjetivas, realizados durante a graduação entre o segundo semestre de 2002 e o segundo semestre de 2003. Os resultados apontaram melhor desempenho e maior competência de aprendizagem dos alunos do Paies. O número de dependências e a taxa de evasão durante o curso foram ligeiramente maiores no grupo do vestibular.

O primeiro estudo significativo sobre o PAS da UnB foi solicitado em 2005. A pesquisa institucional, intitulada “Programa de Avaliação Seriada (PAS): o Balanço de uma década”, de autoria de MATOS (2006), utilizou-se de três bases de dados distintas, entre elas os dados institucionais fornecidos pela administração da UnB com estatísticas básicas de seis subprogramas concluídos entre 1996 e 1998, 2001 e 2003 e 11 vestibulares realizados no mesmo período (1/1999 a 1/2004). Foram aplicados questionários para uma amostra representativa de toda a comunidade escolar do ensino médio do Distrito Federal e entorno composta por pais, professores e gestores com o objetivo de agregar ao estudo uma avaliação da comunidade escolar sobre o Programa. Para captar a percepção dos atores mais diretamente envolvidos no PAS, foram formados grupos focais compostos de membros da Comissão de Acompanhamento, dos Comitês de Revisão dos Objetos de Avaliação e de professores responsáveis pela oferta de cursos no âmbito do Fórum Permanente de Professores. Os dados referentes à percepção da comunidade sobre os efeitos do PAS no ensino médio se mostraram favoráveis ao Programa. O autor também encontra um desempenho melhor dos alunos ingressantes pelo PAS quando comparados aos que ingressaram pelo vestibular. A comparação é feita por meio do IRA de alunos com dois ou mais anos de estudo.

Por meio da análise de dados de 415 alunos do curso de Medicina da UnB, de 1999 a 2004, Sobral e Oliveira (2006) avaliam o efeito do PAS, em comparação ao vestibular sobre o perfil do alunato e o desempenho acadêmico dos ingressantes. Foram comparados indicadores de rendimento acadêmico, baseado em notas e reprovações, de demografia e de valoração do aprendizado. O grupo formado pelos alunos ingressantes pelo PAS apresentou proporções significativamente maiores de alunos oriundos do Distrito Federal e de faixa etária mais jovem. Ainda, o IRA foi significativamente mais elevado desde o segundo até o nono semestre do curso para este grupo. Além disso, os autores observaram que parcela

significativa da variação no desempenho no primeiro ano pode ser atribuída à diferenças de sexo, idade, naturalidade, ano de ingresso e valoração do aprendizado. Os resultados sugerem que o PAS privilegiou a aptidão acadêmica quanto ao rendimento cognitivo e alterou tendências na composição demográfica do corpo discente a partir de 1998 (ano de ingresso da primeira turma do PAS), com efeitos no próprio rendimento estudantil.

Gottschall (2008), por sua vez, aponta os efeitos do PAS sobre o ensino médio, por meio de pesquisa feita em escola particular do Distrito Federal. A autora focou seu estudo na disciplina de matemática e concluiu que o PAS provocou mudanças nas práticas avaliativas e na sequência com que os conteúdos são ministrados. O PAS trouxe dinamização ao ensino da disciplina por meio da contextualização e da interdisciplinaridade. Todavia, houve maior ênfase nos conteúdos em detrimento de outras dimensões da aprendizagem, além de também ter proporcionado uma relação mais estreita entre os conteúdos ensinados no ensino médio e os que são exigidos pelo Programa.

Autores como Ferreira (2002) e Missio (2007), fazem algumas ressalvas quanto à adoção de programas de avaliação seriada. O primeiro afirma que esses programas também culminam em uma prova e que por isso englobam a lógica de exclusão de classificação já presente no vestibular. Para o segundo, a avaliação seriada agrava as características excludentes do vestibular, tendo em vista que a maior parte dos candidatos inscritos fracassa. Esse argumento, no entanto, enfraquece-se perante a análise das taxas de aprovação na avaliação seriada para a UnB, por exemplo, que apresentam valores superiores às taxas de aprovação no vestibular.

2.3 Metodologia

A análise dos efeitos do PAS sobre o desempenho dos alunos que ingressaram por este processo de seleção será feita por meio da metodologia de avaliação de programas descrita a seguir. O método visa encontrar um grupo de controle para o grupo de tratamento que, para este estudo, é formado pelos alunos que entraram na universidade pelo PAS. O tratamento é, para efeitos desse estudo, ter ingressado na universidade pelo Programa de Avaliação Seriada.

O maior desafio da avaliação de impacto é a determinação de o que teria acontecido aos que foram diretamente beneficiados por um programa caso ele não tivesse existido (KHANDKER, KOOLWAL e SAMAD, 2010). Nesse sentido, a forma ideal para avaliação

de impacto resultaria da comparação, em um período de tempo posterior ao tratamento, do resultado (*outcome*) do indivíduo tratado com o resultado deste mesmo indivíduo caso ele não tivesse sido tratado. O resultado do beneficiário na ausência de intervenção denomina-se *contrafactual*. No entanto, nota-se que é impossível mensurar o resultado para um indivíduo em ambas as situações (na ausência e presença do Programa), isto é, o *contrafactual* não é observável, pois se tratam de situações mutuamente excludentes, o que é citado pela literatura como “o problema do *contrafactual*”. O pesquisador irá, portanto, conduzir a sua pesquisa tendo como foco um grupo de tratamento e tentará encontrar um grupo de controle composto por indivíduos que, caso tivessem sido tratados, teriam um resultado, na média, similar ao resultado médio encontrado no grupo de tratamento. Esse grupo seria portanto, o *contrafactual* do grupo de tratamento.

A primeira pergunta que normalmente surge durante a avaliação do efeito de um tratamento relaciona-se à escolha de um grupo de controle adequado. O efeito do tratamento sobre a variável de interesse não pode ser medido por meio de uma subtração simples entre o resultado depois do tratamento para todos os tratados e o resultado para um grupo de não tratados. A justificativa tem estreita relação com o problema de viés de seleção: os indivíduos que buscam se tratar tendem a ser diferentes dos que não buscam o tratamento. Desta forma, a mensuração do efeito tal como mencionado introduziria um viés ao estimador de efeitos de tratamento.

No mesmo sentido, a comparação simples entre os resultados pré e pós-tratamento para um grupo de tratados também não seria uma forma adequada de mensuração do impacto, pois vários outros fatores externos ao Programa que ocorrem concomitantemente à intervenção podem interferir no impacto do Programa. Não controlar por esses fatores poderá atribuir erroneamente ao Programa um efeito que, em verdade, foi originado por outros fatores, o que afetará a validade interna da avaliação.

2.3.1 O Modelo de Resultados Potenciais

O referencial padrão para a discussão sobre avaliação de programas é conhecido como Modelo de Resultados Potenciais, ou Modelo de Roy-Ruben (ROY, 1951; RUBIN, 1974). Nesta seção será formalizada a abordagem de avaliação de programas e seguida a notação de Menezes-Filho (2012).

Quando se busca avaliar o impacto de um programa sobre uma variável de interesse Y , considera-se que um indivíduo, ou unidade de intervenção, possui dois resultados potenciais: $Y_i(1)$ caso ele tenha sido tratado e $Y_i(0)$ caso ele não tenha sido tratado. O termo “potencial” é utilizado justamente porque o indivíduo pode, apenas potencialmente, assumir ambos os resultados. A diferença entre os resultados potenciais $\beta_i = Y_i(1) - Y_i(0)$ fornece o impacto para o indivíduo i .

Especifica-se o modelo de geração de valores dos resultados potenciais por meio do par de equações lineares:

$$Y_i(1) = X_i' \alpha + \beta_i + \varepsilon_i \quad (1)$$

$$Y_i(0) = X_i' \alpha + \varepsilon_i \quad (2)$$

Em que X_i é o vetor de características observáveis (como gênero, cor, idade, escolaridade, educação dos pais etc) e ε_i é componente não observável pelo avaliador que é capaz de influenciar os resultados potenciais do indivíduo i .

Sendo T uma variável *dummy* que assume valor unitário caso o indivíduo seja tratado e zero caso contrário, tem-se que a equação de resultados potenciais é dada por:

$$Y_i = T_i Y_i(1) + (1 - T_i) Y_i(0) = Y_i(0) + T_i [Y_i(1) - Y_i(0)] \quad (3)$$

Pela manipulação de (1), (2) e (3), obtêm-se:

$$Y_i = X_i' \alpha + \beta_i T_i + \varepsilon_i \quad (4)$$

Caso seja válida a hipótese de que o impacto do Programa é o mesmo para todos os indivíduos, tem-se que:

$$Y_i = X_i' \alpha + \beta T_i + \varepsilon_i \quad (5)$$

A estimação do efeito do tratamento por meio da equação acima pode gerar resultados viesados, pois o impacto do Programa pode estar relacionado ao componente não observável. Em verdade, como a decisão de participar do Programa não é, em geral, aleatória, há grandes chances de que os componentes que determinam a decisão de participação no tratamento também influenciem o resultado de interesse.

Em avaliações de impacto, busca-se estimar, em geral, três principais efeitos: o efeito médio do tratamento (ATE), o efeito do tratamento sobre os tratados (ATT), e o efeito do tratamento sobre os não-tratados (ATU).

O ATT fornece o impacto médio do Programa para a parte da população que foi, de fato, tratada:

$$ATT = E[Y_i(1)|T_i = 1] - E[Y_i(0)|T_i = 1] \quad (6)$$

Em que $E[Y_i(1)|T_i = 1]$ é a média populacional da variável de resultado para os tratados e $E[Y_i(0)|T_i = 1]$ é a média contrafactual do resultado para os não tratados caso eles tivessem sido tratados.

Suponha que se utilizem indivíduos que não participaram do Programa como contrafactual para o grupo dos tratados. A diferença observada nos efeitos equivale a:

$$\begin{aligned} & E[Y_i(1)|T_i = 1] - E[Y_i(0)|T_i = 0] \\ &= E[Y_i(1)|T_i = 1] - E[Y_i(0)|T_i = 0] - \{E[Y_i(0)|T_i = 1] - E[Y_i(0)|T_i = 1]\} \\ &= E[Y_i(1)|T_i = 1] - E[Y_i(0)|T_i = 1] + E[Y_i(0)|T_i = 1] - E[Y_i(0)|T_i = 0] \\ &= ATT + Viés \end{aligned}$$

Em que $E[Y_i(1)|T_i = 0]$ é a média para os tratados caso eles não tivessem sido tratados e $E[Y_i(0)|T_i = 0]$ é média populacional do resultado para os não tratados. O estimador $E[Y_i(1)|T_i = 1] - E[Y_i(0)|T_i = 0]$ é denominado “estimador ingênuo”. O viés de seleção decorre do fato de que o processo de decisão sobre participação no tratamento não foi aleatória.

O uso da aleatorização do tratamento (*randomized control trials*) entre os indivíduos é considerado o padrão-ouro dos métodos de avaliação, pois sua aplicação elimina o viés de seleção. Ao se aleatorizar o tratamento, os indivíduos que irão formar o grupo de tratamento e de controle possuem, em média, características observáveis e não-observáveis similares. Devido à dificuldade e até mesmo impossibilidade de se conduzir um experimento aleatorizado, busca-se avaliar um programa por meio de desenhos quasi-experimentais, como o método do pareamento, os quais visam corrigir o viés de seleção e tornar comparáveis as unidades do grupo de tratamento e de controle.

2.3.2 O Método do Pareamento

O método do pareamento foi inicialmente proposto por Rosenbaum e Rubin (1983). Por meio dele é construído um grupo de controle de tamanho suficiente que possua covariadas com distribuição similar à do grupo de tratamento. Neste sentido, cada unidade do grupo de tratamento teria uma unidade correspondente no grupo de controle, que representa o seu contrafactual. Assim, a única diferença existente entre as unidades seria a participação ou não no Programa.

O método possui como principal hipótese a ignorabilidade (ou hipótese de seleção nos observáveis):

$$[Y(0), Y(1)] \perp T | X$$

Pela hipótese acima, tem-se que, condicional ao vetor de características individuais observáveis, a participação ou não do indivíduo no Programa é independente dos resultados potenciais, o que elimina o viés de seleção. Assume-se, portanto, que a seleção ao tratamento se dá apenas com base em características observáveis.

2.3.3 O Uso do Escore de Propensão

À medida que o número de covariadas do vetor X aumenta, amplia-se a dificuldade em se encontrar uma unidade no grupo de comparação que possa ser emparelhada à uma unidade no grupo de tratamento. Todavia, o pareamento, tal como demonstrado por Rosenbaum e Rubin (1983), também pode ser feito por meio de uma função das covariadas X – função *balancing score* – a qual resume toda a informação contida nesse vetor (MENEZES-FILHO, 2012). Por esse motivo, recorre-se ao escore de propensão, definido como a probabilidade de um indivíduo receber o tratamento dado suas características pré-tratamento observáveis:

$$P(X) = Pr(T_i = 1 | X)$$

O método do pareamento por escore de propensão consiste em combinar cada indivíduo tratado a outro indivíduo não tratado que possua características pré-tratamento semelhantes.

Pelo teorema do escore de propensão (ROSENBAUM e RUBIN,1983), se a hipótese de seleção nos observáveis é válida, a independência entre os resultados potenciais e o *status* de tratamento, condicional ao escore de propensão também é válida. Isto é:

$$Y_i(0), Y_i(1) \perp T_i | X_i \Rightarrow Y_i(0), Y_i(1) \perp T_i | P(X_i)$$

A hipótese de seleção nos observáveis acima garante que, condicional àquele escore de propensão, podemos assumir que a atribuição do Programa foi aleatória.

A segunda hipótese, denominada de suporte comum, faz-se necessária para garantir a comparabilidade entre os grupos de tratamento e controle (HECKMAN, LALONDE e SMITH, 1999). Formalmente:

$$0 < P(T_i = 1 | X_i) < 1$$

Obtém-se a região de suporte comum pela eliminação das unidades que estiverem muito abaixo ou muito acima da média dos escores. Essa hipótese assegura que existem indivíduos em ambos os grupos (tratamento e controle) para cada característica em X que se deseja comparar.

Se assumirmos que as duas hipóteses são válidas, o estimador ATT torna-se:

$$ATT = E_{P(x)|D=1}\{E[Y(1)|P(X), T = 1] - E[Y(0)|P(X), D = 0]\}$$

Denotando-se por i a unidade tratada e por j a unidade não tratada, sendo p_i e p_j os seus respectivos escores de propensão; seja T o conjunto de unidades tratadas, C o conjunto de unidades não tratadas, N_T o número de unidades do grupo de tratamento que estão na região de suporte comum e w_{ij} os pesos, que dependem da distância entre p_i e p_j . O estimador típico do pareamento (SMITH e TODD, 2005) por escore de propensão é dado por:

$$ATT^{PSM} = \frac{1}{N_T} \sum_{i \in T} \left(Y_i^T - \sum_{j \in C} w_{ij} Y_j^C \right)$$

2.3.4 Implementação do Método de Pareamento

Para proceder com o método do escore de propensão é preciso, primeiramente, escolher dentre os diversos algoritmos de pareamento presentes na literatura. Por meio deles se definirá a proximidade do escore de propensão dos indivíduos tratados em relação ao escore de

propensão dos indivíduos não-tratados (MENEZES-FILHO, 2012). Como a probabilidade de se observar duas unidades com o mesmo escore de propensão é, em princípio, nula, já que $P(X)$ é uma variável contínua, métodos foram propostos para superar essa questão (BECKER e ICHINO, 2002). Os principais algoritmos encontrados na literatura são: (i) estratificação ou pareamento estratificado; (ii) pareamento pelo vizinho mais próximo e (ii) pareamento radial, (ii) Kernel e (iii) *Linear Local Regression Matching*. As descrições e as notações utilizadas a seguir baseiam-se em Becker e Ichino (2002).

O algoritmo de pareamento estratificado divide o escopo da variação do escore de propensão em K blocos igualmente espaçados, para os quais se testa se há diferença estatisticamente significativa entre a média dos escores de propensão entre tratados e não tratados em cada intervalo. Caso essa média seja diferente, o intervalo é dividido ao meio e um novo teste é realizado. O procedimento é repetido até que os indivíduos tratados e não tratados tenham na média o mesmo escore de propensão para cada intervalo. A estimação final do ATT é dada pela média dos efeitos de tratamento em cada bloco, em que os pesos são dados pela fração de unidades tratadas existentes em cada bloco.

Uma das desvantagens do método de estratificação reside no fato de que são descartadas as observações tratadas que não encontram pares em seu bloco. Métodos que procuram, dentro de todo o grupo de controle, uma ou mais unidades para se efetuar o pareamento, surgem como alternativas ao algoritmo de estratificação desde que cumpram algumas condições que visam garantir a qualidade do pareamento.

No pareamento pelo vizinho mais próximo (*nearest neighbor matching*), busca-se uma unidade no grupo de comparação cujo escore de propensão esteja próximo do escore da unidade tratada, isto é, busca-se minimizar a distância entre os escores de propensão das unidades tratadas e não tratadas. Seja $C(i)$ o conjunto formado pelas unidades do grupo de controle pareadas à unidade i , as quais possuem um escore de propensão p_i , sendo o escore de propensão da unidade do grupo de controle. Assim, o pareamento é dado por

$$C(i) = \min ||p_i - p_j||,$$

o qual é, na maior parte das vezes, um conjunto unitário (BECKER e ICHINO, 2002). Pode-se escolher o número de vizinhos mais próximos que serão utilizados no pareamento, bem como optar por um *match* com ou sem reposição – opções que envolvem um *trade-off* entre viés e variância (CALIENDO e KOPEINIG, 2008).

No pareamento radial (*radius matching*), utilizam-se todas as unidades de controle com escore de propensão a uma distância inferior a r do escore da unidade tratada. Formalmente:

$$C(i) = \{p_j \mid \|p_i - p_j\| < r\}$$

Seja N_i^C o conjunto formado pelos controles pareados à observação i do grupo de tratamento, definem-se os pesos $w_{ij} = \frac{1}{N_i^C}$ se $j \in C(i)$ e $w_{ij} = 0$ caso contrário. Assim, o estimador de pareamento pelo método do vizinho mais próximo e radial pode ser escrito como:

$$ATT^{VMP/R} = \frac{1}{N^T} \sum_{i \in T} Y_i^T - \frac{1}{N^T} \sum_{j \in C} w_j Y_j^C$$

Em que os pesos w_j são definidos por $w_j = \sum_i w_{ij}$.

Os algoritmos discutidos até o momento utilizam de poucas observações do grupo de comparação para a construção do grupo de controle. Os estimadores não-paramétricos de *Kernel Matching* e *Local Linear Regression Matching* (LLM) diferem dos métodos anteriores por utilizarem as médias ponderadas de todos os indivíduos no grupo de controle para obtenção do resultado para os não tratados.

O estimador de pareamento por kernel é dado por:

$$ATT^K = \frac{1}{N^T} \sum_{i \in T} \left\{ Y_i^T - \frac{\sum_{j \in C} Y_j^C G\left(\frac{P_j - P_i}{h_n}\right)}{\sum_{k \in C} G\left(\frac{P_k - P_i}{h_n}\right)} \right\}$$

Onde $G(\cdot)$ é a função kernel e h_n é o parâmetro da janela (*bandwidth*). Se a função kernel e a janela estão sob condições padrão, tem-se que $w(i, j) = \frac{\sum_{j \in C} Y_j^C G\left(\frac{P_j - P_i}{h_n}\right)}{\sum_{k \in C} G\left(\frac{P_k - P_i}{h_n}\right)}$ é um estimador consistente do resultado contrafactual. Como apontado por Smith e Todd (2005), o pareamento por Kernel pode ser visto como uma regressão ponderada do resultado contrafactual sob um intercepto. Os pesos $w(i, j)$ são inversamente proporcionais à distância entre o indivíduo do grupo de controle e o indivíduo que está sendo pareado no grupo de tratamento para o qual se deseja estimar o contrafactual. O intercepto estimado fornece uma estimativa da média do contrafactual.

Para a estimação da função de densidade Kernel é necessário, portanto, a escolha da função kernel (K) e do parâmetro de janela. Na prática, a escolha da função Kernel é tida como importante (DINARDO e TOBIAS, 2001). A escolha da janela é mais relevante, pois envolve um *trade-off*: por um lado, uma janela maior gera funções de densidade mais suaves, o que leva a uma melhor estatística t e a uma menor variância. Por outro lado, a escolha de um parâmetro maior também aumenta o viés.

Heckman, Ichimura e Todd (1997) propuseram uma versão generalizada do kernel matching, denominada *local linear matching*, em que a função de ponderação é dada por:

$$W(i, j) = \frac{G_{ij} \sum_{k \in I_0} G_{ik} (p_k - p_i)^2 - [G_{ij} (p_j - p_i)] [\sum_{k \in I_0} G_{ik} (p_k - p_i)]}{\sum_{j \in I_0} G_{ij} \sum_{k \in I_0} G_{ij} (p_k - p_i)^2 - (\sum_{k \in I_0} G_{ij} (p_k - p_i))^2}$$

Pelo método LLM, estima-se uma regressão não paramétrica localmente ponderada do resultado do grupo de comparação na “vizinhança” de cada observação tratada (ICHIMURA, e TODD 1997). A diferença entre Kernel e *Local Linear Matching* decorre do fato de o último incluir um termo linear no intercepto, o que representa uma vantagem em relação ao Kernel quando as observações do grupo de comparação estão distribuídas de forma assimétrica em torno da observação tratada, *e.g.* nos *boundary points*, ou quando há lacunas na distribuição do escore de propensão. Entre as vantagens do *local linear matching* estão uma taxa de convergência mais rápida perto dos pontos de fronteira e maior robustez perante diferentes densidades dos dados (FAN, 1992).

A literatura aponta que os algoritmos de PSM baseados em kernel tendem a ser mais eficientes e consistentes, pois utilizam informação de toda a amostra (CALIENDO e KOPEINIG, 2008). A desvantagem deste método reside no fato de também utilizar pares ruins (*bad matches*). Logo, a correta imposição do suporte comum é de grande importância para ambos os estimadores em questão. Heckman, Ichimura e Todd (1997) derivam distribuição assintótica dos estimadores.

Para se estimar o efeito médio do tratamento sobre os tratados é necessário estimar o escore de propensão por meio de modelos de variável binária discreta, tais como *probit* ou *logit*, tendo como variável dependente binária a aprovação no Programa e como variáveis independentes o vetor de características pré-tratamento X.

Sabe-se que a qualidade do pareamento é sensível às variáveis observadas pré-tratamento utilizadas para estimar a probabilidade de aprovação no PAS, o que é averiguado por meio da verificação da condição de balanceamento (ROSENBAUM e RUBIN, 1983). Caso essa condição seja satisfeita, podemos afirmar que a distribuição conjunta das covariadas observáveis para tratamento e controle é semelhante.

Em relação à hipótese de suporte comum, a literatura indica algumas formas de se avaliar se tal condição é válida para determinada amostra. A forma mais direta é feita por meio da análise visual da densidade da distribuição do escore de propensão em ambos os grupos (LECHNER, 2002). Caliendo e Kopeinig (2008) apresentam alguns procedimentos que podem facilitar a averiguação do suporte comum. Pelo critério dos valores máximos e mínimos, compara-se a amplitude dos escores de propensão nos grupos, com a posterior eliminação das unidades cujos escores de propensão se situam fora da amplitude comum aos dois grupos. Apesar de este ser um procedimento apenas necessário para estimação do efeito médio do tratamento, para efeitos da estimação do ATT, ele garante que cada unidade tratada pode ser pareada a uma unidade não tratada.

A estimação do erro padrão é normalmente realizada por meio do procedimento de *bootstrapping* (LECHNER, 2002), técnica de reamostragem que consiste na replicação de várias amostras selecionadas. Ainda que haja controvérsia acerca da evidência que justifique o uso do procedimento (IMBENS, 2004), seu uso é bastante amplo na literatura de avaliação de impacto.

Existem alguns pontos atinentes à escolha das covariadas para o cômputo do escore de propensão que devem ser levados em consideração. Parte dos autores argumenta que apenas as variáveis que influenciam simultaneamente a decisão de participação no Programa e a variável de resultado devem ser incluídas no modelo. Ademais, a análise sobre quais variáveis incluir deve levar em consideração a teoria econômica, literatura específica e as especificidades do Programa (CALIENDO e KOPEINIG, 2008). Heckman, Ichimura e Todd (1997) mostram que a omissão de variáveis importantes pode levar a estimações enviesadas. Ainda, apenas variáveis que não são afetadas pela participação devem ser incluídas no modelo.

2.3.4 O Método de Pareamento e o Viés de Estimação

O viés surge quando os resultados médios do grupo de controle para os que não foram tratados diferem dos resultados médios do grupo de tratados caso não tivessem sido tratados, pois o grupo dos não tratados não é um bom contrafactual para o resultado potencial, não observado, dos tratados na ausência de tratamento.

De acordo com Heckman, Ichimura e Todd (1997), pode-se dividir o viés em três componentes. O primeiro surge pela falta de suporte comum, sendo esta a principal fonte de viés para os autores; o segundo decorre de viés nas características observáveis; o terceiro, denominado viés de seleção, decorre de diferenças nos não observáveis, o qual surge quando há variáveis não observadas que influenciam tanto os condicionais quanto as variáveis observáveis. A realização do pareamento na região de suporte comum elimina a primeira fonte de viés. O pareamento, feito de forma cuidadosa, é capaz de eliminar o segundo componente do viés. Os métodos de pareamento, como visto, apenas lidam com características observáveis, restando o problema de heterogeneidade latente que pode resultar em viés na estimação do efeito do Programa. Dessa forma, é importante ressaltar que o método do escore de propensão permite reduzir, mas não eliminar, o viés causado por não observáveis (RESENDE e OLIVEIRA, 2008).

2.4 Investigação Empírica

Após a descrição do método a ser implementado, parte-se para a apresentação da base de dados, com a descrição do questionário que originou a primeira base e as informações que foram retiradas da segunda. Em seguida, apresentam-se a definição do grupo de tratamento e controle, das covariadas escolhidas e a exposição das estatísticas descritivas. Por fim, são definidas as variáveis de interesse deste capítulo.

2.4.1 Base de Dados

Francis e Tannuri-Pianto (2012) conduziram pesquisa na UnB entre 2007 e 2008 com alunos que ingressaram na universidade entre 2/2003 e 2/2005, a qual será denominada pelo acrônimo PSEU. Os questionários foram respondidos em entrevistas ou pela internet, e os entrevistados foram escolhidos aleatoriamente a partir da população de alunos da UnB ingressantes nos anos supracitados. As entrevistas ocorreram no Departamento de Economia

da UnB e 2814 estudantes foram entrevistados. O questionário completo era composto por aproximadamente 200 questões, divididas em seis blocos de perguntas:

- I. *Demografia e História*: questões referentes à história do aluno e *background* familiar – nível de escolaridade dos pais, número de irmãos, se é solteiro etc – bem como questões referentes ao *status* socioeconômico.
- II. *Educação pré-universitária*: perguntas como o tipo de escola na qual estudou, perfil das notas no ensino básico e fundamental e quantidade de horas de estudo semanal.
- III. *Admissão universitária*: tempo dedicado ao preparo para exames de seleção, número de universidades para as quais se inscreve, número de aprovações em exames de seleção.
- IV. *Educação Universitária*: tempo dedicado ao estudo, perfil e evolução das notas, fonte de financiamento etc.
- V. *Emprego*: perfil do emprego – se é ou não remunerado, remuneração líquida, número de horas dedicadas ao trabalho.
- VI. *Futuro e Expectativas*: expectativas em relação à formatura, vida pessoal, estudos futuros, emprego, rendimento esperado, etc.
- VII. *Raça e Sociedade*: identidade racial do aluno e dos pais, descendência e perguntas referentes ao sistema de cotas – se aplicou, se era elegível etc.

A segunda base utilizada neste capítulo refere-se às informações sobre o perfil acadêmico dos alunos dentro da universidade. A base, doravante denominada de SAA, possui dados sobre o Índice de Rendimento Acadêmico (IRA). Esta base será utilizada para o cálculo do IRA médio do aluno ao longo do curso. Como a base foi extraída no segundo semestre de 2011, apenas há a informação sobre o IRA até este período. Como uma pequena parte da amostra ainda não havia se formado até este semestre, para esses alunos – 49 ao total – o IRA médio irá corresponder à média dos IRAs obtidos até o último período de informação disponível. As bases fornecidas pelo SAA também possuem dados relativos ao curso e opção¹⁰, semestre de ingresso e de saída, se o aluno ainda está ativo na universidade e forma de saída e histórico escolar dos alunos ativos.

¹⁰ Um curso pode ter mais de uma opção, a depender do número de turnos em que ele é ofertado (diurno e noturno), se o curso possui mais de uma habilitação (bacharelado e licenciatura) e se há mais de uma habilitação específica dentro de um curso, como é o caso Letras e Artes. Música, por exemplo, possui habilitação de bacharelado em Canto e em Clarineta, Composição, Contrabaixo, Fagote, Piano etc, além de licenciatura em Educação Artística e em Música. O curso de Letras possui habilitações de bacharelado em

A partir da informação sobre o período de ingresso e de saída, calcula-se o tempo de permanência do aluno na universidade. Os alunos que formaram durante o período do semestre de verão tiveram seu período de saída modificado para o semestre posterior, pois a existência dos semestres de versão comprometia a análise comparativa do tempo de permanência dos alunos na universidade.

Havia 12 alunos ainda ativos na universidade em 2013/1. A retirada desses alunos da amostra poderia gerar viés tendo em vista que uma das variáveis de interesse é o número de semestres cursados até a formatura. Por meio de planilha também fornecida pelo SAA, analisou-se o histórico dos alunos ativos, os quais puderam ser divididos em três grupos. O primeiro é composto pelos quatro estudantes que já haviam cumprido todos os pré-requisitos para a sua formatura, mas que cursavam disciplinas de outro curso/habilitação. Nesses casos, foi considerado como período de saída o último semestre necessário para se cumprir os requisitos para formatura no curso de ingresso na UnB. O segundo grupo é composto pelos 2 prováveis formandos em 2013/1, alunos que se formariam caso aprovados em todas as disciplinas em que estavam matriculados. Nestes casos, considerou-se 2013/1 como semestre de formatura. No terceiro grupo estão os seis alunos cujos requisitos para formatura não poderão ser cumpridos em 2013/1. Para estes estudantes não foi possível projetar a data de formatura, tendo os mesmos permanecidos na base como alunos ativos.

2.4.2 Grupo de Tratamento e de Controle

Como o objetivo deste capítulo é avaliar o efeito de ter ingressado na UnB pelo PAS sobre o desempenho acadêmico, taxa de evasão e tempo de permanência dos estudantes na universidade por meio da metodologia de avaliação de programas descrita na seção anterior, os alunos do PAS irão formar o grupo de tratamento, enquanto o grupo de não tratados, ou grupo controle, será formado a partir dos alunos que ingressaram na universidade pelo vestibular. O tratamento corresponde ao ingresso na UnB pelo PAS.

Para garantir a comparabilidade entre os grupos, apenas os alunos que ingressaram na universidade no primeiro semestre de 2004 e 2005 irão constar da amostra, pois apenas nesses semestres há o ingresso de alunos via PAS. Tal como no capítulo 1, os alunos que prestaram o

exame vestibular do segundo semestre não são bons controles por se tratarem de candidatos com características bastante distintas daqueles que o prestam no primeiro semestre. A Tabela 2.1 apresenta a distribuição dos alunos do grupo de controle e tratamento por semestre de ingresso. A amostra é composta, portanto, por 1051 alunos. Observa-se que o número de ingressantes em cada semestre é semelhante entre os alunos do PAS e do vestibular.

Tabela 2.1 – Distribuição dos alunos por processo seletivo e por semestre

	PAS	Vestibular	Total
1/2004	279	261	540
1/2005	234	277	511
<i>Total</i>	513	538	1051

Fonte: Pseu.

Sabe-se que os indivíduos que formam o grupo de controle não foram definidos aleatoriamente. Para verificar se esse grupo forma um bom controle para o grupo de tratamento, faz-se necessário testar se há diferenças significativas entre os grupos para um conjunto de variáveis observáveis.

2.4.3 Estatísticas Descritivas e Teste de Médias

A Tabela 2.2 fornece as estatísticas básicas em relação aos alunos do PAS e do vestibular que responderam ao questionário Pseu, bem como a diferença nas variáveis e o seu respectivo p-valor. Dividiram-se as variáveis em cinco grupos: características dos alunos, *background* familiar, características socioeconômicas, educação pré-universitária e raça e sociedade. Optou-se pela descrição da base previamente à escolha das variáveis que serão de fato utilizadas na avaliação por se acreditar que tais dados também possuem relevância para a literatura da área.

Das 68 características investigadas, há diferença estatisticamente significativa a 5% entre os grupos de controle e tratamento em 14 casos. Em particular, destacam-se as diferenças na proporção entre homens e mulheres na amostra de alunos do PAS, bem como na média de idade entre os grupos: como esperado, os alunos ingressantes pelo PAS são mais jovens, já que este processo seletivo se restringe a alunos que ainda estão no ensino médio. Quanto à educação dos pais, pode-se observar que os grupos diferem apenas em relação à proporção dos pais que possuem até 8 anos de estudo.

Tabela 2.2 - Estatísticas descritivas e teste de médias

Variáveis	PAS	Vestibular	Diferença	P-valor
Características dos alunos				
Masculino	40,90	51,80	-10,90	0,000
Idade	17,87	20,67	-2,80	0,000
<i>Religião</i>				
Católica	36,11	42,59	-6,48	0,220
Evangélica	11,67	10,49	1,18	0,730
Outra religião	17,77	17,90	-0,13	0,976
Nenhuma religião	34,44	29,01	5,43	0,282
Background familiar				
<i>Educação do Pai</i>				
Até 8 anos de estudo	8,90	13,60	-4,70	0,017
Ensino médio incompleto	4,35	4,80	-0,45	0,257
Ensino médio completo	25,00	25,60	-0,60	0,825
Ensino superior	61,74	56,00	5,74	0,061
<i>Educação da Mãe</i>				
Até 8 anos de estudo	5,41	12,89	-7,48	0,000
Ensino médio incompleto	4,10	3,32	0,78	0,503
Ensino médio completo	30,03	28,71	1,32	0,638
Ensino superior	60,45	55,07	5,38	0,079
Idade com que mãe teve primeiro filho	25,11	24,76	0,34	0,512
Morava com mãe e pai aos 14 anos	73,60	74,00	-0,40	0,920
<i>Número de irmãos</i>				
Nenhum	3,71	4,29	-0,58	0,637
Um a dois irmãos	68,59	69,59	-1,00	0,725
Três ou mais	27,69	26,12	1,57	0,565
Características Socioeconômicas				
Possui acesso à internet em casa	94,73	90,71	4,02	0,012
Pais possuem automóvel	93,75	92,46	1,29	0,421
Número de automóveis que pais possuem	2,16	2,04	0,12	0,307
Pais possuem geladeira duplex	70,42	73,31	-2,89	0,312
Casa possui mais que três banheiros	62,30	56,25	6,05	0,263
Pais contrataram empregada doméstica	63,20	55,78	7,42	0,014
Mora em residência própria	83,24	84,38	-1,14	0,777
Mora com os pais	82,77	74,69	8,08	0,067
<i>Cidade de residência</i>				
Brasília	53,16	52,43	0,73	0,814
Demais cidades satélites	35,50	28,84	6,66	0,021
Entorno do DF	2,60	3,70	-1,10	0,306
Estados	8,73	15,01	-6,28	0,002
Educação Pré-universitária				
Ensino fundamental em escola particular	64,80	62,96	1,84	0,724
<i>Média das notas durante ensino fundamental</i>				
Maioria entre de 9 e 10	53,93	46,58	7,35	0,177
Maioria entre 7 e 8	43,26	44,10	-0,84	0,876
Maioria entre 5 e 6	2,81	9,31	-6,50	0,011
Ensino médio em escola particular	77,57	60,91	16,66	0,000
<i>Média das notas durante ensino médio</i>				
Maioria entre de 9 e 10	26,17	22,24	3,93	0,139
Maioria entre 7 e 8	60,37	57,87	2,50	0,412
Maioria entre 5 e 6	12,90	18,50	-5,60	0,013
Teve notas abaixo de 5	0,56	1,38	-0,82	0,176
Horas de estudo semanais durante dois primeiros anos do ensino médio - de 1(muitas horas) a 5 (poucas horas)	2,89	3,12	-0,23	0,093

Variáveis	(conclusão)			
	PAS	Vestibular	Diferença	P-valor
Média de horas de estudo semanais durante o 3º ano - de 1(muitas horas) a 5 (poucas horas)	2,37	2,72	-0,35	0,012
<i>Disciplina favorita no ensino médio</i>				
Matemática	27,93	27,77	0,16	0,975
Ciências	34,64	30,86	3,78	0,459
História ou estudos sociais	23,46	28,39	-4,93	0,299
Português, Inglês ou literatura	10,61	9,87	0,74	0,823
Outra disciplina	3,35	3,09	0,26	0,890
Horas semanais se preparando para exames de seleção	34,46	34,59	-0,13	0,912
Fez cursinho pré-vestibular	31,65	73,71	-42,06	0,000
Número de universidades nas quais se inscreveu	1,53	2,23	-0,70	0,000
Tinha UnB como primeira opção de universidade	92,22	90,06	2,16	0,482
Raça e Sociedade				
<i>Identidade racial</i>				
Branca	53,21	54,15	-0,94	0,764
Parda	38,25	39,13	-0,88	0,773
Preta	5,11	3,70	1,41	0,290
Asiática	1,89	1,98	-0,09	0,923
Indígena	0,44	0,90	-0,46	0,447
Autodeclara-se negro	12,10	11,74	0,36	0,861
<i>Identidade racial do Pai</i>				
Branca	55,53	55,75	-0,22	0,944
Parda	33,77	31,74	2,03	0,488
Preta	7,25	8,13	-0,88	0,595
Asiática	2,10	2,18	-0,08	0,927
Indígena	1,33	2,18	-0,85	0,301
<i>Identidade Racial da Mãe</i>				
Branca	63,44	64,37	-0,93	0,7571
Parda	29,92	28,74	1,18	0,676
Preta	3,03	3,74	-0,71	0,053
Asiática	1,70	1,37	0,33	0,670
Indígena	1,89	1,77	0,12	0,883
Tentou sistema de cotas em 1/2005	12,35	10,81	1,54	0,759
<i>N</i>	<i>538</i>	<i>513</i>		

Fonte: Pseu. Nota: p-valores acima de 5% foram destacados para negrito. *Utiliza-se conceito amplo de Brasília que inclui Asa Norte, Asa Sul, Lago Sul, Lago Norte, Sudoeste, Estrutural e Varjão¹¹.

Em relação às características socioeconômicas avaliadas, o grupo de alunos do PAS possui maior acesso à internet e contam, em maior parcela, com empregada doméstica. Ainda, a porcentagem de alunos oriunda de outros estados é superior no vestibular, o que já era esperado tendo em vista que o vestibular é a forma de acesso mais tradicional também existente em vários estados e, portanto, mais acessível àqueles que não moram no Distrito Federal.

¹¹ Como o questionário que originou a variável “Cidades” possuía o mesmo código para as regiões administrativas Asa Norte, Asa Sul, Lago Norte, Lago Sul, Estrutural e Varjão, não foi possível retirar as duas últimas cidades deste grupo.

Em relação às características sobre educação pré-universitária, observa-se que uma maior proporção dos alunos do PAS estudou em escola privada durante o ensino médio, e afirmou ter passado mais horas durante o terceiro ano estudando para os exames de seleção, apesar de esta diferença na quantidade de horas não ser estatisticamente significativa quando se analisa a dedicação no primeiro e segundo ano do ensino médio. A diferença mais notória refere-se a parcela de alunos do grupo de controle que fez cursos pré-vestibulares: enquanto 73,71% dos ingressantes pelo vestibular afirmaram terem feito tal curso, essa proporção reduz-se para 31,65% entre os ingressantes pelo PAS, algo também já esperado tendo em vista que a decisão de entrar em um curso preparatório é mais comum entre os que já se formaram no ensino médio e que precisam, portanto, se prepararem mais intensivamente para o exame vestibular.

2.4.4 Covariadas

As variáveis pré-tratamento que serão empregadas para cálculo do escore de propensão foram escolhidas a partir da lista de variáveis originadas do questionário da Pseu. Utilizou-se, todavia, apenas as variáveis geradas a partir de perguntas respondidas por todos os alunos - como no Censo, parte das perguntas não foi feita a todos os indivíduos da amostra. As especificações finais foram obtidas, primeiramente, pela seleção das covariadas que influenciam a participação no Programa e a variável de interesse, tal como discutido na metodologia. A partir deste conjunto, testou-se uma série de especificações para checar se as mesmas satisfaziam a condição de balanceamento por meio do comando “pscore”. As covariadas utilizadas para o cálculo do escore de propensão estão listadas no Quadro 2.3.

Optou-se pela inclusão de variáveis de cada um dos grupos de características descritos, bem como pela exclusão de variáveis que apresentam informações similares como “pais possuem computador” e “pais possuem internet”. A renda é controlada por meio das seguintes variáveis: “carro”, “paiscompc” “duplex” e “empregada”. Utiliza-se a variável “ate19anos” para captar diferenças que resultam da restrição do PAS aos alunos que ainda estão no ensino médio. Como a variável “horas de estudo” inclui horas em sala de aula, 45 horas semanais de estudo implicam uma média de 25 horas em sala de aula (jornada-padrão nas escolas de ensino médio) e 20 horas de estudo fora da sala. Assim, a *dummy* “ate45horasestudosemanal” busca captar o efeito da quantidade de estudo além da uma jornada padrão de estudo em sala de aula.

Quadro 2.3 – Variáveis independentes para cálculo do escore de propensão

Variável	Descrição
Masculino	1 se homem, 0 caso contrário
ate19anos	1 se possui até 19 anos no ano em que ingressou na UnB, 0 caso contrário
Branca	1 se branco, 0 caso contrário
Brasilia	1 se reside em Brasília, 0 caso contrário
Paiensinosup	1 se pai possuía ensino superior, 0 caso contrário
Carro	1 se pais possuem carro, 0 caso contrário
Paiscomp	1 se possui computador em casa, 0 caso contrário
Duplex	1 se pais possuem geladeira duplex, 0 caso contrário
Empregada	1 se pais possuem empregada, 0 caso contrário
Escolaprivmedio	1 se estudou apenas em escola privada durante o ensino médio, 0 caso contrário
ate45hestudosemanal	1 se passou até 45 horas semanais se preparando para exames de seleção
Cursinhovest	1 se fez curso preparatório para o vestibular, 0 caso contrário
Inscricaoapenasunb	1 se apenas se inscreveu para a UnB, 0 caso contrário
Admitidoapenasunb	1 se foi admitido apenas na UnB, 0 caso contrário
sem20041	1 se ingressou na unb em 20041, 0 caso contrário

Fonte: Elaboração própria.

A variável “inscricaoapenasunb” busca controlar por alunos que objetivam apenas entrar na UnB e que por isso não prestaram outros exames. A variável “admissaoapenasunb” é importante, em particular, para a avaliação do impacto sobre a decisão de evasão. Entende-se que o aluno que é admitido apenas em uma universidade, após se controlar pelo número de universidades nas quais se inscreveu, tenderia a evadir menos tal como discutido em Sampaio et al (2011). Foi incluída a *dummy* “sem20041”, já que os grupos de controle e tratamento são formados por alunos ingressantes em dois períodos distintos. Apesar de a literatura apontar que as notas obtidas nos exames de seleção para a universidade representam bons preditores do desempenho dos alunos na universidade, optou-se pela não inclusão das mesmas entre as covariadas, pois o exame vestibular e PAS possuem estruturas distintas, com número de questões, divisão das matérias em questões, tipo de questões exigidas e escala de notas diferentes, o que prejudica a comparabilidade das notas.

2.4.5 Variáveis de Interesse

O objetivo do uso da metodologia de escore de propensão neste trabalho é o de permitir uma comparação adequada entre os alunos do PAS e do vestibular quanto às variáveis de

interesse, que por sua vez visam compreender várias esferas do desempenho do aluno dentro da universidade.

A primeira variável de interesse a ser avaliada, como descrito anteriormente, refere-se ao desempenho acadêmico dos alunos. O desempenho pode ser medido por meio da construção de indicadores utilizando-se as menções obtidas pelos alunos nas disciplinas cursadas. A UnB, bem como várias universidades federais, utiliza o Índice de Rendimento Acadêmico (IRA) como principal indicador de desempenho dos alunos ao longo do curso, o qual é critério para matrícula em disciplinas, concessão de bolsas e seleção para programas de intercâmbio. Para o cômputo do IRA, que varia de 0 a 5, são consideradas as menções obtidas e o número de créditos das disciplinas cursadas pelo aluno, bem como o trancamento de matérias optativas e obrigatórias. A fórmula do IRA utiliza o número de disciplinas obrigatórias trancadas (DTb); o número de disciplinas optativas trancadas (DTp) e o número de disciplinas matriculadas (DC). As menções recebem os seguintes pesos Pi: SS=5, MS=4, MM=3, MI=2, II=1 e SR=0. CRi refere-se ao número de créditos da disciplina e Pei ao período em que uma dada disciplina foi cursada, obedecendo a seguinte limitação: $\min | 6, \text{período} |$. A fórmula do IRA é dada, então, por:

$$IRA = \frac{[1 - (0,6 * DTb + 0,4 * DTp)]}{DC} * \frac{\sum Pi * CRi * Pei}{\sum CRi * Pei}$$

Como o IRA é um índice semestral, utiliza-se como variável de interesse a média dos IRAs obtidos ao longo da graduação. A segunda limitação decorre do fato de a base SAA possuir informações apenas até o segundo semestre de 2011. Portanto, para parte da amostra que não havia se formado até este período, não foi possível o cálculo do IRA semestral de todo o curso. Até o primeiro semestre de 2013, 46 alunos ainda não haviam se formado. Entende-se que o IRA é um indicador que sintetiza diversas informações e que, por isso, possui algumas limitações. Alunos que possuem perfil distinto de menções podem ter IRA similar já que outras informações não relacionadas ao desempenho acadêmico medido em notas também são utilizadas no cálculo do IRA. Esse indicador é, contudo, o principal indicador de desempenho da universidade e por isso optou-se por sua utilização como variável de interesse.

Para avaliar o tempo de permanência dos alunos na universidade, criou-se a variável “número de semestres além do recomendado” resultante da subtração da variável referente à

quantidade de semestres que a universidade recomenda para que o aluno termine o seu curso¹² e o número de semestres que o aluno precisou permanecer na universidade até a sua formatura¹³.

O terceiro resultado de interesse refere-se à evasão dos cursos, variável binária gerada a partir da variável “Forma de Saída”. A variável evasão agrega todas as modalidades de desligamento bem como a saída por reprovação repetida na mesma disciplina. A Tabela 2.4 apresenta a distribuição dos alunos por forma de saída e sistema.

Tabela 2.4 – Distribuição dos alunos por sistema de seleção e tipo de saída

Tipo de Saída	Vestibular	PAS	Total
Aluno ativo	0,39	0,74	0,57
Desligamento - não cumpriu condição	7,06	7,81	7,44
Desligamento – abandono	6,08	6,13	6,11
Desligamento Voluntário	2,55	2,42	2,48
Formatura	82,16	81,6	81,87
Novo vestibular	0,59	0,74	0,67
Reprovou três vezes na mesma disciplina	1,18	0,56	0,86

Fonte: SAA.

Observa-se que o não cumprimento de condição é a principal causa da saída dos alunos que não se formam. Essa modalidade de desligamento é aplicada ao aluno que, tendo sido anteriormente identificado como provável desligado por rendimento acadêmico¹⁴ ou por tempo de permanência, não tenha cumprido no decorrer do período fixado a condição que lhe foi imposta pelos órgãos colegiados (UNB, 2004). O desligamento por abandono de curso aplica-se ao estudante que por dois períodos letivos consecutivos não tenha efetivado matrícula em disciplinas ou que, embora matriculado, não tenha cursado disciplina. Por fim, o

¹² Para concluir o curso, o estudante deve ser aprovado nas disciplinas obrigatórias do curso e integralizar o número de créditos exigidos. O número de semestres recomendados origina-se de um fluxo de disciplinas definidas para cada semestre fixado por cada departamento. Caso o aluno siga o fluxo, garante-se que as disciplinas obrigatórias serão cursadas no número de semestres recomendados. A informação sobre os requisitos para a formatura encontram-se no *website* da Secretária de Administração Acadêmica e o número de semestres recomendados por curso/opção no *website* de matrícula da UnB.

¹³ Não foi possível subtrair do tempo de permanência do aluno na UnB o número de semestres trancados, pois esta informação estava disponível apenas até o primeiro semestre de 2011.

¹⁴ Pela Resolução CEPE/UnB n.41/2004 (UNB, 2004), a identificação do aluno como provável desligamento é automática quando o estudante: (a) obtiver média de disciplinas cursadas com aproveitamento, por período letivo, inferior a duas; (b) for reprovado, numa mesma disciplina obrigatória de seu curso, duas vezes, consecutivas ou não; (c) obtiver média de créditos por período letivo inferior ao mínimo exigido para seu curso; (d) apresentar número de créditos a obter incompatível com o tempo disponível para conclusão do curso.

desligamento voluntário é aplicado ao aluno que por iniciativa própria tenha desistido do vínculo com seu curso.

2.5 Resultados

Primeiramente, apresenta-se o modelo de variável binária estimado. Em seguida, descrevem-se os procedimentos para o pareamento, passando-se posteriormente para a análise dos efeitos do tratamento para as variáveis de interesse. Esta análise é feita em relação a três aspectos: a magnitude do valor estimado para o ATT; a direção do efeito; e sua significância estatística. Apresenta-se, ainda, uma comparação entre os algoritmos utilizados.

2.5.1 Estimação do Modelo *Logit*

As probabilidades de participação no Programa foram preditas a partir de um modelo *logit*. A Tabela 2.5 apresenta o resultado para o modelo estimado.

Tabela 2.5 – Resultados do Modelo *Logit*

Variáveis	Coefficiente	Desvio Padrão	Estatística t	
Masculino	-0,6485	0,1688	-3,84	**
Idade até 19 anos	3,7038	0,4944	7,49	**
Cor branca	-0,4098	0,1707	-2,4	**
Pai com ensino superior	-0,1592	0,1986	-0,8	
Pais possuem computador em casa	0,5794	0,4755	1,22	
Pais possuem carro	-0,4895	0,4371	-1,12	
Pais possuem geladeira duplex	-0,4991	0,1987	-2,51	**
Pais contrataram empregada doméstica	0,2547	0,1930	1,32	
Cursou todo o ensino médio em escola particular	0,0396	0,2170	0,18	
Dedicou até 45 horas semanais para estudo preparatório	0,2439	0,2060	1,18	
Apenas fez inscrição para processos seletivos da UnB	0,7286	0,2790	2,61	**
Apenas foi admitido na UnB	0,6428	0,2861	2,25	**
Reside em Brasília	0,0659	0,1846	0,36	
Fez curso preparatório para exames de seleção	-1,8163	0,1783	-10,19	**
<i>Dummy</i> para semestre de ingresso na UnB em 20041	-0,3017	0,1664	-1,81	*
Constante	-2,4677	0,6777	-3,64	**
<i>Pseudo R</i> ²	0,3070			
<i>N</i>	932			

Fonte: Elaboração própria. Nota: * significante a 10%. ** significante a 5%.

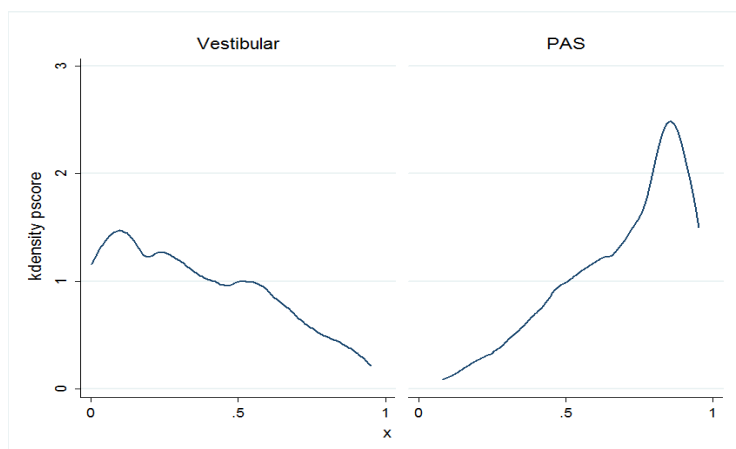
A análise dos resultados da estimação revela que algumas variáveis não são estatisticamente significativas. Segundo Rubin e Thomas (1996) apud Shadish, Cook e Campbell et al. (2002, p. 162), mesmo que algumas covariadas não sejam estatisticamente significativas, aconselha-se que elas sejam incluídas no modelo de escore de propensão.

Assim sendo, uma covariada apenas deve ser excluída caso haja consenso de que a mesma não está relacionada ao resultado de interesse. Em face do que foi exposto acerca da relevância das covariadas para modelar a decisão de participação no Programa, optou-se pela não exclusão das variáveis não significativas deste modelo.

2.5.2 Estimação do Efeito do PAS

Para a análise dos efeitos do ingresso pelo PAS sobre as variáveis de rendimento acadêmico, número de semestres além do recomendado e taxa de evasão aplicou-se primeiramente o método de pareamento kernel, que está entre os mais utilizados pela literatura. Para avaliar a robustez dos resultados, comparam-se os efeitos médios do tratamento sobre os tratados, por meio de dois outros métodos, utilizando-se o mesmo conjunto de covariadas para o cômputo do escore de propensão. Se os métodos resultarem em estimativas similares, pode-se concluir que os resultados encontrados são consistentes (KHANDKER, KOOLWAL e SAMAD, 2010). O uso do *Local Linear Matching* foi escolhido por ser um método indicado quando as observações do grupo de controle estão distribuídas assimetricamente em relação às observações tratadas. A distribuição dos escores de propensão é apresentada a seguir. Como indicado por Nichols (2007), utiliza-se a função de densidade kernel para obtenção do Gráfico 2.6. Verifica-se, visualmente, tanto a assimetria das distribuições quanto o fato das mesmas não ocorrerem em torno dos mesmos valores de escore de propensão.

Gráfico 2.6 – Densidade kernel da probabilidade predita para vestibular e PAS



Fonte: Elaboração própria.

Khandker, Koolwal e Samad (2010) apresentam, ainda, outra forma para checar a robustez, que consiste em utilizar o pareamento pelo vizinho mais próximo. Caso os resultados encontrados sejam similares, assume-se que os mesmos são mais confiáveis. Realizou-se, assim, o pareamento pelo método em destaque utilizando-se duas especificações distintas em relação ao *caliper* - à saber, 0,01 e 0,005, para avaliar a consistência dos resultados.

Os métodos kernel e LLR foram aplicados utilizando-se duas janelas - 0,01 e 0,06. Ambos geraram resultados semelhantes e por isso optou-se pela apresentação de apenas uma das especificações. Os erros-padrão foram estimados pela técnica de *bootstrapping* com 1.000 repetições.

Para o indicador de desempenho acadêmico IRA, observa-se que os resultados, mostrados na Tabela 2.6, apresentam o mesmo padrão: o efeito do ingresso pelo PAS sobre a variável é positivo, de magnitude que varia entre 0,1052 a 0,1765, a depender do método adotado. Nota-se também que o número de observações analisadas no pareamento pelo método do vizinho mais próximo é inferior ao dos métodos não paramétricos. Todos os algoritmos geraram resultados estatisticamente significativos, o que significa que, em média, os estudantes que ingressaram pelo PAS possuem um índice de rendimento de 2,8% a 4,7% superior ao rendimento dos que ingressaram na Universidade de Brasília pelo PAS. Tendo em vista que o IRA é um índice bastante denso e que sintetiza diversas informações, um impacto de 0,1 a 0,18 pode ser considerado de magnitude relevante.

Tabela 2.6 – Efeito do Tratamento sobre o IRA

Índice de Rendimento Acadêmico	Kernel (0,01)	Local Linear Regression (0,01)	Vizinho mais próximo (0,005)
Tratamento	3,9280	3,9281	3,9209
Controle	3,8228	3,8096	3,7444
ATT	0,1052	0,1185	0,1765
T-stat	1,74*	2,77**	1,79*
Obs. (Trat./Contr.)	477/446	477/446	419/446

Fonte: Elaboração própria. Nota **Estatisticamente significante a 5%. * Estatisticamente significante a 10%.

Os efeitos do Programa sobre o tempo de permanência na universidade são apresentados na Tabela 2.7. Ainda que haja uma diferença negativa em relação ao número médio de semestres cursado por cada grupo de estudantes, pela qual se verifica que os alunos

ingressantes pelo PAS tendem a permanecer por menos semestres na universidade, essa diferença não foi estatisticamente significativa em nenhuma das estimações realizadas.

Tabela 2.7 – Efeito do Tratamento sobre o número de semestres cursados além do recomendado

Número de semestres cursados além do recomendado	Kernel (0,01)	Local Linear Regression (0,01)	Vizinho mais próximo (0,005)
Tratamento	1,2490	1,2494	1,2358
Controle	1,2750	1,2779	1,2955
ATT	-0,0260	-0,0588	-0,0597
T-stat	-0,08	-0,43	-0,19
Obs. (Trat./Contr.)	385/365	385/365	332/365

Fonte: Elaboração própria. Nota: **Estatisticamente significativa a 5%. * Estatisticamente significativa a 10%.

Adicionalmente, o impacto da decisão de entrada na UnB pelo Programa sobre outras variáveis de interesse também relacionadas ao tempo de permanência foi avaliado usando a metodologia de escore de propensão descrita. Primeiramente, estimou-se o impacto para a variável binária “trancamento” e para a variável discreta “número de períodos trancados”. Por meio desta análise, buscou-se identificar qual grupo de alunos utilizava com maior frequência o recurso da suspensão das atividades acadêmicas sem a perda do vínculo com a universidade. Como a informação sobre trancamento apenas está disponível até o primeiro semestre de 2011, a análise se restringiu à amostra para os alunos que haviam saído da universidade neste semestre. O método de Kernel revelou diferença de 4,9% nas taxas de trancamento: enquanto 18,9% do alunos do PAS já haviam trancado algum semestre entre 2004/1 e 2011/1, 23,8% dos alunos ingressantes pelo vestibular já haviam recorrido ao trancamento. Os resultados, contudo, não apresentaram significância estatística para nenhum dos algoritmos de pareamento. Em relação à quantidade de semestres trancados, os alunos do PAS trancaram, em média, 1,32 semestres. Os alunos do vestibular utilizaram o trancamento, em média, 1,47 vezes. A diferença estimada entre tratados e controles por este mesmo método foi de 0,15, a qual também não teve significância estatística. A estimacão do ATT pelo método de LLR e Vizinho Mais Próximo para as duas variáveis apresentaram resultados levemente distintos, porém os efeitos possuíam o mesmo sinal dos efeitos estimados por kernel.

Uma nova variável para avaliar permanência denominada “permanência relativa”, dada pela divisão do número de semestres cursados pelo número de semestres recomendados, foi criada para melhor compreender o impacto da permanência estendida na universidade de

acordo com a duração do curso. Pelos métodos de Kernel, LLR e Vizinho mais Próximo também não foram encontrados diferenças estatisticamente significativas para essa variável. Conclui-se que, portanto, não há efeitos significativos do ingresso pelo PAS no tempo de permanência na universidade quando esses alunos são comparados aos alunos do vestibular.

Por fim, analisam-se os efeitos do tratamento sobre a taxa de evasão dos alunos. A Tabela 2.8 apresenta os resultados encontrados. Os resultados mostram que o efeito do ingresso pelo PAS sobre a taxa de evasão não é estatisticamente significativo em nenhum dos métodos adotados, sendo a taxa de abandono dos tratados muito próxima a dos não tratados para todos os métodos de pareamento aplicados.

Tabela 2.8 – Efeito do Tratamento para a variável Taxa de Evasão

Taxa de Evasão	Kernel (0,01)	Local Linear Regression (0,01)	Vizinho mais próximo (0,0005)
Tratamento	0,1614	0,1614	0,1623
Controle	0,1601	0,1371	0,1523
ATT	0,0013	0,0243	0,0095
T-stat	0,03	0,05	0,23
Obs. (Trat./Contr.)	446/486	446/486	419/446

Fonte: Elaboração própria. Nota **Estatisticamente significativa a 5%. * Estatisticamente significativa a 10%.

O método de pareamento por *Local Linear Regression* apresentou o maior efeito, o que indica que a taxa de evasão dos alunos do PAS é superior à taxa de evasão dos alunos do PAS apesar de ele não ser estatisticamente significativo. Não se pode dizer, portanto, que ter ingressado na Universidade de Brasília pelo PAS possui impactos sobre a proporção de alunos que decidem não concluir a graduação.

2.6 Conclusão

O Programa de Avaliação Seriada da Universidade de Brasília enquadra-se na categoria de processos seletivos que buscam amenizar a passagem do estudante de ensino médio pelo vestibular, por meio de provas aplicadas ao final de cada ano do ensino médio. O exame tem como característica o caráter interdisciplinar, com conteúdos demandados de forma contextualizada, os quais são definidos em fóruns dos quais participam docentes do ensino médio e universitário, o que contribui para adequar o que é ensinado em sala de aula às habilidades e capacidades demandadas do aluno dentro da universidade. A necessidade de

avaliação desse tipo de programa visa compreender em que medida essa diferença no exame impacta o grupo de alunos que entra na universidade. A existência da forma tradicional de seleção, o vestibular, gera um grupo de alunos que é, em geral, utilizado nas comparações acerca do desempenho dos alunos do PAS, comparações estas nem sempre feitas da melhor forma, tendo em vista que os alunos que entram por cada processo seletivo possuem diferenças em várias características, como origem socioeconômica, grau de escolaridade dos pais, bem como em diversas outras características.

Este capítulo buscou avaliar os efeitos do PAS sobre três variáveis: desempenho acadêmico, permanência na universidade e evasão. Adotou-se o Método de Pareamento por Escore de Propensão para medir os efeitos do Programa sobre as variáveis de interesse. O Método consiste na seleção de um grupo de comparação adequado, dentre o grupo de alunos do vestibular, para tentar reproduzir o resultado contrafactual, tendo em vista a impossibilidade de se averiguar o resultado do Programa sobre as variáveis de interesse para os alunos do PAS caso tais alunos não tivessem sido selecionados pelo Programa. Caso as hipóteses do modelo de escore de propensão sejam válidas, assegura-se que o resultado médio do grupo de comparação pode ser utilizado como contrafactual do resultado para o grupo de tratamento formado pelos alunos do PAS. Dentre os métodos de pareamento existentes, foram utilizados os algoritmos de kernel, bem como *local linear regression* e Vizinheiro Mais Próximo, optando-se pela apresentação dos três métodos para avaliar a robustez das estimações. A análise dos resultados estimados revelou que o PAS possui impacto positivo sobre o desempenho acadêmico, medido por meio do índice de rendimento acadêmico (IRA). Conclui-se, portanto, que os alunos ingressantes pelo PAS possuem, em média, rendimento acadêmico superior aos alunos que ingressaram pelo vestibular. Não foram encontrados resultados estatisticamente significativas sobre a permanência na universidade e sobre a taxa de evasão dos cursos.

REFERÊNCIAS BIBLIOGRÁFICAS

- BAI, C.; CHI, W. Determinants of Undergraduate GPAs in China: College Entrance Examination Scores, High School Achievement, and Admission Route. *MPRA*, n. 31240, 2011. Disponível em: <http://mpra.ub.uni-muenchen.de/31240/1/MPRA_paper_31240.pdf>. Acesso em: 10 jan. 2013.
- BECKER, S. O.; ICHINO, A. Estimation of average treatment effects based on propensity score. *Stata Journal*, v. 2, n. 4, p. 358-377, 2002.
- BETTS, J. R.; MORELL, D. The Determinants of Undergraduate Grade Point Average: The Relative Importance of Family Background, High School Resources, and Peer Group Effects. *The Journal of Human Resources*, v. 34, n. 2, p. 268-293, 1999.
- CALIENDO, M.; KOPEINIG, S. Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*, v. 22, n. 1, p. 31-72, 2008.
- CAMERON, A., TRIVEDI, P. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press, 2005.
- CENTRO DE SELEÇÃO E PROMOÇÃO DE EVENTOS (CESPE). O que é o PAS – Princípios Orientadores. Disponível em: <http://www.cespe.unb.br/pas/PAS_oque_principios.aspx> Acesso em: 1 fev. 2012.
-
- Edital nº 1 – PAS /
- UnB – Subprograma 2012, de 20 de setembro de 2012. Disponível em: <http://www.cespe.unb.br/pas/arquivos%5CEd_1_2012_PAS_1_etapa.pdf> Acesso em: 20 mar. 2013.
- CUNHA, A. M.; TUNES, E.; SILVA, R. R. Evasão do curso de química da Universidade de Brasília: a interpretação do aluno evadido. *Química Nova*, v. 24, n. 2, p. 262-280, 2001.
- DAYIGLU, M.; TURUT-ASIK, S. Gender Differences in Academic Performance in a Large Public University in Turkey. *ERC Working Papers in Economics*, 2004. Disponível em: <<http://www.erc.metu.edu/menu/series04/0417.pdf>>. Acesso em: 30 jan. 2013.
- DINARDO, J.; TOBIAS, J. Nonparametric Density and Regression Estimation. *Journal of Economic Perspectives*, v. 15, n. 4, p. 11-28, 2001.
- EMILIO, D. R.; JUNIOR, W. B.; ALVES, D. C. O. Uma análise econométrica dos determinantes do acesso à Universidade de São Paulo. *Pesquisa Planejamento Econômico*, v. 34, n. 2, 2004.
- FAN, J. Design adaptive nonparametric regression. *Journal of the American Statistical Association*, v. 87, n. 420, p. 998-1004, 1992.
- FERREIRA, V. C. *O programa de avaliação seriada da UnB – PAS: Sua influência sobre as práticas avaliativas de uma escola de ensino médio no DF*. 2002. 145 f. Dissertação (Mestrado em Educação) – Faculdade de Educação, Universidade de Brasília, Distrito Federal.

FRANCIS, A. M.; TANNURI-PIANTO, M.. The Redistributive Equity of Affirmative Action: Exploring the Role of Race, Socioeconomic Status, and Gender in College Admissions. *Economics of Education Review*, v. 31, n. 1, p. 45-55, 2012.

FREITAS, L. P. *Os fatores condicionantes e determinantes do êxito no vestibular — o caso da Fundação Universidade Federal do Piauí em 1978*. 1979. 251 f. Dissertação (Mestrado em Educação) – Departamento de Educação, Pontifícia Universidade Católica, Rio de Janeiro.

GOTTSCHALL, L. A. *A influência do programa de avaliação seriada-PAS/UNB no ensino médio: o caso da matemática*. 2008. 109 f. Dissertação (Mestrado em Educação) – Programa de Pós-Graduação em Educação, Universidade Católica de Brasília, Distrito Federal.

GUIMARÃES, D. B.; ARRAES, R. A. Status Sócio-Econômico, Background Familiar, Formação Educacional e as Chances de Sucesso dos Candidatos ao Vestibular da UFC. *Revista Contemporânea de Economia e Gestão*, v. 8, n. 2, p. 81-94, 2010.

GUIMARÃES, J.; SAMPAIO, B. *The Influence of Family Background and Individual Characteristics on Entrance Tests Scores of Brazilian University Students*. In: ENCONTRO REGIONAL DE ECONOMIA, 12., 2007, Fortaleza. Disponível em: <http://www.anpec.org.br/encontro2007/artigos/A07A092.pdf>. Acesso em: 14 mar. 2013.

GUIMARÃES, J.; SAMPAIO, B.; SAMPAIO, Y. What is behind university dropout decision in Brazil? A Bivariate Probability Model. *The Empirical Economics Letters*, v. 9, n. 6, p. 601-608, 2010.

HECKMAN, J.; ICHIMURA, H.; TODD, P. Matching as an Econometric Evaluation Estimator. *Review of Economic Studies*, v. 65, n. 2, p. 261-294, 1997.

HECKMAN, J.; LALONDE, R.; SMITH, J. The economics and econometrics of active labor market programs. In: ASHENFELTER, O., CARD, D. (Eds.) *The Handbook of Labor Economics*. Amsterdam: North Holland, v. 3A, part. 6, cap. 31, p. 1865-2097, 1999.

HONORATO, L.; WING, N. H. Determinantes socioeconômicos na probabilidade de aprovação no exame vestibular: uma análise entre os campi da Universidade Federal de Pernambuco. *Planejamento e Políticas Públicas*, IPEA, n. 37, p. 97-131, 2011.

HOSMER, D. W.; LEMESHOW, S. A goodness-of-fit test for the multiple logistic regression model. *Communications in Statistics*, A10, p.1043-1069, 1980.

HOSMER, D. W.; LEMESHOW, S. *Applied logistic regression*. 2nd. ed. New York: John Wiley & Sons, 2000.

IMBENS, G. W. Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review. *The Review of Economics and Statistics*, v. 86, n. 1, p. 4-29. 2004.

KHANDKER, S. R.; KOOLWAL, G. B.; SAMAD, H. A. *Handbook on Impact Evaluation: Quantitative Methods and Practices*. The International Bank for Reconstruction and Development, 2010. Disponível em:

<<https://openknowledge.worldbank.org/bitstream/handle/10986/2693/520990PUB0EPI11010fficial0Use0Only1.pdf?sequence=1>> Acesso em: 12 jan. 2013.

LECHNER, M. Some Practical Issues in The Evaluation of Heterogeneous Labour Market Programmes by Matching Methods. *Journal of the Royal Statistical Society - Series A*, v. 165, n. 1, p. 59-82, 2002.

LIMA, C. F. Rendimento Acadêmico, o que Prediz (e o que não Prediz): O Caso dos Alunos de Ciências Econômicas da UnB. *MPRA*, n. 36131, 2011. Disponível em: <http://mpra.ub.uni-muenchen.de/36131/1/MPRA_paper_36131.pdf>. Acesso em: 12 jan. 2013.

JUNIOR, L. P.; SILVEIRA, F. L.; OSTERMANN, F. Análise de sobrevivência aplicada ao estudo do fluxo escolar nos cursos de graduação em física: um exemplo de uma universidade brasileira. *Revista Brasileira de Ensino de Física*, v. 34, n. 1, 2012. Disponível em: <<http://www.sbfisica.org.br/rbef/pdf/341403.pdf>>. Acesso em: 10 mar. 2013.

MACIEL, H. M.; LOPES, M. I. P. *Avaliação Comparativa do Desempenho dos Ingressos Via PEIES e Vestibular no Curso de Engenharia Civil da UFSM*. In: CONGRESSO BRASILEIRO DE ENSINO DE ENGENHARIA, 29., 2001, Porto Alegre. Disponível em: <<http://www.abenge.org.br/CobengeAnteriores/2001/trabalhos/APP031.pdf>>. Acesso em: 09 jan. 2013.

MANZANO, M. E. *Vestibular Seriado: estado da arte e a percepção docente sobre o tema*. 2011. 75 p. Dissertação (Mestrado em Ensino de Ciências) – Faculdade de Educação, Universidade de São Paulo, São Paulo.

MATOS, B. P. *Programa de avaliação seriada (PAS): balanço de uma década*. Brasília: Universidade de Brasília, mai. 2006. (Relatório de pesquisa)

MENDES JUNIOR, 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*. 2013. 94 p. Dissertação (Mestrado em Economia), Faculdade de Economia, Universidade Federal Fluminense, Rio de Janeiro.

MENEZES-FILHO, N. (Org.) *Avaliação econômica de projetos sociais*. 1. ed. São Paulo: Dinâmica Gráfica e Editora, 2012. 186 p.

MISSIO, L. “Futuro, Qual é o seu?” *Um estudo sobre a relação ensino médio x PEIES a partir da fala docente em Santa Maria/RS*. 2007. 116 p. Dissertação (Mestrado em Educação) – Faculdade de Educação, Universidade Federal de Santa Maria, Rio Grande do Sul.

MURTAUGH, P. A.; BURNS, L. D.; SCHUSTER, J. Predicting the retention of university students. *Research in Higher Education*, v. 40, n. 3, p. 355-371, 1999.

NICHOLS, A. Causal inference with observational data. *Stata Journal*, v. 7, n. 4, p. 507-541, 2007. Disponível em: <<http://www.stata-journal.com/sjpdf.html?articlenum=st0136>>. Acesso em: 19 jan. 2013.

- OLIVEIRA, J. H. *Determinantes do sucesso no vestibular 79 da Universidade Federal de Viçosa – Minas Gerais*. 1980. 196 f. Dissertação (Mestrado em Educação) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro.
- OLIVEIRA, I. S. V.; SILVA, M. V. B.; SIQUEIRA, L. B. O. Determinantes do desempenho dos estudantes no vestibular da Universidade Federal da Paraíba. *Economia e Desenvolvimento*, Recife (PE), v. 7, n. 2, p. 286-329.2008. Disponível em <<http://periodicos.ufpb.br/ojs/index.php/economia/article/view/4865/3670>>. Acesso em: 4 jun. 2013
- QUEIROZ, D. M. *Desigualdade no Ensino Superior: Cor, Status e Desempenho*. In: REUNIÃO ANUAL DA ANPED, 27., 2008. Novo Governo. Novas Políticas? Caxambu, 2003. Disponível em: <<http://www.anped.org.br/reunioes/26/trabalhos/delcelemascarenhasqueiroz.pdf>> Acesso em: 17 fev. 2013.
- REIS, S. M. de A. S. *Avaliação comparativa do desempenho de alunos admitidos na Faculdade de Odontologia da Universidade Federal de Uberlândia (FOUFU) via Vestibular e via PAIES, no período de 2000 a 2003*. 2006. 209 p. Dissertação (Mestrado em Odontologia) – Faculdade de Odontologia, Universidade Federal de Uberlândia, Minas Gerais.
- RESENDE, A.C. OLIVEIRA, A.M. Avaliando resultados de um programa de transferência de renda: o impacto do Bolsa-Escola sobre os gastos das famílias brasileiras. *Estudos Econômicos*. 2008, vol.38, n.2, pp. 235-265.
- RODRIGUES, D. G.; PELISOLI, C. Ansiedade em vestibulandos: um estudo exploratório. *Revista de Psiquiatria Clínica*, v. 35, n. 5, p. 171-177, 2008. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-60832008000500001&lng=pt&nrm=iso>. Acesso em: 10 jul. 2013
- ROSENBAUM, P.; RUBIN, D. B. The Central role of the propensity score in observational studies for causal effects. *Biometrika*, v. 70, n. 1, p. 41-55, 1983.
- ROY, A. Some thoughts on the distribution of earnings. *Oxford Economic Papers*, v. 3, n. 2, p. 135-146, 1951.
- RUBIN, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, v. 66, n. 5, p. 688-701, 1974.
- SAMPAIO, B. et al. Desempenho no Vestibular, Background Familiar e Evasão: Evidências da Universidade Federal de Pernambuco. *Economia Aplicada*, v. 15, n. 2, p. 287-309, 2011.
- SCHLICHTING, A. M. S.; SOARES, D. H. P.; BIANCHETTI, L. Vestibular seriado: Análise de uma experiência em Santa Catarina. *Psicologia & Sociedade*, v. 16, n. 2, p. 114-126, 2004.
- SHADISH, W. R.; COOK, T. D.; CAMPBELL, D. T. *Experimental and quasiexperimental designs for generalized causal inference*. New York: Houghton Mifflin Company, 2002. 623 p.

SIANO, L. M. F. *Determinantes do êxito do vestibular na Universidade Federal do Espírito Santo (UFES)*. 1977. 137 f. Dissertação (Mestrado em Educação) – Departamento de Educação, Pontifícia Universidade Católica, Rio de Janeiro.

SILVA, A. O.; FREITAS, W. W. L.; CAVALCANTI, G. A. *Modelos de Sobrevivência aplicados a Evasão dos Alunos de Estatística da UFPB*. In: SIMPÓSIO DE MATEMÁTICA E ESTATÍSTICA DO DELTA, 1., 2012, Parnaíba. Disponível em: <http://www.simed.estatistico.com/trabalhos/poster/SIMED_Poster010.pdf>. Acesso em: 21 jan. 2013.

SMITH, J.; TODD, P. Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators? *Journal of Econometrics*, v. 125, p. 305-353, 2005.

SOBRAL, D. T.; OLIVEIRA, P. G. Avaliação Seriada versus Exame Vestibular: Semelhanças e Diferenças entre Coortes no Curso de Medicina da Universidade de Brasília. *Revista Brasileira de Educação Médica*, v. 30, n. 3, p.181-191. 2006.

STINEBRICKNER, R.; STINEBRICKNER, T. R. Working During School and Academic Performance. *Journal of Labor Economics*, v. 21, n. 2, p. 473-491, 2003.

UNIVERSIDADE DE BRASÍLIA (UNB). Resolução da Reitoria n. 67/1995, de 5 jun. 1995.

_____. Resolução da Reitoria n. 56/1996, de 12 jun. 1996.

_____. Resolução do Conselho de Ensino, Pesquisa e Extensão n. 41/2004. Disponível em: <<http://www.cic.unb.br/~lamar/coordenacao/orientacao/RESOLUCAO041-2004.pdf>>. Acesso em: 01 mai. 2013.

VELLOSO, J. *Curso e Concurso: Rendimento na Universidade e Desempenho em um Vestibular com Cotas da UnB: preliminar*. Brasília, NESUB e Faculdade de Educação da UnB, 2006.

WOOLDRIDGE, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, 2002.

YOUNG, J. W.; FISLER, J. L. Sex Differences on the SAT: An Analysis of Demographic and Educational Variables. *Research in Higher Education*, v. 41, n. 13, p. 401-416, 2000.