



Universidade do Minho
Instituto de Educação

Maria Eugénia Neto Ferrão da Silva

**Contextos e trajetórias do aluno na
explicação do sucesso no Ensino Superior:
A contribuição da modelação multinível**

Contextos e trajetórias do aluno na explicação do sucesso
no Ensino Superior: A contribuição da modelação multinível

Maria Eugénia Neto Ferrão da Silva

UMinho|2018

outubro de 2018



Universidade do Minho
Instituto de Educação

Maria Eugénia Neto Ferrão da Silva

**Contextos e trajetórias do aluno na
explicação do sucesso no Ensino Superior:
A contribuição da modelação multinível**

Tese de Doutoramento em Ciências da Educação
Especialidade em Psicologia da Educação

Trabalho efetuado sob a orientação do
Professor Doutor Leandro da Silva Almeida

outubro de 2018

DECLARAÇÃO DE INTEGRIDADE

Declaro ter atuado com integridade na elaboração da presente tese. Confirmando que em todo o trabalho conducente à sua elaboração não recorri à prática de plágio ou a qualquer forma de falsificação de resultados.

Mais declaro que tomei conhecimento integral do Código de Conduta Ética da Universidade do Minho.

Universidade do Minho, 9 de Outubro de 2018

Nome completo: Maria Eugénia Neto Ferrão da Silva

Assinatura:

M. Eug. Neto Ferrão da Silva

Agradecimentos

À minha família e amigos pelo estímulo na concretização deste projeto e, de forma especial, ao Professor Doutor Leandro S. Almeida pela liberdade crítica que me permitiu alargar horizontes de investigação científica.

Resumo

A Tese desenvolve-se ao longo de seis estudos, recorrendo a métodos quantitativos, tendo como objetivo a explicação do sucesso académico dos estudantes do ensino superior. O Estudo 1 - Reflexões sobre o modelo de valor acrescentado - explora a ligação à política pública em Educação e mostra que a produção científica de língua portuguesa é escassa face à agenda científica imposta pelas potencialidades inerentes ao tema. O Estudo 2 - Statistical methods in higher education research – identifica os métodos estatísticos e as variáveis de resultados considerados na investigação sobre Ensino Superior nos últimos vinte anos. Os modelos de regressão - clássico e multinível – surgem como os mais aplicados e os fenómenos mais frequentemente estudados pelos investigadores são o desempenho académico e o abandono ou persistência. Após os dois primeiros estudos, os restantes envolvem a análise e modelação de dados referentes aos estudantes da Universidade do Minho matriculados pela 1ª vez no 1º ano em 2015/16. O Estudo 3 - Differential effect of university entrance score on first-year students' academic performance in Portugal - apresenta evidência de que a associação entre a nota de ingresso na universidade e o desempenho académico dos estudantes, no final do primeiro ano, varia aleatoriamente entre os cursos, controlando pelas variáveis sociodemográficas, de trajetória escolar e motivacionais. Aplicámos o modelo de componentes de variância e o modelo de coeficientes aleatórios, ambos com dois níveis, estudantes agrupados em cursos. Os resultados indicam que a nota de ingresso é o mais forte preditor do desempenho académico, que 34% da variabilidade do desempenho é devida a diferenças entre os cursos e 80% dessa variabilidade é explicada pelo modelo. O Estudo 4 - Multilevel modeling of persistence in higher education - analisa a associação da persistência no 1º ano com a trajetória escolar e com as condições de ingresso, controlando a origem socioeconómica, a idade e sexo. Aplicámos o modelo de regressão logística multinível. Os resultados sugerem que a reprovação no ensino básico tem efeito de longo prazo, isto é, a probabilidade de persistência é influenciada por esse atributo do indivíduo, mesmo controlando pela respectiva classificação de ingresso. A admissão do estudante no curso de 1ª opção, bem como a classificação de ingresso também estão associadas à probabilidade de persistência. Adicionalmente, os estudantes mais velhos e do sexo masculino parecem ter menor probabilidade de persistência. O Estudo 5 - Personal and institutional factors' impact on persistence in higher education - aprofunda a análise dos fatores individuais que influenciam a persistência. Os resultados sugerem uma forte influência dos métodos de estudo sobre a persistência, bem como um papel importante das expectativas na persistência dos alunos mais velhos. Este estudo também testa o efeito diferencial da nota de ingresso na persistência, tendo-se encontrado evidência da variação aleatória entre os cursos. O Estudo 6 - Acesso e desempenho no ES em Portugal: Questões de género, origem sociocultural, percurso académico, expectativas, escolha do curso e idade dos estudantes – analisa estes atributos através de diversos métodos estatísticos. Os principais resultados sugerem que sexo, origem sociocultural e percurso escolar estão associados à escolha do curso, à nota de candidatura e à opção de entrada. Não foi encontrada associação entre origem sociocultural e persistência. A origem sociocultural dos estudantes parece não influenciar a decisão de abandonar, suspender ou transferir-se de curso.

Palavras-chave: Ensino Superior; Sucesso; Persistência; Modelo de valor acrescentado; Modelo multinível.

Abstract

The Thesis is developed over six studies, based on quantitative methods, and it focuses on students' success in higher education. The study 1 - Reflections on the value-added model - explores the link to educational policy and shows that the Portuguese knowledge production is scarce in view of the potentialities of the subject. The study 2 - Statistical methods in higher education research - identifies the statistical methods and outcome variables considered in Higher Education (HE) research over the last twenty years. Regression models - classical and multilevel - appear as the most applied and the phenomena most often studied by researchers are academic performance and abandonment or persistence. The remaining studies involve the analysis and modeling of data referring to the students of the University of Minho enrolled for the 1st time in the 1st year in 2015/16. The study 3 - Differential effect of university entrance score on first-year students' academic performance in Portugal - presents evidence that the association between the university entrance score and the academic performance of the students at the end of the first year, varies randomly across the courses, controlling for the sociodemographic, educational trajectory and motivational variables. We applied the variance components model and the random coefficients model, both with two levels, i.e. students grouped in courses. The results suggest that the university entrance score is the strongest predictor of academic achievement, that 34% of the performance variability is due to differences between courses and 80% of that variability is explained by the model.

The study 4 - Multilevel modeling of persistence in higher education - analyzes the relationship between 1st year students' persistence with their schooling trajectory and the transition to HE, controlling for the students' socioeconomic status, age and gender. We applied the multilevel logistic regression model. The results suggest that failure in primary, elementary or lower secondary education has a long-term effect, i.e., the probability of persistence is influenced by grade repetition even controlling for university entrance score. Admission of the student to his/her 1st option program, as well as entry score are also associated with the probability of persistence. In addition, older and male students appear to be less likely to persist. The study 5 - Personal and institutional factors' impact on persistence in higher education - deepens the analysis of individual factors that influence persistence. The results suggest a strong influence of study methods as well as an important role of expectations in the persistence of older students. This study also tests the differential effect of the university entrance score, and the random effect across courses was found.

The study 6 - Access and performance in HE in Portugal: issues of gender, sociocultural status, academic program, expectations, program choice and age of students - analyzes these characteristics through various statistical methods. The main results suggest that gender, sociocultural status and the schooling trajectory are associated to the choice of program, the university entry score and the admission option. No association was found between sociocultural status and persistence. The student' sociocultural status does not seem to influence the decision to abandon, suspend or transfer program.

Keywords: Higher Education; Success; Persistence; Value added model; Multilevel model.

Índice

Agradecimentos.....	v
Resumo	vii
Abstract	ix
CAPITULO 1 – Introdução.....	13
1.1 O Ensino Superior como campo de investigação científica.....	15
1.2 A modelação estatística na investigação em Educação.....	19
1.3 A centralidade dos estudos sobre a persistência e o sucesso académico no 1º ano..	23
.....	
1.4 Desenvolvimento por artigos: síntese das questões de investigação e resultados..	27
CAPITULO 2 – Reflexões sobre Modelo de Valor Acrescentado	31
2.1 Introdução.....	32
2.2 Estatística e política pública em Educação	34
2.3 Valor Acrescentado em Educação	36
2.3.1 Definições.....	36
2.3.2 Características, potencialidades e limitações	38
2.3.3 Finalidade: Prestação de contas-responsabilização, escolha da escola ou melhoria da Educação e Ensino Superior	41
2.4 Valor acrescentado na literatura científica de língua portuguesa.....	44
2.5 Considerações Finais	47
CAPITULO 3 – Statistical methods in higher education research	49
3.1 Introduction.....	49
3.2 Scope of the review.....	51
3.3 Summarizing the evidence	52
3.4 Discussion and conclusion	60
CAPITULO 4 – Differential effect of university entrance score on first-year students’ academic performance in Portugal	63
4.1 Introduction.....	63
4.2 The relationship between students’ performance in high school and in college	65
4.3 Differentials of academic performance by area of study	67
4.4 Participants, variables and procedures	68
4.5 Statistical modeling.....	70

4.6 Results	72
4.7 Discussion and Conclusion.....	77
CAPITULO 5 – Multilevel modeling of persistence in higher education	81
5.1 Introduction.....	81
5.2 Methods	85
5.3 Results	88
5.4 Discussion and Conclusion.....	91
CAPITULO 6 – Personal and institutional factors’ impact on persistence in higher Education: A study with first-year Portuguese students.....	95
6.1 Introduction.....	95
6.2 Methods	98
6.3 Results	102
6.4 Discussion and Conclusion.....	106
CAPITULO 7 – Acesso e desempenho no Ensino Superior em Portugal: Questões de género, origem sociocultural, percurso académico, expectativas, escolha do curso e idade dos estudantes	109
7.1 Introdução.....	109
7.2 Métodos	112
7.3 Resultados.....	114
7.4 Discussão	120
CAPÍTULO 8 – Conclusão	125
8.1 Discussão dos principais resultados	125
8.2 Limitações e futuros desenvolvimentos.....	135
REFERÊNCIAS	137
ANEXO 1	155
- Anexo 1 - Table A1. Steps for papers selection	

CAPITULO 1 – Introdução

O virar do século foi marcado pelo crescimento do Ensino Superior (ES) na Europa com taxa de 116% em 2002 face a 1998, crescimento esse verificado em todos os países da Comunidade Europeia com exceção de Itália (99%), Áustria (90%) e Bulgária (88%) (Santos Silva et al., 2005). Considerando 1998 ano base, em Portugal o índice de variação tem sido sempre superior a 100%, aproximando-se deste valor em 2015, tendo atingido o máximo de 115% em 2003. A grande expansão do ES português decorreu após a década de 70 do século passado, concretizando o princípio da igualdade de oportunidades subjacente à revolução de abril de 1974. O número de estudantes a frequentar o ES no início da década de 70 era aproximadamente 60 mil e no ano 2002/3 era de 400 mil, o que representa uma variação superior a 650%. Estes números traduzem a democratização do acesso ao ES. Outro índice relevante para a análise em apreço é o que quantifica a variação do número de estudantes inscritos pela 1ª vez no 1º ano. O número absoluto aumentou sempre até 2011, representando a variação de 161% considerando 1998 o ano base. De acrescentar que a democratização também aumentou a heterogeneidade da população discente, quer quanto ao perfil e trajetória escolar, quanto às regiões urbanas ou rurais de proveniência ou quanto à origem socioeconómica e cultural, entre outras variáveis.

A democratização do acesso assentou fundamentalmente na criação de lugares discentes, com atenção insuficiente às condições e qualidade da sua formação. Alguns autores diferenciam “democratização do acesso” de “democratização do sucesso”. A “democratização do sucesso” no ES é passível de dúvidas e reservas, uma vez que “as taxas de abandono e de insucesso penalizam, sobretudo, os estudantes provenientes dos estratos sociais mais desfavorecidos” (Almeida, Marinho-Araujo, Amaral, & Dias, 2012, p.913). O desempenho do sistema caracteriza-se por considerável taxa de abandono e de reprovação. De acordo com a Direção-Geral de Estatísticas da Educação e Ciência (Engrácia & Baptista, 2018), considerando os estudantes que se inscreveram pela 1ª vez no 1º ano no ano letivo 2011/12 em cursos de 1º ciclo com duração de três anos, 46% deles em 2014/15 tinham obtido o diploma do curso em que se tinham matriculado em 2011, 14% continuavam inscritos no mesmo curso, 11% frequentavam um curso diferente daquele onde se haviam matriculado e os restantes 29% não foram encontrados no ES português.

Ao longo das últimas décadas verificou-se amplo consenso sobre a necessidade da avaliação da qualidade académica, o que se refletiu no aumento da produção de rankings institucionais e de cursos (Dill & Soo, 2005). De acordo com Foley e Goldstein (2012), os rankings de desempenho destinam-se a dois propósitos principais: o primeiro pode ser descrito como responsabilização/prestação de contas; e o segundo é fornecer informação aos utentes e *stakeholders* dos serviços públicos de educação e ES. De acordo com Hazelkorn (2011), após a crise financeira de 2008, os rankings das universidades passaram a simbolizar o impulso para maior responsabilização e transparência, fornecendo a evidência subjacente à agenda de modernização e à adoção de novas práticas de gestão pública. A investigação científica sobre o tema enquadra-se no quarto tema - Qualidade - da estrutura apresentada por Tight (2013, 2014) para a produção de conhecimento sobre ES, com teorias associadas na área da avaliação, classificação e resultados, práticas de monitoramento, listas ordenadas de desempenho e padrões do Sistema. Foi a partir da Declaração publicada pela *American Statistical Association* em 2014 sobre modelos de valor acrescentado (Morganstein & Wasserstein, 2014) que surgiu, como ideia, esta Tese.

A Tese tem, assim, dois pilares temáticos: o da Qualidade onde se insere a revisão de literatura reportada no segundo capítulo, e o da investigação institucional (Altbach, 2014) onde se enquadram tópicos de pesquisa que pertencem ao terceiro tema – Experiência dos alunos – da estrutura apresentada por Malcom Tight (op. cit.) e cujos estudos de investigação são relatados do quarto ao sétimo capítulos. A abordagem metodológica predominante ao longo da Tese recai sobre os modelos de regressão multinível, cuja escolha é justificada adiante. Adicionalmente, a revisão da literatura apresentada no terceiro capítulo sobre quais são os métodos estatísticos adotados na investigação em ES revela que, após o ano 2010, esta foi a abordagem estatística mais escolhida pelos investigadores.

Este capítulo de introdução desenvolve-se em quatro secções. Na primeira seção apresentamos uma reflexão sobre o campo de investigação em ES. Na segunda, elaboramos sobre o papel da modelação estatística na investigação em Educação. Na terceira apresentamos as ideias centrais sobre a relevância do estudo que incide em dados de estudantes do primeiro ano e, especialmente, dos fenómenos de sucesso e persistência no ES. Finalmente, na quarta seção, apresentamos a estrutura e conteúdo dos restantes capítulos da Tese.

Em geral, o estilo bibliográfico *American Psychological Association* (APA), com adaptação para português quando aplicável, foi adotado para citações em texto e referências bibliográficas, com exceção das normas próprias de cada revista em que o artigo foi submetido e/ou publicado.

1.1 O Ensino Superior como campo de investigação científica

O pensamento segundo o qual a investigação científica sobre ES é um campo multidisciplinar, envolvendo investigadores de toda a academia e que utilizam uma ampla gama de abordagens metodológicas, é consensual (Altbach, 2014; Altbach & Engberg, 2001; National Research Council, 2002, 2005; Tight, 2013). Trata-se de um campo de investigação sem um corpo intelectual, metodológico ou disciplinar próprio, e por conseguinte, não obedece às características necessárias para se firmar como uma disciplina académica ou científica autónoma (Altbach, 2014). Trata-se antes de um campo de estudo cuja investigação analisa tópicos diversos, baseado em abordagens metodológicas variadas e em ampla expansão. O thesaurus da organização *Education Resources Information Center* (ERIC) fornece os seguintes sete resultados para o termo "Higher Education", incluindo sinónimos e termos em desuso: "Admission Tests (Higher Education)", "Higher Education", "Higher Education as a Field of Study", "Higher Education Institutions", "Higher Education Marketing", "Private Higher Education" e "Public Higher Education". A *American Educational Research Association* (AERA), a maior e mais antiga associação científica de Educação, está organizada em doze divisões de modo a representar as suas subáreas académicas ou científicas, entre as quais se encontra a divisão de ES (Divisão J), instigando os seus membros com a seguinte pergunta: "How are we as higher educational researchers positioned to steady our national conversation on equity, to furnish ballast necessary for democratic progress and social stability?". Os objetivos da Divisão J são os seguintes: enriquecer o grau de compreensão das políticas e práticas neste nível de ensino; expandir e aprofundar o conhecimento de estruturas e experiências pós-secundárias de modo a encontrar a explicação para a iniquidade e a desvendar os muitos enigmas cognitivos e de desenvolvimento que caracterizam as experiências dos estudantes de ES.

Apesar do recente crescimento da produção científica sobre ES, ainda se nota uma predominância que reflete a realidade contextual dos países desenvolvidos, principalmente dos Estados Unidos da América, onde diversas instituições de ES criaram departamentos e centros

específicos sobre este campo de estudo (Altbach & Engberg, 2001). Saunders, Kolek, Williams, e Wells (2016) estudaram a distribuição da afiliação institucional dos autores que publicaram em três “top-tier” periódicos nos Estados Unidos da América (EUA) entre 2006 e 2010, concluindo que a produção científica é fortemente concentrada no que se refere a autores e a determinadas instituições, levantando questões fundamentais sobre os processos sociais que governam a produção e disseminação do conhecimento sobre ES.

Tight (2013; 2014) analisou a relação entre disciplina e a metodologia na investigação sobre ES, bem como a relação entre disciplina e teoria, através da análise de 567 artigos publicados em quinze periódicos especializados, editados e/ou publicados na Australasia, Europa e América do Norte, durante o ano de 2010¹. No que concerne ao país de afiliação do primeiro autor, em 74% dos artigos analisados o primeiro autor reside num país onde o Inglês é a língua oficial, isto é, nos Estados Unidos da América ou no Canadá (33%), na Austrália ou Nova Zelândia (24%) e no Reino Unido (17%), sendo que os demais países europeus e resto do mundo representam 16% e 10%, respetivamente. Em ambas as publicações, se afirma a investigação sobre ES como campo de estudo multidisciplinar, desenvolvida principalmente por académicos inseridos em departamentos de educação, mas também por académicos provenientes de outros departamentos – fundamentalmente das ciências sociais. A tal respeito, há evidência de que as contribuições para o desenvolvimento do campo são provenientes de diversas disciplinas e que, principalmente entre os países emergentes, tal como a China, as mudanças nos últimos anos são acentuadas (Altbach, 2014; Altbach & Engberg, 2001) e promissoras. Tight (2013) procurou identificar quais são as disciplinas envolvidas e com que metodologias se produz novo conhecimento no ES. No que se refere às disciplinas, aferidas através do departamento onde os autores são afiliados, 59.8% dos casos válidos são provenientes de outras áreas disciplinares que não a Educação, variando entre 1.3% na Engenharia e 17.2% em outras Ciências Sociais. Apesar da maioria da produção ser de outras áreas que não a Educação, é a Educação que atinge maior percentual de autores/coautores (40.2%). Malcom Tight também apresenta evidência de grande diversidade quanto às metodologias adotadas, sendo os métodos multivariados (*surveys* e análise multivariada) a

¹ Assessment and Evaluation in Higher Education, Higher Education, Higher Education Management and Policy, Higher Education Policy, Higher Education Quarterly, Higher Education Research & Development, Innovative Higher Education, Journal of College Student Development, Journal of Higher Education, Journal of Higher Education Policy and Management, Research in Higher Education, Review of Higher Education, Studies in Higher Education, Teaching in Higher Education and Tertiary Education and Management.

escolha mais frequente. O enquadramento teórico dos 567 artigos analisados reparte-se por oito temas: (1) Ensino-aprendizagem; (2) Planeamento curricular; (3) A experiência dos alunos; (4) Qualidade; (5) Política do sistema; (6) Gestão institucional; (7) Trabalho académico; (8) Conhecimento e investigação. O autor associa a estes temas as respetivas áreas e teorias (em desenvolvimento), indicando, quando aplicável, as áreas de onde são importadas. Ao longo da Tese iremos apresentar estudos de investigação sobre a experiência dos alunos e sobre qualidade, isto é, tratam-se de estudos que se inserem nos temas (3) e (4). No que concerne ao tema (3) Experiência dos alunos, Tight (2014) refere como domínios: acesso ao ES; as escolhas do aluno; ampliação da participação; ajuda financeira; a experiência no curso; sucesso e não conclusão; estudantes internacionais; a experiência de pós-graduação. Os tópicos em desenvolvimento são os seguintes: Transição para o ES; aculturação; *engagement*; atrito, abandono ou evasão; *mentoring*; plágio; rendimento; empregabilidade; transição para o mercado de trabalho.

No tema (4) Qualidade, os domínios enunciados são os seguintes: Avaliação do curso; classificação e resultados; práticas nacionais de monitoramento; listas ordenadas de desempenho (*league tables*) e padrões do sistema. Quanto aos tópicos em desenvolvimento são: Perceção e satisfação do aluno; métodos de avaliação e classificação; avaliação de qualidade do ensino; sistemas de classificação e ordenação (pp. 101-102).

Ainda sobre o avanço científico-disciplinar, uma das questões a que o relatório sobre investigação científica em educação *National Research Council* pretende responder é a seguinte: “Does knowledge from scientific education research progress as it does in the physical, life, or social sciences?” (National Research Council, 2002, p. 29). A resposta considera dois argumentos principais: o primeiro de que os achados ou descobertas da investigação em Educação progrediram ao longo do tempo e tal desenvolvimento permitiu fornecer contribuições importantes sobre políticas e práticas; o segundo argumento é de que o desenvolvimento disciplinar baseado na investigação científica, quer seja na Educação ou nas restantes áreas científico-disciplinares, é feito de avanços e recuos à medida que os métodos, teorias e descobertas empíricas interagem e evoluem de forma não linear, e que esse processo requer investimentos substanciais de tempo e de financiamento. Segundo os autores, esse desenvolvimento conta com a crítica e a controvérsia associada à apropriação das descobertas científicas pelos pesquisadores, mentores e agentes de política e prática, bem como por cidadãos, i.e., “The path to scientific knowledge wanders through contested terrain

as researchers, as well as the policy, practice, and citizen communities critically examine, interpret, and debate new findings and it requires substantial investments of time and money.” (op.cit., p. 29). Adicionalmente, os autores lembram que os princípios orientadores da investigação científica na Educação são os mesmos das outras áreas científico-disciplinares, mas que a sua aplicação pode depender da natureza dos fenómenos investigados. De facto, a natureza dos fenómenos nas ciências exatas ou físicas é diferente da natureza nas ciências sociais. Cabe então aos académicos estabelecer, na sua própria área, quais são os padrões adequados para a aplicação de tais princípios, seguindo o pensamento de Jared Diamond, ou seja, “Scholars working in a particular area establish the traditions and standards for how to most appropriately apply the guiding principles to their area of study” (citado em National Research Council, 2002, p. 80). De acordo com Walvoord et al. (2000),

A discipline is a set of assumptions and tools for viewing the world in a certain way, addressing certain kinds of questions, and valuing certain kinds of evidence and insight. Disciplines are not stable; they splinter into subdisciplines, and they combine with other disciplines. (op.cit., p. 25)

Educação como disciplina, dividida em subdisciplinas e combinada com outras disciplinas – como é o caso da Estatística, tem um caminho para a sua compreensão científica feito de avanços e recuos, à semelhança da investigação em todas as disciplinas científicas. Ou seja, “Its advancement is choppy, pushing the boundaries of what is known by moving forward in fits and starts as methods, theories, and empirical findings evolve” (National Research Council, 2005, p. 29).

Este relatório do *National Research Council*, apresenta exemplos da ligação da Educação a outras disciplinas, tais como a Economia ou a Psicologia, que contribuem para a produção de conhecimento sobre Educação. Segundo os autores, tal natureza multidisciplinar tem três implicações que importa reter: (1) As diversas perspetivas disciplinares focam diferentes partes do sistema e, conseqüentemente, criam legitimidade à diversidade de enquadramentos e métodos de investigação, podendo conduzir a conclusões aparentemente contraditórias. Neste caso, o desafio necessário é o da integração de teorias e de descobertas empíricas; (2) Os avanços da investigação em Educação dependem parcialmente dos avanços nas disciplinas que lhe estão associadas; (3) A proliferação de enquadramentos teóricos e metodológicos,

juntamente com a miríade de campos que contribuem para a compreensão da Educação, tornam o desenvolvimento profissional extremamente complicado. “A complexidade e diversidade dos temas abrangidos, bem como os múltiplos enquadramentos epistemológicos e metodológicos são quase impossíveis de cobrir adequadamente em um único programa de graduação” (National Research Council, 2002, p. 93), o que constitui um grande desafio tanto para as instituições de formação e educação como para os académicos e investigadores responsáveis pela substituição geracional.

1.2 A modelação estatística na investigação em Educação

Os últimos anos são marcados por grande avanço nos métodos quantitativos, tanto na perspectiva do seu desenvolvimento metodológico como da aplicação ao estudo de fenómenos complexos da vida real (Goldstein, 1998), com forte impacto na produção de conhecimento e, em particular a modelação estatística, com forte impacto no avanço disciplinar (Rodgers, 2010).

A diversidade de abordagens metodológicas na investigação em Educação é grande e, possivelmente em parte, também se deve à diversidade de tipos de instituições que a realizam. Morris (2002) identifica quatro possíveis fontes de diversidade assente no tipo de instituição, ou seja, investigação de base disciplinar usualmente realizada por universidades, investigação aplicada e realizada nas próprias escolas ou organizações de formação profissional, estudos sobre o mercado de trabalho e necessidades dos empregadores e investigação realizada por organismos oficiais frequentemente associada a políticas ou programas governamentais. Adicionalmente, enuncia três propósitos principais: “Investigação destinada a aprofundar a compreensão de questões e conceitos, para ampliar o corpo de conhecimento e teorias subjacentes; investigação orientada à prática; e investigação orientada à política. Discorrendo sobre a complexidade do processo de investigação em Educação, a autora apresenta uma proposta de estrutura como se tratando da sequência de três fases: (1) Planeamento; (2) Investigação; (3) Recomendação². A proposta de Morris (2002) tem um domínio amplo, pois

² Tradução livre. “Planning phase (Analysing needs, Establishing priorities, Surveying existing research, Designing the project); Investigating phase (Analysing existing knowledge, Identifying resources, Gathering data, Analysing data, Establishing findings); Influencing phase (Identifying potential for change, Interpreting findings, Distilling messages, Disseminating messages, Working with agents of change)” (op. cit. p. 8, 9).

seu objetivo é encontrar uma forma de concepção da investigação para a política e prática. A diversidade na investigação em Educação também é reconhecida por Dill e Soo (2005) especialmente na implicação que tem na comparação institucional da qualidade académica. O Presidente da norte-americana *National Academy of Sciences*, Bruce Alberts, acrescenta que “researchers who often disagree along philosophical and methodological lines nonetheless share much common ground about the definition and pursuit of quality” (National Research Council, 2003, p. viii), justificando a necessidade de se estabelecerem princípios científicos orientadores da investigação em Educação. No que se refere ao desenvolvimento da investigação descrita nesta Tese, foram seguidos os seis princípios científicos descritos nesse documento e, em particular a operacionalização do princípio número três, segundo o qual “Devem ser usados métodos que permitam a pesquisa direta da questão em análise” (op. cit., p. 62), reforçando a necessidade da adoção dos métodos estatísticos em particular da modelação estatística.

A classe de modelos de regressão multinível ou hierárquicos (Bryk & Raudenbush, 1992; Ferrão, 2003; Goldstein, 2011; Plewis, 1997) tem um papel central na investigação em Educação (Aitkin & Longford, 1986; Goldstein, 1997; Goldstein, 1998; Pike & Rocconi, 2012), desde logo por permitir a modelação explícita dos contextos e a sua influência nos indivíduos. A atenção ao contexto também sugere que o avanço da compreensão dos fenómenos em contextos educacionais complexos e diversos pode exigir uma estreita coordenação entre pesquisadores e profissionais, o trabalho interdisciplinar e a interação entre diversas formas de pesquisa educacional. Além disso, implica maior ênfase na diversidade inerente à experiência educacional e seus resultados para diferentes populações de estudantes (AERA, APA, & NCME, 1999). Também a *British Academy* se refere à importância de a modelação estatística ser usada na investigação sobre ES para descrever diferenças institucionais que permitam apurar o desempenho diferencial para diferentes grupos de estudantes, ou seja,

More appropriate statistical analysis models should be used to describe institutional differences that allow for differential performance for different groups of students. In particular a shift away from the comparison of individual institutions towards research that helps to identify modifiable factors that appear to be related to good performance. (Foley & Goldstein, 2012, p. 34)

Os modelos multinível ou hierárquicos incorporam a estrutura de agrupamento da população educacional (alunos agrupados em turmas, turmas em cursos, cursos em departamentos, departamentos em faculdades, faculdades em universidades, universidades em países...) que está patente nos dados. Johnston, Manley, Jones, Harris, e Hoare (2016) apresentam a aplicação à investigação da transição do ensino secundário para o ES no Reino Unido através do estudo da relação entre as classificações de ingresso e a classificação de desempenho no ES. Os autores criticam as análises realizadas pelo Departamento de Educação e apresentam resultados contraditórios aos reportados pelo Departamento sobre uma questão de investigação central: a classificação do aluno no *General Certificate of Secondary Education* (GCSE) é ou não tão bom preditor do desempenho do aluno no ES (com classificação muito bom) quanto é a classificação *Advanced Subsidiary* (AS) level? Os dados foram reanalisados por Johnston e colegas através da aplicação de modelos de regressão logística multinível e concluíram que a classificação *AS-level* é um preditor mais fiável do que o GCSE.

Sobre as principais características que nos conduzem à adoção da modelação multinível, mantém-se atual o que escrevi em 2003 no livro “Introdução aos modelos de regressão multinível em educação” publicado no Brasil:

Muitos conjuntos de dados refletem a estrutura da população em que foram recolhidos. Na sociedade, os indivíduos estão naturalmente organizados em grupos: o indivíduo na família, as famílias agrupadas em bairros, estes organizados em municípios, etc. Nas escolas, os alunos estão agrupados em turmas, e um conjunto de turmas forma a escola. No sistema eleitoral, o eleitor é agrupado em distritos eleitorais e estes, por sua vez, agrupados em estados. No que diz respeito aos sistemas de saúde, principalmente aqueles em que existe o médico de família, podemos identificar quatro níveis a estrutura de agrupamento da população: o indivíduo pertence a determinada família, um conjunto de famílias é acompanhada por um médico de família, o grupo de médicos organizados em centros de saúde ou hospitais. Estas estruturas organizativas da população e, por conseguinte, dos dados, também podem ser designadas estruturas multinível ou hierárquica.

Até recentemente, a análise estatística dos dados não levava em conta a sua estrutura e, portanto, dos sistemas em estudo. Isto devia-se fundamentalmente à falta de métodos e pacotes computacionais que viabilizassem a generalização de propostas metodológicas já existentes, para lidar com dados hierárquicos. [...] A principal

dificuldade subjacente à modelagem dos dados educacionais, ou quaisquer outros de diferente natureza, mas com estrutura hierárquica, é a escolha da unidade de análise sobre a qual o estudo irá incidir. [...] Em síntese, precisamos de um instrumento de análise do contexto e da sua interação ao nível individual. (Ferrão, 2003, pp. 29-30)

E mais adiante acrescento:

As vantagens da utilização do modelo de regressão multinível podem ser resumidas em quatro pontos:

- Possibilita a obtenção de estimativas eficientes dos coeficientes da regressão;
- Leva em conta a informação sobre o agrupamento dos dados, tornando erros-padrão, intervalos de confiança e testes de hipóteses corretos;
- Permite o uso de variáveis explicativas mensuradas em cada nível da hierarquia, permitindo ao analista a exploração detalhada do impacto e da contribuição de cada nível para a variabilidade da variável resposta;
- Permite estabelecer listas comparativas do desempenho institucional, que vem ganhando relevo na educação e na saúde, evidenciando a impossibilidade técnica de estabelecer listas ordenadas das escolas tal como elas têm sido apresentadas em alguns países. (op.cit., p. 31)

Diversas especificações dos modelos serão apresentadas nos capítulos seguintes, em função das respetivas questões de investigação enunciadas. Em síntese, são usadas as seguintes especificações: Modelo de componentes de variância (Capítulo 4), Modelo de coeficientes aleatórios (Capítulo 4), Modelo de regressão logística multinível (Capítulo 5, Capítulo 6). Diversas estatísticas descritivas e métodos inferenciais são usados (por exemplo no Capítulo 7), tais como o teste de hipóteses para a independência entre variáveis, usualmente designado por teste de qui-quadrado, o teste de hipóteses *ANOVA* e o teste para a comparação da média em amostras independentes.

Nos quinze anos que decorreram entretanto, a modelação estatística ganhou ainda maior centralidade decorrente do grande volume de dados disponíveis. Este fenómeno deverá acentuar-se no futuro próximo, já que cada vez maior número de países aderem à política de

transparência na governação através do conceito *open government data*, observando o princípio de que os dados públicos, i.e. todos os dados não personalizados que são recolhidos pelo Estado devem ser *open source* para a criação de valor digital e conhecimento (Ferrão, 2015; Ritter & Mayer, 2018). Deparamo-nos face a uma autêntica revolução decorrente de “big data” (Singer, 2018). No ES, a emergência dos dados e a necessidade de perícia em analisá-los também é referida por Altbach, Bozeman, Janashia, e Rumbley (2007) que alertaram para a sua imprescindibilidade sobre todos os aspetos do ES e para a melhor compreensão da natureza das instituições académicas, ponto crítico para o futuro da Universidade. O autor refere que os responsáveis pelo planeamento e administração das instituições de ES começam então a reconhecer a necessidade de recorrer a dados e à sua interpretação para suportar a tomada de decisão. À medida que as instituições se expandem, precisam de mais informação disponível sobre a tendência das matrículas, indicadores de desempenho dos alunos, e dados sobre docentes e/ou investigadores, etc. A recolha e análise desses dados é usualmente designada por investigação institucional e é focada numa só instituição, mas que pode ser relevante para uma audiência bastante ampla (Altbach, 2014; Altbach & Engberg, 2001).

1.3 A centralidade dos estudos sobre a persistência e o sucesso académico no 1º ano

A investigação institucional (Altbach, 2014; Altbach & Engberg, 2001) é fortemente marcada pelo interesse de pesquisa em fenómenos relacionados com a transição do ensino secundário para o ES, bem como dos fenómenos preditores do sucesso do estudante no ES (Baron & Norman, 1992). O projeto relatado por Almeida e Castro (2016) encontra-se nessa linha de investigação, no âmbito do qual foram recolhidos os dados usados nos capítulos quatro, cinco, seis e sete, e cujo principal objetivo de pesquisa é contribuir para identificar os determinantes da transição e adaptação dos estudantes ao ES visando o aumento do sucesso do aluno na conclusão do curso e obtenção do seu 1º diploma no ES.

A taxa de conclusão e obtenção do 1º diploma em Portugal era de 41% em 2012 (OECD, 2014) mantendo-se estável em 2015 (OECD, 2017). Esta caracterização vai ao encontro dos resultados obtidos na Direção-Geral de Estatísticas da Educação e Ciência,

[...] a percentagem de alunos que, ao fim de 4 anos, já se encontravam diplomados, ser bastante semelhante nos quatro subsistemas considerados, rondando os 45%. Portanto, entre todos os alunos que ingressaram em cursos de licenciatura, com duração teórica de três anos, no ano letivo 2011/12, menos de metade terminaram o seu curso nos quatro anos seguintes. (Engrácia & Baptista, 2018, p. 3)

Os autores também referem o abandono no ES, havendo grande variação entre as áreas de formação, sendo a Informática uma das que apresenta menor taxa de conclusão em cursos de três anos pois 22% dos alunos conseguiram concluir o curso em quatro anos. E apresentam detalhes por áreas de formação,

Além da Informática, as áreas de formação em que os alunos menos conseguiram concluir as suas licenciaturas em quatro anos foram as das Engenharias⁶, da Arquitetura e Construção e das Ciências Físicas, com percentagens de conclusão entre os 33% e os 38%. Estas são áreas caracterizadas por taxas de abandono significativas, na casa dos 20%, elevadas taxas de transferência, também da ordem dos 20%, e números substanciais de alunos que continuam inscritos nos seus cursos iniciais, passados quatro anos, sem os terem concluído. (op. cit., p. 10)

Também verificámos que, com base no método de *cross cohort*³, o indicador “*How many students complete tertiary education?*” é 65% para Portugal comparando com 47% na Eslovénia, 48% no Brasil, 94% na Irlanda, 92% no Japão ou 85% na Coreia. O mesmo indicador calculado com base no método *true cohort* apresenta o valor de 71% para o Reino Unido, 43% para a França ou para a Finlândia (OECD, 2016).

³ “O método de *cross cohort* (coorte cruzada, tradução da autora) requer dados sobre o número de novos participantes para um dado nível CITE e o número de graduados N anos depois, onde N corresponde à duração teórica do programa ou curso. Sob o pressuposto de fluxos constantes de estudantes, a taxa de conclusão da coorte cruzada está mais próxima de uma taxa de conclusão total (ou seja, independentemente do tempo que os estudantes levaram para se formar). Assim, em países onde uma grande parte dos estudantes não se graduam no período de tempo adequado (dentro da duração teórica do programa), a taxa de conclusão da coorte cruzada pode ser mais comparável a prazos mais longos na conclusão da verdadeira coorte (*true cohort*). As taxas de conclusão de ambos os métodos, *cross cohort* e *true cohort* são calculadas como o número de graduados dividido pelo número de participantes N or ou N + 2 anos antes (onde N é a duração teórica do programa)” (e.g. OECD, 2017, p. 160).

A desagregação da taxa de conclusão por idade mostra que a taxa é 34.6% no grupo de alunos com idade menor do que 30 anos, e é 5.93% no grupo com idade igual ou maior do que 30 anos. A este propósito, OECD (2014, p. 78) refere que, considerando os países da OCDE, em média um estudante obtém seu primeiro diploma universitário aos 27 anos de idade, mas a idade de graduação varia muito entre os países. Por exemplo, os estudantes na Bélgica, Luxemburgo, México, Holanda e Reino Unido formam-se antes de completarem 25 anos e que, em média, os estudantes no Brasil, Finlândia, Islândia, Israel e Suécia recebem seu primeiro diploma universitário após completarem 29 anos (Gráfico A3.1). Em Portugal, a idade média de obtenção do 1º diploma é 26 anos (OECD, 2016, Tabela A3.4) e a idade média de ingresso no curso de 1º ciclo é 22 anos (OECD, 2014, Tabela C2.1a), refletindo em parte a performance do subsistema de ensino não-superior. De acordo com OECD (2017, p. 165) apenas 49% dos alunos de ensino secundário terminaram esse nível de ensino no tempo teórico de três anos (59% dos alunos terminaram-no em 5 anos).

Bruinsma e Jansen (2009) realizaram uma análise de sobrevivência com dados referentes a 3,879 alunos do 1º ano de uma universidade holandesa, a Universidade de Groningen, com a finalidade de obter evidência sobre um conjunto de fatores, preditores do momento no tempo em que os alunos passam do primeiro para o segundo ano e concluíram que 69% dos estudantes não obtiveram o diploma do primeiro ano dentro da duração teórica de 12 meses. Além disso, no final do período observado de 24 meses, 61% dos alunos obtiveram o seu diploma de primeiro ano. A estimativa da mediana foi atingida passados 23 meses. Os autores acrescentam que a percentagem parece bastante baixa mas não é incomum na Holanda, porque as estatísticas da Associação de Universidades na Holanda mostram que as percentagens de estudantes que obtiveram seu diploma de primeiro ano dentro de 48 meses diminuíram de 68% em 1989/1990 para 59% em 1999/2000.

Brahm, Jenert, e Wagner (2017) analisaram os dados de uma amostra de 820 estudantes de uma universidade suíça para estudar o papel da motivação no desempenho ao longo do primeiro ano. Os autores referem que em alguns cursos a taxa de abandono é aproximadamente de 30%. Por sua vez, outros autores (Clerici, Giraldo, & Meggiolaro, 2015; Meggiolaro, Giraldo, & Clerici, 2015) debruçaram-se sobre os fatores de desempenho e quiseram saber se esses fatores atuam de maneira diferente de acordo com as áreas de estudo dos cursos de graduação em que os alunos estão matriculados. Em particular, os fatores que influenciam o abandono, transferência para outro curso, reprovação, ou a conclusão, em

curso de três anos na Universidade de Pádua são analisados a partir da perspectiva dos campos de estudo. Os autores examinaram os registros de mais de 32,000 alunos matriculados dos anos académicos de 2002/3 a 2005/6 em 84 cursos de primeiro ciclo, agrupados em quatro diferentes áreas de estudo e concluíram que as variáveis associadas à maior probabilidade de abandono e menor probabilidade de conclusão de curso, são os seguintes: ser um estudante estrangeiro, ser proveniente de escolas profissionais, ter notas escolares baixas no ensino secundário e trajetória escolar irregular. Os autores também reportam especificidades associadas aos cursos. Por exemplo, na área das ciências humanas e sociais, os estudantes que não se inscreveram imediatamente após deixarem a escola secundária constituem um grupo em risco e nos cursos de ciências exatas (estudos científicos), os estudantes do sexo masculino têm menor probabilidade de conclusão. No artigo Meggiolaro et al. (2015), os autores usaram um modelo de regressão multinível, cujos resultados destacam a relevância dos fatores contextuais. Segundo os autores, as ações e os serviços de apoio devem ser devidamente diferenciados de acordo com o contexto e de acordo com as características individuais dos alunos. Por exemplo, em cursos caracterizados por maior risco de atraso, devem ser fornecidos serviços de apoio e de aconselhamento ao estudo individual e o desenvolvimento de métodos educacionais. Em cursos com maior risco de transferência, sessões de orientação mais detalhadas no momento da matrícula e o aconselhamento psicopedagógico poderiam ser organizadas durante a experiência inicial do aluno no ES.

Pike e Saupe (2002) analisaram dados de 8,674 estudantes universitários de 124 escolas secundárias que se matricularam numa universidade dos Estados Unidos da América entre 1996 e 1999 e consideraram o *Grade Point Average* (GPA) do primeiro ano dos alunos como variável dependente. Eles também incluíram na análise a pontuação do *American College Testing* (ACT) como parte da política de admissão na universidade, classificação percentual da classe do ensino médio e uma variável binária que representa se o aluno cumpriu ou não os novos requisitos do curso implementados em 1997. Estudaram a relação entre ter cumprido os novos requisitos do curso principal e o GPA do primeiro ano, fornecendo evidência da adequação dos novos requisitos de admissão para os alunos terem sucesso no ES. Em relação ao efeito do score composto da avaliação do ACT e do percentil de classe alta no GPA, ambos os coeficientes foram positivos e estatisticamente significativos. Fuller, Wilson e Tobin (2011) exploraram em que medida as pontuações transversais e longitudinais da *National Survey of Student Engagement* (NSSE) constituem preditores do GPA final de um estudante universitário.

Seus modelos incluem pontuações ACT e GPA do ensino médio como indicadores das aptidões pré-universitárias dos alunos. Com base na modelação longitudinal de dados, os autores concluem que 27.8% da variância explicada poderia ser atribuída aos indicadores académicos pré-universitários. Além disso, McKenzie e Schweitzer (2001) realizaram uma investigação da relação entre fatores académicos, psicossociais, cognitivos e demográficos e o desempenho académico dos alunos, avaliados pela classificação de um aluno no final do primeiro semestre. Os autores sugeriram que, como os índices de ingresso na universidade são moderadamente correlacionados com o GPA. A literatura também mostra que o desempenho dos alunos ao longo do primeiro ano tende a determinar seu desempenho nos anos seguintes (Berg & Hofman, 2005; Yorke & Longden, 2007).

Considerando o exposto, as experiências académicas e de integração do aluno no 1º ano do curso são essenciais para o seu percurso até à conclusão e obtenção de diploma. Diversas variáveis de resultado têm sido tratadas na literatura. Nos capítulos seguintes desta Tese, consideraremos como fenómeno de interesse a média das classificações obtidas pelo aluno no final do 1º ano (Capítulo 4, Capítulo 7) e também a probabilidade de persistência no 1º ano (Capítulo 5, Capítulo 6, Capítulo 7). A modelação destes fenómenos de interesse é função de variáveis individuais e específicas do curso frequentado, tal como desenvolvido em detalhe nos respetivos capítulos.

1.4 Desenvolvimento por artigos: síntese das questões de investigação e resultados

Em cada um dos próximos seis capítulos desenvolve-se a temática específica já antecipada nas secções anteriores. Tal desenvolvimento reflete as diversas etapas de investigação. Tendo como motivação inicial o artigo publicado pela *American Statistical Association* em 2014 sobre modelos de valor acrescentado (Morganstein & Wasserstein, 2014), deu origem ao artigo com título “Estatística educacional e política pública: A propósito dos modelos de valor acrescentado” que se apresenta no segundo capítulo. Trata-se de um ensaio sobre a revisão do tópico e a ligação à política pública em Educação. O tratamento do tema “valor acrescentado na literatura científica de língua portuguesa”, levou-nos à constatação da

diminuta produção e disseminação face à agenda científica imposta pelas potencialidades inerentes ao tema.

Em seguida quisemos saber quais são os métodos estatísticos usados na investigação sobre ES, donde resultou o título: “Statistical methods in higher education research: A review” que se apresenta no terceiro capítulo. Este estudo examina quais são os métodos estatísticos e as variáveis de resultados que foram considerados na investigação sobre ES nos últimos vinte anos. A recolha de artigos para revisão foi realizada usando mecanismos de busca que abrangem a literatura de Educação e de Estatística, tais como o *Education Resources Information Center (ERIC)* e o *Current Index to Statistics (CIS)*. Adicionalmente, a recolha foi conduzida em seis periódicos científicos de ES. Após a eliminação de duplicações, restaram 122 artigos de autoria ou coautoria de 290 pesquisadores que foram publicados em 46 periódicos e abrangem autores de 27 países. Os modelos de regressão - modelos clássicos e multiníveis - são os mais aplicados na pesquisa em ES e os fenómenos mais frequentemente estudados pelos investigadores são o desempenho académico e o abandono escolar ou tópicos relacionados.

Centrámos então a atenção sobre o projeto “Observatório dos percursos académicos dos estudantes da UMinho” descrito pelos seus coordenadores (Almeida & Castro, 2016) e que está em linha com as características de uma investigação institucional (Altbach, 2014; Altbach et al., 2007; Altbach & Engberg, 2001). No quarto capítulo apresentamos o estudo com título “Differential effect of university entrance score on first-year students’ academic performance in Portugal” cujo principal objetivo é mostrar que a associação entre a nota de ingresso na universidade e o desempenho académico dos alunos no final do primeiro ano varia aleatoriamente entre os cursos, controlando pelas variáveis sociodemográficas, de trajetória escolar e motivacionais dos alunos. A amostra é constituída por 2,697 alunos do primeiro ano matriculados em 54 cursos no ano letivo 2015/16. A modelação multinível do desempenho académico considera o modelo de dois níveis com estudante como unidade de nível um e curso como unidade de nível dois. Os resultados corroboram que a nota de ingresso é o mais forte preditor do desempenho académico no primeiro ano. Adicionalmente, os resultados do modelo sugerem que 34% da variabilidade do desempenho académico é devida a diferenças entre os cursos e que 80% dessa variabilidade é explicada por quatro variáveis a saber: área de estudo; se a universidade é a primeira escolha do aluno; o nível de escolaridade dos pais ou responsáveis, sexo e idade dos alunos.

As taxas de abandono ou evasão no ES são hoje preocupação social e das instituições, justificando a implementação de políticas públicas de prevenção do fenómeno. Estas políticas carecem de estudos sobre as variáveis mais determinantes do risco de abandono ou evasão. Nesta linha de preocupações, o objetivo central do estudo apresentado no capítulo quinto, com título “Multilevel modeling of persistence in higher education” é analisar o impacto da trajetória escolar prévia do estudante e das condições de ingresso no ES, controlando a origem socioeconómica, a idade e o género dos estudantes. Aplicámos o modelo de regressão logística multinível, considerando alunos agrupados por curso, aos dados dos estudantes ingressados pela primeira vez no 1º ano na Universidade do Minho no ano letivo 2015/16. Os resultados do modelo sugerem que a reprovação no ensino básico tem efeito de longo prazo, isto é, a probabilidade de persistência durante o 1º ano é influenciada por esse atributo da trajetória escolar do indivíduo, mesmo controlando pela respetiva classificação de ingresso na universidade. De acordo com as estimativas obtidas, os alunos que declaram não ter reprovado no ensino básico têm maior razão de vantagens na persistência do que os estudantes que autodeclaram ter reprovado no ensino básico. A admissão do aluno no curso de 1ª opção, bem como a classificação de ingresso na universidade também estão associadas à probabilidade de persistência. Os resultados também mostram que os estudantes mais velhos e do sexo masculino têm menor probabilidade de persistência. No sexto capítulo aprofundámos a análise dos fatores individuais que influenciam a persistência do aluno no 1º ano através do estudo com título “Personal and institutional factors’ impact on persistence in higher education: A study with first-year Portuguese students”. Ou seja, pesquisámos sobre a associação entre a persistência do estudante com as expectativas e os métodos de estudo, controlando o histórico dos alunos, as características sociodemográficas e as condições de acesso ao ES. Os resultados sugerem uma forte influência dos métodos de estudo sobre a persistência, bem como um papel importante desempenhado pelas expectativas sobre a persistência dos alunos mais velhos. Além disso, este estudo testa a especificidade do curso através do efeito diferencial da nota de ingresso na persistência, tendo-se encontrado evidencia estatisticamente significativa da variação aleatória entre os cursos. No sétimo capítulo retomamos os argumentos apresentados por Almeida, Guisande, Soares, e Saavedra (2006) e analisam-se, uma década depois, as características sociodemográficas e percurso escolar dos alunos inscritos pela 1ª vez no 1º ano de 1º ciclo no ano letivo 2015/16 e a sua associação à escolha do curso, às expectativas sobre o ES, condições de ingresso, bem como a associação com o desempenho ao longo do 1º ano. Os principais resultados indicaram que o sexo, a origem

sociocultural das famílias e o percurso escolar estão associados à escolha do curso, à nota de candidatura e à opção de entrada. A análise de variância multivariada da classificação no final do 1º ano sugere que, ao nível de significância de 5%, o termo fixo de área científico-disciplinar e os termos de interação entre opção do curso e área científico-disciplinar são estatisticamente diferentes de zero. Não foi encontrada associação estatisticamente significativa de origem sociocultural com persistência no ES. A origem dos estudantes parece não influenciar a decisão de abandonar, suspender ou transferir-se de curso, pelo menos durante o 1º ano de estudos. Os resultados associados ao percurso escolar sugerem a resiliência e/ou a ação institucional na promoção da equidade no ES.

CAPITULO 2 – Reflexões sobre Modelo de Valor Acrescentado⁴

RESUMO

O artigo apresenta uma revisão sobre modelos de valor acrescentado e ligação à política pública em educação e ES, omitindo os aspectos técnico-metodológicos que são por natureza complexos. Assim, o artigo destina-se prioritariamente a pesquisadores e actores de política educativa. Também procede à revisão sistemática da literatura científica publicada em periódicos do Brasil e de Portugal registados na plataforma SciELO, donde resulta a escassa produção face à agenda científica imposta pelas potencialidades inerentes à melhoria da educação e ES e, conseqüentemente, à prosperidade da sociedade. Por fim, são enunciadas três etapas principais para o desenvolvimento de um sistema de avaliação que inclua o indicador de valor acrescentado.

Palavras-chave: Valor acrescentado. Avaliação. Responsabilização. Indicadores de desempenho.

ABSTRACT

The article presents a review on value added models and connection to public policy in education and higher education, omitting the technical and methodological aspects that are complex. Thus, the article is primarily aimed at researchers and actors of educational policy. It also proceeds to the systematic review of the scientific literature published in Brazilian and Portuguese journals registered in the SciELO platform, which results the scarce production in view of the scientific agenda imposed by the potentialities inherent to the improvement of education and higher education and, consequently, to the prosperity of society. Finally, three main steps are set out for the development of an evaluation system that includes the value added indicator.

Keywords: Value added; Evaluation; Accountability; Performance indicators.

RÉSUMÉ

L'article présente un aperçu des modèles à valeur ajoutée et la connexion à la politique publique dans l'éducation et l'enseignement supérieur, en omettant les aspects techniques et méthodologiques qui sont complexes par nature. Ainsi, l'article est destiné principalement aux chercheurs et acteurs de la politique éducative. Réalise également l'examen systématique de la littérature scientifique publiée dans des revues du Brésil et du Portugal enregistrés dans la plate-forme SciELO, d'où la faible production en raison de l'agenda scientifique imposé par le potentiel inhérent à l'amélioration de l'éducation et de l'enseignement supérieur et, par conséquent, à la prospérité de la société. Enfin, ils ont mis trois étapes principales dans le développement d'un système d'évaluation qui comprend l'indicateur de la valeur ajoutée.

Mots-clés: Valeur ajoutée; Évaluation; Responsabilité; Indicateurs de performance.

⁴ Baseado no artigo Ferrão, M. E. (2017). Estatística educacional e política pública: A propósito dos modelos de valor acrescentado. *Educação & Sociedade*, 39(142), 19–38. doi:10.1590/es0101-73302017176230.

2.1 Introdução

A vertente transdisciplinar da Estatística tem-se desenvolvido ao longo dos últimos anos com novos temas e desafios. As últimas duas décadas foram portadoras de metodologia que nos permite desvendar a complexidade do mundo real (Goldstein, 1998), em especial na subárea da Estatística Educacional que é fundamental ao desenvolvimento disciplinar e científico da Educação. Os avanços tecnológicos e computacionais, notadamente no que se refere aos métodos computacionais e estatísticos para lidar com *large and complex surveys* e dados longitudinais, potenciam um salto gigantesco na produção de conhecimento nesse campo. Ora, a Declaração da *American Statistical Association* (ASA) sobre modelos de valor acrescentado, suscita uma nova agenda de investigação científica em estatística educacional, mas também em política educativa (Morganstein & Wasserstein, 2014). Neste caso, uma das principais exigências é produzir indicadores úteis para os atores políticos nos vários níveis da tomada de decisão e, simultaneamente, demonstrar que a qualidade de tais indicadores é objectivamente superior e mais abrangente do que a que está disponível através de outras abordagens. A Declaração é parte integrante da missão da ASA, enunciada da seguinte forma: “Part of the ASA’s mission is promoting sound statistical practice to improve public policy and improve human welfare” (citado em Morganstein & Wasserstein, 2014, p. 108). Tal como se mostrará adiante, a definição de valor acrescentado ou valor agregado é muito intuitiva. Contudo, a complexidade envolvida na especificação dos modelos, no procedimento de estimação e outros aspetos técnico-metodológicos, afasta a atenção dos mentores e atores de política educativa do benefício que tais modelos e respectivos indicadores representam para programas de melhoria da educação e do ES. As barreiras entre a pesquisa de base quantitativa e o processo político podem ser mitigadas se os pesquisadores, avaliadores, formuladores e/ou analistas de política educativa tirarem o máximo proveito da interdisciplinaridade inerente ao objecto de estudo. A Declaração também refere que a comunidade estatística norte-americana está envolvida na compreensão e melhoria dos modelos estatísticos que podem ser usados em educação e que a comunidade tem um papel fundamental na colaboração com as organizações educativas que desejam usar dados para melhorar a educação. De facto, o debate em torno dos modelos de valor acrescentado emerge da onda crescente na implementação de sistemas de avaliação educacional que têm como principal propósito a responsabilização - prestação de contas com consequências de alto risco

para indivíduos e instituições. Cabe aqui uma nota para referir que o termo 'alto risco' constitui a tradução livre de '*high stakes*' (e.g. Ferrão & Klein, 2013), geralmente usado em avaliações que têm muita influência na vida de estudantes, de professores e/ou nos planos de actividades e de acção das instituições. Alguns autores brasileiros têm preferido a tradução 'alto impacto'. Por exemplo, a responsabilização - prestação de contas (*accountability*) foi um dos princípios enunciados para a boa governança das organizações da União Europeia, a par de outros quatro: abertura, participação, eficácia e coerência (Commission of the European Communities, 2001), princípios vertidos na legislação dos Estados-membros.

Por todo o mundo, tem-se verificado a implementação, em ritmo acelerado, de sistemas de avaliação que visam a comparação do desempenho profissional e institucional, apesar de ainda nos depararmos com desafios metodológicos e técnicos consideráveis para que os resultados da avaliação contribuam efectivamente para a melhoria da educação e para a prosperidade da sociedade (Ferrão, 2016). A comparação do desempenho pode ser bastante produtiva para os envolvidos quando é bem feita. Quando é mal feita, pode ser muito custosa e não apenas inútil, mas prejudicial e até mesmo destrutiva (Bird et al., 2005). De acordo com Foley e Goldstein (2012), nos últimos anos o ES. Com o número de universidades crescendo drasticamente e a introdução e/ou elevação das taxas de frequência, a elaboração de rankings foi formulada em termos de um "mercado" crescente no ES. Os autores procedem à revisão que estabelece as características no monitoramento do desempenho e dos rankings no ES, bem como dão recomendações para melhorar o seu uso e abordar os aspectos mais problemáticos desse sistema de avaliação.

Contribuições recentes têm mostrado que é possível obter estimativas da eficácia educacional (eficácia escolar e eficácia docente) na promoção das aprendizagens. Apesar das limitações de tais abordagens, alguns mentores de política educativa têm sido excessivamente entusiastas na incorporação dessas estimativas como indicadores de desempenho nos sistemas de avaliação educacional dos seus países. As propriedades associadas a esses indicadores são cruciais para a legitimidade desses sistemas, particularmente quando são usados com consequências de alto risco. A maioria dos pesquisadores em estatística está consciente quanto às implicações da assunção de determinados pressupostos, ou da sua violação, nos resultados obtidos; dependendo do uso dado àqueles indicadores, assim serão os efeitos nas pessoas e instituições envolvidas.

Este artigo tem como principal objetivo geral contribuir para a divulgação do conhecimento sobre modelos de valor acrescentado e respectivos indicadores junto de uma audiência não interessada na complexidade dos aspectos técnico-metodológicos. Principalmente, uma audiência formada por pesquisadores e mentores e decisores de política educativa. O artigo está estruturado em quatro secções além desta introdução. Na secção 2 abordam-se alguns aspectos sobre a ligação entre a estatística e a política pública em educação e ES na perspetiva da concepção e implementação de políticas orientadas pelo conhecimento de base científica. Na secção três apresentam-se, ao nível introdutório, uma revisão sobre valor acrescentado em educação, considerando a análise de diferentes definições registadas na literatura especializada, características com implicação de ordem metodológica e finalidade. Na secção três analisa-se o tema a partir da literatura científica publicada no Brasil ou em Portugal e, na secção quatro, apresentam-se algumas considerações finais.

2.2 Estatística e política pública em Educação

O conhecimento de que dispomos sobre valor acrescentado em Educação é o resultado de, pelo menos, 45 anos de investigação científica. O caminho percorrido mostra avanços e recuos. Especialmente nas Ciências Sociais, o conhecimento científico assenta em cânones que o caracterizam como sistemático, replicável, generalizável, entre outros. Obedecer a esses cânones requer tempo. É a dinâmica da Ciência. Uma dinâmica lenta face à premência de soluções com a cientificidade desejável. Entre os marcos do conhecimento científico que traçaram o sentido irreversível do progresso científico- disciplinar da Educação, encontra-se aquele que deu lugar ao título *School Matters* (Mortimore, Sammons, Stoll, Lewis, & Ecob, 1988) e que está intrinsecamente associado ao desenvolvimento metodológico de valor acrescentado em educação, em especial ao desenvolvimento dos modelos de regressão multinível (Ferrão, 2011; Mortimore, Sammons, & Thomas, 1994; Saunders, 1999). Os modelos de valor acrescentado são caracterizados por elevada complexidade, podendo comprometer a sua interpretabilidade tão desejável à tomada de decisão política. Para ser confiável e amplamente aceite, o modelo deve ser parcimonioso, transparente e interpretável. O caminho a percorrer faz-se de redes multidisciplinares, estabelecendo *pipelines* de comunicação e acção, que permitam transformar a sociedade a partir de conhecimento científico. Ou seja, a ligação entre a estatística educacional e a política pública em Educação

visa a essa transformação no sentido da melhoria. Algo semelhante ao que acontece coma sub-área de avaliação educacional,

a ideia de melhoria educacional e, de modo mais amplo, melhoria social está implícita a qualquer processo de avaliação. Pois tem-se verificado em ritmo acelerado sistemas de avaliação que visam à comparação de desempenho profissional e institucional, apesar de ainda nos depararmos com desafios metodológicos e técnicos consideráveis para que os resultados efectivos da avaliação contribuam efectivamente para a melhoria da educação e para a prosperidade da sociedade. (Ferrão & Klein, 2013, pp.7-8)

O relatório *Getting Value Out of Value-Added* publicado pela *National Academy of Sciences* (Braun, Chudowsky, & Koenig, 2010), reporta o parecer de alguns especialistas sobre as vantagens e uso dos modelos de valor acrescentado, tais como:

Adam Gamoran and Robert Gordon, among others, focused on what they saw as the advantages of these models over indicators based on student status. Gordon observed that although many important technical issues still need to be resolved, it is not realistic to think that policy makers will wait 20 years until all of the difficulties are worked out before making use of such methods. Decisions about schools and teachers are being made, and, as Jane Hannaway noted, there is enormous demand from the policy side, to which the testing and research communities need to respond as quickly as possible. (op. cit., pp. 55-56)

Apesar da vantagem dos modelos de valor acrescentado face a indicadores de status ou de modelos de resultados contextualizados, o seu uso com o propósito de responsabilização - prestação de contas, com consequências de alto risco ao nível dos indivíduos e das instituições ainda deve ser encarado com cautela. Se não é realista pensar que os mentores e decisores de política educativa esperarão mais 20 anos até que todas as limitações e dificuldades estejam resolvidas para então fazerem uso desses modelos, urge que a academia e pesquisadores respondam a tais desafios com a rapidez necessária, sendo, desse modo, crucial reforçar a comunicação entre os dois grupos profissionais, isto é, entre a academia/pesquisadores e os mentores/decisores de política educativa.

2.3 Valor Acrescentado em Educação

Esta seção apresenta diversas definições de modelo de valor acrescentado (MVA) em Educação, aspectos comuns que representam potencialidades, mas também limitações, e debate as principais finalidades do seu uso por meio de exemplos da literatura científica.

2.3.1 Definições

Na literatura científica publicada em língua portuguesa, encontram-se as designações sinónimas “valor acrescentado” e “valor agregado”. Apesar de considerarmos relevante a análise da terminologia usada numa perspetiva filológica, ela ultrapassa o domínio deste artigo. Por uma questão de simplicidade doravante usaremos o termo “valor acrescentado” por ter sido o primeiro a ser usado na literatura científica publicada em língua portuguesa.

Na actualidade há consenso amplo relativamente aos requisitos metodológicos mínimos que devem ser satisfeitos para apurar o valor acrescentado em Educação. O relatório *Measuring Improvements in Learning Outcomes, best practices to assess the value-added of schools* (OECD, 2008) define MVA como pertencendo à classe de modelos estatísticos que permitem estimar a contribuição das escolas para o progresso dos alunos. Isto assume particular relevância no que se refere aos objetivos educativos estabelecidos, sendo o progresso medido, pelo menos, em duas ocasiões no tempo. Noutros termos, mas com significado semelhante, a *National Academy of Sciences* (EUA) no relatório *Getting Value Out of Value-Added* (Braun et al., 2010), menciona que a abordagem se refere à quantificação do efeito da instituição, do professor ou do programa no desempenho dos alunos, considerando as diferenças entre os alunos no que se refere ao conhecimento prévio e também a outras variáveis extra-escolares que influenciam o desempenho dos alunos. O volume 26 do *Handbook of Statistics*, no capítulo dedicado à modelação de valor acrescentado (Braun & Wainer, 2007), apresenta-os como sendo a família de modelos estatísticos que são usados para inferir sobre a eficácia das unidades educativas, usualmente instituições e/ou professores. Estas definições são convergentes num conjunto de características sobre as quais se refletirá nas próximas secções. Todavia, parece-nos importante registar que alguns desses conceitos e definições, tais como o de *efeito escola*, têm evoluído ao longo do tempo. Também no que se refere à operacionalização e quantificação de alguns deles têm-se verificado desenvolvimentos, muitos

deles decorrentes das possibilidades metodológicas ao dispor dos pesquisadores. O artigo sobre os resultados da pesquisa de eficácia escolar em países de língua portuguesa, especificamente Brasil e Portugal (Ferrão, 2014) apresenta uma breve síntese de definições em eficácia escolar e efeito escola. Entre outros autores, cita Teddlie e Reynolds (2000) para referir que nos últimos 50 anos surgiram pelo menos seis definições de efeitos escolares: efeito absoluto da escolarização (grupo experimental versus controle); média não ajustada da proficiência dos alunos por escola; impacto da escolarização na média da proficiência dos alunos por escola, ajustada por variáveis de antecedentes familiares e/ou conhecimento prévio; quantificação da proporção da variação dos resultados escolares, entre escolas, face à variação total; medida do efeito exclusivo de cada escola sobre os resultados dos alunos; mensuração do impacto das escolas sobre o desempenho dos alunos ao longo do tempo. Outros detalhes podem ser encontrados em Teddlie e Reynolds (2000, pp. 65-69). As repercussões da escolha de uma ou outra definição e/ou operacionalização no âmbito de um sistema de avaliação educacional são, em parte, debatidas em Ferrão (2012, p. 460), do seguinte modo:

Estudiosos em avaliação educacional são unânimes em reconhecer a importância de dispor de vários indicadores para formar juízo de valor relativamente ao(s) critério(s) formulado(s) no sistema de avaliação e, assim, suportar adequadamente decisões de política educativa. Não se espera que as “melhores” escolas de acordo com um indicador sejam também as “melhores” de acordo com outro indicador. Cada um deles pode revelar uma faceta diferente do mesmo objecto. Por exemplo, se o objetivo é demonstrar que os alunos atingem determinado nível de desempenho estabelecido como meta, então o indicador de status (reflecte o desempenho, resultados escolares, num determinado momento do tempo) é o mais apropriado. Convém, todavia, levar em conta que algumas das escolas, onde os alunos atingem níveis de desempenho elevado, pouco contribuem para esse desempenho. Geralmente, são escolas frequentadas por alunos com elevado nível de desempenho à entrada (escolas selectivas). Essas crianças e jovens são normalmente provenientes de grupos sociais favorecidos. Em contrapartida, as escolas que recebem alunos com baixo desempenho e onde, no final do período sob avaliação, o desempenho se mantém aquém da meta, podem, apesar disso, ter contribuído fortemente para alavancar a aprendizagem desses alunos e, portanto, apresentam elevado valor acrescentado. A sua contribuição

não está reflectida no indicador de status, mas sim no indicador de valor acrescentado. Este é o caso típico das escolas cuja população discente é extremamente desfavorecida do ponto de vista socioeconómico. Um sistema de incentivos orientado pelo indicador de status tende a premiar escolas com reduzido valor acrescentado e a desvalorizar o serviço educativo prestado por aquelas que trabalham junto de populações onde a escola marca a diferença.

No que se refere ao ES, ao longo das últimas três décadas, verificou-se o consenso internacional sobre a necessidade de medição da qualidade académica, o que se reflectiu no aumento da produção de rankings institucionais e de cursos (Dill & Soo, 2005). De acordo com Foley e Goldstein (2012), os rankings de desempenho destinam-se a dois propósitos principais: o primeiro pode ser descrito como responsabilização- prestação de contas e o segundo é fornecer informação aos usuários dos serviços públicos de educação e ES. De acordo com Hazelkorn (2011), após a crise financeira de 2008, os rankings das universidades passaram a simbolizar o impulso para maior responsabilização e transparência, fornecendo a evidência subjacente à agenda de modernização e à adoção de novas práticas de gestão pública.

Foley e Goldstein (2012) recomendam que qualquer ranking deve ser contextualizado, através da abordagem de valor acrescentado. Adicionalmente, Goldstein (2014) lembra que apesar do uso de modelos de valor acrescentado estar razoavelmente bem estabelecido na educação não superior, tal não acontece no ES porque é difícil obter variáveis de controle mensuradas na unidade estatística aluno e ter os estudantes associados às respectivas instituições de ES nas bases de dados disponíveis para análise.

No que concerne à aplicação de modelos de valor acrescentado no ES brasileiro alguns pesquisadores defendem-na como sendo a abordagem metodológica mais adequada, uma vez que “a efetividade de uma instituição de ensino deve ser medida pelo que acrescenta ao conhecimento do aluno durante sua permanência nela, e não pela nota final obtida por ele em determinado exame” (Soares, Ribeiro, & Castro, 2001, p. 365).

2.3.2 Características, potencialidades e limitações

As definições de valor acrescentado anteriormente apresentadas têm três características em comum. A primeira característica é a da existência duma família ou classe de modelos. Assim,

a expressão “o modelo de valor acrescentado” denota uma escolha entre diversos modelos estatísticos possíveis. Na recente revisão sistemática elaborada por Everson (2016) sobre MVA para efeito de responsabilização-prestação de contas, a autora refere o seguinte: “A literatura selecionada para esta revisão mostra a diversidade nas escolhas de modelagem que são feitas, refletindo a complexidade e grande variedade de possibilidades na criação de modelos VA” (p. 48). Assim, tal como defendemos em 2012, a escolha do modelo estatístico deve ser realizada em concordância com o critério de comparação estabelecido pelo *designer* do sistema de avaliação e de modo adequado ao contexto educativo, social, económico e cultural em que o modelo vai ser aplicado (Ferrão, 2012a). Naturalmente, verificam-se diferenças entre os indicadores VA resultantes da aplicação de cada um dos modelos possíveis, e, à luz dessas diferenças, deverá ser realizada a respectiva interpretação dos resultados obtidos (Ferrão, 2010). O princípio da transparência recomenda que qualquer que seja a abordagem de valor acrescentado usada, os métodos estatísticos e *software* devem ser completamente descritos, testados e documentados (Morganstein & Wasserstein, 2014). Embora os estudos de consistência, e da correlação forte entre as estimativas obtidas por modelos com diferentes especificações, sejam encorajadores do seu uso como critério de avaliação educacional (Ferrão, 2010; Ferrão & Goldstein, 2009), é necessário ter em mente que a classificação da instituição (ou do professor) através da aplicação do modelo pode depender da escolha do modelo. De acordo com Everson (2016), a inclusão ou a exclusão de covariáveis no modelo deve depender antes de tudo da resposta a algumas questões de natureza filosófica, pois na comparação inerente ao processo de avaliação, nem sempre se sabe com quem um professor ou os alunos de uma escola devem ser comparados. Sem responder a tais questões, as decisões podem continuar a ser tomadas com alguma arbitrariedade e apenas baseadas nos dados disponíveis ao invés de levar em consideração os propósitos ou finalidades do sistema de avaliação. Na mesma linha de pensamento, Ferrão (2012, p.460) considera que:

[...] as decisões e escolhas inerentes à especificação funcional do modelo estatístico devem consubstanciar, tanto quanto possível, as opções de natureza política e/ou programática estabelecidas pelos governos. Assim, de modo estruturado e formal, o sistema de avaliação educacional, concebido nas dez dimensões – definição; função; objecto; tipo de informação a recolher face a cada objecto; critério para formar juízo de valor; a quem se destina a avaliação (*stakeholders*); processo; método; quem é avaliador e constituição das equipas de avaliação; padrões que regem uma avaliação

(Nevo, 1983, 2006) –, acomoda diversas e complementares vertentes de avaliação, tais como a avaliação das aprendizagens, de métodos de ensino, de materiais, do desempenho de pessoal docente, da avaliação institucional, da auto-avaliação, da avaliação externa, da avaliação de programas e de políticas.

E, portanto, “o critério no âmbito da auto-avaliação da escola ou da avaliação de materiais pode não coincidir com o critério no âmbito da avaliação institucional” (Ferrão, 2012, p. 461).

A segunda característica está relacionada com os cuidados a ter para tornar comparável a acção educativa e formativa das instituições envolvidas já que cada uma delas serve população discente com características diferentes. Para estimar a “contribuição da escola” ou a “eficácia das unidades educativas” é absolutamente necessário ter em mente que a população discente que frequenta a rede pública de escolas é muito heterogénea e que essa heterogeneidade também se verifica entre instituições. Algo similar ocorre no ES onde a necessidade de contextualização é referida pela maioria dos pesquisadores (e.g. Foley & Goldstein, 2012; Goldstein, 2014; Soares et al., 2001). Logo, é preciso que o modelo contemple variáveis extra-escolares ou de contexto que têm influência no desempenho dos alunos e, assim, garanta que a estimativa da “contribuição da escola” ou “contribuição da universidade” esteja livre dessa influência.

A terceira característica envolve a noção de “progresso”. O progresso está associado a um atributo que varia no tempo. Para quantificar essa mudança ao longo do tempo e, por conseguinte, o progresso, são necessários dados longitudinais – de tal maneira que seja possível captar as alterações decorrentes da aprendizagem realizada de permeio. Por um lado, a infra-estrutura técnica para a realização de estudos longitudinais é exigente. Por outro, a construção da métrica que permita quantificar o progresso do aluno não é trivial. Neste caso, não se trata de uma dificuldade específica ou limitação dos modelos de valor acrescentado, mas antes do processo de aferição das aprendizagens. Geralmente, usa-se a classificação dos alunos em testes padronizados como medida dos resultados escolares. A tal respeito, alguns pesquisadores norte-americanos (Lockwood et al., 2007) mostraram que a escolha dos instrumentos de aferição (testes) causaram maior instabilidade nas estimativas de valor acrescentado do que a escolha do modelo estatístico, sugerindo que o problema detectado poderia ser ainda maior se tivessem sido usados instrumentos desenvolvidos por diferentes

produtores de testes. Problema similar é também tratado por McCaffrey, Sass, Lockwood, e Mihaly (2009).

Alguma resistência à abordagem VA radica na crítica à (suposta) unidimensionalidade do indicador. As metodologias estatísticas permitem considerar múltiplas dimensões dos resultados escolares, quer elas sejam aferidas através de variáveis contínuas ou discretas. Encontram-se estudos que consideram também o comportamento disciplinar dos alunos, a sua assiduidade, a auto-estima, etc. Aquela crítica está, portanto, desfocada. Levanta-se então a questão: que resultados escolares devem ser considerados nos modelos de valor acrescentado? O propósito da Educação é assegurar que crianças e jovens se tornem cidadãos aptos nas dimensões pessoal, social, política e económica. Nestas dimensões há aprendizagens que são da estrita esfera da Escola enquanto outras não o são. A valia da abordagem VA e o seu uso ao serviço da melhoria da qualidade da Educação depende, assim, da definição do papel da Escola e dos objetivos educativos que são da sua inteira responsabilidade e quais são os da responsabilidade dos demais intervenientes. Depende, também, da garantia de níveis aceitáveis de validade, fiabilidade e comparabilidade das medidas de resultados escolares. Só então valerá a pena elaborar sobre a finalidade ou o uso do valor acrescentado em Educação (Ferrão, 2011, 2012b).

2.3.3 Finalidade: Prestação de contas-responsabilização, escolha da escola ou melhoria da Educação e Ensino Superior

No sentido do avanço gradual da autonomia das escolas, espera-se cada vez menos regulação nos *inputs* e nos processos e cada vez mais monitorização e avaliação do sistema educativo e do desempenho das escolas. Quer isto dizer que aspectos organizacionais, de liderança e de natureza pedagógica, por exemplo, devem ser cada vez mais da responsabilidade da escola (ao nível da qual a auto-avaliação ocupa um papel determinante). Sabendo que o indicador de valor acrescentado representa a contribuição da escola na aprendizagem realizada durante um certo período de tempo pelos seus alunos, defendemos que ele é indispensável a um sistema de monitorização e avaliação que tenha como desígnio a melhoria da qualidade da Educação. Aliás, algo consensual em termos internacionais, pois diversos países têm vindo a contemplá-lo nos respectivos sistemas, nem sempre coincidindo quanto à finalidade da sua utilização. De um modo geral, as principais finalidades são as seguintes: 1) Incorporar a lógica

da responsabilização e da prestação de contas nas instituições de educação/formação; 2) Fornecer informação a pais e encarregados de educação para a escolha da instituição de educação/formação; 3) Fomentar a implementação de programas de melhoria. Estas três finalidades diferem quanto ao foco e quanto aos programas de desenvolvimento. Todavia, elas podem não ser mutuamente exclusivas dependendo, entre outros aspectos, do enquadramento conceptual onde se insere o programa de desenvolvimento. A prestação de contas deve ser globalmente entendida nas suas dimensões política, pública, de gestão e administração, profissional e pessoal. A responsabilização - e, por conseguinte, a respectiva prestação de contas - atribuída ao governo, departamentos centrais e regionais/federais, escolas/instituições de ES, encarregados de educação e alunos, deve ser tal que, articulada e cumulativamente, sejam criados os meios, o contexto e o apoio para que as escolas possam maximizar as aprendizagens efectuadas por cada um dos seus alunos. Assim, o indicador de valor acrescentado ajuda a rastrear as escolas que pertencem sistematicamente a grupos extremos de desempenho (Ferrão & Couto, 2014). Servindo como critério objetivo para estabelecer prioridades de intervenção nas escolas com valor acrescentado deficitário, a sua principal finalidade pode ser a de identificar, a tempo, unidades educativas que carecem de acompanhamento externo no processo de mudança ou de melhoria. Ou seja, no pormenor das práticas educativas em todos os domínios da prestação do serviço educativo visando o desenvolvimento de um programa específico para a superação das disfunções detectadas.

Do mesmo modo, o critério ajuda a identificar escolas de referência, das quais podem ser extraídas lições sobre boas práticas generalizáveis através de programas de melhoria. Nestes termos, o seu uso para determinar as prioridades de atuação e a elaboração, implementação e acompanhamento de programas de melhoria, encerra simultaneamente dimensões de política, de gestão e administração - profissional e pessoal - da prestação de contas, com a cautela decorrente do impacto nas pessoas e nas instituições (Bird et al., 2005; Ferrão, 2012b). Em alguns países, tais dimensões estão apostas à dimensão profissional, por meio da atribuição de prémios de desempenho, de mérito, ou de outros incentivos individuais. Estes mecanismos de prestação de contas ganharam muita popularidade entre governantes e mentores de política. Os argumentos a favor são principalmente de ordem racional, não havendo até à data a demonstração da existência de uma relação causa-efeito entre a prática de incentivos do tipo *performance-related pay* e a melhoria das aprendizagens.

Nesta linha, mesmo supondo que os incentivos individuais se baseiam em múltiplas medidas do desempenho ou do mérito, 100% justas e 100% fiáveis, convém levar em consideração a natureza da profissão e o efeito de competitividade gerada entre professores pelo mecanismo de incentivos individuais. Ora a competitividade, que é de salutar em outros sectores de actividade, pode ser um obstáculo ao trabalho colaborativo na Educação, sendo este considerado fator determinante do sucesso educativo. Note-se que o processo de criação de valor, produção e transmissão de conhecimento é acentuadamente marcado por externalidades positivas resultantes da cooperação e difusão de conhecimento e práticas de aprendizagem colectiva e não da retenção desse conhecimento como estratégia de reforço do posicionamento individual dos profissionais perante avaliações futuras. Adicionalmente, a avaliação de desempenho de professores com consequências de alto risco deve ser feita com muito cuidado para evitar situações tais como a relatada por Anderson (2013). De acordo com o relato, um profissional foi despedido na sequência da avaliação de desempenho com consequência de alto risco, tendo-se posteriormente concluído que havia um erro no apuramento. A pontuação resultante da abordagem de valor acrescentado representava 35% na avaliação de desempenho dos professores. Os argumentos de ordem racional radicam essencialmente na necessidade de valorização da profissão docente, considerando que a qualidade da prática docente é muito importante para as aprendizagens dos alunos. Logo, a valorização da profissão contribuirá para atrair e reter os melhores para a docência. A este propósito, Clotfelter, Ladd, Vigdor, & Diaz, (2004) apresentam evidência de que, comparando o período posterior à entrada em vigor do sistema de prestação de contas com o período anterior, as escolas da Carolina do Norte cuja população discente se caracterizava por maior proporção de alunos com défice educativo (ou menor desempenho) passaram a ter maior dificuldade em atrair e reter “bons” professores. Mesmo admitindo que tal tivesse ocorrido ao acaso, verificou-se a diminuição da estabilidade do corpo docente nessas escolas o que, por si só, já tem um custo educativo a ser considerado. A este respeito, estudos sobre a estabilidade do corpo docente mostram que é mais provável um professor bem preparado (aferição pelo *National Teacher Exam* ou instrumento similar) abandonar a profissão do que um mal preparado.

Goldstein (2014) refere-se à utilidade dos rankings, quando os dados adequados à sua produção estão disponíveis. O autor alerta para a incerteza que lhes está associada e que pode ser representada através de intervalos de confiança. Acrescenta, ainda, que talvez o uso mais

eficaz dos rankings institucionais seja como instrumento de rastreamento que possa sugerir onde os problemas estão ocorrendo, ao invés de diagnósticos de quais são os problemas.

2.4 Valor acrescentado na literatura científica de língua portuguesa

A pesquisa dos artigos na área de Educação publicados em língua portuguesa foi realizada na base de dados SciELO considerando, de acordo com o teor da plataforma, o método de pesquisa integrada, na colecção de periódicos do Brasil e de Portugal, com as palavras “valor acrescentado” ou “valor agregado” em todos os índices e com filtro WoS Áreas Temáticas: Educação e pesquisa educacional. A pesquisa foi efectuada no dia 13 de Fevereiro de 2017 e teve como resultado oito artigos. Na primeira fase de análise foram excluídos dois deles por estarem fora do âmbito deste artigo. Um deles, publicado num periódico de ensino da física, trata de um modelo de simulação em processos de polimerização. O outro, publicado no periódico Educação & Sociedade, trata do valor acrescentado, na perspetiva profissional, no mercado das escolas de estilismo. Entre os seis artigos restantes, o critério de inclusão estabelecido foi o artigo visado apresentar explicitamente a definição e operacionalização do conceito de valor acrescentado. Os cinco artigos incluídos na análise encontram-se resumidos na Tabela 1, que simplificadamente enuncia a contribuição do artigo, o propósito ou finalidade, bem como o enquadramento ou referencial teórico em que o artigo se desenvolve, através das principais referências bibliográficas usadas para suportar metodologicamente o respectivo artigo.

Qualquer dos cinco artigos traz ao debate as diferentes possibilidades de modelação na produção de indicadores de desempenho institucional e profissional. Dois deles (Ferrão & Couto, 2013a; Soares, Bonamino, Brooke, & Fernandes, 2017) abordam o tema a partir do modelo de resultados contextualizados e do de valor acrescentado aplicados aos dados do GERES 2005, norteando a utilização com vantagem no apuramento do valor acrescentado da escola. Todavia, torna-se mais exigente uma vez que obriga à existência de dados longitudinais. O debate em torno do desempenho profissional com consequências de alto risco é tratado em detalhe por Braun (2013) invocando os achados de pesquisa científica norte-americana. Outra ideia partilhada em todos os trabalhos é a de que um bom sistema de avaliação deve conter múltiplos indicadores. Nas palavras de Braun (2013, p. 127), “multiple indicators are not only necessary for fair evaluations, but also for the accountability system to contribute to the

systematic and sustained improvement of teaching and learning”. A necessidade de prestação de contas através da informação à sociedade também é lembrada por (Soares et al., 2017).

Os autores referem o seguinte:

Acreditamos que a principal informação a ser prestada à sociedade deve ser os níveis de proficiências alcançados pelos alunos ao final das etapas escolares, explicados por meio de uma interpretação pedagógica e, ainda, contextualizados por meio de medidas como as da condição socioeconômica dos alunos e, também, por meio de medidas relevantes de processos escolares. Mas, para amortizar os conflitos que os resultados brutos trazem, e não ser injusto com o esforço da comunidade escolar, nos mais variados contextos sociais e políticos, uma medida direta da qualidade da escola pode ser uma informação útil, senão para a comunidade escolar se espelhar nos resultados, pelo menos para que os gestores possam formular suas políticas públicas adequadamente. (Soares et al. 2017, p. 83)

Tabela 1. Aspectos tratados nos artigos sobre valor acrescentado

Artigo	Contribuição	Propósito ou finalidade	Enquadramento Valor Acrescentado
(Ferrão, 2012a)	Revisão e reflexão sobre modelo de valor acrescentado: definição, características, especificações e requisitos metodológicos, uso e finalidades, limitações.	Ligação entre valor acrescentado e sistema de avaliação educacional.	(Braun et al., 2010; Braun & Wainer, 2007; Goldstein & Spiegelhalter, 1996; OECD, 2008)
(Braun, 2013)	Revisão e reflexão sobre modelo de valor acrescentado e política de prestação de contas-responsabilização com consequências de alto risco.	Resumir os principais resultados de pesquisa nos Estados Unidos; sugerir modos de uso para a melhoria do ensino.	(Braun, 2005; Braun et al., 2010; Reardon & Raudenbush, 2009)
(Ferrão & Couto, 2013)	Debate e comparação de modelo de resultados contextualizados (<i>status</i>) e de valor acrescentado.	Mostrar evidência comparativa do indicador de <i>status</i> e de valor acrescentado, estudo das propriedades de consistência e estabilidade.	(Braun & Wainer, 2007; OECD, 2008; Sammons, Thomas, & Mortimore, 1997; L. Saunders, 1999)
(Brooke, Fernandes, Miranda, & Soares, 2014)	Comparação de duas abordagens de modelação do valor acrescentado, considerando variáveis de alunos, turmas, escolas.	Identificar factores que influenciam a aprendizagem ano a ano e os padrões de crescimento ao longo dos anos.	Não consta.
(Soares et al., 2017)	Comparação de modelo de resultados contextualizados (<i>status</i>) com modelo de valor acrescentado.	Verificar a possibilidade de quantificar a eficácia escolar através do modelo de <i>status</i> , comparando com o de valor acrescentado. Estudo de estabilidade.	(Braun & Wainer, 2007; Ferrão, 2014; Ferrão & Couto, 2013; Reardon & Raudenbush, 2009; Rubin, Stuart, & Zanutto, 2004)

Fonte: elaborada pela autora.

Todos os artigos elencam os desafios e limitações que justificam uma agenda de pesquisa muito exigente,

Os resultados apresentados estão ainda limitados pela ocorrência de erros não amostrais que ultrapassam o âmbito deste artigo mas que é fundamental investigar em trabalhos futuros. Um é o do erro da medida, e o outro é o do erro de representatividade ou viés de selecção devido ao atrito. (Ferrão & Couto, 2013, p. 155)

Contudo, os autores chamam a atenção para outros fenómenos revelados pela modelação, tal como “a importância da informação dada pelo modelo referente ao desempenho relativo de grupos minoritários” (op. cit., p. 155) e que, por serem recorrentes, reclamam a atenção dos mentores de política educativa. Ainda sobre as limitações da investigação reportada, Ferrão (2012, p.462) cita Cronbach e Suppes (1969, p.15) para lembrar boas práticas na condução de investigação científica, isto é,

Qualquer que seja a natureza do estudo, se é rigoroso, o investigador antecipadamente estabeleceu as questões pertinentes, instituiu controles nas etapas de recolha de informação e o raciocínio associado a essas questões, de modo a evitar possíveis fontes de erro. Se os erros não puderem ser evitados, o investigador deve levá-los em consideração, discutindo as limitações das suas conclusões (Tradução livre da autora).

De facto, o reduzido número de artigos publicados, visíveis na plataforma SciELO, nas colecções de Brasil e Portugal é a manifestação de uma extensa agenda de pesquisa e avaliação educacional que está por concretizar. A comunidade científica tem a sua quota de responsabilidade na melhoria da Educação, especificamente na transferência do conhecimento científico (Ferrão, 2009, p.52). Todavia, importa invocar a ideia de avaliação educacional como uma actividade que acontece no contexto político (Weiss, 1975). A produção e comunicação de conhecimento científico que supere as limitações e dificuldades enunciadas ao longo deste artigo está fortemente dependente da política para a ciência, em especial da que concerne ao incentivo/financiamento das subáreas em que os avanços determinantes são obstaculizados pela falta de massa crítica (Ferrão, 2012a). O reduzido número de artigos e falta de massa crítica sobre estas matérias e afins é ainda mais visível na perspetiva comparativa internacional. Note-se, como exemplo, que o artigo de revisão sobre modelos de

valor acrescentado publicado *online* em 2016 (Everson, 2017) no periódico *Review of Educational Research* (com SJR no 1º percentil superior ao longo dos últimos anos na área de Ciências Sociais e subárea Educação) contém 112 referências bibliográficas. Inclui apenas um dos pesquisadores que constam na Tabela 1 anteriormente apresentada. Importa referir que a capacidade de pesquisa científica instalada está condicionada à disponibilidade de dados longitudinais que satisfaçam os requisitos metodológicos enunciados ao longo do artigo.

2.5 Considerações Finais

A Declaração da *American Statistical Association* (Morganstein & Wasserstein, 2014) sobre MVA, suscita uma nova agenda de pesquisa científica em estatística educacional, tal como foi preconizado por Ferrão e Klein (2013), mas também em política educativa. Ao longo deste artigo apresentou-se uma revisão sobre MVA e respectivos indicadores, destinando-se prioritariamente a pesquisadores e a mentores/decisores de política educativa, isto é, a uma audiência não interessada na complexidade dos aspectos técnico-metodológicos dos modelos. Também se procede à revisão sistemática na literatura científica publicada em periódicos do Brasil e de Portugal, registados na plataforma SciELO, donde resulta a diminuta produção face à agenda científica imposta pelas potencialidades inerentes ao tema.

Finalmente, em jeito de síntese e conclusão, vale a pena registar, de acordo com OECD (2008), as etapas necessárias para um país dispor de um sistema de monitorização e avaliação que inclua o indicador de valor acrescentado deliniam-se em seguida: (1) Decidir quanto ao propósito do indicador de valor acrescentado e outros na perspetiva dos objetivos de política e de programas de melhoria da educação; (2) Estabelecer o referencial passível de monitorização e avaliação, e seleccionar o modelo estatístico mais adequado ao contexto social e educacional; (3) Assegurar a infraestrutura necessária à implementação do modelo estatístico. Isto inclui um sistema de aferição dos resultados que obedeça técnica e metodologicamente aos padrões recomendados, que permita a recolha de dados longitudinais considerando o estudante como unidade estatística de observação ao longo da sua trajectória; inclui também a capacidade técnica e científica para a análise e modelação desses dados, depois de anonimizados, bem como para a disseminação de resultados que promova actividades *feedback* da avaliação, envolvendo as previstas no âmbito dos programas de melhoria acima referidos.

Embora seja necessário desenvolver trabalho adicional de modelagem, uma necessidade mais urgente é a de encontrar formas para melhorar a qualidade dos dados e, especialmente, evitar que inferências despropositadas sejam extraídas de qualquer apresentação excessiva dos resultados (Goldstein, 2014).

CAPITULO 3 – Statistical methods in higher education research

Abstract

This study examines the statistical methods and outcomes that have been considered in higher education research in the last twenty years. A review was undertaken using search engines covering the education and statistics literatures, such as the Education Resources Information Center (ERIC) and the Current Index to Statistics (CIS). In addition, the search was conducted for six leading higher education journals. After deleting duplicates, 122 papers authored or co-authored by 290 researchers remained. Approximately 20% were non-research papers and were not considered for review. The final list includes 98 articles and forms the subject of the analysis in the remainder of this paper. These articles have been published in 46 journals and cover 27 countries. Regression models – classic and multilevel models – are those most often applied in higher education research. Academic performance and dropout or related topics are the outcomes most frequently studied by researchers.

Keywords: academic performance; dropout; higher education research; statistical methods.

3.1 Introduction

There is a wide consensus regarding universities' need to uphold their students' success since many of these individuals do not achieve their goals or optimally benefit from their higher education experience, representing a waste of human talent and potential. This consensus – alongside the expansion of tertiary education in recent decades – is perhaps the main reason behind the growth of higher education research (Vincent-Lancrin, 2009). Given the development scenarios in tertiary education for the coming years (OECD, 2008), such growth is expected to continue with increased diversity and complexity (e.g., Altbach, 2014). Although higher education is considered a research field without a clear methodological center (e.g., Altbach, 2014; Tight, 2014), recent findings show that quantitative methods continue to dominate (Hutchinson & Lovell, 2004; Tight, 2013; Wells, Kolek, Williams, & Saunders, 2015). According to Tight (2013, 2014), who analyzed 567 papers published in 15 leading higher

education journals in 2010, three of the methods identified – multivariate analyses (including surveys), documentary analyses and interview-based studies – account for 91% of the papers studied, and the largest proportion (44%) were papers based on multivariate analysis. A detailed analysis of data presented in both papers reveals that most papers using multivariate analysis (35%) have ‘Education’ as the first author’s disciplinary background, followed by other ‘Social science’, with 25%. Tight also showed that 39% of authors with an ‘Education’ background used surveys and multivariate analyses in their articles, 30% used documentary analyses, and 26% used interviews; for those with ‘Social sciences’ backgrounds, the corresponding percentages were 63%, 18%, and 9%, respectively. These figures suggest that multivariate methods are the first choice among higher education researchers in general.

A decade later, Wells et al. (2015) replicated the study conducted by Hutchinson & Lovell (2004). Both papers used content analysis to identify the types of research designs and analytical approaches most often used in articles published by three leading higher education journals – *Journal of Higher Education*, *Research in Higher Education*, and *The Review of Higher Education* – in the 1996-2000 (Hutchinson & Lovell, 2004) and 2006-2010 periods (Wells et al., 2015). The authors compared the contents of the journals to determine if studies reported in each journal differed in their research approaches and levels of methodological sophistication. Their findings showed that approximately 60% of research studies published in both periods used a correlational design, while the most prominent change across the decade was the increased use of longitudinal designs with a decreased use of descriptive quantitative and case study designs. Regarding statistical methods, Wells and colleagues mentioned that the “three most frequently used methods remained the same: descriptive statistics, multiple regression, and factor or principle components analysis” (p. 182) and showed that the largest change in the methods applied over time was the increasing use of logistic regression and multilevel regression, with the sharp decreasing use of bivariate correlation (from 39% to 17%), chi-square tests (from 18% to 12%), and ANOVA (from 22% to 12%).

Several authors (e.g., Huisman & Tight, 2013) have indicated that although most articles make explicit use of methods, hardly any attention is paid to the advantages and disadvantages of the methods used for the research questions, i.e., the choice of the method seems to be based on taken-for-granted norms without any reflection on its appropriateness and assumptions. Wells et al. (2015) reported that 67.5% of articles that used intermediate or advanced statistics included at least a minimal discussion of the statistical assumptions underpinning their

methods. In addition, the use of quantitative methods in the field has been strongly conditioned either by the availability of data or by their quality (c.f. Bratti, McKnight, Naylor, & Smith, 2004; Foley & Goldstein, 2012). According to Wells et al. (2015), among the 408 articles they revised, forty-nine (16%) had no missing data, but among the remaining 257 studies, 125 (48.6%) did not indicate how missing data were handled and 71 (27.6%) reported the missing data treatment with no justification for its use.

Given the recent trends in the statistical methods applied in higher education research, their relevance to the development of the field (e.g., Smart, 2005), the need for a dialectic between theory and methods, the use of research methods “to interrogate and develop the theories” (Ashwin & Case, 2012, p. 271), and the scarcity of research on methodological characteristics of the published literature (Hutchinson & Lovell, 2004), two research questions arise: (1) which statistical methods have been used in higher education research over the last twenty years, and (2) which outcomes have been addressed in higher education research in this period? The review consists of four parts: (1) a description of the paper’s scope; (2) a descriptive characterization of the evidence; (3) a content analysis of statistical methods and outcomes; (4) a critical discussion and a presentation of implications for future research in the higher education field.

3.2 Scope of the review

The heart of this work is a review of the relevant quantitative research in higher education. The primary interest of this study is to identify the outcomes and statistical methods that have been used in higher education research in the last 20 years. This paper focuses on studies published between 1998 and 2017. The rest of this section explains how the education and statistics literature was searched. The main objective of the review process was to identify as many relevant and high-quality articles as possible. Thus, our strategy was to search a wide variety of sources and then systematically eliminate those that did not meet the criteria for content or relevance. The first step of the review was to conduct a search for peer-reviewed papers published between 1998 and 2017 using two search engines covering the education and statistics literatures: the Education Resources Information Center (ERIC) and the Current Index to Statistics (CIS). In the ERIC database, we searched for papers that listed “higher education”

as a keyword or included the term in the abstract and that also included any of the following terms in the abstract: “data”, “dependent variable”, “response variable”, “outcome”, or “performance”. We found 903 papers. Then, the search was limited to papers classified with “OT – Statistical Analysis”, which yielded a total of 126 (step 2). We reviewed each of these papers individually (step 3), and based on the abstracts, we selected those that were potentially relevant. The papers selected did not include contents such as teaching methods, Massive Open Online Course (MOOC), e-learning experiences, validation of assessment tests, economics of education and theory of human capital, theory and ethnography, job and pay-for-performance, personnel evaluations, essays or reviews. The final number of papers remaining after step 3 was 70. A similar process was conducted for the CIS database, with “higher education” as the unique keyword. The number of papers listed was 100 (step 4). As in step 3, we reviewed each of these papers based on the abstract and eliminated those on statistics education, race theory in education, or theories and ethnographies (step 5). Two papers listed in step 4 were not available online. The final number of papers remaining after step 5 was 89. In addition, we looked for other relevant papers published in leading higher education journals, focusing the search on statistical methods and longitudinal data. For this purpose, we considered listings in both the ISI Web of Knowledge Citation reports and in the Scimago database. The search was conducted by considering the list of “Q1” journals obtained from the “Scimago\Social Sciences\Education” webpage and excluding specific journals, such as Internet and Higher Education and Teaching in Higher Education. The list of six journals includes Research in Higher Education, Journal of Higher Education, Studies in Higher Education, Review of Higher Education, Assessment and Evaluation in Higher Education, and Higher Education Research and Development. The search was conducted with “statistical methods” or “statistical models” and “longitudinal data” as keywords or as words included in the abstract. As a result, we obtained 199 papers. After eliminating repeats, the final count was 122 with 290 authors or co-authors. Table A1 in the annex summarizes the steps of searching and selecting papers.

3.3 Summarizing the evidence

The majority of the papers were methodologically explicit, i.e., for 85% of them, it was possible to identify the statistical method and the outcome based on the abstract. For the minority that

were not methodologically explicit, the methods and outcomes were identified through a careful reading of the full paper. The papers addressed various research questions or hypotheses in 27 different countries, but most of them dealt with the US (46%). Among the 122 selected papers, 24 were reviews, essays or biographical narratives. For the purpose of the current work, these papers were excluded, so the remaining 98 constituted the object of the analysis. Table 1 shows the distribution of the number of papers by year of publication. Seventy percent of the selected papers were published after 2010.

Table 1. Distribution by year of publication.

Year	Number	%
[1998,....,2005]	15	15.31
]2005,....,2010]	14	14.29
]2010,....,2015]	35	35.71
]2015,....,2017]	34	34.69
Total	98	100.00

The data used in some of the selected papers (42%) were collected by surveys conducted in only one institution, so the results they report may lack generalizability. We distinguished papers into “one-institution studies” and “multi-institution studies” and split the analysis accordingly. We classified statistical methods mentioned in eight categories (Table 2). The first category, descriptive statistics, includes frequencies, cross-tabulations, averages, variances, correlations, and graphs and represents 5% of the selected papers. The second category includes several basic tests, such as t-tests, ANOVA, contingency tables and chi-square tests, representing 17%. The classic regression models category includes logistic and probit regressions and represents 35% of selected papers. The fourth category includes multilevel regression models and includes 21% of all the papers. Factorial analysis and structural equation models constitute the fifth category, representing 4%. The sixth category includes survival analysis and event history models, with 7%, and finally, a residual category with other methods (less frequent methods such as data envelopment analysis, propensity score analysis, stochastic frontier analysis, and simulation) represents 10%. The majority of the papers are based on regression models (56%). While institutional case studies use classical regression models more often (36.59%), multilevel

regression models are prevalent in multi-institutional studies (31.58%). More institutional case studies apply basic tests (26.83%) or survival analysis (12.20%).

Table 2. Percentage of methods applied given the type of study

Institution case study	Descriptive statistics	Basic tests	Classic regression models	Multilevel regression models	Factorial, Structural eq. models	Survival Analysis	Other
No	5.26	10.53	33.33	31.58	3.51	3.51	12.28
Yes	4.88	26.83	36.59	7.32	4.88	12.20	7.32
Total	5.10	17.35	34.69	21.43	4.08	7.14	10.20

Table 3 presents the percentage distribution of papers over time, given the category of methods applied. Almost 74% of the selected papers were published during or after 2010, and the majority of these were based on more sophisticated methods, with the exception of those published in 2004, which appeared mostly in the Journal of the Royal Statistical Society or in the Journal of Econometrics. Thus, after 2010, there is a greater diversity in statistical methods in higher education research. For example, 80% of papers exclusively using a kind of descriptive statistics were published before 2010. Until 2010, the most often applied methods included descriptive statistics, basic tests or classic regression models, with the exception mentioned above.

Table 3. Year of publication by method applied (%)

Methods	Year			
	[1998,....,2005]]2005,....,2010]]2010,....,2015]]2015,....,2017]
Descriptive statistics	60.00	20.00	20.00	0
Basic tests	0	5.90	47.00	47.00
Classic regression models	20.60	11.80	26.40	41.20
Multilevel regression models	14.30	14.30	47.60	23.80
Factorial / Struct eq models	0	0	50.00	50.00
Survival analysis	0	28.60	57.20	14.30
Other	20.00	30.00	10.00	40.00
Total	15.30	14.40	35.70	34.70

From a geographical point of view, the empirical evidence originates from 27 countries distributed across five continents. The largest percentage of evidence is US related, that is, 52.6% of multi-institutional studies and 36.6% of institutional case studies included an analysis

of US datasets. In general, papers concerned with the US are based on more sophisticated methods, i.e., methods classified in the categories of multilevel regression models (67%), survival analysis (57%), and other (70%, including data envelopment analysis, propensity score analysis, stochastic frontier analysis, and simulation).

Statistical models and outcomes

We identified the outcomes or response variables addressed in the most representative categories of statistical methods mentioned above. The categories of basic tests, classic regression models, and multilevel regression models represent 74% of all papers. We separately analyzed the articles whose datasets concern more than one institution from those focused on only one institution.

Multi-institutional studies

The large majority of articles classified as "Multi-institution studies" use classic or multilevel regression models (65%), and their outcome variables are related to the following phenomena: student dropout (Arbona & Nora, 2007; Bratti et al., 2004; Chen, 2012; Georg, 2009; Ishitani, 2016; Jaeger & Eagan, 2011; Gary R Pike & Graunke, 2015; Rhee, 2008; Schudde, 2011); educational choice, admission and participation (Chowdry, Crawford, Dearden, Goodman, & Vignoles, 2013; Davis, Smither, Zhu, & Stephan, 2017; English & Umbach, 2016; Gormley & Murphy, 2006; Pustjens, Van de gaer, Van Damme, & Onghena, 2004; Szelenyi, Denson, & Inkelas, 2013); progression, learning, and academic achievement (Draper & Gittoes, 2004; Fuller, Wilson, & Tobin, 2011; Kilgo, 2017; Ngo & Kosiewicz, 2017; Yamada, 2014); value added (Liu, Bridgeman, & Adler, 2012; T. Rodgers, 2005; Shavelson et al., 2016; Steedle, 2012); institutional or student commitment (Kodama & Dugan, 2013; Niehaus, Campbell, & Inkelas, 2014; R. Stinebrickner & Stinebrickner, 2004; Strauss & Volkwein, 2004), students' political identification (Astin & Denson, 2009); other institutional issues (Lang, 2015; López, Rojas, López, & López, 2015); and earnings (Black, Sanders, & Taylor, 2003; Giani, 2016; Moretti, 2004).

Despite the diversity of response variables addressed, issues related to student dropout represent researchers' first choice, followed by papers related to educational choice, admission

to higher education and participation.

Arbona and Nora (2017) applied a logistic regression model to examine characteristics that are believed to impact college persistence and undergraduate degree attainment among Hispanic high school graduates. Only Hispanic students who participated in the NELS: 88 from 1990 to 2000 and who enrolled in either a community college (n = 517) or a four-year college (n = 408) in 1992 or early 1993 were included in their analysis. Three outcomes were examined: (a) initial enrollment in a four-year college or university versus enrollment in a two-year community college; (b) bachelor's degree attainment by 2000 among those who first enrolled in a community college; and (c) bachelor's degree attainment by 2000 among those who first enrolled in a four-year institution. The findings reported that precollege characteristics and college experiences predicted bachelor's degree attainment among Hispanic high school graduates who initially attended a two- or a four-year institution and that fewer than half of the students who started college at four-year institutions obtained a bachelor's degree. This finding suggests that a large proportion of Hispanic students in four-year colleges encounter significant challenges in completing their studies. Thus, according to the authors, the implications of these results are relevant for policymakers, who "must begin to focus on institutional practices in community colleges ...", and for research development through an intervention plan to enhance students' aspirations and preparedness to attend four-year colleges. Bratti et al. (2004) applied probit regression models to examine university performance indicators based on data on students collected by the First Destination Supplement (FDS) and the Higher Education Statistics Agency (HESA) for 1997-1998. The population of students responding to FDS is not necessarily representative of the full population of those who left each higher education institution. For instance, university leavers who did not obtain qualifications and those who did not respond to the FDS were not included within the official analysis. The authors compared the performance indicators based on alternative approaches, addressing the potential non-random nature of this subgroup of university leavers. The authors gave two recommendations for policy and practice. First, they suggest that the performance indicators for universities in the area of graduate labor market outcomes should be interpreted with caution due to the misleading and low statistical confidence in such performance rankings, and second, they recommend that official employment indicators would be more valuable if they were based on data on graduates' employment status three years or more after graduation. In order to examine which institutional characteristics contribute to reducing the risk of student dropout,

Chen (2012) applied a multilevel event model to longitudinal hierarchical data that were collected from 1996 to 2001 and that involved 5762 students and 400 institutions. The main finding obtained was that institutional expenditure on student services (student affairs) is negatively associated with student dropout occurrence. Policy implications to promote equality in US higher education are given at the student and at institutional levels. In particular, the author mentioned that “programs and practices that aim at reducing dropout can probably maximize their effects by placing more emphasis on student persistence in the first year” (op. cit., p. 500). In terms of its contribution to research development, this article identifies very promising paths for the analysis of longitudinal data. Specifically, it allows for a better understanding of the student's experience over time that culminates in dropout. Georg (2009) applied a multilevel logistic regression model to data collected by the Konstanz Student Survey to examine the influence of institutional and individual variables on student dropout. According to the author, to better understand the processes leading to dropout, a methodological shift from cross-sectional data to longitudinal data is necessary. Ishitani (2016) applied multilevel logistic models to data on 7,571 first-time freshman students enrolled in public and private 4-year institutions in order to investigate the first two years of student college persistence, with a focus on individual and institutional characteristics. Data were drawn from the 2004-2009 Beginning Postsecondary Students Longitudinal Study (BPS: 04/09), and the article explores the influence of academic and social integration, financial aid, housing arrangement, and first-year college GPA on retention. According to the authors, due to a lack of information on such variables for students' second year, the study assumed that the values for these student characteristics did not change in the second year. Therefore, the authors indicated that the findings for the second year should be interpreted with caution and asked for additional research to confirm the results. Jaeger and Eagan (2011) applied logistic regression models to student data to examine the effects of contingent faculty use on first-year student retention for six institutions within a public higher education system. The study showed that high levels of exposure to part-time faculty in the first year of college appear to have significant negative relationships with student retention into the second year. According to the authors, the study's implications are focused on hiring policies and practices “in order to balance budgetary challenges with overarching concern of student success” (p. 530) and on data collection. The data availability challenges represent “significant limitations for institutions that want to better understand the role that contingent faculty play on their respective campuses”. In terms of research development, the authors mentioned that “Institutional researchers who are unable

to link faculty and course data are also not able to provide critical information to decision makers” (p. 531). Rhee (2008) applied a multilevel multinomial model to data on 3947 students to explore the influence of four-year institutions’ climate for diversity on different types of student withdrawal (dropout, stop out and transfer). Data were collected by the Cooperative Institutional Research Program from 1985 to 1989. The author justified the dataset choice with its unique inclusion of measures of student perceptions of institutional diversity, alongside distinct measures of student withdrawal behavior. However, in the limitations section, he warned the reader that “the CIRP dataset shows weakness with regard to the measures available for use in a study of withdrawal behaviors” (p. 172), i.e., a limitation that affects the core of the study. Schudde (2011) applied logistic regression models to data collected by the Educational Longitudinal Study (ELS: 2002) and the Integrated Postsecondary Education Data System (IPEDS: 2003-2004 & 2004-2005) to show the association between living on campus and student retention. The study’s implication for further research is the inclusion of campus-level variables that cannot be derived from the ELS or IPEDS, and its implication for policy is the provision of trustworthy information to guide college administrators’ decisions.

One-institution case studies

Most articles classified as "one-institution studies" use basic hypotheses tests or classic regression models (63%). Articles based on basic tests are focused mainly on the analysis of issues related to academic achievement (Amin, Hanawi, Mohamed, Saad, & Sahari, 2013; Clerici, Castiglioni, Grion, Zago, & Re, 2014; Schutz et al., 2015), dropout (Danaher, Bowser, & Somasundaram, 2008; Gutierrez & Dantes, 2009; Naidoo & Lemmens, 2015), perceptions of innovation (Kasule et al., 2015), and students’ commitment and attitudes (Howard & Davies, 2013; Leach, 2016). In papers based on regression models, the outcome variables are as follows: student performance and course grades (Beenstock & Feldman, 2018; Chingos, 2016; Masui, Broeckmans, Doumen, Groenen, & Molenberghs, 2014; McKenzie & Schweitzer, 2001; Shih, Tillett, & Lawrence, 2012; Zimmermann, Brodersen, Heinemann, & Buhmann, 2015); students’ motivation and attitudes (Brahm et al., 2017); admissions, financial aid, and earnings after graduation (Arcidiacono, 2005); participation (Crawford, Wang, & Andrews, 2016); conflicts and perceived quality of education (Mukhtar, Islam, & Siengthai, 2011); affirmative action admissions (Park, 2009); degree classification and employment status (Pietro, 2017);

dropout (Montmarquette, Mahseredjian, & Houle, 2001; Stinebrickner & Stinebrickner, 2012); and the monetary value of alumni contributions (Durango-Cohen & Balasubramanian, 2015).

Students' achievement and performance are the main choices of topic among researchers in this subgroup of studies. Amin et al. (2013) applied ANOVA analysis to data on 144 computer science students during their 1st, 2nd and 3rd years in order to explore students' academic achievement over time. Clerici et al. (2014) applied descriptives, principal component analysis, and analysis of variance. Data were collected from 210 graduates by the University of Padova in 2007-2008 to provide the research conditions for a qualitative study in the field of educational sciences. Schutz et al. (2015) analyzed data gathered from 1559 full-time and adjunct faculty of a Midwestern community college (US) in 2012 to explore the degree of self-perception regarding factors that are thought to influence grade assignment. These authors used mainly analysis of variance. Masui et al. (2014) applied chi-square tests, ANOVA and regressions to data collected from 168 freshmen in business economics at Hasselt University in order to examine the role of study time in academic performance. Beenstock & Feldman (2016) applied regression models to longitudinal data gathered between 2003 and 2008 from social sciences students at the Hebrew University of Jerusalem to decompose course grades into rewards for student ability and those for instructor quality at universities. This study comprised 1217 students. Chingos (2016) applied regression models to data for 8654 students collected from 2008 to 2011 to examine the extent to which student learning varies across different sections of the same courses at a large community college campus in California. The results obtained suggest "the instructor effect was a powerful predictor of student performance on the final exam, but not on the likelihood that the student would take the final or pass the course" (op. cit., p. 103). This intriguing finding let the author to call for further research. The author also highlighted the difficulties in conducting research on student learning in higher education due data availability restrictions. Considering a sample of 197 first-year students, McKenzie and Schweitzer (2001) used ANOVA and regression models to study student-level factors that influence university grades. Shih et al. (2012) used chi-square tests and regression models to examine the factors affecting 114 foreign students' satisfaction or dissatisfaction in adapting and adjusting to life in Taiwan and how these factors may affect their academic performance. Zimmermann et al. (2015) analyzed 81 variables from 171 student records from a bachelor's program and a master's program in Computer Science at the ETH Zurich in order to identify which indicators of undergraduate achievements may predict graduate-level

performance. The methods chosen included regression models and the bootstrap method.

3.4 Discussion and conclusion

Our review of the statistical methods and outcomes considered in higher education research over the last twenty years is based on content analysis of papers selected according to the systematic method described above. As a comprehensive review, this study may have limitations associated with the inevitable bias introduced in the streamlining process. Although the method for paper selection was systematically planned and the respective analysis carefully conducted, sources of bias such as selection bias, publication bias, and language of publication bias may be introduced (Butler, Deaton, Hodgkinson, Holmes, & Marshall, 2005) since we decided to use readily accessible literature. Every step of the study is fully described in order to minimize the impact of eventual bias and to support the replicability of the study.

Patterns can be drawn regarding the statistical methods and outcomes used most often in higher education research papers. From a geographical point of view, the empirical evidence originates from 27 countries distributed across five continents. Our findings suggest the prominence of American scientific production in this field of study, as manifest not only in terms of a larger number of articles but also in terms of the use of more sophisticated statistical methods. This finding corroborates the evidence found by Tight (2013), who examined the relation between discipline and methodology in higher education research by analyzing a database with 567 papers published in 2010 in 15 leading higher education journals. The author reported a high degree of geographic variation, with North American researchers exhibiting a preference for the use of more complex multivariate techniques. We found that there are two possible reasons for this trend over the last two decades. The first is the degree of development of higher education academic groups – departments or centers with postgraduate research programs – and the second is the longer and wider availability of complex datasets for analysis. We found that recent years have seen the sharp growth of publications in the field (74% of selected papers were published after 2010) that made use of more sophisticated statistical methods and corroborated previous research (e.g., Wells et al., 2015).

Our results suggest that until 2010, the majority of papers selected made use of a limited

panoply of methods, i.e., descriptive statistics, basic tests, or classic regressions, with the exception of those published in 2004, which appeared mostly in the Journal of the Royal Statistical Society or in the Journal of Econometrics and were thus more methodologically centered. These findings corroborate those presented in Table 5 by Wells et al. (2015). We found that more methods were in use after 2010, including factorial analysis, multilevel regression models, structural equation models, survival models, data envelopment analysis, propensity score analysis, stochastic frontier analysis, and simulation.

Additionally, we found diversity in the choice of outcomes and response variables. Nevertheless, researchers' preferred topics appear to be student dropout (or its equivalents: persistence, retention, and attrition) and variables related to students' academic performance. These papers show a rather diverse focus, i.e., the study of phenomena among minority groups, data challenges with implications for official statistics, the role of institutional characteristics and their influence on student dropout, the effects of contingent faculty use on retention, and the impact of living on campus on retention, among others.

In some countries, these themes are central for public policy, especially in regard to minority groups and socio-economic disadvantages. For instance, Arbona and Nora (2017) indicated that true access to higher education cannot be defined as the enrollment of Latino students in postsecondary institutions; instead, degree attainment should be considered. According to the evidence found by Fry (2002), the discrepancy between initial enrollment and subsequent graduation rates suggests that the gap for Hispanics is due largely to factors related to these students' experiences in US higher education. These topics have a long tradition in higher education research (c.f. Tinto & Cullen, 1973). It is possible that the long-term effect of authors and institutions (Saunders et al., 2016) alongside the persistence of dropout in countries other than the US contribute to researchers' primary choice of the abovementioned topics as primary phenomena studied all over the world.

It seems that the higher education literature has experienced some delay in appropriating the most sophisticated statistical methods. From the perspective of the diffusion of innovations, Rodgers (2010) called this phenomenon the "quiet revolution", i.e., the transition from the basic test paradigm to a paradigm based on building, comparing, and evaluating statistical models. In fact, the transdisciplinary strand of statistics has developed in the last few years, featuring new themes and challenges. Particularly in the subarea of educational statistics, the innovative methodologies that have appeared in the last two decades have increased the

research capacity to uncover the complexity of the real world (Goldstein, 1998), which is fundamental to the disciplinary and scientific development of all fields of education. Technological and computational advances, notably those regarding computational and statistical methods to deal with large and complex surveys and longitudinal data, have powered a giant leap in the production of knowledge in education and, thus, to the advancement of the discipline. In the era of big data, the “quiet revolution” is indeed a very promising trend in higher education research. In addition, this review shows that several authors attributed the limitations of the interpretation and extrapolation of their findings to restrictions or limitations imposed by data availability and data quality. These factors represent an obstacle to the development of the field that may only be overcome with the support of agencies or government bodies responsible for the promotion and development of science and higher education.

CAPITULO 4 – Differential effect of university entrance score on first-year students' academic performance in Portugal⁵

Abstract

The main goal of this study is to show that the association between the university entrance score and first-year students' academic performance varies randomly across courses after controlling for students' sociodemographic, schooling trajectory and motivational variables. The sample consists of 2,697 first-year students who were enrolled in 54 courses at a Portuguese public university in 2015/16. Multilevel modeling of academic performance suggests that 34% of variability in grade point average is due to differences among courses and that 80% of such variability is explained by the field of study; whether the university is the student's first choice; and the student's gender, age and parents' level of education. In addition, the results corroborate that the university entrance score is the stronger predictor of first-year academic performance.

Keywords: academic achievement; first-year experience; higher education; multilevel model.

4.1 Introduction

Portuguese higher education is organized as a binary system, with university education oriented toward the provision of solid academic training, combining the efforts and responsibilities of both teaching and research units, while polytechnic education concentrates on vocational and advanced technical training that are professionally oriented (Assembleia da República, 2007). From 1971 to 2011, the number of students enrolled in higher education increased by 700%, ranging from 49,461 in 1971 to 396,268 in 2011. Dias (2015) presented a detailed analysis of official statistics to investigate how this strong increase was achieved and raised the question of whether such an increase ("massification") led to greater equity in terms not only of access but also of success in higher education. In fact, this growth implied an increase in the heterogeneity of the student population in terms of gender, socioeconomic level, age group,

⁵ Based on the paper Ferrão, M. E., Almeida, L. S. (2018). Differential effect of university entrance score on first-year students' academic performance in Portugal. *Assessment and Evaluation in Higher Education*. Accepted.

and past academic trajectories. There is a broad consensus on the need for universities to promote students' success, since many do not achieve their goals or fully realize their potential during their university experience, thus wasting talent and human potential. The rates of failure and abandonment have increased in recent years in Portugal, leading to questions regarding how institutions can support student retention (Dias, Marinho-Araújo, Almeida, & Amaral, 2011; Ferreira, Vidal, & Vieira, 2014).

Access to higher education in Portugal is organized by a system of *numerus clausus* defined as a national public tender to give priority of admission to students with higher admission scores – according to official statistics, in 2015/16 approximately 50% of students were placed in their first option for course and institution. Portela et al. (2008) suggested that students implicitly rank courses and institutions when making their choices, reinforcing the need to quantify the effect of first-option admission on students' individual performance in undergraduate courses. The interaction between student and university starts when the student chooses the university and course s/he intends to attend. For instance, Almeida, Guisande, Soares, and Saavedra (2006, p. 512) explored students' sociodemographic characteristics related to the choice of courses at the University of Minho (a public university in northern Portugal), and concluded that students from less advantaged social classes were more likely to enroll in social sciences courses and that “higher sociocultural status and female students had higher entrance exam scores and averaged better academic achievement at the end of their first year at the university” (p. 506). Several authors noted that students' socioeconomic and cultural capital are associated with the preference for more prestigious courses or institutions (e.g. Tavares, Tavares, Justino, & Amaral, 2008; Tavares & Ferreira, 2012). According to Amado Tavares et al. (2008), students from disadvantaged family backgrounds have stronger preferences for courses related to teacher training and management, while students from advantaged family backgrounds are more likely to choose courses in the areas of law, fine arts, health and science. Considering, for example, the health sector, the authors found that the parents of 73.2% of students who were attending a university course in medicine hold higher education degrees, while 73.0% of the students enrolled in a health care course at a polytechnic institute came from families with no higher education degrees. Such findings focus the equity debate on the transition from high school to higher education and on the rules regarding access to public higher education, since the evidence appears to support the idea of strong socioeconomic selectivity during the process of students' admission to higher education.

Several studies suggested the relevance of a student's high school achievement or university entrance score as predictors of the student's academic performance in higher education (McKenzie & Schweitzer, 2001; Almeida, 2002; Soares, Guisande, Almeida, & Páramo, 2009; Johnston, Manley, Jones, Harris, & Hoare, 2016) and as predictors of academic, social and institutional adjustment (Páramo Fernández, Araújo, Vacas, Almeida, & González, 2017). Nonetheless, McKenzie and Schweitzer (2001, p. 29) mentioned that "university entry scores need to be interpreted with caution, as they explain less than half of the variance in GPA". Vulperhorst, Lutz, de Kleijn, & van Tartwijk (2018) compared the predictive validity of high school scores on academic success in university between the Dutch pre-university (314 graduates) and the International Baccalaureate (n=113 graduates) diploma. According to the authors, the most valid high school score to be used for selection procedures in higher education remains unclear, and it seems to depend on the diploma students enter with.

Although several studies have examined these issues around the world, little is known about how university/course choice mediate the influence of prior achievement on academic performance in the context of higher education in Portugal. To our knowledge, this relationship has not been analyzed in the literature as a differential effect that is dependent on the course. Therefore, this work has three main objectives: (1) to quantify the effect of first-option admission (to the course or the university) on students' academic performance, controlling for sociodemographic variables and admission score; (2) to estimate the variability of students' academic performance at the end of the first year by the hierarchical levels of student and course; and (3) to estimate the effect of students' admission scores on their academic performance, controlling for sociodemographic and schooling trajectory variables (sex, age, socioeconomic status, and failure in primary or secondary education) and quantifying the amount of the effect due to differences among students and differences across courses.

4.2 The relationship between students' performance in high school and in college

International research shows that tests for university admission, such as the Scholastic Aptitude Test (SAT) or American College Testing (ACT), or the high school grade point average (GPA) are assumed to be the best predictors of students' academic performance in higher education if academic performance is assessed by the GPA based on marks from curricular course units (Baron & Norman, 1992). Those three predictors taken together can account for 25% or more

of the variance in the mean scores during the first year (Mouw & Khanna, 1993; Robbins et al., 2004). Other variables related to students' personality (O'Connor & Paunonen, 2007; Poropat, 2009), motivation (Eccles & Wigfield, 2002; Pintrich, 2004) and methods of study are also important (Diseth, Pallesen, Brunborg, & Larsen, 2010), but the meta-analysis performed by Richardson, Abraham, and Bond (2012) with all these variables confirms the higher predictive capacity of the SAT, ACT and high school GPA.

Pike & Saupe (2002) analyzed data for 8,674 first-time college students from 124 high schools who enrolled at a midwestern USA university between 1996 and 1999 and considered the students' first-year GPA as a dependent variable. They also included in the analysis the ACT score as part of the university's admission policy, high school class percentile rank, and a binary variable representing whether the student had or had not met new core course requirements implemented in 1997. Their results show a significant positive relationship between having met the new core course requirements and first-year GPA, providing evidence of the appropriateness of the new admission requirements for students to succeed in higher education. Regarding the effect of the ACT assessment composite score and high school class percentile rank on GPA, both coefficients were positive and statistically significant. Fuller, Wilson, and Tobin (2011) explore to what extent the cross-sectional and longitudinal National Survey of Student Engagement (NSSE) scores predict a college student's final GPA. Their models include ACT scores and high school GPA as indicators of incoming students' precollege abilities. Based on longitudinal data modeling, the authors conclude that 27.8% of the variance explained could be attributed to the precollege academic indicators. In addition, McKenzie and Schweitzer (2001) conducted an investigation of the relationship between academic, psychosocial, cognitive and demographic factors and students' academic performance, assessed by a student's classification at the end of the first term in college. The authors suggested that as university entrance scores are moderately correlated with a high GPA, entry to courses with a high level of difficulty should be established at a realistic level to avoid undue problems for students with lower university entrance scores. The literature also shows that students' performance over the first year tends to determine their performance in the following academic years (Berg & Hofman, 2005; Yorke & Longden, 2008).

4.3 Differentials of academic performance by area of study

Masui, Broeckmans, Doumen, Groenen, and Molenberghs (2014) aimed to clarify the role of study time in academic performance by examining its capacity in the prediction of course grades, beyond relevant student characteristics (e.g., prior domain knowledge, prior study delay), and evaluating the predictive role of course characteristics in study time investment. The study is based on a sample of 168 (out of 196) freshmen in business economics at the Hasselt University. According to the authors, their findings emphasize the relevance for future research of examining relationships at the course level, since the influence of students' prior knowledge and skills on their academic performance seems to depend on their match or mismatch with the course requirements. They suggested that studies on the topic should take into account the field of study because differential grading induces the mismeasurement of academic performance. In line with this view, Beenstock and Feldman (2016) applied regression models to longitudinal data gathered between 2003 and 2008 from 1,217 social sciences students (majoring in psychology, sociology, political science, international relations, statistics, economics, business studies, accounting, communications, politics, philosophy, economics and geography) at the Hebrew University of Jerusalem with the purpose of decomposing course grades into the rewards to student ability at the university and the rewards to instructor quality. The authors reported a correlation of 0.41 between a student's high school ability and his/her performance at the university and shown that there were grade differentials that were probably established by policies determined at the departmental level. Their findings suggested that students majoring in economics obtained higher grades in non-economics courses after controlling for students' high school ability. Considering data from algebra courses at a large community college in California, Chingos (2016) applied regression models to the longitudinal data of 8,654 students collected from 2008 to 2011 with the objective of examining to what extent student learning varies across different sections. The author found that student learning varies systematically across instructors, who appeared to influence student learning beyond their impact on course completion rates.

In general, students with higher socioeconomic status and women have higher levels of academic performance as shown by GPA (Dennis, Phinney, & Chuateco, 2005; Laforge, 2003; Robbins et al., 2004; Smith & Naylor, 2001; Voyer & Voyer, 2014), as these socioeconomic factors are present in the explanation of variance in predictors (SAT, ACT or high school GPA)

as well as in the average classifications during the first year of college (Sackett, Kuncel, Arneson, Cooper, & Waters, 2009). Regarding students' age, in some studies, older students show a slight superiority in academic performance (Clifton, 1997), but in other studies, the youngest students perform better (Farsides & Woodfield, 2007). In a meta-analysis by Richardson and collaborators (2012), the correlations between GPA and socioeconomic background, gender and age suggested that students from higher socioeconomic status, female students and older students obtained higher GPA. In spite of being statistically significant, the effect size estimates were small.

This study sought to answer the following research questions: Q1. Does the university or course choice matter for the student's academic performance? Q2. How much is the variance of first-year students' academic performance due to the variability across courses? Q3. Is there any evidence that students' university entrance score is predictive of the first-year academic performance? Q4. Does the relationship between university entrance score and first-year academic performance vary across courses? Q5. How great are the differences in grading among the areas of study? Thus, the first and second research questions are linked respectively to the first and second research objectives above mentioned, the remaining research questions are related to the third objective.

We are confronted with hierarchically structured data in which students are nested within courses. According to some authors (e.g. Pike & Rocconi, 2012) much of the research in higher education has failed to properly account for the hierarchical nature of the data. However, multilevel modeling has become increasingly popular, as Wells, Kolek, Williams, and Saunders (2015) showed.

Multilevel models (Bryk & Raudenbush, 1992; Goldstein, 2003) allow us to simultaneously consider multiple levels of effects, as many as the hierarchical structure includes. To obtain empirical evidence to answer our research questions, we used two types of multilevel models: the variance component model (VCM) and the random coefficients model (RCM).

4.4 Participants, variables and procedures

The sample consists of 2,697 first-year students who were enrolled in undergraduate studies at the University of Minho in the academic year 2015/16, the same year that it became a public

foundation under private law (article 129 of the Legal regime for higher education institution, RJIES) (Guellec, Larrue, Roy, & Weko, 2018). The University of Minho is located in the North of Portugal and offers graduation courses and postgraduate, including masters and doctorate degrees, for a total number of students around 20,000 in all fields of study. Data used for the purpose of this paper concern the following student attributes: academic performance measured by the average score (GPA) at the end of the first year; university entrance score; first-option admission (yes/no) to the institution and the undergraduate course; gender; age; schooling trajectory assessed by the early experience of repetition (yes/no); and parents' education as proxy for the student's socioeconomic status. The area of the course was also included as an explanatory variable.

Most of the students (57%) were female and were enrolled in his/her first-choice course (59%) and first-choice higher education institution (72%). Almost half of the students (47%) were registered in a science, technology, engineering and mathematics (STEM) course; 17% in economics, management or public administration; 13% in humanities, 10% in social sciences; 8% in health (medicine or nursing); and 5% in law. The students' ages ranged from 16 to 61 years, with an average of 18.9 (SD 3.6), and 91.9% were full-time students. The average university entrance score was 152.4 (SD 18.9). Most of the students (83.3%) declared that they had never repeated a year along their schooling trajectory, and 13.9% declared at least one repetition. The distribution of students by parents' education shows that 34.4% of their fathers and mothers had no more than a lower secondary education, and for 15.7%, both parents had a higher education degree. The academic performance data were available for 72% of the students, which is 1940. The remaining cases (28%) were attributable mainly to student dropout and transfers to other undergraduate courses, departments or institutions. The descriptive statistics for the subset of students without an academic score at the end of the first year showed that they are older, more likely to have repeated a year during primary or secondary education, and had a lower university entrance score than the students who composed the full sample. The university entrance score is missing for 6% of students, the majority of them got a special via of access to higher education (Ministério da Educação e Ciência, 2014) and were not submitted to the national exam. In such situation could be immigrant people or older than 23. In addition, there are no more than 4% of missing data in the remaining variables, which we consider to occur completely at random (Little & Rubin, 2002). Students were informed about the objectives of the study and gave their written

agreement, which included the possibility of accessing their classifications (GPA) at the end of the first semester and at the end of the first year (second semester). The remaining data were collected at the moment of enrollment in the university. For statistical modeling, we decided to standardize the variables GPA and university entrance score; thus, the results are interpreted in terms of standard deviations from the mean.

4.5 Statistical modeling

We applied multilevel (Goldstein, 2003) or hierarchical (Bryk & Raudenbush, 1992) models, specifically the Variance Components Model (VCM) and the Random Coefficient Model (RCM), since they have several advantages when the data are hierarchically structured, as is the population in the study. That is, students are grouped in undergraduate courses, so two-level models are applied, with students (indexed by i) at level 1 and courses (indexed by j) at level 2. The advantages of this class of models include the ability to model differential effects and to improve the efficiency of the estimators, i.e., with narrower confidence intervals and shrunken estimators, which are more efficient, especially when the group size is small or there is a large within-group to between-group variance ratio (e.g. Longford, 1993; Ray, McCormack, & Evans, 2009).

For each student i in undergraduate course j , Y represents the first-year academic performance measured by the student's average score (GPA) obtained at the end of the academic year, and the observed scores are y_{ij} , where $i = 1, \dots, n_j$ and $j = 1, \dots, J$, with n_j representing the number of students in course j and J representing the total number of undergraduate courses offered. The event Y is assumed to follow a normal distribution with mean XB and variance Ω , i.e., $Y \sim N(XB, \Omega)$, where XB is the additive linear predictor or deterministic component of the model; X is the matrix of independent variables described above; B is the vector of fixed parameters; and Ω is the variance-covariance matrix, which includes the random parameters of the model. The two-level VCM to address objective 1 is written as follows:

$$y_{ij} = \beta_0_j + \sum_{p=1}^6 \beta_p X_{p_{ij}} + e_{ij} \quad (1.1)$$

$$\beta_0_j = \beta_0 + u_{0j} \quad (1.2)$$

Equation (1.1) is defined at level 1 and consists of the linear predictor and the level 1 random component (e_{ij}), and equation (1.2) is defined at level 2, with the term of intercept (β_0) and

the level 2 random component (u_{0j}). The fixed parameters of the model are $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$, and β_p , where p denotes the number of covariates. The probability distribution assumptions for the random components are $e_{ij} \sim N(0, \sigma_e^2)$ and $u_{0j} \sim N(0, \sigma_{u_0}^2)$, respectively. That is, σ_e^2 is the random parameter at level 1 (the level 1 variance) and $\sigma_{u_0}^2$ is the random parameter at level 2 (the level 2 variance). According to the model defined by equations (1.1) and (1.2) and by the assumptions, the intercept of the model varies randomly across programs (level 2 units).

We used this model to analyze the relationship between academic performance and the student's set of background covariates. Specifically, we were interested in understanding the relationship between the dependent variable and the student's sociodemographic characteristics, such as gender, family sociodemographic and economic background, schooling trajectory assessed by the experience of retention, and score on tests for access to higher education. We also wanted to determine the amount of academic performance variance explained by the background variables and how much is due to the undergraduate program. For this purpose, we compared the estimates of the random components obtained from the full model with those obtained from the null model. Two statistics are used for that purpose: the coefficient of determination and the variance partition coefficient (Goldstein, Browne, & Rasbash, 2002).

The two-level RCM is defined by equations (2.1), (2.2) and (2.3), meaning that the linear relation between university entrance score and academic performance (β_1) varies randomly across courses:

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \sum_{k=2}^p \beta_k X_{kij} + e_{ij} \tag{2.1}$$

$$\beta_{0j} = \beta_0 + u_{0j} \tag{2.2}$$

$$\beta_{1j} = \beta_1 + u_{1j} \tag{2.3}$$

The probability distribution assumptions for the random components are now

$$e_{ij} \sim N(0, \sigma_e^2),$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{u_0}^2 & \sigma_{u_{01}} \\ \sigma_{u_{01}} & \sigma_{u_1}^2 \end{bmatrix} \right).$$

This model includes two additional random parameters for level 2. The variance σ_{u1}^2 quantifies the variability of β_1 across courses, and σ_{u01} quantifies the respective covariance between the intercept and β_1 .

Briefly, the empirical evidence for question Q1 (Does the university or course choice matter for the student's academic performance?) is provided by a RCM; for question Q2 (How much is the variance of first-year students' academic performance due to the variability across courses?) is provided by a VPC; for question Q3 (Is there any evidence that students' university entrance score is predictive of the first-year academic performance?) is given either by VPC or RCM; and for questions Q4 (Does the relationship between university entrance score and first-year academic performance vary across courses?) and Q5 (How great are the differences in grading among the areas of study?) the evidence is based on a RCM.

We used the iterative generalized least squares (IGLS) algorithm as an estimation procedure implemented in MLwiN (Rasbash, Browne, Healy, Cameron, & Charlton, 2014). In addition, the variance partition coefficient (Goldstein et al., 2002) was applied to quantify the amount of variance distributed by each hierarchical level.

4.6 Results

Tables 1 and 2 present the estimates of the fixed and random parameters for the VCM and RCM, respectively. Asterisks are used to identify statistically significant effects based on a critical t value of ± 1.96 , i.e., a level of significance of 5%, and double asterisks when the critical t value is ± 1.65 , i.e., a level of significance of 10%. The variance partition coefficient applied to the null model parameter estimates (Table 1) shows that the variability of GPA decomposes 34% across courses and 66% among students or within course. Such estimates allow us to provide empirical evidence that supports objective 2 and research question Q2. The column M1 in Table 1 contains the estimates for university entrance score as a unique covariate, and M2 includes sociodemographic variables (sex, age, socioeconomic status), school trajectory (failure in primary or secondary education) as well. The column M3 in Table 2 adds the random coefficient parameters keeping the M2 set of explanatory variables, and M4 includes also the covariates related to university and programme choice, and areas of study. Thus, the objective 1 and research question Q1 is addressed by the results obtained from model M4 and the final

model (Table 2); empirical evidence that supports the research questions underlying objective 3 is provided by all fitted models.

Both models produced estimates for fixed coefficient and standard error, indicating that there is a positive statistically significant relationship between GPA and university entrance score ($b_1=0.55$; $se=0.03$). The coefficient of determination (R^2) indicates that the university entrance score explains 25% of the GPA total variance, and, focusing on level 2, 42% of the variability of GPA across courses. The results for model M2 show that after controlling for sociodemographic variables and schooling trajectory, the marginal effect of university entrance score on GPA remains stable ($b_1=0.55$), and the total variance explained increases to 28% (3 more percentage points than before) and increases to 47% at level 2 (5 more percentage points than before).

Table 1. Parameter estimates for variance component models

	Null model		M1		M2	
	Estimate	SE	Estimate	SE	Estimate	SE
<i>Fixed Parameters</i>						
Constant	-0.075	0.081	0.027	0.065	0.031	0.067
University entrance score			0.554*	0.028	0.553*	0.030
Male vs. female					-0.128*	0.041
Age>19					0.169*	0.082
Parents' education: no more than lower secondary					0.118*	0.040
Parents' education: higher education					-0.045	0.052
Repetition at primary(yes vs. no)					0.164	0.150
Repetition at secondary(yes vs. no)					-0.009	0.062
<i>Random Parameters</i>						
Level 2: course						
Var(constant)	0.320*	0.068	0.198*	0.043	0.181*	0.040
Level 1: student						
Var(constant)	0.661*	0.022	0.552*	0.019	0.542*	0.018
Number of level 2 units: course	54		53		53	
Number of level 1 units: student	1940		1826		1797	
Estimation procedure	IGLS		IGLS		IGLS	
-2*log-likelihood:	4847.948		4223.197		4121.568	

(*) Statistically significant at the level of 5%.

The results also suggest that a student's GPA is related to gender, the student's age, and the parents' level of education. When the influence of university entrance score on GPA is compared to the influence of socioeconomic variables, either the variance explained or the estimates for

the fixed parameters indicate that the university entrance score has a stronger influence. M3 adds to the previous findings by allowing the university entrance score coefficient to vary across courses. The change in the -2 log-likelihood value, which is also the change in deviance, has a chi-square distribution on 2 degrees of freedom under the null hypothesis that the extra two random parameters (variance and covariance) have population values of zero. The deviance is 20.365 (p value<0.001). Thus, the statistically significant random effect indicates that the relationship between GPA and university entrance score varies randomly across courses, i.e., the linear relation is stronger or weaker depending on the course.

Table 2. Parameter estimates for random coefficients models

	M3		M4		Final	
	Estimate	SE	Estimate	SE	Estimate	SE
Constant	0.019	0.066	-0.207	0.073	-0.275	0.062
University entrance score	0.531*	0.038	0.580*	0.046	0.567*	0.037
Male vs. female	-0.131*	0.040	-0.131*	0.040	-0.126*	0.040
Age>19	0.147**	0.082	0.131	0.084	—	—
Parents'educ.: < lower secondary	0.122*	0.040	0.129*	0.040	0.135*	0.038
Parents'educ.: higher education	-0.043	0.051	-0.028	0.052	—	—
Repetition at primary (yes vs. no)	0.179	0.149	0.126	0.147	—	—
Repetition at secondary (yes vs. no)	-0.038	0.062	-0.053	0.062	—	—
Course first choice (yes vs. no)	—	—	0.072**	0.042	0.078**	0.042
University first choice (yes vs. no)	—	—	0.139*	0.045	0.141*	0.045
Area of study: ECO vs. STEM	—	—	-0.178	0.117	—	—
Area of study: HUM vs. STEM	—	—	0.722*	0.134	—	—
Area of study: SOC vs. STEM	—	—	-0.152	0.144	—	—
Area of study: HEALTH vs. STEM	—	—	0.391*	0.197	—	—
Area of study: LAW vs. STEM	—	—	-0.795*	0.281	—	—
Area of study ECO X University entrance score	—	—	-0.036	0.102	—	—
Area of study HUM X University entrance score	—	—	-0.269*	0.107	—	—
Area of study SOC X University entrance score	—	—	-0.083	0.122	—	—
Area of study HEALTH X University entrance score	—	—	0.391	0.242	—	—
Area of study LAW X University entrance score	—	—	0.004	0.191	—	—
Area of study: HUM vs. STEM+ECO+SOC	—	—	—	—	0.791*	0.130
Area of study: HEALTH vs. STEM+ECO+SOC	—	—	—	—	0.424*	0.194
Area of study: LAW vs. STEM+ECO+SOC	—	—	—	—	-0.735*	0.249
Area of study HUM X University entrance score	—	—	—	—	-0.271*	0.104
<i>Random Parameters</i>						
Level 2: Course						
Var(constant)	0.167*	0.039	0.061*	0.018	0.062*	0.018
Cov(university entrance score/constant)	-0.037*	0.016	-0.013*	0.009	-0.013*	0.009
Var(university entrance score)	0.028*	0.013	0.019*	0.011	0.021*	0.011
Level 1: student						
Var(constant)	0.530*	0.018	0.520*	0.018	0.520*	0.018
Number of level 2 units: course	53		51		51	
Number of level 1 units: student	1797		1746		1747	
Estimation procedure	IGLS		IGLS		IGLS	
-2*log-likelihood:	4101.203		3890.304		3914.561	

(*) Statistically significant at the level of 5%; (**) Statistically significant at the level of 10%.

The results for M4 suggest that the first-choice university and area of study are relevant variables in the GPA analysis. At the level of significance of 5%, our results do not show a statistically significant effect of first-choice course on GPA unless the variable first-choice university is excluded from the model. In such a situation, the effect seems to be due to the omitted variable, meaning a spurious relationship between the variables. However, the fixed parameter is statistically different from zero at the level of 10%.

Regarding the area of study, on average, students enrolled in humanities, health, or law courses have a significantly different GPA than their colleagues who enroll in any STEM, economics, and social sciences course. According to our estimates, grading in humanities or health is higher than in STEM courses, and in law courses, grading is lower. In economics or social sciences, it is not possible to say whether, on average, grading is significantly different than in STEM courses. For that reason, in the final model, the baseline for areas of study comprises the STEM courses, economics, management, public administration, and social sciences.

The estimates for the interaction terms between area of study and university entrance score suggest a statistical significant term for humanities course with negative estimate, meaning that the humanities students with higher entrance score have their GPA considerably reduced, having added the principal terms related to university entrance score and the dummy term for humanities. Such interaction represents the narrowing in the dispersion of classifications that occurs with greater incidence in the humanities courses.

The final model includes all explanatory variables that coefficients revealed to be statistically significant in models M1, M2, and M3. The coefficient of determination applied to the final model variances indicates that, on average, the model explains 42% of the GPA total variance; focusing on level 2, it explains 82% of the GPA variability across courses and 21% of the GPA variability at the student level.

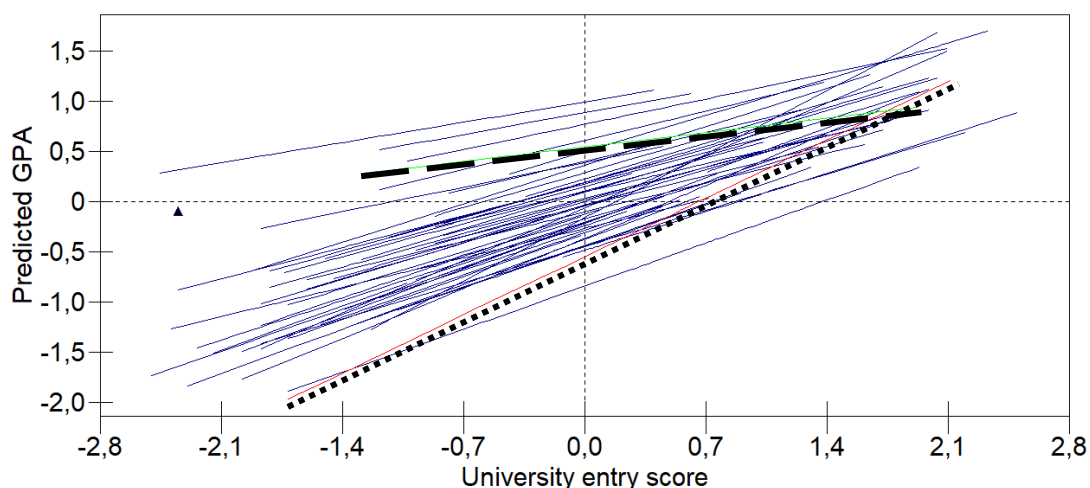
The estimates obtained confirm the interpretations presented above; that is, there is evidence in favor of an effect of first-option admission to the university on students' academic performance, controlling for sociodemographic variables and admission score. Those students for whom the University of Minho was their first choice have their GPA increased by 0.141 standard deviation on average compared to their colleagues. The results do not provide enough evidence, at the level of significance of 5%, to allow us to say that the first-choice course influences students' GPA.

According to the model, male students have their GPA reduced by 0.126 standard deviation, and students whose parents both have no more than a lower secondary education (i.e., a lower level of education) have their GPA increased by 0.135, on average, compared to their colleagues whose parents have any other level of education.

The final model replicates what Figure 1 shows for the relationship between a student's entrance score and his/her GPA, that is, the differential effect between the two variables is represented by one line fitted for each course. It can be observed that the positive relationship between the variables is stronger or weaker depending on the course and, in general, for negative values of student's entrance score the course lines dispersion is greater than for positive values. Such phenomenon is quantified by the covariance between the intercept and university entrance score slope, which has a negative estimate. Consequently, the respective correlation is approximately -0.4, suggesting that at the course level, the higher the GPA intercept, the lower the coefficient between university entrance score and GPA.

This correlation can be observed in the graph (Figure 1). As example, two course lines were highlighted with dashed and dotted lines. The dashed line presents a higher intercept (average GPA) and lower slope (effect of university entrance score on GPA), while the dotted line shows a lower intercept and higher slope. The slope for the dotted line course is higher than the slope for the dashed line course, suggesting a stronger effect on GPA for the dotted course than for the dashed. We need more research on this topic to explore in detail the influence of the curricular prerequisites on the slope of the lines, that is, for the explanation of the degree of the effect of the university entrance score on students' academic success.

Figure 1. Graph based on final model estimates



4.7 Discussion and Conclusion

In this study, we applied VCMs and RCMs to data collected from 1,940 Portuguese first-year students enrolled at a public university in 2015/16, considering first-year GPA as a dependent variable. The partitioning of the first-year GPA variance suggests that 66% is accounted for at the student level and 34% at the course level. The amount of variance of first-year students' academic performance due to variability across courses is approximately 34%. Part of this variability is due to the differences in the mean of the classification by scientific area of study. When STEM courses were considered as a baseline, it was verified that on average, the GPA is not different from that in courses in economics and social sciences. However, courses in the humanities and health areas have a statistically higher mean GPA than STEM courses, and law courses tend to present a statistically lower mean GPA than STEM courses. These findings appear to be in concordance with those of Masui et al. (2014), who demonstrated differential grading by field of study, and support the idea that differential grading is possibly induced by departmental norms (e.g., Beenstock & Feldman, 2016).

The estimates obtained suggest that the relationship between university entrance score and GPA at the end of the first year is on average 0.567 per standard deviation and varies randomly across courses after controlling for sociodemographic variables (gender, parents' education), area of study, and whether it was the first-choice university. The magnitude of the fixed parameter estimate is larger when compared to other marginal effects included in the linear predictor. In addition, the model explains 42% of the GPA variability and 82% at level 2; i.e., it explains 82% of the GPA variability across courses and 21% of the GPA variability at the student level. Our findings corroborate the literature on the predictive power of prior achievement. For instance, Fuller et al. (2011) reported that 27.8% of the variance explained was attributable to the precollege indicators. Moreover, our findings suggest that some courses establish stronger conditions for prior knowledge prerequisites than others, with implications for curriculum development. This evidence supports the theoretical model proposed by Alarcão (2000), granting to curriculum a status similar to that granted to the student, teacher or institution. Our findings also suggest that those students for whom the University of Minho was not their first choice have a lower GPA. The results do not provide evidence at the level of significance of 5% to allow us to say that the first-choice course matters in their academic performance. However, according to the results obtained by Ferrão & Almeida

(2018b), the admission of the student to his/her 1st option programme is important for his/her persistence over the first year of studies in Higher Education. The estimates presented suggest that the odds ratio in the persistence of students who were admitted in their 1st option programme is 1.85. This means that the probability of persistence over the probability of dropping out, suspending or transferring is almost twice as high in the group of students who were admitted to the 1st option course compared to students not admitted in their first choice. Those authors add “the conditions of access to higher education, namely the condition underlying the entrance classification, can contribute to the student's choice being conditioned” (p. 674).

The statistical model for first-year academic performance shows that the students' socioeconomic level has influence through the increased motivation of the subgroup of students whose parents have no more than a lower secondary education. This effect seems to represent the willpower of students from socioeconomically disadvantaged groups who were admitted to the university and strive to achieve good results, even because the good achievement is the requirement for keeping the scholarship. This finding may contribute to the literature on social equity in higher education and to the debate on the motivational component of social status (e.g. Eckland, 1964; Tinto & Cullen, 1973; Cantwell, Archer, & Bourke, 2001) but needs to be confirmed with more cohorts of entrants.

Such results also give quantitative support to the conjecture advanced by Engrácia & Baptista (2018), according to which

...children of parents with low schooling who, despite the contrary winds, can finish high school and enter higher education, will be, on average, young people with school performance and levels of resilience above average among students from disadvantaged backgrounds. Therefore, their subsequent performance in higher education is expected to be more comparable to the performance of students from favored socio-economic backgrounds. (p. 18)

To our knowledge, for the first time in Portugal, the results show the advantage in the GPA that students of disadvantaged sociocultural origin have compared to their colleagues, even after controlling for the remaining variables.

We believe that our study constitutes a meaningful contribution to the available research by using microdata from a large sample of first-year students, by considering their schooling trajectory through retrospective data, and by taking into account the hierarchical structure of the data in statistical modeling. However, care should be taken not to generalize the results. The study concerns the entering cohort of 2015/16. Had different cohorts been selected, the results might have been affected. A full investigation of different cohorts was outside the scope of the current study, but it would be very interesting in the near future to investigate the sensitivity of these results to new nationwide cohorts of students.

CAPITULO 5 – Multilevel modeling of persistence in higher education⁶

Abstract

The dropout or evasion rates in higher education are now a social and institutional concern, justifying the implementation of public policies to prevent this phenomenon. These policies need studies on the most determinant variables of the risk of dropout. The main objective of this study is to analyze the student's persistence in undergraduate courses, and the relationship with the student's previous school trajectory and with the conditions of entrance into higher education, controlling for students' sociodemographic characteristics, such as gender and age. We applied multilevel logistic regression models to data of 2,697 freshmen enrolled in a Portuguese public university in the academic year 2015/16. The results suggest that failure in basic education (ISCED 2) has a long-term effect. According to the estimates obtained, students who declare not having failed in basic education have odds ratio of persistence 2.7 times higher than students who declare having failed in basic education. The conditions of student's admission to the course he/she attends are relevant variables to persistence in Higher Education, for example, whether s/he was admitted to her/his first option course and the student's university entrance score. The results also show that older and male students have lower probability of persistence.

5.1 Introduction

In Portugal, the annual number of admissions to undergraduate courses has been around 60,000 students in recent years. From 1971 to 2011, the variation in the number of students enrolled in higher education was 700%, ranging from 49,461 in 1971 to 396,268 in 2011. Dias (2015) presented a detailed analysis of official statistics on how this strong increase was achieved, and raised the question of whether such an increase ("massification") led to greater equity not only in terms of access but also success in higher education. In fact, this growth implied an increase in the heterogeneity of the student population, in terms of gender,

⁶ Baseado no artigo Ferrão, M. E., Almeida, L. S. (2018). Multilevel modelling of persistence in higher education. *Ensaio: Aval.Pol.Públ.Educ.*, 26 (100). Retrieved from <http://dx.doi.org/10.1590/s0104-40362018002601610>.

socioeconomic level, age group and past academic trajectories, not always accompanied by changes in attitudes and practices by teachers and institutions.

There is a broad consensus on the need for universities to promote their students' success, since many of them do not achieve their goals or do not fully realize their potential during university experience, which is a waste of talent and human potential. In the last years, dropout rates have increased in Portugal, and institutions have been questioned about their policies to support student retention (Dias et al., 2011; Ferreira et al., 2014), in particular with regard to the goals of the Europe 2020 strategy (European Commission, 2013) targeting 40% of young adults with a higher education diploma. Several authors (for example, Tinto & Pusser, 2006) report that despite the abundant research and debate on dropout (attrition, withdrawal process or evasion), less attention has been given to the development of a model for action aimed at increasing the persistence and, as such, improving the students' success. The authors propose a model that includes institutional factors as a complement to the students' variables. Such a perspective, already present in Tinto and Cullen (1973), for example, considers that each student exists in a particular context that shapes his/her probability of success in post-secondary education. A specific context for student success is shaped by a variety of contextual factors, including demographics, history, culture, state or locality characteristics, the institutional matrix, available resources, and existing policies. Within and around a given context, competition from interest groups generates demands for institutional policies that can change the likelihood of student success. Also Braxton, Hirschy and McClendon (2004) argue that the responsibility for success in higher education does not lie solely with the individual student. Instead, the process of dropout is understood as an interaction between the student and the university. According to Georg (2009), students do not consider abandoning due to stress or lack of ability, but mainly due to poor commitment to their course or area of study. According to the results presented by the author, the institutional influence on the student's propensity for dropping out is therefore modest. In the same sense, the results of the research presented by Pike and Graunke (2015) emphasize student variables in particular vocational and career issues. Tinto and Cullen (1973) report that evidence-based studies of dropout in only one institution have serious limitations. As institutions differ from one another, they diverge in dropout rates and associated factors, even after taking into account the characteristics of the students. Different institutions present differential effects on the retention of students from different sociocultural backgrounds. Therefore, according to the authors, studies based on

more than one institution are “not only difficult to interpret but also difficult to compare to single institutional studies even when institutional characteristics are taken into “account”” (p. 10). It should be borne in mind that such review – one of the most cited in the literature on higher education – is based on a set of 130 articles, books or reports published in the USA between 1951 and 1973. The empirical evidence included in those publications does not use multilevel or hierarchical regression models that are currently more widely disseminated and accessible. Today, multilevel modeling gives us the possibility of analyzing the search for unprecedented empirical evidence, as can be seen in more recent papers (Draper & Gittoes, 2004; Soares et al., 2001).

On the other hand, the phenomenon of dropout has some complexity in terms of collecting data to ensure the reliability of the measure. Among the studies considered in the review, diversity was considerable in terms of the definition of dropout and its operationalization through the collection of data. According to those authors, the definitions of dropout can be divided into two main groups: (a) referring to the individuals who leave the higher education institution where they are enrolled; (b) referring only to individuals who, having entered higher education, did not complete their training in any higher education institution and never received a degree.

Studies based on single institution data are classified in the first group and, generally, this group also includes studies that consider as dropping out those students who left the institution where they were enrolled, and decided to transfer to another higher education institution. The higher education system works dynamically. It is a system where institutions and students coexist, where institutions have diversified offers and individuals, with varied characteristics and different expectations, take an active role in the search for higher education (vocational choices and professional career perspectives may be the reason for course or institution changes). It is also usual to classify in this group individuals who temporarily suspend their attendance at higher education.

However, we find in the literature concepts that aim to improve the approach to the phenomenon that occurs in reality. The study of phenomena related to dropout, such as the concept of retention (Permzadian & Credé, 2016; Tinto, 2010; Willson & Hughes, 2006), persistence or permanence (Gross, Hossler, Ziskin, & Berry, 2015; Ishitani, 2016; Johnson, 2008; Rhee, 2008) contributes, among other things, to improving the reliability of the measure concerning the phenomenon under research. Empirical studies have shown that the determinants affecting temporary cessation and transfer are different from those that influence

dropout (Rhee, 2008). For example, the study of Pascarella et al. (1981) is illustrative of the differential effects of individual factors for each type of dropout. Specifically, they sought to determine the role of pre-entry characteristics in predicting persistence, suspension, and dropout. The authors found that the patterns of discrimination among the three phenomena were considerably different. In addition to the preentry characteristics, the process of admission to higher education, including the conditions of access and choice of institution and course, also has an effect on student persistence and success (Almeida et al., 2006; English & Umbach, 2016; Soares et al., 2001). Some authors (Bittencourt & Mercado, 2014) refer that the main causes for dropping out from distance education courses may be related to students' dissatisfaction with the tutor/teachers or with technological problems, in particular with the platform. The limited number of publications on student's persistence in distance education courses shows that the subject has received little attention from researchers in education. Such specificity goes beyond the scope of this article dealing with evasion in face-to-face courses. The access to Portuguese higher education is guided by a *numerus clausus* system established through a national public tender, in order to prioritize the admission of students with higher access classification. Approximately 40% of students are placed in their first option on course and institution. Portela et al. (2008) have suggested that students rank courses and institutions implicitly when making their choices in the access to the university. This raises questions about the influence of first choice-admission on individual student performance throughout his/her undergraduate course. Thus, the interaction between student and university begins when the student chooses the university and the course he/she intends to attend. Several authors have emphasized that the student's socioeconomic and cultural capital is associated with a preference for more prestigious courses or institutions (Tavares et al., 2008; Tavares & Ferreira, 2012). According to Tavares et al. (2008), students who come from disadvantaged backgrounds have a greater preference for courses related to teacher training and management, while students who have a privileged family background are more likely to choose courses in law, art, health, science and technology.

Studies on Portuguese higher education provide important descriptive analysis, but are mostly based on aggregated data. In the scope of this study, we will use microdata and the concept of student persistence, an attribute measured through the observation of students enrolled for the first time in the first year at the University of Minho in the academic year 2015/6 and who reached the end of the first year with GPA (Grade Point Average). This means the student did

not give up during the first year, did not suspend their studies, did not transfer to another course/institution or did not finish the school year without at least one curricular unit approved. We will use the students' microdata to model the phenomenon of interest. To our knowledge, this is the first study to apply multilevel regression models to Portuguese microdata on persistence in higher education.

As suggested by the literature, the risk of dropping out is higher during the first year (Bardagi & Hutz, 2009; Chen, 2012; Georg, 2009). Programs and practices aimed at reducing dropout can maximize their effects by placing more emphasis on the permanence of students in the first year. Given the hierarchical structure of the data, we will apply a multilevel logistic regression model, similar to other studies (Draper & Gittoes, 2004), where student is the level 1 unit and course is the level 2 unit. This study aims to answer the following research questions: (1) Is the first choice/admission in the university or in the course important for the persistence of the student enrolled for the first time in the first year? (2) Is there any empirical evidence in Portugal that the classification of university entrance is predictive of the persistence of the student enrolled for the first time in the first year? (3) Is there any evidence of the long-term effect of the school trajectory, namely repetition in primary education and/or secondary education, in the student's persistence in higher education?

5.2 Methods

5.2.1 Participantes e variáveis

The sample consists of 2,697 first year students who were enrolled in the University of Minho, a public university in the north of Portugal, in the academic year 2015/2016. The data used for the purpose of this work relates to the following attributes of the student: persistence (1: if the student reached the end of the first year with academic performance assessed through an average score; 0: otherwise), university entrance classification ; first option admission (yes/no) in the institution and in the undergraduate course; sex (1: male; 0: female); trajectory of schooling assessed by the experience of early repetition (0: yes; 1: no); education of parents as a proxy for student 'socioeconomic status. Most of the students are female (57%), enrolled in their first choice course (59%) and in the first choice college (72%). Almost half of the students (47%) enrolled in a STEM (Science, Technology, Engineering and Mathematics) course, 17% in Economics, Administration or PublicAdministration, 13% in Humanities, 10% in Social

Sciences, 8% in Health or nursing and 5% in Law. The students' ages range from 16 to 61 years, with an average of 18.9 ($SD = 3.6$) and 91.9% are full-time students. The average entrance classification at the university is 152.4 ($SD = 18.9$). Most students (83.3%) state that they have been always promoted throughout their schooling trajectory and 13.9% mentioned at least one repetition. The distribution of students by parents' education shows that 34.4% of students have parents with no more than basic education (ISCED 2) and 15.7% of students have parents with a higher education (at least ISCED 4) diploma. Student's academic performance (GPA) data is available for 72% of the students. The other cases (28%) are not available due to different situations, such as non-achievement in all curricular units, dropout or transfer to other institutions. Descriptive statistics for this subset of students with no academic scores at the end of the first year show that they are generally older, most likely to have been repeaters in primary (ISCED 2) or secondary education (ISCED 3), and have, on average, lower grades of university entrance.

5.2.2 Procedures

The students were informed about the objectives of the Observatory study and we guaranteed the confidentiality of the data. The students gave their informed written consent, also allowing access to their classifications (GPA) at the end of the first semester and at the end of the first year. The remaining data was collected at the time of the student's registration at the university. For statistical modeling purposes, we decided to standardize the variable that represents the university entrance score. The GPA of the first year is missing for 757 students, of whom 57% are in the "dropout" situation from the University of Minho, 17% did not enroll in any course, and the remaining 27% may have suspended their studies temporarily or did not get approval in any curricular unit. The majority (53%) are female students. Statistical modeling aims to contribute to defining the profile of students who enter for the first time the University of Minho and reach the end of the year with GPA, a proxy for persistence, compared to students who for any reason – suspension, dropout, transfer or failure – did not obtain an average grade at the end of the first year.

5.2.3 Statistical modeling

For the modeling of persistence, we considered a two-level logistic regression model (students grouped in courses), with a logit link function. The variable of interest, Y , represents the situation of the student in relation to the demonstration of persistence in the course in which he/she enrolled. It is a binary variable with a value of 1 if the student reached the end of the first year and obtained GPA score and 0 otherwise.

The probability of persistence of student i in course j is denoted by $P(y_{ij} = 1)$ with $i=1, \dots, n_j$, $j=1, \dots, J$, where J is the number of courses ($J=54$) and n_j is the number of students in course j . The equation of the multilevel logistic regression model is written as follows,

$$\log \left[\frac{P(y_{ij}=1)}{1-P(y_{ij}=1)} \right] = \beta_{0j} + \beta_1 x_{1(ij)} + \dots + \beta_9 x_{9(ij)}, \quad (1)$$

$$\beta_{0j} = \beta_0 + u_{0j},$$

$$u_{0j} \sim N(0, \sigma_{u0}^2),$$

where there are additive terms referring to nine explanatory variables, and the function is designated by logit function, or by logite in Portugal, logito in Brazil (Paulino et al., 2011). Details about the model can be found, for example, in Goldstein (2011) or in the synthesis presented by Ferrão (2015b). The explanatory variables, X1 through X9 in equation (1), are defined as follows. The variables X1 and X2 represent individual characteristics of students, gender and age, respectively. The variables X3 and X4 are dummies and refer to the level of education of parents or guardians, that is, X3 represents the group of students whose fathers and mothers do not have more than basic education, and X4 represents the group of students whose fathers and mothers hold a higher education diploma. In this design, the reference group consists of the group of students whose parents hold any other combination in terms of level of education. With such a design, we have attempted to represent in X3 the lower tail of the parents' education distribution and in X4 the upper tail. Variables X5 and X6 represent the student's situation in relation to the experience of repetition in primary and secondary education, respectively. The variable X7 represents the student's university entrance score. The conditions of admission to the university are also obtained through variables X8 and X9, indicating if the student was admitted to the course and the university, respectively, in his/her first option.

We used the MLwiN 2.31 statistical software (Rasbash et al., 2014) and the estimation procedure PQL2 (Goldstein & Rasbash, 1996). We considered missing values occurring completely at random (Little & Rubin, 2002).

5.3 Results

Table 1 contains the estimates of the fixed and random parameters of the multilevel logistic regression model. Model M1 considers the individual variables of the student and the family background, the M2 model additionally includes the variables of the previous schooling trajectory and the M3 model adds the variables that represent the conditions of admission to the university. Finally, the model was adjusted with variables whose coefficient parameters are statistically different from zero at the significance level of 10% and the results are presented in Table 2.

The relationship of the individual variables remained statistically significant in all models, indicating that females and individuals up to 19 years of age were more likely to persist in the first year of studies. The estimates obtained in the final model suggest that the odds ratio in the persistence of students aged up to 19 years is almost six times (5.94) higher when compared to students over 19 years of age.

The association between family background variables that assess the parents' educational level and the probability of persistence is not statistically different from zero in any of the M1, M2, M3 models, so they were not included in the final model. This result shows that we did not find evidence of the influence of the socioeconomic status of the student's family, as measured by the parents' schooling, in the probability of persistence.

Table 1. Estimates of fixed and random parameters

	M1		M2		M3	
	Estimate	SE	Estimate	SE	Estimate	SE
<i>Fixed parameters</i>						
Constant	1.709	0.174	0.881	0.365	0.471	0.385
Male vs. Female	-0.456*	0.124	-0.416*	0.128	-0.336*	0.131
Age>19	-1.856*	0.161	-1.701*	0.169	-1.771*	0.176
Parents' Education: lower than 9th grade	-0.129	0.118	-0.121	0.128	-0.140	0.130
Parents' Education: higher education	-0.099	0.167	-0.201	0.173	-0.186	0.176
Repetition at basic education (no vs. yes)			0.939*	0.319	0.955*	0.323
Repetition at Secondary Education (no vs. yes)			0.301	0.186	0.299	0.188
Entrance HE classification score			0.224*	0.085	0.148**	0.088
First option course (yes vs. no)					0.591*	0.132
First option university (yes vs. no)					0.070	0.142
<i>Random parameters</i>						
Level 2: Course						
Variance	1.102*	0.251	1.084*	0.253	1.102*	0.261
Level 1: Student						
Variance	1.000	0.000	1.000	0.000	1.000	0.000
Number of level 2 units: Course	55		55		55	
Number of level 2 units: Student	2438		2378		2328	
Estimation Procedure:	RIGLS (PQL2)		RIGLS (PQL2)		RIGLS (PQL2)	

Source: own calculation. * p value < 0.05; ** p value < 0.10.

Table 2. Estimates of fixed and random parameters in the final model

	Final model		
	Estimate, $\hat{\beta}$	SE	$\text{Exp}(\hat{\beta})$
<i>Fixed parameters</i>			
Constant	0.374	0.365	1.454
Male vs. Female	-0.333*	0.129	0.717
Age>19	-1.782*	0.172	0.168
Parents'Education: Lower than9th grade	—	—	—
Parents'Education: Higher Education	—	—	—
Repetition at Basic Education (No vs. Yes)	1.008*	0.318	2.740
Repetition at Secondary Education (No vs. Yes)	—	—	—
Entrance HE classification score	0.140**	0.084	1.150
First option course (Yes vs. No)	0.614*	0.126	1.848
First option university (Yes vs. No)	—	—	—
<i>Random parameters</i>			
Level 2: Course			
Variance	1.102	0.264	
Level 1: Student			
Variance	1.000	0.000	
Number of level 2 units: Course	55		
Number of level 2 units: Student	2392		
Estimation Procedure:	RIGLS (PQL2)		

Source: own calculation. * p value < 0.05; ** p value < 0.10.

We found a statistically significant association between the situation regarding repetition in basic education and the probability of persistence in the first year of higher education. The estimate obtained in the final model, 1.008 with a standard error of 0.318, suggests that students who declare not to have failed in primary, elementary and lower secondary education present odds ratio in the persistence of 2.7 times higher than students who declare to have failed in any of the basic education levels. No statistically significant association was found between the failure in secondary education and the probability of persistence in the first year. The influence of the university entrance score is positive and statistically significant, indicating that the higher the score, the greater the probability of persistence, but in model M3 and the final model, this is verified only for the level of significance of 10%. That is, when the variable

that captures the admission condition in the undergraduate course is included in the model, the influence of the entrance score loses magnitude and is no longer statistically significant at the 5% level. The odds ratio in the persistence of students who are admitted in the first option course is 1.85, for example, the probability of persistence over the probability of dropping out, suspending or transferring is almost twice as high in the group of students who are admitted in the first option course compared to what happens in the group of students who are not admitted in their first choice.

5.4 Discussion and Conclusion

In this study, we used the concept of student's persistence, measured by the observation of 2,697 students enrolled for the first time in the first year of the University of Minho in the academic year 2015/16 and who reached the end of the first year. We applied a multilevel logistic regression model, considering student as the level 1 unit and course as the level 2 unit, with multiple objectives: to quantify the effect of the admission conditions in the course and university in the behavior of the student concerning the persistence; to verify whether university entrance score is predictive of student's persistence; to verify the long-term effect of the schooling trajectory, namely repetition in primary education and/or secondary education, in student's persistence.

According to the results, the admission of the student in the course chosen as first option is important for his/her persistence. The estimates obtained suggest that the odds ratio in the persistence of students who are admitted in the first option course is 1.85. This means that the probability of persistence over the probability of dropping out, suspending or transferring is almost twice as high in the group of students who are admitted to the first option course compared to students not admitted in their first choice. The conditions of access to higher education, namely the condition underlying the entrance classification, can contribute to the student's choice being conditioned. The interaction between the first option and the entrance classification is shown by comparing the results of the model M2 and the model M3, where the relationship between classification and persistence probability loses magnitude and is no longer statistically significant at the significance level of 5%. However, that ratio remains statistically significant at the 10% level. Even if conditioned to the university entrance score, the effect of admission to the first option course seems to catch the motivation and commitment of the

student to finish the course successfully. This corroborates the literature according to which the higher the motivation and commitment, the less likely the student is to drop out of higher education (English & Umbach, 2016; Tinto & Cullen, 1973; Tinto & Pusser, 2006). The research conducted by Ferrão and Almeida (2018a) showed that the university entrance score is the best predictor of the GPA obtained at the end of the first year at the University of Minho. The authors refer that the student's prior knowledge and preparation are essential to his/her success in higher education, confirming the international literature. By extension, if students are not well prepared for the course they enroll in, they are not successful regardless of persistence. The more students learn and value this learning, the more likely they are to stay in courses and get a diploma. In this sense, an institutional program that aims to increase the rate of persistence in undergraduate courses through early intervention should pay special attention to the group of students at risk. According to the results reported above, such group consists mainly of students who were admitted to a course other than the first option and/or students who entered higher education over the age of 19. For the time being, our discussion leaves open the analysis of student's success and persistence as a result of successful learning in the classroom. The student's success and persistence is a consequence of student-institution interaction that aims at academic progression. One of the results to highlight in this interaction is the role of the teaching staff for the institutional efforts to promote student success (Umbach & Wawrzynski, 2005). Some research findings (for example, Andriola, Andriola, and Moura, 2006) indicate that most teachers and coordinators sustain the opinion that student's evasion could be reduced if the role of the mentoring teacher or tutor was rescued. Regarding the long-term effect of the student's school trajectory – specifically the repetition in basic and/or secondary education – in his/her persistence in higher education, our results suggest that the group of students who declare to have failed in basic education level have lower probability of persistence. According to the estimates obtained, those who declare that they did not fail in basic education have odds ratio 2.7 times higher than students who declare to have failed. This result refers to the marginal effect associated with the student's age, which suggests that the student entering higher education aged 19 years or younger has a persistence odds ratio almost six times higher than his/her older colleagues. The delay in the age of admission may be due to failure in primary or secondary education, although it may also be due to the postponement of the university application after obtaining a high school diploma. Thus, there is a cumulative effect associated with age and failure in primary education. We noted that if the age variable was not included in the linear predictor of the model, the odds ratio in persistence between the

groups – non-failure vs. failure in basic education – would increase to 4. This study is the very first to report the longterm effect of early repetition on higher education, and reinforces the argument presented by Ferrão (2015a) according to which the early repetition is a strong predictor of late repetition. According to the author, based on PISA 2012 data, the probability of early repetition is 35 times higher in the first tenth of the distribution of students' socioeconomic conditions compared to the tenth higher, indicating a lack of social equity in the Portuguese educational system. She adds "These results state a problem of equity in the Portuguese educational system. The distribution of educational resources does not guarantee the full development of the children or young adult regardless of their socioeconomic level" (p. 14). This kind of evidence brings to the debate the theme of the selection made by the education system and the process of access/admission to higher education (Chiroleu, 2012; Tavares et al., 2008; Travitzki, Ferrão, & Couto, 2016) jeopardizing future social mobility. However, studies on affirmative action programs in other countries, such as Brazil, show that the rate of persistence of the beneficiary students (quota holders) is higher, which corresponds to the higher graduation rate (Mendes Junior, 2014). That is, despite the fact that these students enter the State University of Rio de Janeiro with a lower average score and obtain the diploma with lower GPA, the graduation rate in this group was always higher in 2009, 2010 and 2011 compared to non-beneficiaries.

The "Third Commitment", enunciated by Raymond and Negassi (2015), "aims at equity of exposure to fundamental knowledge and skills, sometimes called equal treatment or equal opportunities", in other words, from the point of view of curriculum development, the system must ensure that all Portuguese young people "are ready to embrace the challenges of higher education or to achieve a viable job". However, the goals of learning for secondary education have been mostly focused on general education, giving access to higher education. Although the growth of alternative vocational training has provided greater freedom of choice to students of secondary education level, the learning goals for vocational studies are more heterogeneous, often linked to specific professional skills and knowledge of the craft. Specifically, this basis can improve the low levels of completion [of secondary education] – about 50 per cent – that are taking place in Portugal today and improve the wage differential between vocational training students and those who have followed a path of general education throughout their careers (Raymond & Negassi, 2015, p. 17).

A new strand of research on equity may be opened as one confronts the alternatives of education and training offered by the education system and the results obtained in this work, particularly concerning the conditions of access to higher education. Therefore, the implications of this study for future research fall on the detail of the area of study associated with the first option of the course and the quantification of its predictive capacity in the probability of persistence. We believe that this study contributes in an innovative way to the research in the sub-area of higher education, specifically on the subject of student's persistence: by analyzing and modeling the microdata of a large sample of students enrolled for the first time in the first year, taking into account the hierarchical structure of the data and thus contemplating simultaneously the statistical units of student and course, considering the long-term effect of student's schooling trajectory, namely the effect of school failure, and taking into account the conditions of admission to the university. However, some caution is needed not to generalize the results. The study concerns the enrollment cohort of 2015/16. If other cohorts had been selected, the results could be slightly different. In this perspective, it is very important to develop a broader research project involving samples from different cohorts of newcomers and multiple institutions in order to obtain national results.

CAPITULO 6 – Personal and institutional factors’ impact on persistence in higher Education: A study with first-year Portuguese students

Abstract

The dropout rates in higher education are now a social and institutional concern, justifying the implementation of public policies to prevent this phenomenon. Educational research has shown that student background characteristics, such as parental socio-economic status and prior academic ability, influence student performance and persistence in Higher Education. This paper presents a study on the relationship of student’s persistence with expectations and study methods, controlling for students’ background, sociodemographic characteristics and for conditions of access to higher education. Multilevel logistic regression models were applied to data of 2,967 first-year undergraduate students enrolled in 55 courses at a public university in Portugal. The results suggest a strong influence of study methods on persistence, as well as an important role played by expectations on older students’ persistence. In addition, this study tests course specificity through the differential effect of student’s university entrance score on persistence, which was found to vary at random across courses. Implications of the results for institutional practices and future research are discussed.

Keywords: Higher Education, Persistence, Expectations, Study Methods, Multilevel logistic regression.

6.1 Introduction

The dropout from higher education (HE) by students runs counter to the need for greater academic qualification of individuals in response to the growing complexity of society and the labor market. In the face of public investment in higher education, dropout represents an economic loss for society and families (Aina, 2013; Montmarquette, Mahseredjian, & Houle, 2001). Due to this fact, dropout is a reason for growing concern among academic institutions that meet one of the quality criteria and/or the amount of their public funding in the rate of persistence and completion of graduation.

Dropout in the context of HE takes different forms, each differentiated in the constellation of inherent motives (e.g. van Rooij, Jansen, & van de Grift, 2017). If abandonment by exchange of course or institution can be associated with the students' vocational choices, for example the search for the training that best corresponds to professional career projects, the two-year or three-year dropout may be related to personal needs for new experiences or the necessity of taking on new family and/or professional responsibilities. Finally, the most worrying scenario is the permanent abandonment of students who have not adapted or left as an investment in HE for an added value for their training and qualification for the labor market. When the initial expectations of the students do not materialize, the inherent frustration translated to less involvement and availability to fulfill the academic challenges and to benefit from the resources available from the institution, which may also explain the higher dropout rates in the first year (James, 2002; Nicholson, Putwain, Connors, & Hornby-Atkinson, 2013).

Research on dropout highlights the impact of a heterogeneous set of personal and institutional variables, and, above all, their interaction (Tinto, 1993). By listing some personal variables, the studies show that students from families with better socioeconomic status have a higher rate of permanence and completing their graduation, and this situation may be associated with the fact that they bring more academic competencies from secondary education, more easily surpassing the curricular requirements of HE. Alternatively, it may be associated with attending courses with more social prestige or with students that do not need to reconcile studies with work activities necessary to finance attending of their studies (e.g. Letkiewicz et al., 2014). In Portugal, with a system of *numerus clausus* for filling vacancies by course/institution, students from more favored socioeconomic groups access to HE at higher rates. This permits them to attend courses of their first option or choice, that is, the more socially valued courses (Fonseca, Dias, Sá, & Amaral, 2014). In turn, attending a first-option college is related to higher student GPAs. A first-option course increases the probability of student's permanence in HE (Ferrão & Almeida, 2018b). Other studies report the impact of the gender variable, suggesting that women complete their academic education to a greater extent, and here too the possibility of being better prepared academically from secondary education should be noted, because some dropout occurs after the initial examinations and classifications are less positive.

Educational research has shown that prior academic performance may be a strong predictor for persistence in HE (e.g. Pinxten, Van Soom, Peeters, De Laet, & Langie, 2017). Considering a sample of 3,418 first-year students attending the University of Montréal (Montmarquette et

al., 2001), the parameter estimates of probit models for persistence over the first period suggest that the student's scores in pre-university achievement is statistically significant, at a level of significance of 10%, in the restricted sample, but is not statistically different from zero for the full sample of students. According to the authors, "the sample is restricted to students enrolled in a program with an entrance quota" (p. 479) and consists of 2,328 students. Since in the full sample the coefficient is not statistically significant, it raises the question whether course specificity mediates the relationship between the student's prior achievement or student's university entrance score and the probability of persistence in HE. Also, Ferrão e Almeida (2018b) found an association between university entrance score and student persistence, at a level of significance of 10%. In other words, considering the estimates of fixed parameters, at the commonly used 5% level of significance, both sources of evidence would let us conclude not to reject the null hypothesis for the association between student's prior achievement and persistence. This study contributes to the topic by showing that such association varies from course to course, i.e., course specificity mediates such an association.

As this prior academic achievement is marked socially, it is suggested that the democratization of access for students from more disadvantaged sociocultural groups does not correspond to the success rates and completion of higher education by non-traditional students (Lohfink & Paulsen, 2005). The growing social diversification of university students is thus a challenge for institutions that must structure ways of acting and services to support students with less familiar social and personal references to HE (Dias et al., 2011; Shaw, 2013). Considering the demands and challenges of the university context, it is necessary to promote the development and mobilization of the students' personal resources for their academic success (Cabras & Mondo, 2018; Nicholson et al., 2013; Robbins et al., 2004). Regarding institutional variables, Tinto (1993, 1997) argues that dropout is associated with a set of factors that do not favor the student's academic and social integration. Classrooms, laboratories, libraries, services, among other contexts of students' academic life, have a direct impact on academic and social integration in the institution, leading students to rebuild their initial goals and commitments, thus influencing their involvement and academic achievement, and their decision to leave the HE. This influence tends to be mediated by the student's effort, their satisfaction and their sense of belonging to the university and to a particular course (Adams & Proctor, 2010; Ostrove & Long, 2007; Tinto, 1997). Teachers' teaching methods and relational proximity to students are relevant to academic success and thus affect the risk of dropout (Hébert, 2018; Pascarella,

Pierson, Wolniak, & Terenzini, 2004). When such teaching practices and timely feedback favor the participation of students, they attend more classes and perform the curricular tasks requested by their teachers, thus obtaining better academic performance (Credé, Roch, & Kieszczynka, 2010; Johnson, Richeson, & Finkel, 2011; Pascarella et al., 2004). It is also important to consider eventual difficulties related to participating in social and leisure activities. Scholarships or other forms of financial aid favor persistence and completion rates (Bettinger, 2015; Dynarski & Scott-Clayton, 2013; Ross et al., 2012), particularly for the students from lower-income families (Cabrera, Stampen, & Hansen, 1990; Nora, Cabrera, Hagedon, & Pascarella, 1996). Without the need to maintain a job, these students can become more involved in the academic and social activities of the campus, which is not always the case for older students who need to reconcile working hours with classes and study activities.

This study aims to answer the following research questions: (1) Is the student's method of study associated with the probability of persistence in higher education? (2) Are the student's expectations related to the probability of persistence? (3) Are the conditions of entrance to higher education (university entrance score and first-option course admission) predictive variables for the probability of student's persistence in the first year?

6.2 Methods

Ethical statement

The study protocol was in accordance with ethical standards (OECD, 2007). The students were informed of the objectives and they gave informed written consent, also allowing access to their grades (GPA) at the end of the first semester and at the end of the first year Almeida and Castro(2016). No incentives were provided for participation in the study.

Participants

The sample consists of 2,697 first year students who were enrolled in a public university located in the north of Portugal. The data used for the purpose of this work relate to the following student attributes: persistence (1: if the student reached the end of the first year with academic performance assessed through an average score; 0: otherwise), university entrance classification; first-option admission (yes/no) in the institution and in undergraduate course;

sex (1: male; 0: female); trajectory of schooling assessed by the experience of early repetition (0: yes; 1: no); education of parents as a proxy for student' socioeconomic status. Most of the students are female (57%), enrolled in their first-choice course (59%) and in their first-choice college (72%). Almost half of the students (47%) enrolled in a STEM (Science, Technology, Engineering and Mathematics) course, 17% in Economics, Administration or Public administration, 13% in Humanities, 10% Social sciences, 8% Health or nursing and 5% in Law. The students' ages range from 16 to 61 years, with an average of 18.9 (SD = 3.61) and 91.9% are full-time students. The average entrance classification at the university is 152.4 (SD = 18.93). Most students (83.3%) state that they have been always promoted throughout their schooling trajectory and 13.9% mentioned at least one repetition of a grade level. The distribution of students by parents' education shows that 34.4% of students have parents with no more than basic education and 15.7% of students have parents with a higher education diploma. The students' academic performance (GPA) data are available for 72% of students. The other cases (28%) are not available due to different situations, such as failure in all curricular units, dropout or transfer to other institutions. Descriptive statistics for this subset of students with no academic scores at the end of the first year show that they are generally older, most likely to have been repeaters in primary or secondary education, and have, on average, lower grades for university entrance. Further descriptive characteristics are presented in Table 1.

Table 1. Descriptive Characteristics of Students

	Number of valid cases	Min	Max	Mean	SD
University entrance score	2,430	105	200	152.39	18.91
Study Methods (z scores)	2,459	-4.60	1.92	0	1
Expectations (z scores)	2,332	-6.44	1.70	0	1
Age	2,690	16	61	18.85	3.59
Age>19	2,690	0	1	13.1	-
First-option course	2,450	0	1	58.7	-
Male	2,695	0	1	43.5	-

Instruments

Academic Perceptions Questionnaire - Expectations (Almeida, Deaño, et al., 2012). This instrument explores students' beliefs and aspirations in the transition to higher education,

namely, what they expect to find and to develop. Its items combine cognitive and motivational aspects of academic experience and covered several dimensions: (a) Training for job and career development; (b) Personal and social development; (c) International student mobility; (d) Political engagement and citizenship; (e) Social pressure; (f) Training quality in the course; and (g) Living together and social interaction. Students rate their agreement with item content on a 6-point Likert-type scale ranging from 1 (*completely disagree*) to 6 (*completely agree*). From the intercorrelations among the seven dimensions, a general factor of expectations was estimated by the first principal component given by principal component analysis (Jolliffe, 2002; Nicholson et al., 2013).

Study Methods Grid. Considering their experience in secondary school, students evaluate their study routines described in six situations (take notes in classes, have a weekly schedule, perform the jobs in the defined timings, etc.). Students rate their agreement with each situation on a 6-point Likert-type scale ranging from 1 (*completely disagree*) to 6 (*completely agree*). For the purpose of this study, the first principal component given by the technique of principal component analysis (Jolliffe, 2002) was used as a global score for study methods.

Socio-academic questionnaire. Students answered several short questions reporting personal information related to, for example, previous academic trajectory, parents' academic degrees, if the course and the college were their first option.

Procedures

The students were informed about the objectives of the study and were guaranteed the confidentiality of their data. The students gave their informed written consent, also allowing access to their grades (GPA) at the end of the first semester and at the end of the first year. The remaining data were collected at the time of student registration at the university. For statistical modeling purposes, we decided to standardize the variable that represents the university entrance score.

The GPA of the 1st year is missing for 757 students, of whom 57% are in a "dropout" situation from the university, 17% did not enroll in any course, and the remaining 27% may have suspended their studies temporarily or did not get approval from any curricular unit.

Statistical modeling

Statistical modeling aims to contribute to define the profile of students who enter for the first time in the university and reach the end of the year with GPA, a proxy for persistence, compared to students who for any reason - suspension, dropout, transfer or failure - did not obtain an average grade at the end of the first year.

Consider the basic structure of educational organization, in this case courses, indexed by j and students within courses indexed by i . The response variable is student's persistence. It is a binary variable with a value of 1 if the student reached the end of the 1st year and obtained GPA score, and 0 otherwise. We applied a two-level random coefficient model with a logistic function which is the one most commonly used in the social sciences. The probability of persistence of student i in course j is denoted by $P(y_{ij} = 1)$, with $i = 1, \dots, n_j, j = 1, \dots, J$, where J is the number of courses ($J = 54$) and n_j is the number of students in course j . We write the statistical model as follows,

$$\log \left[\frac{P(y_{ij}=1)}{1-P(y_{ij}=1)} \right] = \text{logit} (P(y_{ij} = 1)) = \beta_{0j} + \beta_1 x_{1(ij)} + \dots + \beta_{4j} x_{4(ij)} + \dots + \beta_{10} x_{10(ij)}, (1)$$

$$\beta_{0j} = \beta_0 + u_{0j}, (2)$$

$$\beta_{4j} = \beta_4 + u_{4j}, (3)$$

$$\begin{pmatrix} u_{0j} \\ u_{4j} \end{pmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{u0}^2 & \sigma_{u04} \\ \sigma_{u04} & \sigma_{u4}^2 \end{bmatrix} \right).$$

In the first line of the model, the predictor variables, denoted by X1 through X10, constitute additive terms where the respective coefficients $\beta_1, \dots, \beta_{10}$ represent the fixed parameters related to each of the explanatory variables. These are defined as follows: X1 represents the student's study method, X2 represents the student's expectations, X3 is the 1st order interaction term between expectations and age, X4 represents the student's university entrance score, X5 and X6 represent the individual characteristics of students, gender (1: Male; 0: Female) and age (1: age > 19; 0: otherwise), respectively. X7 is a dummy variable indicating whether the student was admitted to the course in his/her 1st option and the variable X8 represents the student's experience of repetition in basic education (1: No; 0: Yes). The variables X9 and X10 are dummies and refer to the level of education of parents or guardians, that is, X9 represents the group of students whose father and mother does not have more than basic education, and

X10 represents the group of students whose father and mother hold a higher education diploma. According to this design, the reference group consists of students whose parents hold any other combination in terms of level of education. With such a design, we have attempted to represent in X9 the lower tail of the parents' education distribution and in X10 the upper tail.

In the second and third lines of the model the coefficients related to the intercept and the variable X4 (student's university entrance score) are set to vary randomly across courses according to a bivariate normal distribution with zero mean and variance-covariance matrix

defined as
$$\begin{bmatrix} \sigma_{u0}^2 & \sigma_{u04} \\ \sigma_{u04} & \sigma_{u4}^2 \end{bmatrix}.$$

We used the MLwiN 2.31 statistical software (Rasbash et al., 2014) with the estimation procedure PQL2 (Goldstein & Rasbash, 1996). The parametrization of level 1 variance allows for extra-binomial variation (McCullagh & Nelder, 1989). According to Williams (1982), the residual variation in the logistic regression model may be greater than a binomial sampling variation assumed by the model. In such situation we may postulate a source of extra-binomial random variation between observations. We considered missing values occurring completely at random (Little & Rubin, 2002).

6.3 Results

The Table 2 presents the estimates of the fixed and random parameters of the multilevel logistic regression model. The fixed parameter estimates suggest that at the significance level of 5%, the student' study methods influence his/her probability of persistence (estimate = 0.139, S.E. = 0.065).

There is an interaction effect between expectations and age, i.e. the main effect of expectations is not statistically different from zero in association with the probability of persistence (estimate = 0.001, S.E. = 0.072), but its interaction with student's age is statistically different from zero (estimate = 0.298, S.E. = 0.147).

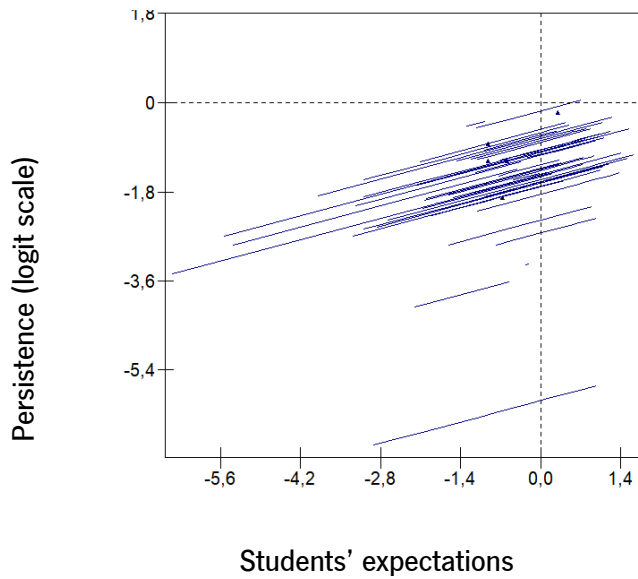
Table 2. *Estimates of Fixed and Random Parameters*

	Estimates	S.E.
<i>Fixed Parameters</i>		
Constant	0.402	0.395
Study Methods	0.139*	0.065
Expectations	0.001	0.072
Interaction: Expectations X Age >19	0.298*	0.147
Z-entrance score	0.172	0.139
Male vs. Female	-0.195	0.139
Age > 19	-1.626*	0.191
Course 1st option (Yes vs. No)	0.581*	0.130
Repetition at Basic Educ (No vs. Yes)	0.886*	0.345
Parents' education	-0.140	0.130
<i>Random Parameters</i>		
Level 2: Course		
Var(constant)	1.174*	0.304
Var(Z-entrance score)	0.498*	0.187
Covar(Z-entrance score / Constant)	0.083	0.167
Level 1: Student		
Extra-binomial	0.944*	0.029
Estimation:	RIGLS (PQL2)	

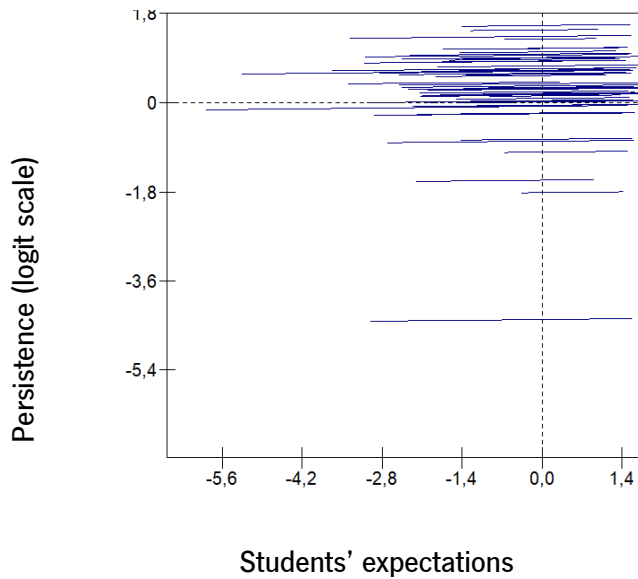
* p value < 0.05.

That is, although students aged over 19 years have a lower probability of persistence (estimate = -1.626; S.E. = 0.191), those who have high expectations, such effect is attenuated. For instance, students over 19 and whose expectations are at a standard deviation higher from the mean, the negative effect on the probability of persistence is reduced to -1.398 (-1.626 + 0.298). This means that high expectations are relevant to older students since such expectations protect them and increase their likelihood of persistence compared to younger students or to students with lower expectations. The graphs 1 and 2 illustrate the relationship between expectations and persistence (logit scale), based on the predictive model, for students older than 19 years (Graph 1) and for younger students (Graph 2). Each line represents a given course.

Graph 1 Persistence given expectations: age > 19



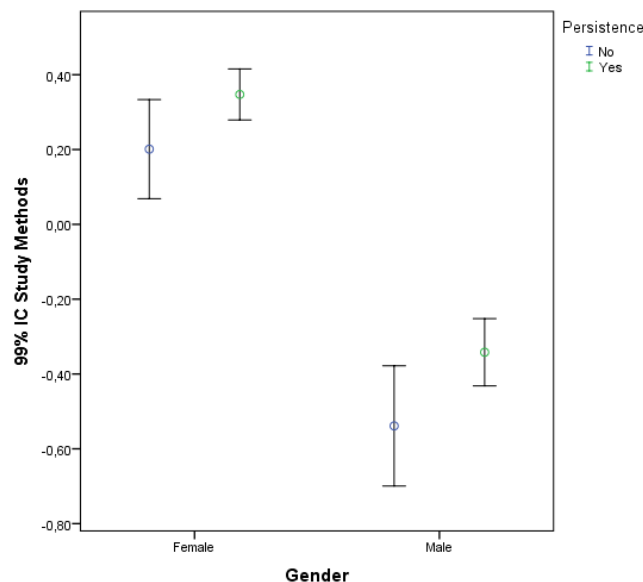
Graph 2 Persistence given expectations: age ≤ 19



When the variable study methods is included in the model, there is no statistically significant gender difference in the likelihood of persistence. That is, the probability of persistence is influenced by the variable study methods; in turn there is a relationship between methods of study and gender, as observed in Graph 3. The graph illustrates the average and 99% interval confidence for study methods by gender and persistence. It can be observed that the average

of study methods by gender presents the similar pattern of the average of study methods by persistence. Therefore, if the variable study methods is omitted from the model, the gender variable catches the difference that, in fact, is attributable to study methods. This is a very important result since "study methods" are likely to intervene in the short term not only to increase the likelihood of persistence in higher education, but to reduce the gender difference reported in the literature.

Graph 3 Multivariate analysis of Study methods, Persistence and Gender



The estimates obtained also confirm the effect of choice of course in the first option (estimate = 0.581, S.E. = 0.13) and the long-term effect of failure in primary education (estimate = 0.886, S.E. = 0.345) on student's persistence.

Regarding the random parameters, the estimates show that the mean of the persistence probability varies from course to course (β_0 level two variance estimate = 1.174). In addition, the influence of the university entrance score on the probability of persistence also varies across courses. The random parameter, or the variance associated with the slope coefficient of the university entrance score is statistically greater than zero (β_4 level two variance estimate = 0.498). The estimate of the extra-binomial parameter is close to 1 (estimate= 0.944;

SE=0.029), meaning that the assumption of binomial variance for the persistence given the predictor variables is adequate.

6.4 Discussion and Conclusion

In this study, we applied multilevel logistic regression models to data collected from 2,967 first-year Portuguese students enrolled at a public university in 2015-2016, considering first-year persistence as the dependent variable. The results are consistent with theoretical and empirical research concerning the influence of precollege characteristics on student's achievement and performance in HE. The estimates show the relationship between university entrance scores based on secondary school classifications combined with access examinations to the university and student persistence, which corroborates the evidence reported by other researchers (Ferrão & Almeida, 2018b; Montmarquette et al., 2001; Naidoo & Lemmens, 2015). In addition, data suggested that such a relationship varies randomly across courses after controlling for sociodemographic variables (age, gender, parents' education) and for first-choice course variable. Course specificity is in accordance with the evidence given by Ferrão e Almeida (2018a) that the entrance score influences student academic performance, assessed by grade point average at the end of the first year, and the magnitude of such association depends on the course and the area of study the student attends. Such results are also in line with those of Masui et al. (2014), who demonstrated differential grading by field of study, and support the idea that differential grading is possibly induced by departmental norms (e.g. Beenstock & Feldman, 2016). In turn, our results appear to agree with those obtained by Montmarquette et al. (2001). They showed that students enrolled in any course with an entrance quota, which are in general courses more demanding from the point of view of academic requirements and are mainly professional programs (in law, business or medicine), the probability of perseverance is significantly higher than would be if the students were enrolled in sociology, anthropology or economics programs. Reinforcing the role of course specificity, also Georg (2009) asserts that students do not consider abandonment due to stress or lack of ability, but mainly due to poor commitment to their course or area of study. In turn, our results also show the strong effect of being admitted to the first-option course on persistence. Such evidence may capture the motivation and commitment of the student to finish the course successfully, and

corroborates findings reported in the literature, according to which the higher is the motivation and commitment, the less likely is the student to drop out from higher education (English & Umbach, 2016; Tinto & Cullen, 1973; Tinto & Pusser, 2006). Students' preferences, motivation, commitment and expectations have always been intricately interwoven. Since students' expectations must have some bearing on their motivations, expectations must in turn influence the quality of higher education (James, 2002; Lavigne, Vallerand, & Miquelon, 2007), and, therefore, student retention. The direct implication of our finding is the need for universities to respond older students' expectations through rethinking the undergraduate curriculum in order to meet the students' developmental expectations and place the best interests of students at the heart of the teaching.

Regarding the relevance of study methods to persistence, our results appear to accord with the findings presented by Ishitani (2016); i.e., the evidence reported supports the hypothesis that academic integration plays a vital role in student persistence. The assessment of academic integration conducted by Ishitani was based on how often students participated in study groups, met with an academic advisor, or talked with faculty about academic matters outside class. Although it is not the same assessment instrument, there appears to be the same underlying construct related to study methods. Since the score we used is based on students' self-declarations about their past routines of study in secondary school, the dynamic of student integration in higher education may slightly change such routines. If so, study methods may play an even greater role in student retention, which we hope to clarify in future research.

This study contributes in an innovative way to research in the sub-area of higher education, specifically in regard to the subject of student persistence: by analyzing and modeling the microdata of a large sample of students enrolled for the first time in the first year, taking into account the hierarchical structure of the data and thus simultaneously contemplating the statistical units of student and course, considering both individual variables and the conditions of admission to HE as controlling variables that mediate the association between expectations, study methods and the probability of persistence. However, some caution is needed in not generalizing the results. The study concerns the enrollment cohort of 2015-2016, and it is very important to develop broader research involving samples from different cohorts of newcomers and multiple institutions in order to obtain national results.

CAPITULO 7 – Acesso e desempenho no Ensino Superior em Portugal: Questões de género, origem sociocultural, percurso académico, expectativas, escolha do curso e idade dos estudantes

Resumo

O artigo tem o propósito de caracterizar e contribuir para o debate sobre a democratização do ES português, quer no que se refere ao acesso como ao desempenho dos estudantes matriculados numa universidade pública. Analisam-se características sociodemográficas e percurso escolar dos 2697 alunos inscritos pela 1^a vez no 1^o ano no ano letivo 2015/16 e a sua associação à escolha do curso, às expectativas sobre o ES, condições de ingresso, bem como a associação com o desempenho ao longo do 1^o ano. Os principais resultados indicam que o sexo, a origem sociocultural das famílias e o percurso escolar estão associados à escolha do curso, à nota de candidatura e à opção de entrada. A análise de variância multivariada da classificação no final do 1^o ano sugere que, ao nível de significância de 5%, o termo fixo de área científico-disciplinar e os termos de interação entre opção do curso e área científica-disciplinar são estatisticamente diferentes de zero. Não foi encontrada associação estatisticamente significativa de origem sociocultural com persistência no ES. A origem dos estudantes parece não influenciar a decisão de abandonar, suspender ou transferir-se de curso, pelo menos durante o 1^o ano. Os resultados associados ao percurso escolar sugerem a resiliência e/ou a Ação institucional na promoção da equidade no ES.

Palavras-chave: Acesso ao ES; persistência; sucesso académico; género; origem sociocultural.

7.1 Introdução

O virar do século foi marcado pelo crescimento do Ensino Superior (ES) na Europa com taxa de 116% em 2002 face a 1998, crescimento esse verificado em todos os países da Comunidade Europeia com exceção de Itália (99%), Áustria (90%) e Bulgária (88%) (Santos Silva et al., 2005). Considerando 1998 ano base, em Portugal o índice de variação tem sido sempre superior a 100%, aproximando-se deste valor em 2015, e tendo atingido o máximo de

115% em 2003. A grande expansão do ES português decorreu após a década de 70 do século passado, concretizando o princípio da igualdade de oportunidades subjacente à revolução de abril de 1974. O número de estudantes a frequentar o ES no início da década de 70 era aproximadamente 60 mil e no ano 2002/3 era de 400 mil, o que representa uma variação superior a 650%. Estes números traduzem a democratização do acesso ao ES.

Outro índice relevante para a análise em apreço é o que quantifica a variação do número de estudantes inscritos pela 1ª vez no 1º ano. O número absoluto aumentou sempre até 2011, representando a variação de 161% considerando 1998 o ano base. De acrescentar que a democratização também aumentou a heterogeneidade da população discente, quer quanto ao perfil e trajetória escolar, quanto às regiões urbanas ou rurais de proveniência ou quanto à origem socioeconómica e cultural, entre outros aspetos.

A democratização do acesso assentou fundamentalmente na criação de lugares discentes, com atenção insuficiente pela qualidade da formação. Alguns autores diferenciam democratização do acesso de democratização do sucesso. A “democratização do sucesso” no ES é passível de dúvidas e reservas, uma vez que “as taxas de abandono e de insucesso penalizam, sobretudo, os estudantes provenientes dos estratos sociais mais desfavorecidos” (Almeida, Marinho-Araujo, et al., 2012, p. 913). A performance do sistema caracteriza-se por elevada taxa de abandono e de reprovação. De acordo com a Direção-Geral de Estatísticas da Educação e Ciência (Engrácia & Baptista, 2018), considerando os estudantes que se inscreveram pela 1ª vez no 1º ano no ano letivo 2011/12 em cursos de 1º ciclo com duração de três anos, 46% deles em 2014/15 tinham obtido o diploma do curso em que se tinham matriculado em 2011, 14% continuavam inscritos no mesmo curso, 11% frequentavam um curso diferente daquele onde se haviam matriculado em 2011/12 e os restantes 29% não foram encontrados no ES português.

Os principais fatores explicativos do (in)sucesso no ES reportam-se a processos fortemente interdependentes como são o da aprendizagem, do ensino e da avaliação (Almeida, 2002). Assim, o insucesso diferencia-se em função dos cursos e/ou disciplinas, habitualmente nos primeiros anos, com maior incidência no âmbito das ciências exatas e engenharias. Num outro estudo (Ferrão & Almeida, 2018b) aplicou-se um modelo de regressão logística multinível aos dados da coorte de estudantes inscritos pela 1ª vez no 1º ano na Universidade do Minho no ano letivo 2015/16. O objetivo do artigo foi o de aprofundar o conhecimento sobre o fenómeno da persistência e a sua relação com atributos da trajetória escolar prévia dos estudantes, com

as condições de ingresso no ES, controlando tal relação pelas características sociodemográficas dos estudantes ingressantes, tais como idade, sexo e nível socioeconômico. Os resultados reportados sugerem o efeito de longo-prazo da reprovação, em qualquer ano de escolaridade do ensino básico (ISCED 2), na probabilidade de persistência do estudante no ES (ISCED 5), bem como o efeito das condições de ingresso no ES (a nota de candidatura ao ES e se o estudante foi colocado no curso de 1ª opção). Os resultados também mostraram que os sujeitos do sexo masculino e com idade maior do que 19 anos têm menor probabilidade de persistência no 1º ano do ES. Por outro lado, a nota de candidatura ao ES e se o estudante foi colocado na universidade de 1ª opção influenciam a classificação média obtida pelo estudante no final do 1º ano e foi encontrada evidência de que a relação linear entre a nota de candidatura e a classificação média no final do 1º ano varia aleatoriamente entre cursos – sugerindo a especificidade de curso – mesmo após controlar pelo efeito fixo da área de estudo (Ferrão & Almeida, 2018a). Estes resultados sugerem que o abandono no ES é multivariado, destacando-se o impacto de um conjunto heterogêneo de variáveis pessoais e institucionais e, acima de tudo, a sua interação (Tinto, 1993).

Estudos empíricos, tal como Rhee (2008), mostraram que os determinantes que afetam a interrupção temporária e a transferência são diferentes daqueles que influenciam o abandono. Por exemplo, o estudo de Pascarella, Duby, Miller, and Rasher (1981) é ilustrativo dos efeitos diferenciais dos fatores individuais no abandono, interrupção temporária ou transferência de curso/instituição. Especificamente, eles procuraram determinar a relação das características prévias à entrada no ES na predição da persistência, da suspensão e do abandono. Os autores verificaram diferenças nos padrões de discriminação entre os três fenômenos. Além das características prévias à entrada no ES, o processo de admissão ao ES, incluindo as condições de acesso e a escolha da instituição e do curso, também têm efeito na persistência e no sucesso do estudante (Almeida et al., 2006; English & Umbach, 2016; Soares et al., 2001).

Em síntese, ao relacionar algumas variáveis pessoais, os estudos mostram que estudantes de famílias com maior nível socioeconômico apresentam maior taxa de permanência e conclusão de graduação, e essa situação pode estar associada ao fato de esses jovens trazerem mais competências acadêmicas do ensino secundário, superando com maior facilidade as exigências curriculares no ES. Alternativamente, pode estar associada à frequência de cursos com maior prestígio social (e.g. Almeida et al., 2006) ou a estudantes que não precisam conciliar estudos com atividades de trabalho necessárias para financiar a frequência de seus

estudos (Letkiewicz et al., 2014). A tal respeito, considerando a população discente da Universidade do Minho, um estudo (Almeida et al., 2006) sugere que a escolha do curso tem um papel determinante na reconstituição e reprodução da hierarquia social estabelecida, já que a “a abertura generalizada e a aparente democratização do acesso ao ES acabam por integrar diferenciações relativamente aos trajetos escolares anteriores, à escolha do curso e às taxas de sucesso associadas ao género e à origem sociocultural dos estudantes” (p. 513). Passada mais de uma década sobre a coorte de dados cuja análise suporta essa afirmação, retomamos os objetivos daquele estudo para verificar se há e quais foram as alterações verificadas na década de permeio.

Nestes termos, retomamos a análise da influência do género e da origem sociocultural dos estudantes sobre a nota de candidatura ao ES (que pondera as classificações do ensino secundário com os resultados nos exames de acesso ao ES), sobre a área do curso e sobre a média obtida pelo estudante no final do 1º ano. Alargamos o âmbito da análise acrescentando-lhe variáveis individuais e de resultados, tais como a idade de ingresso, as expectativas dos estudantes à entrada do ES, e a persistência do estudante no 1º ano. Desta forma, entre outros objetivos específicos, procuramos evidência que permita responder às seguintes questões de pesquisa: O acesso do estudante ao ES é influenciado por fatores de natureza sociodemográfica tais como origem sociocultural do aluno ou sexo? O sucesso do estudante no ES é influenciado por fatores de natureza sociodemográfica tais como origem sociocultural do aluno ou sexo?

7.2 Métodos

Participantes e variáveis

A amostra é composta por 2697 estudantes do primeiro ano que se matricularam na Universidade do Minho, uma universidade pública no norte de Portugal, no ano letivo 2015/2016. Os dados utilizados para o propósito deste trabalho dizem respeito a atributos individuais e do percurso escolar do estudante, bem como a variáveis de resultado no 1º ano no ES. Concretamente, quanto às variáveis de resultado, considerámos: Persistência (1: se o estudante atingiu o final do 1º ano com desempenho académico avaliado através de média da pontuação obtida nas unidades curriculares; 0: caso contrário), 72% dos estudantes obtiveram média no final do 1º ano; Média da classificação no final do 1º ano (13.48, $DP = 1.66$).

Procedimentos e Instrumentos

Os estudantes foram informados sobre os objetivos do presente estudo, sendo-lhes garantida a confidencialidade dos dados. Por escrito, os estudantes deram o seu consentimento informado, autorizando ainda o acesso às suas classificações acadêmicas no final do primeiro semestre e no final do primeiro ano. Os dados restantes foram coletados no momento da sua inscrição na universidade.

Questionário de Percepções Acadêmicas – Expectativas (Almeida, Deaño, et al., 2012). Este instrumento explora as crenças e aspirações dos estudantes na transição para o ES, ou seja, o que eles esperam encontrar e desenvolver. Seus itens combinam aspetos cognitivos e motivacionais da experiência acadêmica e abrangem várias dimensões: (a) Treinamento para o desenvolvimento profissional e profissional; (b) desenvolvimento pessoal e social; (c) mobilidade internacional de estudantes; (d) engajamento político e cidadania; (e) pressão social; (f) Qualidade de treinamento no curso; e (g) Convivência e interação social. Os estudantes avaliam sua concordância com o conteúdo do item em uma escala do tipo *Likert* de 6 pontos variando de 1 (discordo completamente) a 6 (concordo completamente). A partir das intercorrelações entre as sete dimensões, um fator geral de expectativas foi estimado pela primeira componente resultante da análise de componentes principais (Jolliffe, 2002; Nicholson et al., 2013).

Questionário sócio-acadêmico. Os estudantes responderam a várias perguntas curtas relatando informações pessoais, por exemplo, gênero e idade, trajetória acadêmica anterior, graus acadêmicos dos pais, se o curso e a universidade foram ou não a primeira opção do estudante.

Métodos estatísticos

Os métodos estatísticos foram escolhidos em função da natureza das variáveis envolvidas e da questão de pesquisa. Aplicámos teste de hipóteses de qui-quadrado com base em tabelas de contingência nas situações em que ambas as variáveis são qualitativas. Para variáveis relacionadas com o acesso do estudante ao ES, testámos a associação entre área científico-disciplinar do curso e sexo, origem sociocultural, idade de ingresso; testámos a associação entre opção de entrada no curso e sexo, origem sociocultural, idade de ingresso. Para a variável

relacionada com o sucesso do estudante – persistência no 1º ano – testámos a associação com sexo, origem sociocultural e idade de ingresso.

Aplicámos o teste de hipóteses para a comparação de medias em amostras independentes, com estatística de teste *t-Student*, para explorar a relação entre variáveis em que uma delas é quantitativa e a outra é qualitativa com duas categorias. Ou seja, expectativas do estudante ou nota de candidatura com sexo e opção de entrada. Também usámos o mesmo teste de hipóteses para comparar a média final por sexo, por idade de ingresso ou por opção de entrada.

Por fim, usámos ANOVA com um fator e ANOVA multivariada para testar os efeitos fixos de 1ª ordem e respetivas interações das variáveis das variáveis sexo, origem sociocultural, idade de ingresso, área científico-disciplinar do curso e opção do curso na média final de 1º ano.

7.3 Resultados

Os resultados são agrupados em duas subseções, de maneira a que na primeira se inclui a caracterização dos estudantes e das condições de ingresso no ES e na segunda subseção se inclui a caracterização do desempenho académico dos estudantes no 1º ano.

Caracterização dos estudantes e das condições de ingresso

Quanto às variáveis individuais e do percurso escolar, a maioria dos estudantes são do sexo feminino (57%), foram matriculados no seu curso de 1ª escolha (59%) e na instituição de ES de 1ª escolha (72%). Quase metade dos estudantes (47%) inscreveu-se em um curso STEM (ciência, tecnologia, engenharia e matemática), 17% em Economia, Administração ou Administração Pública, 13% em Humanidades, 10% Ciências Sociais, 8% Saúde (medicina ou enfermagem), e 5% em Direito. A idade dos estudantes varia de 16 a 61 anos, com média de 18,9 (DP = 3,6), sendo que 91,9% não possuem qualquer atividade profissional a tempo inteiro ou parcial, ou sejam assumem-se como estudantes em tempo integral.

Como pode ser observado na Tabela 1, a média da nota de candidatura à universidade, numa escala de pontuações entre 0 e 200, situa-se em 152,4 (DP=18,9). A maioria dos estudantes (83,3%) declaram que nunca foram reprovados ao longo da sua trajetória escolar e 13,9% mencionaram pelo menos uma repetência nos ensinamentos básico e/ou secundário. A distribuição dos estudantes pela educação dos pais mostra que 34,4% dos estudantes, tanto o pai como a

mãe não detêm mais do que o ensino básico e 15,7% dos estudantes, tanto o pai como a mãe possuem diploma de ES.

Tabela 1. Estatísticas descritivas do rendimento académico

Variável	Mínimo	Máximo	Média	Desvio Padrão
Nota de candidatura	105	200	152,39	18,91
Persistência	0,00	1,00	0,72	0,45
Média final de 1ºano	10,00	18,92	13,48	1,57

Os dados de desempenho académico durante o 1º ano na universidade encontram-se disponíveis para 72% dos estudantes. Assim, a classificação média do 1ºano está em falta para 757 estudantes, dos quais 57% encontram-se na situação de “abandono” da Universidade do Minho, 17% não se matricularam em qualquer curso e os restantes 27% podem ter suspenso os estudos temporariamente ou não obtiveram aproveitamento em qualquer unidade curricular. Consideramos aluno persistente aquele que não abandonou, que não suspendeu os estudos ou se transferiu para outro curso. As estatísticas descritivas para o subconjunto de estudantes sem pontuação académica no final do 1º ano mostram que em geral são mais velhos, têm maior probabilidade de ter reprovado no ensino primário e/ou secundário e apresentam, em média, menor classificação de entrada na universidade. A Tabela 2 apresenta a distribuição dos estudantes por cada uma das áreas de cursos segundo o sexo e a origem sociocultural. As frequências apresentadas na Tabela 2 mostram que se mantém o padrão de género na escolha dos cursos, i.e., enquanto a maioria dos rapazes provenientes de qualquer origem sociocultural frequenta cursos de engenharia (entre 64,18% e 67,49%), nas ciências sociais e económicas são a minoria em qualquer grupo de origem sociocultural (entre 32,82% e 36,96%).

Tabela 2. Distribuição dos estudantes por área de estudo segundo sexo e origem sociocultural

Origem sociocultural	Sexo	CEN		ENG		SOC+ECO		HUM		SAU		Total
		n	%	n	%	n	%	n	%	n	%	
Baixa	Masculino	48	36,36	164	67,49	85	32,82	25	24,51	4	8,51	326
	Feminino	84	63,64	79	32,51	174	67,18	77	75,49	43	91,49	457
Média	Masculino	94	52,81	267	64,18	119	34,39	32	26,45	17	18,48	529
	Feminino	84	47,19	149	35,82	227	65,61	89	73,55	75	81,52	624
Alta	Masculino	27	55,10	97	64,67	34	36,96	7	41,18	11	23,40	176
	Feminino	22	44,90	53	35,33	58	63,04	10	58,82	36	76,60	179

Tal padrão é ainda mais acentuado nos cursos de humanidades onde, nos grupos de origem “baixa” e “média”, a razão de percentagens é aproximadamente 3. Ou seja, em média, há aproximadamente três mulheres para um homem a frequentar esses cursos. Nos cursos de ciências da saúde a vantagem também é a favor do sexo feminino. Entre os 47 estudantes do grupo de origem sociocultural baixa, 91,49% são do sexo feminino; entre os 92 do grupo de origem sociocultural média, 81,52% são mulheres e, finalmente, no grupo “alta”, a percentagem de mulheres é 76,6% entre os 47 estudantes. Uma análise mais detalhada das frequências segundo origem e sexo por cada um dos cursos permite observar que parte daquela vantagem percentual se deve ao curso de enfermagem, principalmente no grupo de origem sociocultural baixa, no qual 65,1% das mulheres frequentam o curso de graduação em enfermagem e 34,9% o mestrado integrado em medicina. Por outro lado, a distribuição dos estudantes por área de estudo segundo a idade, apresentada na Tabela 3, mostra que entre os estudantes com idade de ingresso no ES menor do que 20 anos, 35% frequentam curso de engenharia, enquanto tal percentagem é de 18,5% entre os estudantes com idade igual ou maior do que 20 anos. Neste grupo de estudantes quase 70% frequenta um curso de ciências sociais, económicas ou humanas.

Tabela 3. Distribuição empírica dos estudantes por área de estudo segundo idade

Idade	Área de estudo					Total
	CEN	ENG	SOC+ECO	HUM	SAU	
Idade < 20	340 (15,6%)	759 (34,8%)	665 (30,5%)	222 (10,2%)	195 (10,2%)	2 181 (100%)
Idade => 20	24 (7,4%)	60 (18,5%)	119 (36,6%)	108 (33,2%)	14 (4,3%)	325 (100%)
Total	364 (14,5%)	819 (32,7%)	784 (31,3%)	330 (13,2%)	209 (8,3%)	2 506 (100%)

Para reforçar a análise descritiva aplicámos o teste de hipóteses não paramétrico de qui-quadrado para testar a não independência entre as variáveis sexo e área de curso ($\chi^2 = 274.78$, $p < .001$), origem e área de curso ($\chi^2 = 46.52$, $p < .001$), e idade e área de curso ($\chi^2 = 161.03$, $p < .001$), sendo sempre estatisticamente significativas as associações observadas.

No que se refere às expectativas dos estudantes, o teste de hipóteses para a comparação da média em amostras independentes foi aplicado para comparar os grupos definidos pela variável sexo e pela variável idade. Quanto à primeira comparação, a estatística de teste t-

Student, $t = 11.16$ ($p < .001$) e quanto à segunda obtivemos $t = 7.58$ ($p < .001$). Ou seja, ambos os testes de hipóteses indicam diferenças estatisticamente significativas, ao nível de 1%, das expectativas por sexo ou por idade. Quanto ao teste de hipóteses para verificar a diferenciação das expectativas por origem sociocultural, aplicámos o teste ANOVA com um fator, cujo resultado é $F = 2.174$ ($p = .114$) indicando a não rejeição da hipótese nula, segundo a qual os grupos de origem sociocultural não se diferenciam quanto às expectativas. Já o teste ANOVA aplicado à relação entre expectativas e área de estudo, apresenta a estatística $F = 5.73$ ($p < .001$), sugerindo que há diferenças estatisticamente significativas das expectativas dos estudantes por área de estudo. Os resultados da ANOVA sugerem, ainda, a diferenciação da nota de candidatura por grupo de origem sociocultural ($F = 78.73$, $p < .001$) e confirma-se a diferenciação por área científica-disciplinar ($F = 120.24$, $p < .001$). O teste de hipóteses para a diferença das médias nas notas de candidatura indica a vantagem estatisticamente significativa a favor das mulheres ($t = 4.98$, $p < .001$), ao mesmo tempo que a média de candidatura é estatisticamente maior no grupo de estudantes com idade de ingresso inferior a 20 anos ($t = 2.42$, $p < .001$). Por último, o teste de hipóteses de qui-quadrado sugere a associação entre sexo e opção de entrada no curso ($\chi^2 = 17.81$, $p < .001$), entre origem sociocultural e opção de entrada ($\chi^2 = 6.145$, $p < .05$), e entre idade e opção de entrada ($\chi^2 = 9.349$, $p < .01$).

Desempenho dos estudantes no 1º ano: persistência e média

Não encontramos associação estatisticamente significativa entre grupo de origem sociocultural e persistência ($\chi^2 = 3.8$, $p = .15$), contudo os resultados sugerem que persistência é um fenómeno marcado por diferenciação de género ($\chi^2 = 4.39$, $p < .05$). A variável idade de ingresso também parece estar associada à probabilidade de persistência. Enquanto 78% dos estudantes com idade menor de 20 anos são persistentes, 67% dos estudantes com idade igual ou maior a 20 anos, não são persistentes ($\chi^2 = 299.21$, $p < .001$). O teste de hipóteses sugere que persistência é um fenómeno marcado por diferenciação quanto à opção de entrada no curso. Ou seja, ao nível de significância de 5% rejeita-se a hipótese nula segundo a qual as variáveis opção de entrada no curso e persistência são independentes ($\chi^2 = 28.28$, $p < .001$). Por conseguinte, também encontramos associação entre a área científica-disciplinar do curso que o estudante frequenta e persistência ($\chi^2 = 46.2$, $p < .001$). No que se refere às expectativas

dos estudantes, o teste de hipóteses para a comparação da média em amostras independentes foi aplicado para comparar o grupo de estudantes que persiste com o que não persiste. Os resultados obtidos indicam diferenças estatisticamente significativas das expectativas ($t = 4.39, p < .001$).

Reportando-nos à media das classificações no final do 1º ano, o teste de hipóteses para a comparação da média entre os estudantes dos dois géneros indica uma diferença estatisticamente significativa ($t = 8.52, p < .001$). Quanto ao teste de hipóteses para verificar a diferenciação da média final por origem sociocultural, aplicámos o teste ANOVA com um fator, registando-se uma diferença com significado estatístico ($F = 8.39, p < .001$). As análises de variância realizadas para testar se idade de ingresso (menor de 20 anos vs. igual ou maior), área científico-disciplinar e opção de entrada no curso também se constituem como fatores de diferenciação da média final do 1ºano. Os resultados da estatística F são apresentados na Tabela 4.

Tabela 4. Teste de hipóteses ANOVA com um fator

Fator	F	Valor de prova
Idade > 19	0,152	0,696
Área de curso	85,88	<0,001
Opção de entrada	81,33	<0,001

A análise de variância com estes três fatores, bem como com origem sociocultural e sexo, e as respetivas interações resulta em efeitos fixos estatisticamente significativos para os coeficientes associados a sexo (com $\alpha=10\%$), a área científico-disciplinar (com $\alpha=1\%$), associados aos termos de interação entre opção do curso e área científica-disciplinar (com $\alpha=5\%$) e interação entre opção do curso, origem sociocultural e área científica-disciplinar (com $\alpha=10\%$).

Os gráficos apresentados nas figuras 1, 2 e 3 ilustram, a título de exemplo, a especificidade da interação entre sexo, opção de curso e origem sociocultural na média final de 1º ano. Através deles podemos constatar que a associação entre sexo e opção de curso é mediada pela origem sociocultural. Em qualquer dos grupos de origem sociocultural, a média final dos estudantes aumenta quando comparamos a média daqueles que frequentam cursos de 1ª opção com os que os não frequentam. Podemos ainda verificar que estudantes do sexo

feminino em cursos que não são de 1ª opção têm média final superior aos estudantes do sexo masculino e que essa vantagem também ocorre nos cursos de 1ª opção, com exceção do que ocorre no grupo de origem sociocultural baixa, que os sujeitos do sexo masculino que frequentam cursos de 1ª opção atingem média final superior aos estudantes do sexo feminino. Qualquer dos gráficos ilustra que a diferença de média final entre homens e mulheres é maior nos cursos que não são 1ª opção do que nos cursos de 1ª opção.

Figura 1

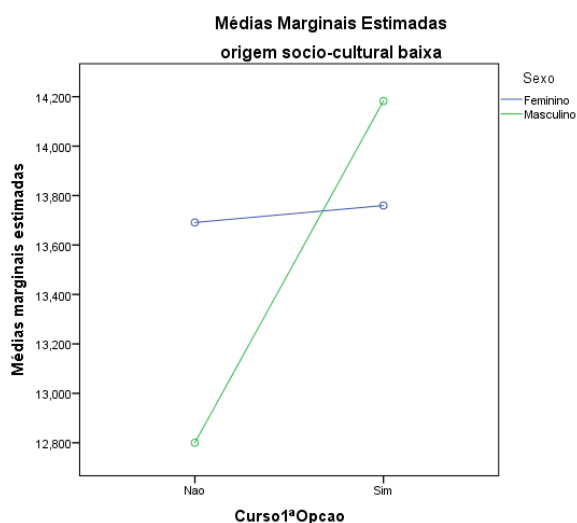


Figura 2

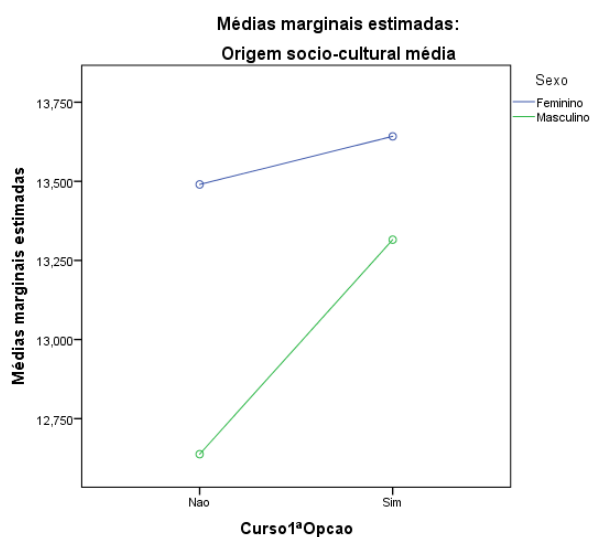
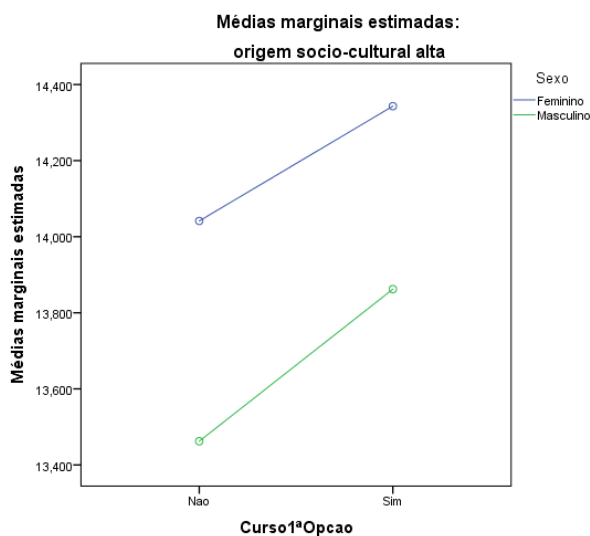


Figura 3



7.4 Discussão

Ao longo deste artigo as condições sociodemográficas e expectativas de uma coorte de estudantes matriculados pela 1ª vez no 1º ano são analisadas com a finalidade de explorar a “democratização” do ingresso e da qualidade dos estudos desenvolvidos no ES. A Tabela 5 sintetiza os principais resultados obtidos com a análise e modelagem estatística apresentados ao longo desta seção. Isto é, no que se refere a variáveis associadas ao ingresso no ES, tais como a escolha da área científico-disciplinar, expectativas, nota de candidatura e opção de entrada no curso, encontrámos evidência de diferenças estatisticamente significativas por sexo, origem sociocultural, qualidade e idade do estudante. Tais resultados vão ao encontro do registado na literatura há mais de dez anos, pois Almeida et al. (2006) ao explorar as condições de acesso e de sucesso encontraram “diferenças consideráveis, considerando o género, o capital sociocultural das famílias de origem e a qualidade do percurso escolar dos estudantes, na escolha do tipo de cursos a frequentar, na nota de candidatura e nas classificações escolares obtidas [...]” (p. 512). Os resultados agora encontrados mostram que a tendência de género na escolha de cursos por área científico-disciplinares mantém e, adicionalmente, se verifica também em cursos de humanidades, a favor das mulheres onde a

percentagem é aproximadamente 75%. A vantagem é ainda maior nos cursos da área de saúde, mas em grande medida fica a dever-se ao curso de enfermagem.

No que se refere às expectativas do estudante à entrada da universidade, parecem não variar de acordo com o grupo sociocultural de origem do estudante. No que se refere às variáveis de desempenho no 1ºano, verificámos que as análises bivariadas sugerem que a média final do estudante varia em função da variável sexo, origem sociocultural, área científico-disciplinar e opção de entrada no curso. Já persistência do estudante está estatisticamente associada a todas as variáveis individuais estudadas com exceção de origem sociocultural. Estes resultados estão parcialmente em linha com os reportados na literatura (Almeida et al., 2012; Dias et al., 2011). Não encontramos associação estatisticamente significativa de persistência no ES com origem sociocultural. Ou seja, a origem dos estudantes da Universidade do Minho parece não influenciar a decisão de abandonar, suspender ou transferir-se de curso, pelo menos durante o 1º ano, confirmando os resultados obtidos através da aplicação do modelo de regressão logística multinível em que o fenómeno de associação persistência-origem sociocultural é controlado por variáveis individuais tais como sexo, idade, reprovação até ao 9º ano (ISCED 2), nota de entrada, curso de 1ª opção (Ferrão & Almeida, 2018b). Tal resultado, a confirmar-se ao longo do curso, pode constituir a manifestação de uma política institucional – explícita ou implícita – de atenuação de desigualdades sociais que importa assinalar, pois contraria resultados prévios segundo os quais o fenómeno de abandono e insucesso no ES penaliza sobretudo os estudantes oriundos de estratos sociais desfavorecidos (Almeida et al., 2006; Engrácia & Baptista, 2018). Estes resultados estão em concordância com os apresentados por Ferrão e Almeida (2018a). Usando a mesma coorte de estudantes da Universidade do Minho, os autores aplicaram modelos de regressão multinível à classificação média final do 1ºano, incluindo como variáveis explicativas a nota de candidatura, sexo, idade, 1ª opção da universidade e área científico-disciplinar, e verificaram que o grupo de estudantes com origem sociocultural baixa têm, em média, a classificação no final do 1ºano acrescida 0.14 desvios padrão comparativamente com os seus colegas provenientes de outros estratos socioculturais. Ora, tais resultados, obtidos por métodos de análise diferentes, dão suporte quantitativo à conjectura avançada por Engrácia e Baptista (2018), segundo a qual

...os filhos de pais com baixa escolaridade que, apesar dos ventos contrários, conseguem terminar o ensino secundário e ingressar no ensino superior serão, em média, jovens com desempenhos escolares e níveis de resiliência francamente acima

da média entre os estudantes oriundos de contextos desfavorecidos. Assim sendo, é expectável que o seu posterior desempenho no ensino superior seja mais comparável ao desempenho dos estudantes oriundos de meios socioeconómicos favorecidos. (Engrácia & Baptista, 2018, p. 18)

De fato, os resultados apresentados neste artigo e em Ferrão e Almeida (2018a), pela primeira vez em Portugal, mostram a vantagem na média final de 1º ano que os estudantes de origem sociocultural desfavorecida têm comparativamente com os colegas, mesmo após controlar pelas restantes variáveis acima discriminadas. Ora, considerando que 34% dos estudantes da coorte em análise pertencem ao grupo com origem sociocultural baixa (tanto o pai como a mãe têm como nível de escolaridade mais elevado o 9ºano de escolaridade (no máximo nível ISCED 2)) e que na população portuguesa com idade 15-64 anos tal grupo representa 52% (Pordata, 2018); que na coorte de estudantes 15% pertencem ao grupo com origem sociocultural alta (tanto o pai como a mãe têm nível de escolaridade mais elevado o ES (4<ISCED<=6)) e que tal grupo na população portuguesa com idade 15-64 representa 22%, parece-nos que além da conjectura “níveis de resiliência dos estudantes” há, adicionalmente a considerar em pesquisas futuras, o papel da universidade numa ação afirmativa a favor da atenuação das desigualdades sociais e, portanto, na promoção da equidade social.

Tabela 5. Síntese dos principais achados

		Sexo	Origem socio-cultural	Idade	Área científico-disciplinar	Expectativas	Opção
Ingresso	Área científico-disciplinar	Sim	Sim	Sim	—	Sim	—
	Expectativas	Sim	Não	Sim	Sim	—	Sim
	Nota de candidatura	Sim	Sim	Sim	Sim	Sim	—
	Opção de entrada	Sim	Sim	Sim	—	—	—
Desempenho no 1ºano	Persistência	Sim	Não	Sim	Sim	Sim	Sim
	Média final	Sim	Sim	Não	Sim	—	Sim

Os resultados que se referem ao padrão da escolha do curso por grupo de origem sociocultural vão ao encontro do que vários autores sublinharam sobre a influência do capital

socioeconômico e cultural do estudante na preferência por cursos ou instituições mais prestigiadas (Tavares et al., 2008; Tavares & Ferreira, 2012). De acordo com Tavares et al. (2008) estudantes provenientes de contextos familiares desfavorecidos têm maior preferência por cursos relacionados à formação de professores e gestão, enquanto estudantes que têm origem familiar privilegiada apresentam tendência em escolher cursos na área de direito, artes plásticas, saúde, ou ciência e tecnologia.

É necessária alguma cautela para não generalizar os resultados aqui apresentados pois o estudo diz respeito à coorte de ingressantes da Universidade do Minho no ano 2015/16. Se tivessem sido selecionadas outras coortes e/ou outras instituições os resultados poderiam ser ligeiramente diferentes. Nesta perspetiva, é muito importante desenvolver uma investigação mais ampla que contemple amostras provenientes de diferentes coortes de ingressantes e de múltiplas instituições de ES de modo a obter resultados de âmbito nacional.

CAPÍTULO 8 – Conclusão

A Tese desenvolveu-se por artigos de dois tipos: artigos de revisão que sintetizam investigação anterior em Educação e artigos de investigação que apresentam contribuições originais em Educação, com ênfase na subárea de ES. Os artigos de revisão, segundo e terceiro capítulos, sintetizam o conhecimento sobre o uso de modelos de valor acrescentado em educação, com enfoque na produção científica publicada em periódicos de língua portuguesa e sintetizam o conhecimento sobre os modelos estatísticos usados na investigação sobre ES ao longo de 20 anos, respetivamente. Os artigos de investigação, quarto a sétimo capítulos, inserem-se na categoria de investigação institucional (Altbach, 2014) e apresentam resultados inovadores sobre a experiência e sucesso, aqui circunscrito ao sucesso nas aprendizagens curriculares (Almeida, 2002) dos estudantes no primeiro ano de estudos no ES.

Cada um dos artigos está estruturado de modo a evidenciar a relevância do tema de revisão ou do tópico de investigação, como é que cada um deles se relaciona com estudos anteriores, quais são os aspetos que potencialmente poderiam conduzir a algum viés nos resultados, a interpretação e contextualização desses resultados, a discussão e conclusões, assim como as limitações inerentes a cada um. Neste capítulo final procedemos à unificação dos principais achados cuja síntese se apresenta na seção 8.1 e as limitações e sugestões para investigação futura na seção 8.2.

8.1 Discussão dos principais resultados

Os artigos de revisão incluem contribuições inovadoras sobre modelos de valor acrescentado na literatura científica de Língua Portuguesa, no segundo capítulo, e, no terceiro capítulo, sobre os modelos estatísticos usados na investigação de ES aos longo dos últimos vinte anos. Parte do estudo apresentado no segundo capítulo baseia-se em artigos publicados em periódicos incluídos nas coleções do Brasil e Portugal da base SciELO, donde resulta o reduzido número de artigos publicados sobre o tema em revisão, manifestação de uma extensa e promissora agenda de investigação e avaliação educacional. A produção e comunicação de conhecimento científico que supere as limitações e dificuldades enunciadas à concretização dessa agenda está fortemente dependente do contexto político (Weiss, 1975), concretamente das políticas

públicas para o desenvolvimento científico, em especial da que concerne ao incentivo/financiamento das subáreas em que os avanços determinantes são obstaculizados pela falta de massa crítica (Ferrão, 2012a). O reduzido número de artigos sobre estas matérias e afins é ainda mais visível na perspetiva comparativa internacional. Note-se, como exemplo, que o artigo de revisão sobre modelos de valor acrescentado publicado *online* em 2016 (Everson, 2017) no periódico *Review of Educational Research* (com SJR no centésimo superior ao longo dos últimos anos na área de Ciências Sociais e subárea Educação) contém 112 referências bibliográficas. Inclui apenas um dos autores com produção na literatura científica de Língua Portuguesa. Importa referir que a capacidade de investigação científica instalada está também condicionada à disponibilidade de dados longitudinais que satisfaçam os requisitos metodológicos enunciados no segundo capítulo.

As principais descobertas da revisão apresentada no terceiro capítulo sobre os métodos estatísticos e fenómenos de interesse (variáveis dependentes) consideradas na investigação sobre o ES nos últimos vinte anos é baseada na análise de conteúdo de trabalhos selecionados de acordo com o método sistemático. Como uma revisão abrangente, o estudo pode ter limitações associadas ao inevitável viés introduzido no processo de simplificação de etapas ou eliminação de etapas desnecessárias. Ou seja, apesar do método de seleção de artigos ter sido sistematicamente planeado e as respetivas análises cuidadosamente conduzidas, fontes de viés tais como o viés de seleção, o viés de publicação e o de linguagem de publicação podem ter ocorrido (Butler, Deaton, Hodgkinson, Holmes, & Marshall, 2005), uma vez que decidimos usar a literatura disponível online. Contudo, cada passo do estudo é totalmente descrito, a fim de minimizar o impacto de eventuais fontes de viés e apoiar a replicabilidade do estudo.

Os padrões que podem ser traçados com relação aos métodos estatísticos e variáveis dependentes usadas com maior frequência em trabalhos de investigação do ES são resumidos seguidamente. Do ponto de vista geográfico, a evidência empírica provém de 27 países distribuídos pelos cinco continentes. Os nossos resultados sugerem a proeminência da produção científica norte-americana neste campo de estudo, como manifesto não apenas em termos de um número maior de artigos, mas também em termos do uso de métodos estatísticos mais inovadores e sofisticados. Este achado corrobora as evidências encontradas por Tight (2013), que examinou a relação entre disciplina e metodologia na pesquisa em ES, analisando um banco de dados com 567 trabalhos publicados em 2010 em 15 periódicos de ES. O autor reportou elevada variação geográfica, com pesquisadores norte-americanos

exibindo uma preferência pelo uso de métodos multivariados mais complexos. Parece haver duas razões possíveis para essa tendência nas últimas duas décadas. A primeira é o grau de desenvolvimento de grupos académicos de ES - departamentos ou centros com programas de pesquisa de pós-graduação - e o segundo é a disponibilidade mais longa e mais ampla de conjuntos de dados complexos adequados às análises. Nos últimos anos houve um crescimento acentuado das publicações na área (74% dos trabalhos selecionados foram publicados depois de 2010) recorrendo a métodos estatísticos mais sofisticados (por exemplo, Wells et al., 2015).

Os nossos resultados sugerem que até 2010, a maioria dos artigos selecionados fez uso de uma panóplia limitada de métodos, ou seja, artigos onde os investigadores recorrem exclusivamente a estatística descritiva, testes básicos ou regressões clássicas, com exceção das obras publicadas em 2004, que apareceram principalmente no *Journal of the Royal Statistical Society* ou no *Journal of Econometrics* e, portanto, com enfoque metodológico. Estes achados corroboram os apresentados por Wells et al. (2015). O estudo que realizámos também revela que métodos mais sofisticados foram usados após 2010, incluindo análise fatorial, modelos de regressão multinível, modelos de equações estruturais, modelos de sobrevivência, análise envoltória de dados (*data envelopment analysis*), análise de propensão (*propensity score matching*), análise de fronteira estocástica e simulação.

Além disso, encontrámos diversidade na escolha das variáveis dependentes. No entanto, os fenómenos preferidos pelos pesquisadores parecem ser o abandono dos estudantes ou fenómenos afins tais como persistência, retenção, atrito ou evasão, e outros fenómenos relacionados com o desempenho académico dos alunos. Esses trabalhos mostram grande diversidade no seu desenvolvimento, como são exemplos os estudos sobre grupos minoritários, sobre os desafios de dados com implicações para a produção das estatísticas oficiais, sobre o papel das instituições e sua influência na evasão de estudantes, os efeitos do contingente de professores na retenção e o impacto da vivência no campus, entre outros.

A análise dos resultados também sugere que a literatura de ES apresenta algum atraso na apropriação dos métodos estatísticos mais sofisticados. Do ponto de vista da difusão de inovações, Rodgers (2010) designou-a de “revolução silenciosa”, ou seja, a transição do paradigma do teste de hipóteses básico para um paradigma baseado na construção, comparação e avaliação de modelos estatísticos. De fato, a vertente transdisciplinar da estatística desenvolveu-se nos últimos anos, apresentando novos temas e desafios.

Particularmente na subárea da estatística educacional, as metodologias inovadoras que surgiram nas duas últimas décadas aumentaram a capacidade de pesquisa para explorar e revelar a complexidade do mundo real (Goldstein, 1998), que é fundamental ao desenvolvimento disciplinar e científico da Educação. Os avanços tecnológicos e computacionais, notadamente aqueles relacionados a métodos computacionais e estatísticos para lidar com pesquisas grandes e complexas de dados longitudinais, deram um salto gigantesco na produção de conhecimento em educação e, assim, no avanço da disciplina. Na era *big data e data science*, a “revolução silenciosa” é de fato uma tendência muito promissora na pesquisa em ES. Além disso, vários autores atribuíram as limitações da interpretação e extrapolação de seus achados às restrições ou limitações impostas pela disponibilidade de dados e qualidade dos dados. Esses fatores representam um obstáculo ao desenvolvimento do campo que só pode ser superado com o apoio de organizações governamentais responsáveis pela promoção e desenvolvimento da ciência e do ES.

Os artigos de investigação institucional baseiam-se em evidência empírica obtida através da análise e modelação de dados referentes aos 2697 alunos ingressados pela 1ª vez no 1º ano da Universidade do Minho no ano letivo 2015/2016. Os dados foram recolhidos no âmbito do ObservatoriUM da Universidade do Minho (Almeida & Castro, 2016).

De acordo com a caracterização descritiva dos dados, a maioria dos ingressantes são do sexo feminino (57%), escolheram a Universidade do Minho em 1ª opção (72%), matricularam-se no curso de 1ª opção (59%), são estudantes em tempo integral (91.9%), nunca foram reprovados no ensino básico e secundário (83.3%). A média da idade desses alunos é 18.9 (DP 3.6) anos, com mínimo 16 e máximo 61, sendo que 13% têm idade maior ou igual a 20 anos. A distribuição dos estudantes pela educação dos pais mostra que 34.4% dos estudantes, tanto o pai como a mãe não detêm mais do que o ensino básico, registando-se também que em apenas 15.7% dos estudantes, tanto o pai como a mãe possuem diploma de ES.

A nota de candidatura dos alunos ingressados tem média 152.39 (DP = 18.91) pontos (escala 0-200), com mínimo 105 e máximo 200 pontos. Quase metade dos estudantes (47%) inscreveu-se em curso STEM (ciência, tecnologia, engenharia ou matemática), 17% em Economia, Administração ou Administração Pública, 13% em Humanidades, 10% Ciências Sociais, 8% Saúde (medicina ou enfermagem), e 5% em Direito.

No que concerne ao desempenho do aluno, usaram-se duas medidas de sucesso: a média das notas de aprovação nas unidades curriculares do 1º ano, doravante designada por nota final, e situação do aluno face a persistência (o aluno não abandonou o curso, não suspendeu a matrícula, não se transferiu para outro curso ou instituição). A nota final encontra-se disponível para 72% dos estudantes matriculados na Universidade do Minho. A média das notas finais é 13.48 (DP = 1.57) valores (escala 0-20), com mínimo 10 e máximo 18.42 valores. Dos restantes 28%, a maioria dos estudantes (57%) encontram-se na situação de abandono, 17% não estão associados a qualquer curso e os restantes 27% podem ter suspenso os estudos temporariamente ou não obtiveram aproveitamento em qualquer unidade curricular. Esta caracterização vai ao encontro da descrita sobre o sistema de ES português pela Direção-Geral de Estatísticas da Educação e Ciência (Engrácia & Baptista, 2018), considerando os estudantes que se inscreveram pela 1ª vez no 1º ano, no ano letivo 2011/12 em cursos de 1º ciclo com duração de três anos, 46% deles em 2014/15 tinham obtido o diploma do curso em que se tinham matriculado em 2011, 14% continuavam inscritos no mesmo curso, 11% frequentavam um curso diferente daquele onde se haviam matriculado e os restantes 29% não foram encontrados no ES português. Ou seja, uma e outra caracterização reforçam a necessidade de ações institucionais e do sistema que visem a melhoria da eficácia do ES em Portugal.

As análises e modelação multinível da variável nota final do 1ºano e da variável persistência mostram uma realidade sobre a transição e experiência no 1º ano do ES que contribui para a definição da ação institucional que proporcione aos estudantes o seu desenvolvimento pessoal e profissional. Em cada um dos quatro artigos usámos métodos estatísticos diferentes para abordar os objetivos específicos enunciados em cada um deles. Os resultados reforçam-se mutuamente no delinear de uma realidade em mutação, tal como passaremos a enunciar.

A partição da variância das notas finais é 66% entre alunos e 34% entre cursos, ou seja 34% da variabilidade do desempenho dos alunos do primeiro ano é devida a diferenças entre cursos. Estes resultados parecem estar em concordância com os de Masui et al. (2014), segundo os quais a classificação dos alunos varia por área de estudo, e apoiam a ideia exposta por Beenstock e Feldman (2018) segundo a qual tal variabilidade é possivelmente induzida por normas departamentais, quer elas sejam normas formais ou informais. Ora, verificámos que parte dessa variabilidade se deve à área científico-disciplinar onde se insere o curso. Ou seja, considerando os cursos STEM como categoria de referência, verificámos que, em média, a nota final não é estatisticamente diferente nos cursos de Economia e Ciências Sociais. Os

estudantes de cursos nas áreas de Humanidades ou Saúde têm nota final estatisticamente maior do que os cursos STEM, e os estudantes dos cursos de Direito tendem a atingir nota final estatisticamente menor do que os cursos STEM. Estes resultados foram obtidos através da aplicação de um modelo de regressão multinível com alunos (nível 1) agrupados em cursos (nível 2) e incluindo, como variáveis de controle, sexo e idade dos estudantes, educação dos pais, nota de candidatura e universidade de 1ª opção.

O mesmo modelo estatístico indica que, em média, a nota final aumenta quase 0.6 desvios padrão por cada unidade de desvio padrão acrescido na nota de entrada e, também indica, que essa relação varia aleatoriamente entre os cursos após o controle de variáveis sociodemográficas (idade, sexo, escolaridade dos pais), da área científico-disciplinar de estudo e opção de universidade. Importa referir que a magnitude da estimativa do parâmetro fixo é maior quando comparada aos outros efeitos marginais incluídos no preditor linear acima enunciados. A capacidade explicativa deste modelo, aferida pelo coeficiente de determinação, é aproximadamente 40%. A desagregação do coeficiente de determinação pelos níveis hierárquicos do modelo mostra que ao nível dos alunos o modelo explica 21% da variabilidade da nota final e que ao nível dos cursos explica 80%. Isto é, 80% da variabilidade das notas finais dos estudantes que é devida a diferenças entre cursos é explicada pelas variáveis acima enunciadas. Os nossos resultados corroboram a literatura sobre o poder preditivo da nota de candidatura ou da preparação pré-universitária. Por exemplo, Fuller et al. (2011) relataram que 27.8% da variância explicada foi atribuída aos indicadores pré-universitários. Mas além disso, nossas descobertas contribuem com a evidência de que, mesmo depois de considerar os fatores sociodemográficos e de trajetória escolar do aluno, ainda há uma componente aleatória não explicada associada à especificidade de cada curso. Este resultado inovador sugere que alguns cursos requerem a satisfação de condições mais fortes como pré-requisitos, o que tem implicações para o desenvolvimento de currículo. Essa evidência sustenta o modelo teórico proposto por Alarcão (2000), conferindo ao currículo um status semelhante ao concedido ao aluno, professor ou instituição. É necessária investigação adicional para identificar quais são os cursos que impõem condições mais fortes à preparação pré-universitária.

Nossas descobertas sugerem ainda que aqueles estudantes para os quais a Universidade do Minho não foi a sua primeira escolha, em média, têm nota final menor comparativamente com os colegas para os quais esta universidade foi a primeira opção, mas os resultados não

fornece evidência suficiente para nos permitir dizer que a colocação do estudante no curso de primeira opção é relevante no seu desempenho académico aferido pela nota final. Contudo, quer o teste de hipóteses de qui-quadrado quer o modelo de regressão logística para o fenómeno persistência sugerem que tendencialmente os alunos colocados no curso de 1ª opção têm maior probabilidade de permanecer no curso ao longo do 1º ano de estudo. Adicionalmente, o teste de independência do qui-quadrado, realizado ao nível de significância de 5%, indica que a colocação no curso de 1ª opção não é estatisticamente independente de origem sociocultural do aluno, o que corrobora a evidência relatada uma década antes por Almeida et al. (2006), não é estatisticamente independente de sexo, nem da idade do estudante.

As condições de ingresso, isto é, a nota de entrada e a colocação no curso de 1ª opção do estudante, têm marca de género, de idade e de origem sociocultural. Ou seja, as mulheres, os estudantes com idade inferior a 20 anos e os estudantes cujos pais têm maior escolaridade tendem a ingressar na universidade com nota de candidatura mais elevada. Os resultados que se referem ao padrão da escolha do curso por grupo de origem sociocultural vão ao encontro do que vários autores sublinharam sobre a influência do capital socioeconómico e cultural do estudante na preferência por cursos ou instituições mais prestigiadas (Tavares, Tavares, Justino, & Amaral, 2008; Tavares & Ferreira, 2012). De acordo com Tavares et al. (2008), estudantes provenientes de contextos familiares desfavorecidos têm maior preferência por cursos relacionados com a formação de professores e gestão, enquanto estudantes que têm origem familiar privilegiada apresentam tendência em escolher cursos na área de direito, artes plásticas, saúde, ou ciência e tecnologia. O padrão de género e idade mantém-se na modelação estatística, quer quando o fenómeno de interesse é nota final quer quando é probabilidade de persistência.

Não encontramos evidência que nos permita afirmar que a probabilidade de persistência está associada à origem sociocultural do aluno, sendo esta aferida pela educação dos pais, mantendo como variáveis de controle sexo, idade, situação face a reprovação no ensino básico ou secundário, nota de candidatura e opção de colocação. Todavia, a origem sociocultural parece estar associada à nota final dos estudantes, pois o modelo de coeficientes aleatórios cujos resultados são apresentados no quarto capítulo e a análise de variância apresentada no sétimo capítulo, mostram que, incluindo como variáveis explicativas o grupo sociocultural familiar, a nota de candidatura, sexo, idade, 1ª opção da universidade e área científico-

disciplinar, o grupo de estudantes com origem sociocultural baixa tem, em média, a nota no final do 1º ano acrescida 0.14 desvios-padrão comparativamente com os seus colegas provenientes de outros estratos socioculturais. Ora, tais resultados, obtidos por métodos de análise diferentes, dão suporte quantitativo à conjectura avançada por Engrácia e Baptista (2018, p.18), segundo a qual

... os filhos de pais com baixa escolaridade que, apesar dos ventos contrários, conseguem terminar o ensino secundário e ingressar no ES serão, em média, jovens com desempenhos escolares e níveis de resiliência francamente acima da média entre os estudantes oriundos de contextos desfavorecidos. Assim sendo, é expectável que o seu posterior desempenho no ES seja mais comparável ao desempenho dos estudantes oriundos de meios socioeconómicos favorecidos.

De fato, os resultados apresentados no capítulo quatro, pela primeira vez em Portugal, mostram a vantagem na nota final de 1º ano que os estudantes de origem sociocultural desfavorecida têm comparativamente com os colegas, mesmo após controlar pelas restantes variáveis acima discriminadas. Ora, considerando que 34% dos estudantes da coorte em análise pertencem ao grupo com origem sociocultural baixa (tanto o pai como a mãe têm como nível de escolaridade mais elevado o 9º ano de escolaridade (no máximo nível ISCED 2)) e que na população portuguesa com idade 15-64 anos 52% têm nível mais elevado de escolaridade o 9º ano (Pordata, 2018); que na coorte de estudantes 15% pertencem ao grupo com origem sociocultural alta (tanto o pai como a mãe têm nível de escolaridade mais elevado o ES (4<ISCED<=6)) e que na população portuguesa com idade 15-64, 22% têm algum diploma de ES, parece-nos que além da conjectura “níveis de resiliência dos estudantes” há, adicionalmente a considerar em pesquisas futuras, o papel da Universidade do Minho quanto à sua ação afirmativa a favor da atenuação das desigualdades sociais e, portanto, na promoção da equidade social.

A admissão do aluno no curso que escolheu em 1ª opção é importante para a persistência do aluno matriculado pela 1ª vez no 1º ano. As estimativas obtidas sugerem que a razão de vantagem na persistência dos alunos que são admitidos no curso de 1ª opção é 1.85. Isto quer dizer que a probabilidade de persistência sobre a probabilidade de abandonar, suspender ou transferência é quase duas vezes maior no grupo de alunos que são admitidos no curso de 1ª

opção comparativamente com o que ocorre no grupo de alunos que não são admitidos na sua 1ª escolha. Ou seja, a 1ª opção do curso parece ter subjacente a motivação e empenho do aluno para terminar o curso com sucesso, corroborando a literatura no sentido de quanto maior a motivação e o empenho menor é a probabilidade do indivíduo abandonar o ES (English & Umbach, 2016; Tinto & Cullen, 1973; Tinto & Pusser, 2006).

Ao nível de significância de 5% não encontramos evidência da associação entre nota de candidatura e a probabilidade de persistência, quando a relação é controlada pelas variáveis sociodemográficas, de trajetória escolar, educação dos pais e opção de entrada no curso, mas é estatisticamente significativa ao nível de 10%. Os resultados também mostraram que tal associação varia de curso para curso, isto é, a especificidade do curso medeia tal associação o que reforça o argumento anteriormente exposto. Resultados similares foram encontrados por Montmarquette et al. (2001) com base numa amostra de 3,418 alunos do 1º ano que frequentavam a Universidade de Montréal, pois as estimativas dos parâmetros dos modelos *probit* para persistência no primeiro semestre sugerem que as pontuações dos alunos no desempenho pré-universitário são estatisticamente significantes, a um nível de significância de 10%, na amostra restrita, mas não é estatisticamente diferente de zero para a amostra completa de estudantes. Segundo os autores, “a amostra é restrita a alunos matriculados em um programa com cota de ingresso” (p. 479) e é composta por 2,328 alunos. Como na amostra completa o coeficiente não é estatisticamente significativo, parece assim que a especificidade do curso medeia a relação entre a realização anterior do aluno ou a pontuação de admissão da universidade e a probabilidade de persistência.

Quando o modelo estatístico para a probabilidade de persistência inclui no preditor linear variáveis que aferem método de estudo e expectativas do estudante, tal como o que apresentámos no capítulo seis, a nota de candidatura deixa de ter relevância, enquanto aquelas duas variáveis revelam-se estatisticamente significativas quanto à probabilidade de persistência e, adicionalmente, aparece também o termo de interação entre idade e expectativas como sendo estatisticamente significativo.

Os modelos estatísticos para a probabilidade de persistência que apresentámos nos capítulos quatro e cinco revelam um efeito de longo prazo da experiência de reprovação no ensino básico. Os estudantes que declaram ter passado por tal experiência têm em média menor probabilidade de persistência comparativamente com os colegas. Tal resultado remete-nos para o efeito marginal associado à idade do aluno, o qual sugere que o aluno ingressante no

ES com idade igual ou menor do que 19 anos tem razão de vantagem de persistência quase seis vezes maior do que os colegas mais velhos. Ora, o atraso na idade de ingresso pode dever-se a reprovação no ensino básico ou secundário ou pode dever-se ao adiamento da candidatura após a obtenção do diploma de ensino secundário. Assim, há um efeito cumulativo associado à idade e associado à reprovação no ensino básico. Registamos que se a variável idade não fosse incluída no preditor linear do modelo, a razão de vantagens na persistência entre os grupos – não reprovação vs. reprovação no ensino básico – era aproximadamente de 4 para 1. Este estudo é o primeiro a reportar o efeito de longo prazo que a reprovação na escolaridade básica tem no ES e vem reforçar o argumento apresentado por Ferrão (2015a, 2015c) segundo o qual a retenção/reprovação precoce é um forte preditor da retenção/reprovação tardia. De acordo com a autora, com base nos dados do PISA 2012, a probabilidade de retenção/reprovação precoce é 35 vezes maior no décimo inferior da distribuição do nível socioeconómico dos alunos comparativamente com o décimo superior, indicando falta de equidade social no sistema educativo português e menciona “Estes resultados enunciam um problema de equidade no sistema educativo português, ou seja, a distribuição de recursos educativos não está a garantir a toda a criança ou jovem o seu pleno desenvolvimento, independentemente do seu nível socioeconómico” (p. 14). Tais evidências trazem ao debate o tema da seleção realizada pelo sistema educativo e pelo processo de acesso/admissão ao ES (e.g. Dias, 2015; Tavares, Tavares, Justino, & Amaral, 2008), comprometendo a mobilidade social futura. Também o “Terceiro compromisso”, enunciado por Raymond e Negassi (2015, p. 16), “visa a equidade de exposição a conhecimentos e competências fundamentais, às vezes chamado de equidade de tratamento ou igualdade de oportunidades”, ou seja, do ponto de vista do desenvolvimento curricular, o sistema deve assegurar que todos os jovens portugueses diplomados do ensino secundário “estão prontos para abraçar os desafios do ES ou para conseguir um emprego viável”. Ora, as metas de aprendizagem para o ensino secundário têm-se centrado, na sua maior parte, no ensino geral, dando acesso ao ES (e.g. Guellec, Larrue, Roy, & Weko, 2018). Apesar do crescimento das vias alternativas de formação profissional ter proporcionado uma maior liberdade de escolha aos alunos do ensino secundário, as metas de aprendizagem para estudos vocacionais são mais heterogéneas, muitas vezes ligadas a competências profissionais específicas e conhecimentos do ofício. Na realidade de muitas crianças e jovens, a dita “liberdade de escolha” resume-se ao percurso alternativo oferecido por um sistema com défice de equidade e que tem metas a atingir no âmbito da estratégia Europa 2020 (European Commission, 2013). Concretamente, essa “base pode melhorar os

baixos níveis de conclusão [do ensino secundário] – cerca de 50 por cento – que se verificam hoje em Portugal e melhorar o diferencial de salários entre os alunos oriundos da formação profissional e os que seguiram uma via de educação geral, ao longo das suas carreiras (Organization for Economic Cooperation and Development, 2014)” (Raymond & Negassi, 2015, p. 17).

8.2 Limitações e futuros desenvolvimentos

As características que moldam a aplicação dos princípios da ciência à investigação em educação ressaltam o importante papel do contexto (National Research Council, 2002, 2005). Uma implicação específica do papel dos fatores contextuais na pesquisa em educação é que os limites da generalização da pesquisa científica precisam ser cuidadosamente delineados. Para construir teorias, formular questões de investigação, planejar e concretizar estudos, e tirar conclusões no âmbito da investigação científica em educação, o/a investigador(a) deve atender a tais condições contextuais. Essa atenção ao contexto também sugere que o avanço da compreensão em contextos educacionais complexos e diversos pode exigir uma estreita coordenação entre investigadores e profissionais, o trabalho interdisciplinar e a interação entre diversas formas de investigação educacional. Significa também a atenção especial à diversidade inerente à experiência educacional e à generalização das descobertas para diferentes populações de estudantes.

Com efeito, a sequência de artigos de investigação institucional que compõem a Tese integram um projeto mais amplo (Almeida & Castro, 2016), que visa uma melhor compreensão da transição dos estudantes do ensino secundário para o ES e sua adaptação nesta nova etapa. Portanto, desde o início da investigação que nos preocupámos com a generalização, já que a maioria da investigação de base quantitativa apresentada diz respeito a investigação institucional (Altbach, 2014). Para que as descobertas sejam passíveis de generalização, incluímos a caracterização da Universidade do Minho, com o objetivo de explicitamente especificar a população de inferência (Committee on Publication Ethics, 1999; Hedges, 2018). De acordo com as normas, é este tipo de conduta que dá suporte a muitos dos estudos baseados em evidência empírica de uma única instituição de ES (Georg, 2009; Páramo Fernández et al., 2017; Vulperhorst et al., 2018). Nestes termos, é necessária alguma cautela para não generalizar os resultados aqui apresentados a populações diferentes daquela a que

esta série de estudos concerne: coorte de ingressantes da Universidade do Minho no ano lectivo 2015/16. Se tivessem sido seleccionadas outras coortes e/ou outras instituições os resultados poderiam ser ligeiramente diferentes. Nesta perspetiva, é muito importante desenvolver uma investigação mais ampla que reconsidere as questões de investigação aqui tratadas e que contemple amostras provenientes de diferentes coortes de ingressantes e de múltiplas instituições de ES de modo a obter resultados de âmbito nacional. Também seria de grande contribuição para o progresso científico-disciplinar se o âmbito fosse alargado do 1º ano a todos os anos do ciclo de estudos, culminando com a obtenção do diploma. Tal alargamento do âmbito da investigação implicaria hipóteses diferentes, perguntas de pesquisa diferentes, recolha de novos dados, novas análises, e discussões diferentes. Seria o desenvolvimento natural deste trabalho. É meu grande desejo partilhá-lo com outros colegas e, conjuntamente, torná-lo possível, acreditando que com estes desenvolvimentos se possa contribuir para a conceção, planeamento e implementação de políticas e medidas favoráveis à adaptação.

REFERÊNCIAS

- Adams, K. S., & Proctor, B. E. (2010). Adaptation to college for students with and without disabilities: Group differences and predictors. *Journal of Postsecondary Education and Disability*, 22(3), 166–184.
- AERA, APA, & NCME. (1999). *The standards for educational and psychological testing*. (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, Eds.). Washington DC: American Educational Research Association.
- Aina, C. (2013). Parental background and university dropout in Italy. *Higher Education*, 65(4), 437–456. doi:10.1007/s10734-012-9554-z
- Aitkin, M., & Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society. Series A*, 149(1), 1–43. doi:10.2307/2981882
- Alarcão, I. (2000). Para uma conceptualização dos fenómenos de sucesso/insucesso escolares no ensino superior. In J. Tavares & R. Santiago (Eds.), *Ensino superior: (In)sucesso académico* (pp. 13–23). Porto: Porto Editora.
- Almeida, L. S. (2002). Formatar o ensino a pensar na aprendizagem. In A. S. Pouzada, L. S. Almeida, & R. M. Vasconcelos (Eds.), *Contextos e dinâmicas da vida académica* (pp. 239–252). Guimarães: Universidade do Minho.
- Almeida, L. S., & Castro, R. V. (2016). Ser estudante no ensino superior: Observatório dos percursos académicos dos estudantes da UMinho. In L. S. Almeida & R. V. Castro (Eds.), *Ser estudante no ensino superior: O caso dos estudantes do 1º ano* (pp. 1–14). Braga: Universidade do Minho.
- Almeida, L. S., Deaño, M., Araújo, A. M., Costa, A. R., Conde, Á., & Alfonso, S. (2012). *Questionário de perceções académicas: Versão expectativas (QPA-E)*. Braga/Ourense: Universidade do Minho/Universidade de Vigo.
- Almeida, L. S., Guisande, M. A., Soares, A. P., & Saavedra, L. (2006). Acesso e sucesso no Ensino Superior em Portugal: Questões de género, origem sócio-cultural e percurso académico dos alunos. *Psicol. Refl. Crít.*, 19(3), 507–514. doi:10.1590/S0102-79722006000300020
- Almeida, L. S., Marinho-Araujo, C. M., Amaral, A., & Dias, D. (2012). Democratização do acesso e do sucesso no ensino superior: uma reflexão a partir das realidades de Portugal e do Brasil. *Avaliação: Revista Da Avaliação Da Educação Superior (Campinas)*, 17(3), 899–920. doi:10.1590/S1414-40772012000300014
- Altbach, P. G. (2014). The emergence of a field: research and training in higher education. *Studies in Higher Education*, 39(8), 1306–1320. doi:10.1080/03075079.2014.949541
- Altbach, P. G., Bozeman, L. A., Janashia, N., & Rumbley, L. E. (2007). *Higher Education: A worldwide inventory of centers and programs (Revised edition)*. Boston: Sense Publishers.
- Altbach, P. G., & Engberg, D. (2001). *Higher Education: A worldwide inventory of centers and programs*. Phoenix: Oryx Press.
- Amin, H. M., Hanawi, S. A., Mohamed, H., Saad, S., & Sahari, N. (2013). Academic

- performance of pre-graduate students. *International Education Studies*, 6(6). doi:10.5539/ies.v6n6p131
- Anderson, N. (2013). D.C. school officials: 44 teachers were given mistaken performance evaluations. *The Washington Post*, p. December 23. Retrieved from https://www.washingtonpost.com/local/education/dc-school-officials-44-teachers-were-given-mistaken-performance-evaluations/2013/12/23/c5cb9f26-6c0c-11e3-a523-fe73f0ff6b8d_story.html?utm_term=.f4c9e9b81ffc
- Andriola, W. B., Andriola, C. G., & Moura, C. P. (2006). Opiniões de docentes e de coordenadores acerca do fenômeno da evasão discente dos cursos de graduação da Universidade Federal do Ceará (UFC). *Ensaio: Aval.Pol.Públ.Educ.*, 14(52), 365–382. doi:10.1590/S0104-40362006000300006
- Arbona, C., & Nora, A. (2007). The influence of academic and environmental factors on Hispanic college degree attainment. *The Review of Higher Education*, 30(3), 247–269.
- Arcidiacono, P. (2005). Affirmative action in higher education: how do admission and financial aid rules affect future earnings? *Econometrica*, 73(5), 1477–1524. doi:10.1111/j.1468-0262.2005.00627.x
- Ashwin, P., & Case, J. (2012). Questioning theory-method relations in higher education research. *Higher Education Research and Development*, 31(3), 271–272. doi:10.1080/07294360.2012.647390
- Assembleia da República. Lei nº 62/2007, Regime jurídico das instituições de ensino superior. Diário da República, 1ª série, n.174 (2007). Portugal.
- Astin, A. W., & Denson, N. (2009). Multi-campus studies of college Impact : Which statistical method is appropriate ? *Res High Educ*, 50, 354–367. doi:10.1007/s11162-009-9121-3
- Bardagi, M. P., & Hutz, C. S. (2009). “Não havia outra saída”: percepções de alunos evadidos sobre o abandono do curso superior. *Psico-USF*, 14(1), 95–105. Retrieved from <http://dx.doi.org/10.1590/S1413-82712009000100010>
- Baron, J., & Norman, M. F. (1992). Sats, achievement tests, and high-school class rank as predictors of college performance. *Educational and Psychological Measurement*, 52(4), 1047–1055. doi:10.1177/0013164492052004029
- Beenstock, M., & Feldman, D. (2018). Decomposing university grades: a longitudinal study of students and their instructors. *Studies in Higher Education*, 43(1), 114–133. doi:10.1080/03075079.2016.1157858
- Berg, M. N. Van Den, & Hofman, W. H. A. (2005). Student success in university education: A multi-measurement study of the impact of student and faculty factors on study progress. *Higher Education*, 50(3), 413–446. doi:10.1007/s10734-004-6361-1
- Bettinger, E. (2015). Need-based aid and college persistence. *Educational Evaluation and Policy Analysis*, 37(1_suppl), 102S–119S. doi:10.3102/0162373715576072
- Bird, S. M., Cox, D., Farewell, V. T., Goldstein, H., Holt, T., & Smith, P. C. (2005). Performance indicators: Good, bad, and ugly. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168(1), 1–27. doi:10.1111/j.1467-985X.2004.00333.x
- Bittencourt, I. M., & Mercado, L. P. L. (2014). Evasão nos cursos na modalidade de educação

- a distância: estudo de caso do curso piloto de administração da UFAL/UAB. *Ensaio: Aval. Pol. Públ. Educ.*, 22(83), 465–504. doi:10.1590/S0104-40362014000200009
- Black, D., Sanders, S., & Taylor, L. (2003). Measurement of higher education in the census and current population survey. *Journal of the American Statistical Association*, 98(463), 545–554. doi:10.1198/016214503000000369
- Brahm, T., Jenert, T., & Wagner, D. (2017). The crucial first year: A longitudinal study of students' motivational development at a Swiss business school. *Higher Education*, 73, 459–478. doi:10.1007/s10734-016-0095-8
- Bratti, M., McKnight, A., Naylor, R., & Smith, J. (2004). Higher education outcomes, graduate employment and university performance indicators. *J. R. Statist. Soc. A*, 167(3), 475–496. Retrieved from <http://www.jstor.org/stable/3559775>
- Braun, H. (2005). Value-added modeling: What does due diligence require. In R. Lissitz (Ed.), *Value-added models in education: Theory and applications* (pp. 19–38). Maple Grove, MN: JAM Press Books.
- Braun, H. (2013). Value-added modeling and the power of magical thinking. *ENSAIO: Aval. Pol. Públ. Educ.*, 21(78), 115–130. doi:10.1590/S0104-40362013000100007
- Braun, H., Chudowsky, N., & Koenig, J. (2010). *Getting value out of value-added*. *Social Sciences*. Washington, DC: National Academies Press.
- Braun, H., & Wainer, H. (2007). Value-added modeling. In C. R. Rao & S. Sinharay (Eds.), *Handbook of Statistics: Psychometrics* (Vol. 26, pp. 867–892). Amsterdam: Elsevier. doi:10.1016/S0169-7161(06)26027-9
- Braxton, J., Hirschy, A., & McClendon, S. (2004). *Understanding and reducing college student departure*. Indianapolis: Jossey-Bass. (Eric n. ED 501184).
- Brooke, N., Fernandes, N. da S., Miranda, I. P. H., & Soares, T. M. (2014). Modelagem do crescimento da aprendizagem nos anos iniciais com dados longitudinais da pesquisa GERES I. *Educação e Pesquisa*, 40(1), 77–94. doi:10.1590/S1517-97022014000100006.
- Bruinsma, M., & Jansen, E. P. W. A. (2009). When will I succeed in my first - year diploma? Survival analysis in Dutch higher education. *Higher Education Research & Development*, 28(1), 99–114. doi:10.1080/07294360802444396
- Bryk, A., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: SAGE Publications.
- Butler, G., Deaton, S., Hodgkinson, J., Holmes, E., & Marshall, S. (2005). *Quick but not dirty: Rapid evidence assessments as a decision support tool in social policy*. London.
- Cabras, C., & Mondo, M. (2018). Coping strategies, optimism, and life satisfaction among first-year university students in Italy: gender and age differences. *Higher Education*, 75(4), 643–654. doi:10.1007/s10734-017-0161-x
- Cabrera, A. F., Stampen, J. O., & Hansen, W. L. (1990). Exploring the effects of ability to pay on persistence in college. *The Review of Higher Education*, 13(3), 303–336. doi:10.1353/rhe.1990.0020
- Cantwell, R., Archer, J., & Bourke, S. (2001). A comparison of the academic experiences and achievement of university students entering by traditional and non-traditional means.

- Assessment & Evaluation in Higher Education*, 26(3), 221–234.
doi:10.1080/02602930120052387
- Chen, R. (2012). Institutional characteristics and college student dropout risks: A multilevel event history analysis. *Res High Educ*, 53, 487–505. doi:10.1007/s11162-011-9241-4
- Chingos, M. M. (2016). Instructional quality and student learning in higher education: Evidence from developmental algebra courses. *The Journal of Higher Education*, 87(1), 84–114. doi:10.1080/00221546.2016.11777395
- Chiroleu, A. (2012). Políticas de Educación Superior en América Latina en el Siglo XXI: ¿Inclusión o Calidad? *Archivos Analíticos de Políticas Educativas*, 20(13), 1–16.
- Chowdry, H., Crawford, C., Dearden, L., Goodman, A., & Vignoles, A. (2013). Widening participation in higher education: analysis using linked administrative data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(2), 431–457. doi:10.1111/j.1467-985X.2012.01043.x
- Clerici, R., Castiglioni, M., Grion, V., Zago, G., & Re, L. Da. (2014). A participative approach to evaluation of graduates' professional outcomes. *Universal Journal of Educational Research*, 2(6), 454–469. doi:10.13189/ ujer.2014.020603
- Clerici, R., Giraldo, A., & Meggiolaro, S. (2015). The determinants of academic outcomes in a competing risks approach: Evidence from Italy. *Studies in Higher Education*, 40(9), 1535–1549. doi:10.1080/03075079.2013.878835
- Clifton, R. (1997). The effects of social psychological variables and gender on the grade point averages and educational expectations of university students: A case study. *Canadian Journal of Higher Education*, 27(2–3), 67–90. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ559665&site=eho-st-live>
- Clotfelter, C. T., Ladd, H. F., Vigdor, J. L., & Diaz, R. A. (2004). Do school accountability systems make it more difficult for low-performing schools to attract and retain high-quality teachers? *Journal of Policy Analysis and Management*, 23(2), 251–271. doi:10.1002/pam.20003
- Commission of the European Communities. (2001). *European Governance: A white paper*. Brussels.
- Committee on Publication Ethics. (1999). Guidelines on good publication practice. Retrieved from <https://publicationethics.org/files/u7141/1999pdf13.pdf>
- Crawford, I., Wang, Z., & Andrews, G. (2016). Exploring the influence of individual and academic differences on the placement participation rate among international students: A UK case study. *Education + Training*, 58(4), 342–357.
- Credé, M., Roch, S. G., & Kieszczynka, U. M. (2010). Class attendance in college. *Review of Educational Research*, 80(2), 272–295. doi:10.3102/0034654310362998
- Cronbach, L., & Suppes, P. (1969). *Research for tomorrow's schools: Disciplined inquiry for evaluation*. New York: MacMillan.
- Danaher, P. A., Bowser, D., & Somasundaram, J. (2008). The student departure puzzle: Do some faculties and programs have answers? *Higher Education Research & Development*, 27(3), 271–280. doi:10.1080/07294360802183820

- Davis, E., Smither, C., Zhu, B., & Stephan, J. (2017). *Characteristics and postsecondary pathways of students who participate in acceleration programs in Minnesota*. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>
- Dennis, J. M., Phinney, J. S., & Chuateco, L. I. (2005). The role of motivation, parental support, and peer support in the academic success of ethnic minority first-generation college students. *Journal of College Student Development, 46*(3), 223–236. doi:10.1353/csd.2005.0023
- Di Pietro, G. (2004). The determinants of university dropout in Italy: A bivariate probability model with sample selection. *Applied Economics Letters, 11*(3), 187–191. doi:10.1080/1350485042000203832
- Dias, D. (2015). Has massification of higher education led to more equity? Clues to a reflection on Portuguese education arena. *International Journal of Inclusive Education, 19*(2), 103–120. doi:10.1080/13603116.2013.788221
- Dias, D., Marinho-Araújo, C., Almeida, L., & Amaral, A. (2011). The democratisation of access and success in higher education the case of Portugal and Brazil. *Higher Education Management and Policy, 23*(1), 23–42.
- Dill, D. D., & Soo, M. (2005). Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems. *Higher Education, 49*(4), 495–533. doi:10.1007/s10734-004-1746-8
- Diseth, Å., Pallesen, S., Brunborg, G. S., & Larsen, S. (2010). Academic achievement among first semester undergraduate psychology students: The role of course experience, effort, motives and learning strategies. *Higher Education, 59*(3), 335–352. doi:10.1007/s10734-009-9251-8
- Draper, D., & Gittoes, M. (2004). Statistical analysis of performance indicators in UK higher education. *J. R. Statist. Soc. A, 167*(3), 449–474.
- Durango-Cohen, E. J., & Balasubramanian, S. K. (2015). Effective segmentation of university alumni: Mining contribution data with finite-mixture models. *Res High Educ, 56*, 78–104. doi:10.1007/s11162-014-9339-6
- Dynarski, S., & Scott-Clayton, J. (2013). *Financial aid policy: Lessons from research*. Cambridge, MA. doi:10.3386/w18710
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*(1), 109–132. doi:10.1146/annurev.psych.53.100901.135153
- Eckland, B. (1964). Social class and college graduation: Some misconceptions corrected. *American Journal of Sociology, 70*(1), 36–50.
- English, D., & Umbach, P. D. (2016). Graduate school choice: An examination of individual and institutional effects. *The Review of Higher Education, 39*(2), 173–211. doi:10.1353/rhe.2016.0001
- Engrácia, P., & Baptista, J. O. (2018). *Percursos no ensino superior: Situação após quatro anos dos alunos inscritos em licenciaturas de três anos*. Lisboa: Direção-Geral de Estatísticas da Educação e Ciência.

- European Commission. (2013). *Europe 2020 target: early leavers from education and training*. Retrieved from http://ec.europa.eu/europe2020/pdf/themes/29_early_school_leaving.pdf
- Everson, K. C. (2017). Value-added modeling and educational accountability: Are we answering the real questions? *Review of Educational Research*, 87(1), 35–70. doi:10.3102/0034654316637199
- Farsides, T., & Woodfield, R. (2007). Individual and gender differences in good and first-class undergraduate degree performance. *British Journal of Psychology*, 98(3), 467–483. doi:10.1348/000712606X150246
- Ferrão, M. E. (2003). *Introdução aos modelos de regressão multinível em educação*. Campinas: Komedi.
- Ferrão, M. E. (2009). *Modelo de Valor Acrescentado: Algumas evidências a partir do projecto 3EM, com discussão*. Covilhã: Universidade da Beira Interior.
- Ferrão, M. E. (2010, December). Valor acrescentado em Educação. *Educação/Jornal de Letras*, pp. 7–8.
- Ferrão, M. E. (2011). Valor acrescentado em Educação: Finalidade e uso. *Educação/Jornal de Letras, Abril*, 4–5.
- Ferrão, M. E. (2012a). Avaliação educacional e modelos de valor acrescentado: Tópicos de reflexão. *Educação & Sociedade*, 33(119), 455–469.
- Ferrão, M. E. (2012b). Note on parents' choice of schools and the quality of education. *Educational Research Review*, 7(3). doi:10.1016/j.edurev.2012.05.001
- Ferrão, M. E. (2014). School effectiveness research findings in the Portuguese speaking countries: Brazil and Portugal. *Educational Research for Policy and Practice*, 13(1), 3–24. doi:10.1007/s10671-013-9151-7
- Ferrão, M. E. (2015a). Investigação em educação e os resultados do PISA: Análise estatística da retenção. In Conselho Nacional de Educação (Ed.), *Estado da Educação 2014* (pp. 328–337). Lisboa: Conselho Nacional de Educação. Retrieved from <http://www.cnedu.pt/pt/publicacoes/estudos-e-relatorios/estado-da-educacao/1048-estado-da-educacao-2014>
- Ferrão, M. E. (2015b). Modelos de regressão multinível em Educação e Psicologia. In M. C. R. Silva, D. Bartholomeu, C. M. M. Vendramini, & J. M. Montiel (Eds.), *Métodos estatísticos avançados aplicados à Psicologia e Educação* (pp. 25–36). São Paulo: Vetor Editora.
- Ferrão, M. E. (2015c). Tópicos sobre retenção em Portugal através do PISA: Qualidade e equidade. *Education Policy Analysis Archives [Arquivos Analíticos de Políticas Educativas]*, 23(114). doi:1014507
- Ferrão, M. E. (2016). Estatística educacional: A propósito dos modelos de valor acrescentado. *Boletim Sociedade Portuguesa de Estatística, Outono*, 23–25.
- Ferrão, M. E., & Almeida, L. S. (2018a). Differential effect of university entrance score on first-year students' academic performance in Portugal. *Assessment & Evaluation in Higher Education*, *In press*. doi:10.1080/02602938.2018.1525602
- Ferrão, M. E., & Almeida, L. S. (2018b). Multilevel modelling of persistence in higher

- education. *Ensaio: Aval. Pol. Públ. Educ.*, 26(100), 664–683. doi:10.1590/S0104-40362018002601610
- Ferrão, M. E., & Couto, A. (2013). Indicador de valor acrescentado e tópicos sobre consistência e estabilidade: Uma aplicação ao Brasil. *Ensaio: Aval. Pol. Públ. Educ.*, 21(78), 131–164. doi:10.1590/S0104-40362013000100008
- Ferrão, M. E., & Couto, A. P. (2014). The use of a school value-added model for educational improvement: a case study from the Portuguese primary education system. *School Effectiveness and School Improvement*, 25(1). doi:10.1080/09243453.2013.785436
- Ferrão, M. E., & Goldstein, H. (2009). Adjusting for measurement error in the value added model: Evidence from Portugal. *Qual Quant*, 43(6), 951–963. doi:10.1007/s11135-008-9171-1
- Ferrão, M. E., & Klein, R. (2013). Editorial. *Ensaio: Aval. Pol. Públ. Educ.*, 78(21), 7–9.
- Ferreira, C., Vidal, J., & Vieira, M. J. (2014). Student guidance and attention to diversity in the processes of quality assurance in Higher Education. *European Journal of Education*, 49(4), 575–589. doi:10.1111/ejed.12098
- Foley, B., & Goldstein, H. (2012). *Measuring success: League tables in the public sector*. London: British Academy.
- Fonseca, M., Dias, D., Sá, C., & Amaral, A. (2014). Waves of (dis)satisfaction: Effects of the numerus clausus system in Portugal. *European Journal of Education*, 49(1), 144–158. doi:10.1111/ejed.12042
- Fry, R. (2002). *Latinos in higher education: Many enroll, too few graduate*. Washington, D.C.: Pew Hispanic Center.
- Fuller, M. B., Wilson, M. A., & Tobin, R. M. (2011). The national survey of student engagement as a predictor of undergraduate GPA: A cross - sectional and longitudinal examination. *Assessment & Evaluation in Higher Education*, 36(6), 735–748. doi:10.1080/02602938.2010.488791
- Georg, W. (2009). Individual and institutional factors in the tendency to drop out of higher education: A multilevel analysis using data from the Konstanz Student Survey. *Studies in Higher Education*, 34(6), 647–661. doi:10.1080/03075070802592730
- Giani, M. S. (2016). Are all colleges equally equalizing? How institutional selectivity impacts socioeconomic disparities in graduates' labor outcomes. *The Review of Higher Education*, 39(3), 431–461.
- Goldstein, H. (1997). Methods in school effectiveness research. *School Effectiveness and School Improvement*, 8(4), 369–395. doi:10.1080/0924345970080401
- Goldstein, H. (1998). *Models for reality: New approaches to the understanding of educational processes*. London: Institute of Education Publications.
- Goldstein, H. (2003). *Multilevel statistical models* (3rd ed.). London: Edward Arnold.
- Goldstein, H. (2011). *Multilevel statistical models* (4th ed.). Chichester: Wiley.
- Goldstein, H. (2014). Using league table rankings in public policy formation: Statistical issues. *Annual Review of Statistics and Its Application*, 1(1), 385–399. doi:10.1146/annurev-statistics-022513-115615

- Goldstein, H., Browne, W., & Rasbash, J. (2002). Partitioning variation in multilevel models. *Understanding Statistics*, 1(4), 223–231. doi:10.1207/S15328031US0104_02
- Goldstein, H., & Rasbash, J. (1996). Improved approximations for multilevel models with binary responses. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 159(3), 505–513. doi:10.2307/2983328
- Goldstein, H., & Spiegelhalter, D. (1996). League tables and their limitations: Statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 159, 385–443. Retrieved from <http://www.jstor.org/pss/2983325>
- Gormley, I. C., & Murphy, T. B. (2006). Analysis of Irish third-level college applications data. *J. R. Statist. Soc. A*, 169(2), 361–379.
- Gross, J. P. K., Hossler, D., Ziskin, M., & Berry, M. S. (2015). Institutional merit-based aid and student departure: A longitudinal analysis. *The Review of Higher Education*, 38(2), 221–250.
- Guellec, D., Larrue, P., Roy, S., & Weko, T. (2018). *OECD Review of the tertiary education, research and innovation system in Portugal [Preliminary report]*. Paris: OECD.
- Gutierrez, A., & Dantes, J. (2009). Practical model for the community college. *Community College Journal*, 33(11), 958–961. doi:10.1080/10668920903153261
- Hazelkorn, E. (2011). *Rankings and the reshaping of higher education: The battle for world-class excellence*. London: Palgrave Macmillan.
- Hébert, T. P. (2018). An examination of high-achieving first-generation college students from low-income backgrounds. *Gifted Child Quarterly*, 62(1), 96–110. doi:10.1177/0016986217738051
- Hedges, L. V. (2018). Challenges in building usable knowledge in education. *Journal of Research on Educational Effectiveness*, 11(1), 1–21. doi:10.1080/19345747.2017.1375583
- Howard, C., & Davies, P. (2013). Attracting mature students into higher education: The impact of approaches to learning and social identity. *Journal of Further and Higher Education*, 37(6), 769–785. doi:10.1080/0309877X.2012.684038
- Hutchinson, S. R., & Lovell, C. D. (2004). A review of methodological characteristics of research published in key journals in higher education: Implications for graduate research training. *Res High Educ*, 45(4), 383–403.
- Ishitani, T. T. (2016). Time-varying effects of academic and social integration on student persistence for first and second years in college: National data approach. *J. College Student Retention*, 18(3), 263–286. doi:10.1177/1521025115622781
- Jaeger, A. J., & Eagan, M. K. (2011). Examining retention and contingent faculty use in a state system of public higher education. *Educational Policy*, 25(3), 507–537. doi:10.1177/0895904810361723
- James, R. (2002). Students' changing expectations of higher education and the consequences of mismatches with the reality. In *Responding to students expectations* (pp. 71–83). Paris: OECD Publishing.
- Johnson, I. (2008). Enrollment, persistence and graduation of in-state students at a public

- research university: Does high school matter? *Res High Educ*, 49, 776–793.
doi:10.1007/s11162-008-9105-8
- Johnson, S. E., Richeson, J. A., & Finkel, E. J. (2011). Middle class and marginal? Socioeconomic status, stigma, and self-regulation at an elite university. *Journal of Personality and Social Psychology*, 100(5), 838–852. doi:10.1037/a0021956
- Johnston, R., Manley, D., Jones, K., Harris, R., & Hoare, A. (2016). University admissions and the prediction of degree performance: An analysis in the light of changes to the English schools' examination system. *Higher Education Quarterly*, 70(1), 24–42.
doi:10.1111/hequ.12067
- Jolliffe, I. T. (2002). *Principal Component Analysis (2nd ed.)* (Springer S). New York: Springer.
- Kasule, G. W., Wesselink, R., Noroozi, O., Mulder, M., Wilson, G., Wesselink, R., ... Mulder, M. (2015). The current status of teaching staff innovation competence in Ugandan universities : perceptions of managers , teachers , and students. *Journal of Higher Education Policy and Management*, 37(3), 330–343.
doi:10.1080/1360080X.2015.1034425
- Kilgo, C. A. (2017). The estimated effects of service learning on students' intercultural effectiveness. *Journal of College Student Development*, 56(8), 867–871.
- Kodama, C. M., & Dugan, J. P. (2013). Leveraging leadership efficacy for college students: Disaggregating data to examine unique predictors by race leveraging leadership. *Equity & Excellence in Education*, 46(2), 184–201. doi:10.1080/10665684.2013.780646
- Laforge, M. C. (2003). Explanatory style and academic performance among college students beginning a major course of study. *Psychological Reports*, 92(3), 861.
doi:10.2466/PRO.92.3.861-865
- Lang, T. (2015). Campus sustainability initiatives and performance: Do they correlate? *International Journal of Sustainability in Higher Education*, 16(4), 474–490.
- Lavigne, G. L., Vallerand, R. J., & Miquelon, P. (2007). A motivational model of persistence in science education: A self-determination theory approach. *European Journal of Psychology of Education*, 22(3), 351–369. doi:10.1007/BF03173432
- Leach, L. (2016). Exploring discipline differences in student engagement in one institution. *Higher Education Research & Development*, 35(4), 772–786.
doi:10.1080/07294360.2015.1137875
- Letskiewicz, J., Lim, H., Heckman, S., Bartholomae, S., Fox, J. J., & Montalto, C. P. (2014). The path to graduation: Factors predicting on-time graduation rates. *Journal of College Student Retention: Research, Theory & Practice*, 16(3), 351–371.
doi:10.2190/CS.16.3.c
- Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data (2nd ed.)*. New Jersey: John Wiley & Sons Inc.
- Liu, O. L., Bridgeman, B., & Adler, R. M. (2012). Measuring learning outcomes in higher education: Motivation matters. *Educational Researcher*, 41, 352-
doi:10.3102/0013189X12459679
- Lockwood, J. R., McCaffrey, D. F., Hamilton, L. S., Stecher, B., Le, V.-N., & Martinez, J. F. (2007). The sensitivity of value-added teacher effect estimates to different mathematics

- achievement measures. *Journal of Educational Measurement*, 44(1), 47–67.
doi:10.1111/j.1745-3984.2007.00026.x
- Lohfink, M. M., & Paulsen, M. B. (2005). Comparing the determinants of persistence for first-generation and continuing-generation students. *Journal of College Student Development*, 46(4), 409–428. doi:10.1353/csd.2005.0040
- Longford, N. (1993). *Random coefficient models*. New York: Oxford University Press.
- López, D. A., Rojas, M. J., López, B. A., & López, D. C. (2015). Chilean universities and institutional quality assurance processes. *Quality Assurance in Education*, 23(2), 166–183.
- Masui, C., Broeckmans, J., Doumen, S., Groenen, A., & Molenberghs, G. (2014). Do diligent students perform better? Complex relations between student and course characteristics, study time, and academic performance in higher education. *Studies in Higher Education*, 39(4), 621–643. doi:10.1080/03075079.2012.721350
- McCaffrey, D. F., Sass, T. R., Lockwood, J. R., & Mihaly, K. (2009). The intertemporal variability of teacher effect estimates. *Education Finance and Policy*, 4(4), 572–606. doi:10.1162/edfp.2009.4.4.572
- McCullagh, P., & Nelder, J. (1989). *Generalized linear models* (2nd ed.). New York: Chapman & Hall/CRC.
- McKenzie, K., & Schweitzer, R. (2001). Who succeeds at university? Factors predicting academic performance in first year Australian university students. *Higher Education Research & Development*, 20(1), 21–33. doi:10.1080/07924360120043621
- Meggiolaro, S., Giraldo, A., & Clerici, R. (2015). A multilevel competing risks model for analysis of university students' careers in Italy. *Studies in Higher Education*. doi:10.1080/03075079.2015.1087995
- Mendes Junior, A. A. F. (2014). Uma análise da progressão dos alunos cotistas sob a primeira ação afirmativa brasileira no ensino superior: O caso da Universidade do Estado do Rio de Janeiro. *Ensaio: Avaliação E Políticas Públicas Em Educação*, 22(82), 31–56. doi:10.1590/S0104-40362014000100003
- Ministério da Educação e Ciência. (2014). Decreto-Lei nº113/2014. *Diário Da República*, 1ª Série -, 135, 3874–3878.
- Montmarquette, C., Mahseredjian, S., & Houle, R. (2001). The determinants of university dropouts: A bivariate probability model with sample selection. *Economics of Education Review*, 20(5), 475–484. doi:10.1016/S0272-7757(00)00029-7
- Moretti, E. (2004). Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, 121, 175–212. doi:10.1016/j.jeconom.2003.10.015
- Morganstein, D., & Wasserstein, R. (2014). ASA Statement on Value-Added Models. *Statistics and Public Policy*, 1(1), 108–110. doi:10.1080/2330443X.2014.956906
- Morris, A. (2002). *From idea to impact: A guide to the research process*. (A. Morris, Ed.) (Building e). London: Learning and Skills Research Centre.
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School matters*. Somerset: Open Books.

- Mortimore, P., Sammons, P., & Thomas, S. (1994). School effectiveness and value added measures. *Assessment in Education: Principles, Policy & Practice*, 1(3), 315–332. doi:10.1080/0969594940010307
- Mouw, J. T., & Khanna, R. K. (1993). Prediction of academic success: A review of the literature and some recommendations. *College Student Journal*, 27(3), 328–336.
- Mukhtar, U., Islam, Z., & Siengthai, S. (2011). Conflicts in higher education and perceived quality of education: Empirical evidence from Pakistan. *Research in Higher Education Journal*. ERIC n. EJ1068804
- Naidoo, A., & Lemmens, J. (2015). Faculty intervention as support for first-year students. *Journal of Student Affairs in Africa*, 3(2), 17–32. doi:10.14426/jsaa.v3i2.133
- National Research Council. (2002). *Scientific research in education*. (R. J. Shavelson & L. Towne, Eds.). Washington, DC: National Academy Press.
- National Research Council. (2005). *Advancing scientific research in education*. (L. Towne, L. L. Wise, & T. M. Winters, Eds.). Washington, DC: The National Academies Press.
- Nevo, D. (1983). The conceptualization of educational evaluation: An analytical review of the literature. *Review of Educational Research*, 53, 117–128.
- Nevo, D. (2006). Evaluation in education. In I. F. Shaw, J. C. Greene, & M. M. Mark (Eds.), *The Sage handbook of evaluation* (pp. 440–460). London: SAGE Publications.
- Ngo, F., & Kosiewicz, H. (2017). How extending time in developmental math impacts student persistence and success: Evidence from a regression discontinuity in community colleges. *The Review of Higher Education*, 40(2), 267–306.
- Nicholson, L., Putwain, D., Connors, L., & Hornby-Atkinson, P. (2013). The key to successful achievement as an undergraduate student: Confidence and realistic expectations? *Studies in Higher Education*, 38(2), 285–298. doi:10.1080/03075079.2011.585710
- Niehaus, E., Campbell, C. M., & Inkelas, K. K. (2014). HLM Behind the Curtain: Unveiling decisions behind the use and interpretation of HLM in Higher Education research. *Research in Higher Education*, 55(1), 101–122. doi:10.1007/s11162-013-9306-7
- Nora, A., Cabrera, A. F., Hagedon, L. S., & Pascarella, E. T. (1996). Differential impacts of academic and social experiences on college-related behavioral outcomes across different ethnic and gender groups at four-year institutions. *Res High Educ*, 37, 427–451.
- O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences*, 43(5), 971–990. doi:10.1016/j.paid.2007.03.017
- OECD. (2007). *Best practices for ensuring scientific integrity and preventing misconduct*. Paris: OECD Publishing.
- OECD. (2008). *Measuring improvements in learning outcomes*. Paris: OECD Publishing. doi:10.1787/9789264050259-en
- OECD. (2014). *Education at a Glance 2014: OECD Indicators*. *Oecd*. doi:10.1787/eag-2013-en
- OECD. (2017). *Education at a Glance 2017: OECD Indicators*. Paris: OECD Publishing. Retrieved from doi:10.1787/eag-2017-en

- Organization for Economic Cooperation and Development. (2014). *How's life in your region? Measuring regional and local well-being for policy making*. Paris: OECD Publishing. Retrieved from <http://www.oecd.org/regional/how-s-life-in-your-region-9789264217416-en.htm>
- Ostrove, J. M., & Long, S. M. (2007). Social class and belonging: Implications for college adjustment. *The Review of Higher Education, 30*(4), 363–389. doi:10.1353/rhe.2007.0028
- Páramo Fernández, M. F., Araújo, A., Vacas, C. T., Almeida, L. da S., & González, M. S. (2017). Predictors of students' adjustment during transition to university in Spain. *Psicothema, 29*(1), 67–72. doi:10.7334/psicothema2016.40
- Park, J. J. (2009). Taking race into account: Charting student attitudes towards affirmative action. *Res High Educ, 50*, 670–690. doi:10.1007/s11162-009-9138-7
- Pascarella, E., Pierson, C., Wolniak, G., & Terenzini, P. (2004). First-generation college students: Additional evidence on college experiences and outcomes. *The Journal of Higher Education, 75*(3), 249–284. Retrieved from <http://www.jstor.org/stable/3838816>
- Pascarella, E. T., Duby, P., Miller, V., & Rasher, S. (1981). Pre-enrollment variables and academic performance as predictors of freshman-year persistence, early withdrawal, and stopout behavior in an urban, nonresidential university. *Res High Educ, 15*(4), 329–349.
- Paulino, C. D., Pestana, D., Branco, J., Singer, J., Barroso, L., & Bussab, W. (2011). Glossário Inglês-Português de Estatística. Sociedade Portuguesa de Estatística e Associação Brasileira de Estatística. Retrieved from <http://glossario.spestatistica.pt/>
- Permzadian, V., & Credé, M. (2016). Do first-year seminars improve college grades and retention? A quantitative review of their overall effectiveness and an examination of moderators of effectiveness. *Review of Educational Research, 86*(1), 277–316. doi:10.3102/0034654315584955
- Pietro, G. Di. (2017). Degree classification and recent graduates' ability: Is there any signalling effect? *Journal of Education and Work, 30*(5), 501–514. doi:10.1080/13639080.2016.1243230
- Pike, G. R., & Graunke, S. S. (2015). Examining the effects of institutional and cohort characteristics on retention rates. *Res High Educ, 56*, 146–165. doi:10.1007/s11162-014-9360-9
- Pike, G. R., & Rocconi, L. M. (2012). Multilevel modeling: Presenting and publishing the results for internal and external constituents. *New Directions for Institutional Research, 154*, 111–124. doi:10.1002/ir.20017
- Pike, G. R., & Saupe, J. L. (2002). Does high school matter? An analysis of three methods of predicting first-year grades. *Research in Higher Education, 43*(2), 187–207.
- Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*(4), 385–407. doi:10.1007/s10648-004-0006-x
- Pinxten, M., Van Soom, C., Peeters, C., De Laet, T., & Langie, G. (2017). At-risk at the gate: Prediction of study success of first-year science and engineering students in an open-

- admission university in Flanders—any incremental validity of study strategies? *European Journal of Psychology of Education*. doi:10.1007/s10212-017-0361-x
- Plewis, I. (1997). Terminology and definition in multilevel models analysis. *Multilevel Modelling Newsletter*, 9(1), 2–4.
- Pordata. (2018). População residente com 15 e mais anos: total e por nível de escolaridade completo mais elevado. Retrieved March 14, 2018, from <https://www.pordata.pt/Portugal/Popula%C3%A7%C3%A3o+residente+com+15+e+mais+anos+total+e+por+n%C3%ADvel+de+escolaridade+completo+mais+elevado-2101>
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135(2), 322–338. doi:10.1037/a0014996
- Portela, M., Areal, N., Sá, C., Alexandre, F., Cerejeira, J., Carvalho, A., & Rodrigues, A. (2008). Evaluating student allocation in the Portuguese public higher education system. *Higher Education*, 56(2), 185–203. doi:10.1007/s10734-007-9097-x
- Pustjens, H., Van de gaer, E., Van Damme, J., & Onghena, P. (2004). Effect of secondary schools on academic choices and on success in higher education. *School Effectiveness and School Improvement*, 15(3–4), 281–311. doi:10.1080/09243450512331383222
- Rasbash, J., Browne, W., Healy, M., Cameron, B., & Charlton, C. (2014). MLwiN 2.31 [Computer software]. Bristol: Centre for Multilevel Modelling, University of Bristol.
- Ray, A., McCormack, T., & Evans, H. (2009). Value added in english schools. *Education Finance and Policy*, 4(4), 415–438. doi:10.1162/edfp.2009.4.4.415
- Raymond, M. E., & Negassi, Y. (2015). *O quinto compromisso. Desenvolvimento de um sistema de garantia de desempenho educativo em Portugal*. Lisboa: Fundação Francisco Manuel dos Santos.
- Reardon, S. F., & Raudenbush, S. W. (2009). Assumptions of value-added models for estimating school effects. *Education Finance and Policy*, 4(4), 492–519. Retrieved from <http://www.mitpressjournals.org/doi/abs/10.1162/edfp.2009.4.4.492>
- Rhee, B.-S. (2008). Institutional climate and student departure: A multinomial multilevel modeling approach. *The Review of Higher Education*, 31(2), 161–183.
- Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138(2), 353–387. doi:10.1037/a0026838
- Ritter, J., & Mayer, A. (2018). Regulating data as property: A new construct for moving forward. *Duke Law & Technology Review*, 16, 220–277. Retrieved from <https://scholarship.law.duke.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1320&context=dltr>
- Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychological and study skills factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, 130(2), 261–288.
- Rodgers, J. L. (2010). The epistemology of mathematical and statistical modeling: A quiet methodological revolution. *The American Psychologist*, 65(1), 1–12. doi:10.1037/a0018326
- Rodgers, T. (2005). Measuring value added in higher education: Do any of the recent

- experiences in secondary education in the United Kingdom suggest a way forward? *Quality Assurance in Education*, 13(2), 95–106. doi:10.1108/09684880510594355
- Ross, T., Kena, G., Rathbun, A., KewalRamani, A., Zhang, J., Kristapovich, P., & Manning, P. (2012). *Higher education: Gaps in access and persistence study - Statistical analysis report (NCES 2012-046)*. Washington DC.
- Rubin, D. B., Stuart, E. A., & Zanutto, E. L. (2004). A potential outcomes view of value-added assessment in education. *Journal of Educational and Behavioral Statistics*, 29(1), 103–116. doi:10.3102/10769986029001103
- Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary academic performance? *Psychological Bulletin*, 135(1), 1–22. doi:10.1037/a0013978
- Sammons, P., Thomas, S., & Mortimore, P. (1997). *Forging Links: Effective Schools and Effective Departments*. London: Paul Chapman.
- Santos Silva, M., Vaz, D., Loureiro, M. J., Ferrão, M. E., Branco, M. L., Garrido, S., ... Justino, E. (2005). Conceção Estratégica das Intervenções Operacionais no Domínio do Ensino Superior. In O. do QCAIII (Ed.), *Estudos temáticos para preparação do próximo ciclo de intervenções estruturais 2007-2013*. Lisboa: Observatório do QCAIII. Retrieved from http://www.qca.pt/fundos/qren_estudos.asp
- Saunders, D. B., Kolek, E. A., Williams, E. A., & Wells, R. S. (2016). Who is shaping the field? Doctoral education, knowledge creation and postsecondary education research in the United States. *Higher Education Research & Development*, 35(5), 1039–1052. doi:10.1080/07294360.2016.1139552
- Saunders, L. (1999). A brief history of educational “value added”: How did we get to where we are? *School Effectiveness and School Improvement*, 10(2), 233–256. doi:10.1076/sesi.10.2.233.3507
- Schudde, L. T. (2011). The causal effect of campus residency on college student retention. *The Review of Higher Education*, 34(4), 581–610.
- Schutz, K. R., Drake, B. M., Lessner, J., Hughes, G. F., Schutz, K. R., Drake, B. M., ... Hughes, G. F. (2015). A comparison of community college full-time and adjunct faculties’ perceptions of factors associated with grade inflation. *The Journal of Continuing Higher Education*, 63(3), 180–192. doi:10.1080/07377363.2015.1085951
- Shavelson, R. J., Domingue, B. W., Mariño, J. P., Mantilla, A. M., Forero, A. M., & Wiley, E. E. (2016). On the practices and challenges of measuring higher education value added: The case of Colombia. *Assessment & Evaluation in Higher Education*, 41(5), 695–720. doi:10.1080/02602938.2016.1168772
- Shaw, A. (2013). Family fortunes: Female students’ perceptions and expectations of higher education and an examination of how they, and their parents, see the benefits of university. *Educational Studies*, 39(2), 195–207. doi:10.1080/03055698.2012.713549
- Shih, C. P., Tillett, D., & Lawrence, N. (2012). An econometrics analysis on the effect of satisfaction for foreign graduate students’ academic performance in Taiwan. *Journal of Learning in Higher Education*, 8(2), 19–29.
- Singer, J. D. (2018). Even more challenges in building usable knowledge in education. *Journal*

- of *Research on Educational Effectiveness*, 11(1), 22–24.
- Smart, J. C. (2005). Attributes of exemplary research manuscripts employing quantitative analyses. *Res High Educ*, 46(4), 461–477. doi:10.1007/s11162-005-2970-5
- Smith, J., & Naylor, R. (2001). Determinants of degree performance in UK universities: A statistical analysis of the 1993 student cohort. *Oxford Bulletin of Economics and Statistics*, 63(1), 29–60. doi:10.1111/1468-0084.00208
- Soares, A. P., Guisande, A. M., Almeida, L. S., & Páramo, F. M. (2009). Academic achievement in first-year Portuguese college students: The role of academic preparation and learning strategies. *International Journal of Psychology*, 44(3), 204–212. doi:10.1080/00207590701700545
- Soares, J. F., Ribeiro, L. M., & Castro, C. de M. (2001). Valor agregado de instituições de ensino superior em Minas Gerais para os cursos de Direito, Administração e Engenharia Civil. *Dados*, 44(2), 363–396. doi:10.1590/S0011-52582001000200005
- Soares, T. M., Bonamino, A., Brooke, N., & Fernandes, N. da S. (2017). Modelos de valor agregado para medir a eficácia das escolas Geres. *Ensaio: Avaliação E Políticas Públicas Em Educação*, 25(94), 59–89. doi:10.1590/s0104-40362017000100003
- Steedle, J. T. (2012). Selecting value-added models for postsecondary institutional assessment. *Assessment & Evaluation in Higher Education*, 37(6), 637–652. doi:10.1080/02602938.2011.560720
- Stinebrickner, R., & Stinebrickner, T. R. (2004). Time-use and college outcomes. *Journal of Econometrics*, 121, 243–269. doi:10.1016/j.jeconom.2003.10.013
- Stinebrickner, T., & Stinebrickner, R. (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics*. doi:10.1086/666525
- Strauss, L. C., & Volkwein, J. F. (2004). Predictors of student commitment at two-year and four-year institutions. *The Journal of Higher Education*, 75(2), 203–227. doi:10.1353/jhe.2004.0007
- Szelenyi, K., Denson, N., & Inkelas, K. K. (2013). Women in STEM majors and professional outcome expectations : The role of living-learning programs and other college environments. *Res High Educ*, 54, 851–873. doi:10.1007/s11162-013-9299-2
- Tavares, D., Tavares, O., Justino, E., & Amaral, A. (2008). Students' preferences and needs in Portuguese higher education. *European Journal of Education*, 43(1), 107–122. doi:10.1111/j.1465-3435.2007.00331.x
- Tavares, O., & Ferreira, J. B. (2012). Choices and motivations: The why and how of Portuguese students' enrolment choices. *European Journal of Education*, 47(2), 310–326. doi:10.1111/j.1465-3435.2012.01527.x
- Tight, M. (2013). Discipline and methodology in higher education research. *Higher Education Research & Development*, 32(1), 136–151. doi:10.1080/07294360.2012.750275
- Tight, M. (2014). Discipline and theory in higher education research. *Research Papers in Education*, 29(1), 93–110. doi:10.1080/02671522.2012.729080
- Tinto, V. (1997). Classrooms as communities: Exploring the educational character of student persistence. *Journal of Higher Education*, 68(6), 599–623. doi:10.2307/2959965
- Tinto, V. (2010). From theory to action: Exploring the institutional conditions for student

- retention. In J. C. Smart (Ed.), *Higher Education: Handbook of Theory and Research* (pp. 51–89). Springer. doi:10.1007/978-90-481-8598-6_2
- Tinto, V., & Cullen, J. (1973). *Dropout in higher education: A review and theoretical synthesis of recent research*. New York: Columbia University.
- Tinto, V., & Pusser, B. (2006). *Moving from theory to action: Building a model of institutional action for student success*. National Postsecondary Education Cooperative.
- Travitzki, R., Ferrão, M. E., & Couto, A. P. (2016). Educational and socio-economic inequalities of pre-university Brazilian population: A view from the ENEM data. *Education Policy Analysis Archives*, 24. doi:10.14507/epaa.24.2199
- Umbach, P. D., & Wawrzynski, M. R. (2005). Faculty do matter: The role of college faculty in student learning and engagement. *Research in Higher Education*, 46(2), 153–184. doi:10.1007/s11162-004-1598-1
- van Rooij, E. C. M., Jansen, E. P. W. A., & van de Grift, W. J. C. M. (2017). First-year university students' academic success: The importance of academic adjustment. *European Journal of Psychology of Education*. doi:10.1007/s10212-017-0347-8
- Vincent-Lancrin, S. (2009). What is changing in academic research? Trends and prospects. In *Higher Education to 2030* (pp. 145–173). Paris: OECD Publishing.
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140(4), 1174–1204. doi:10.1037/a0036620
- Vulperhorst, J., Lutz, C., de Kleijn, R., & van Tartwijk, J. (2018). Disentangling the predictive validity of high school grades for academic success in university. *Assessment & Evaluation in Higher Education*, 43(3), 399–414. doi:10.1080/02602938.2017.1353586
- Walvoord, B. E., Carey, A. K., Smith, H. L., Soled, S. W., Way, P. K., & Zorn, D. (2000). *Academic departments: How they work, how they change*. Washington, D.C.: Jossey-Bass Publishers. (ERIC n. ED 447 746).
- Weiss, C. H. (1975). Evaluation Research in the Political Context. In M. Struening, Elmer & Guttentag (Ed.), *Handbook of Evaluation Research* (pp. 13–26). SAGE Publications.
- Wells, R. S., Kolek, E. A., Williams, E. A., & Saunders, D. B. (2015). “How we know what we know”: A systematic comparison of research methods employed in higher education. *Journal of Higher Education*, 86(2), 171–195.
- Williams, D. A. (1982). Extra-binomial variation in logistic linear models. *Appl. Statist.*, 31(2), 144–148.
- Willson, V. L., & Hughes, J. N. (2006). Retention of Hispanic/Latino students in first grade: Child, parent, teacher, school, and peer predictors. *Journal of School Psychology*, 44, 31–49. doi:10.1016/j.jsp.2005.12.001
- Yamada, R. (2014). Gains in learning outcomes of college students in Japan: Comparative study between academic fields. *The International Education Journal: Comparative Perspectives*, 13(1), 100–118.
- Yorke, M., & Longden, B. (2008). *The first-year experience in higher education in the UK*. York: The Higher Education Academy. Retrieved from www.heacademy.ac.uk
- Zimmermann, J., Brodersen, K. H., Heinemann, H. R., & Buhmann, J. M. (2015). A model-

based approach to predicting graduate-level performance using indicators of undergraduate-level performance. *Journal of Educational Data Mining*, 7(3), 151–176.

ANEXO 1

Table A1. Steps for papers selection

Steps	Procedures	Number of papers
1	Searching ERIC ⁷ since 1998 with “peer reviewed only” active	903
2	After filtering by Descriptor = Statistical Analysis	126
3	After excluding non-related papers ⁸	70
4	Searching CIS since 1998 for “higher education”	100
5	After excluding non-related papers ⁹	89
6	Identifying the top journals in Higher Education ¹⁰	
7	Research in Higher Education with search set to: ("Statistical methods" OR "Statistical models") AND ("Longitudinal data" OR "Panel data" OR "Repeated observations"); Subdiscipline: “Higher Education”	10
8	Journal of Higher Education with search set to: ("Statistical methods" OR "Statistical models") AND ("Longitudinal data" OR "Panel data" OR "Repeated observations");	6
9	Studies in Higher Education with search set to: ("Statistical methods" OR "Statistical models") AND ("Longitudinal data" OR "Panel data" OR "Repeated observations");	2
10	Review of Higher Education with search set to: ("Statistical methods" OR "Statistical models") AND ("Longitudinal data" OR "Panel data" OR "Repeated observations");	8
11	Assessment and Evaluation in Higher Education with search set to: ("Statistical methods" OR "Statistical models") AND ("Longitudinal data" OR "Panel data" OR "Repeated observations");	6
12	Higher Education Research and Development with search set to: ("Statistical methods" OR "Statistical models") AND ("Longitudinal data" OR "Panel data" OR "Repeated observations");	8
Total number of papers		199

⁷ Search terms abstract: "higher education" and abstract: ("dependent variable" or "response variable" or "outcome" or "performance") and abstract:"data".

⁸ Excluded papers on MOOC, e-learning experiences, teaching methods, validation of assessment tests, assessment of learning, economics of education and theory of human capital, job and pay-for-performance, personnel evaluation, governance and presidents, non-higher education related papers, athletic training students, qualitative research, and experimental studies.

⁹ Excluded papers on statistics education, race theory in education, theories and ethnographies, two papers not available online, non-higher education related papers.

¹⁰ <http://www.scimagojr.com/journalrank.php> Subject areas “Social sciences”; Subject categories “Education”; All regions/countries; year 2016; journals with at least 200 citable docs. (3years); with “Higher Education” in the title.