

UMA ANÁLISE DA TRANSIÇÃO SERIAL NO ENSINO MÉDIO CATARINENSE USANDO MICRODADOS ADMINISTRATIVOS

ÁREA 2: Desenvolvimento econômico

Max Cardoso de Resende, UFPEL. E-mail: max_resende@hotmail.com
Francis Petterini, UFSC. E-mail: f.petterini@ufsc.br

A pesquisa busca identificar as características dos jovens, das famílias, do ambiente escolar e de comportamentos que são determinantes do fluxo escolar no ensino médio, a fim de estimar as probabilidades conjuntas de um estudante aprovar/reprovar e permanecer/evadir do sistema. Para tanto, fez-se uso inédito de uma base de dados longitudinais administrada pela Secretaria de Estado da Educação de Santa Catarina entre anos de 2008 e 2012. Dentre os vários resultados observados, verificou-se que a transição da 1ª para a 2ª série é a de maior risco para a reprovação e a evasão, e que as características familiares são preponderantes no que tange ao rendimento escolar. Também fica claro que a influência familiar decresce com o avançar das séries, e que a aprovação gera custos de oportunidade para a permanência na escola.

Palavras-chave: Ensino Médio. Fluxo escolar. Modelos probabilísticos bivariados.

AN ANALYSIS OF HIGH SCHOOL GRADE TRANSITION IN SANTA CATARINA USING ADMINISTRATIVE MICRODATA

The paper goals is to identify the characteristics of the youngsters, the families, the school environment and behaviors that are determinants of the grade transition in high school, in order to estimate the joint probabilities of a student approve/disapprove and stay/evade of the system. In order to do so, an unprecedented use of a longitudinal database administered by the State Department of Education of Santa Catarina between years 2008 and 2012. Among the several results observed, it was verified that the transition from the 1st to the 2nd year is the one of greater risk for reprobation and evasion, and that the familiar characteristics are preponderant when it comes to the school performance. It is also clear that family influence decreases with the advancement of the series, and that approval generates opportunity costs for staying in school.

Keywords: High school. Grade Transition. Bivariate probability models.

Classificação JEL: A21, C35, I21

1. Introdução

A forma como o Ensino Básico é tratado no Brasil resulta de uma série de discussões que se consolidaram na Constituição Federal de 1988 (CF/88), na Lei de Diretrizes e Bases de 1996 (LDB/96) e, mais recentemente, na Medida Provisória nº 748/2016 – onde surge o chamado “Novo Ensino Médio”. Todos esses normativos buscam ampliar o acesso à escola e a gerar mecanismos que aumentem e estimulem o desejo dos alunos em continuar estudando, principalmente no Ensino Médio. Dessa forma, pretende-se dar condições para uma melhora no bem-estar do cidadão e, por consequência, no desenvolvimento econômico-social do país.

Apesar das várias conquistas relevantes nessas últimas três décadas – como a universalização do Ensino Fundamental, a queda significativa da taxa de analfabetismo e a redução da defasagem idade-série – ainda se observam graves problemas de fluxo e evasão escolar no Ensino Médio. Isso, por sua vez, gera um alto custo social, uma vez que a literatura mostra que estudantes que interrompem os estudos têm maior probabilidade de desemprego, de ter um padrão de vida baixo, de depender dos serviços de assistência pública, de ter mais problemas de saúde etc. – além de causar efeitos negativos na autoestima, na motivação e nas funções cognitivas (RUMBERGER E LIMA, 2008; DE WITTE ET AL 2013; NERI ET AL, 2015).

A expansão do número de matrículas e as melhorias de fluxo escolar no Ensino Fundamental fizeram com que o número de alunos matriculados no Ensino Médio mais que dobrasse nos últimos anos, passando de 3,8 milhões em 1990 para mais de 8 milhões em 2015 (INEP/MEC, 2015). Todavia, o número de alunos que evadem do Ensino Médio ainda é muito elevado, sendo de aproximadamente 12% atualmente. Ao analisar esse indicador por série, nota-se que o maior percentual de evasão ocorre ainda durante o primeiro ano de curso, reportando uma taxa média de evasão cerca de duas vezes maior que as ocorridas no último ano do Ensino Médio. Isso mostra que apesar das conquistas dos últimos anos, ainda há problemas que prejudicam o desenvolvimento educacional do jovem brasileiro, e que precisam ser diagnosticados e tratados de forma acurada.

Essa alta taxa de evasão seria explicada por três motivos principais inter-relacionados (MENEZES-FILHO, 2007; NERI ET AL, 2015): (i) a visão limitada dos alunos em relação aos resultados potenciais da educação sobre a renda futura e outros benefícios; (ii) a falta de interesse dos alunos quanto aos conteúdos pedagógicos lecionados – especialmente por parte dos alunos mais pobres; e, (iii) questões financeiras, como situação de desemprego dos pais, quando a continuidade dos estudos fica em segundo plano diante da necessidade de contornar os efeitos negativos de choques temporários de renda. Ademais, a literatura ressalta que os jovens de grandes regiões metropolitanas têm maiores oportunidades de trabalho, e isso combinado com uma situação socioeconômica desfavorável afeta ainda mais a evasão escolar.

Em termos de repetência no Ensino Médio, os dados do INEP indicam que houve uma melhora: de 30% no início dos anos 1990 para pouco mais de 10% em 2015. Porém, persiste as diferenças históricas entre as taxas de repetência do ano inicial em relação ao ano final. Em 1991, a média da taxa de repetência na 1ª Série foi 41% e de 12% na 3ª Série, enquanto que, em 2015, esses índices recuaram para 15,3% e 4,8%, respectivamente. Para Oliveira e Soares (2012) e Portella, Bussmann e Oliveira (2017), a manutenção de elevadas taxas de repetência é um indicador de ineficiência do fluxo escolar, decorrente tanto da exigência de requerimentos mínimos para que um estudante seja promovido de uma série para a outra quanto do seu uso como instrumento pedagógico, em que a sua combinação com a evasão representa um dos principais entraves do sistema educacional brasileiro. Sob tal problemática, os desenhos de políticas públicas educacionais se veem diante da necessidade de contemplar, ao mesmo tempo, que os alunos aprendam e que progridam em sua escolaridade.

Uma vez que se discutem medidas para aumentar o interesse e a permanência de alunos no Ensino Médio, como atividades em período integral e mudanças de currículo, identificar os fatores associados às transições entre séries – e se tais fatores apresentam o mesmo nível de significância, independentemente da série que o aluno estiver cursando – pode fornecer informações relevantes

para a formulação de políticas públicas educacionais que promovam as melhorias necessárias para atender aos anseios dos estudantes.

Nesse sentido, essa pesquisa tem por objetivo identificar as características individuais dos estudantes e de suas famílias, de ambiente escolar e de comportamento que são decisivas para o fluxo escolar ao longo do Ensino Médio, e por consequência, estimar as taxas de probabilidade de promoção serial e de continuidade nos estudos na rede pública. Para tanto, obteve-se acesso inédito e exclusivo a um rico banco de microdados da Secretaria de Estado da Educação de Santa Catarina (SEDSC), com o qual é possível acompanhar o aluno ao longo de toda sua trajetória no Ensino Médio público e que contém informações de mais de 400 mil estudantes matriculados em 649 escolas na rede pública catarinense entre os anos de 2008 a 2012. Destaca-se que é justamente o uso de dados longitudinais e de origem administrativa que traz um caráter de relevância e contribuição à literatura existente no Brasil (LEE, 2010; RIANI E ASSIS, 2017), já que há poucos estudos que tratam da temática do fluxo escolar em uma perspectiva de acompanhamento do aluno ao longo do ciclo escolar¹.

Destarte, elaboraram-se vários modelos econométricos probabilísticos bivariados, com os quais foi possível encontrar evidências de que aqueles alunos que não apresentam um rendimento escolar satisfatório inicial têm maiores chances de evadir da rede pública de ensino, e que as características familiares e individuais são as mais relevantes para a promoção serial. Além disso, notam-se indicativos de que a infraestrutura da escola pouco impacta na probabilidade de o aluno progredir de série. Em conjunto, os resultados encontrados apontam que qualquer política pública desenhada para se reduzir evasão no Ensino Médio deve ser focada nas primeiras semanas da primeira matrícula do aluno, a fim de evidenciar para ele e sua família as vantagens de continuar estudando, e identificando pontualmente quais argumentos se mostrariam contrários a isso e como a escola poderia contribuir para a permanência do estudante.

Além dessa introdução, o artigo está organizado em mais seis seções. A Seção 2 traz uma revisão da literatura para o Brasil, a Seção 3 apresenta a base de dados, a Seção 4 apresenta a metodologia econométrica, a Seção 5 traz os modelos econométricos propostos, a Seção 6 traz os resultados empíricos e a Seção 7 apresenta as conclusões.

2. Referencial teórico

Para discorrer sobre os determinantes do fluxo escolar, esta seção revisa a literatura sobre fluxo escolar e sua relação com o desempenho acadêmico dos estudantes para identificar algumas das estratégias, temas e metodologias adequadas para a sua análise. Nota-se, sobretudo, que o sucesso acadêmico depende mais das características individuais e da família, e menos de fatores relacionados ao ambiente escolar, qualificação de professores, e interação com o ambiente social no qual esse jovem está inserido.

No âmbito da literatura internacional, um dos primeiros modelos de transição serial foi desenvolvido por Mare (1980), que verificava os efeitos das características familiares no desempenho acadêmico. O autor sugeriu tratar a realização educacional como uma sequência de transições discretas das séries e níveis educacionais iniciais para os finais com base em modelos logit sequenciais. As principais vantagens desse tipo de modelo de transição educacional são: o modelo é invariante às mudanças ao longo do tempo na distribuição geral da educação; se ajusta melhor à interpretação de que a realização educacional é uma sequência de transições — uma sequência de decisões do tipo sim ou não sobre a continuidade da pessoa nos estudos; e, permite estimar os efeitos ao longo do tempo de variáveis, sejam elas de background familiar, ambiente

¹ Este tipo de acompanhamento é impossível de ser feito com os dados do Censo Escolar, PSAE ou qualquer outra base de dados pública fornecida pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) e que conduz a grande maioria dos estudos nessa área. Maiores detalhes sobre as normas de acesso à informação do INEP, em: PORTARIA Nº 467.

escolar ou comportamental na probabilidade de promoções educacionais sucessivas. Entre os principais resultados, o autor identificou uma redução brusca no poder preditivo das características familiares com o avançar dos níveis acadêmicos — de 27% da transição do ensino fundamental para o médio, para apenas 1% da transição da faculdade para a pós-graduação — e que os jovens que têm maior probabilidade de continuar os estudos após o Ensino Médio são aqueles que os pais apresentam um nível socioeconômico mais elevado em relação à média da população.

Ademais, Eckstein e Wolpin (1999) elaboraram modelos probabilísticos que buscam explicar o fluxo escolar para o Ensino Médio nos Estados Unidos, o qual seria função do histórico escolar, das características intrínsecas (habilidade, motivação) e quantidade de horas dedicadas ao trabalho para cada estudante. Nessa perspectiva, observaram que devido à diversidade de experiências e históricos familiares, é pouco provável que os jovens iniciem o ensino médio com o mesmo conjunto de preferências, habilidades, motivações em relação à escola e ao trabalho e expectativas quanto ao valor esperado do diploma. E, mesmo que haja alguma alteração dinâmica nessas individualidades, a importância do perfil do jovem no momento de entrada no Ensino Médio é grande e persistente, o que vai de encontro as suposições de Mare.

Cameron e Heckman (2001) estimaram a probabilidade de transição escolar desde o início do Ensino Médio até a entrada na universidade para homens negros, latinos e brancos entre 15 e 24 anos nos Estados Unidos através de um modelo logístico multinomial de escolha discreta utilizando como variáveis explicativas a renda familiar, escolaridade de pai e mãe, renda per capita nacional, zona residencial, entre outras. Os autores concluíram que as características familiares têm maior poder preditivo sobre negros e latinos para a continuidade dos estudos do que para brancos e que a renda familiar tem um efeito muito maior para a continuidade e diplomação no Ensino Médio do que no Ensino Superior.

Pesquisas que utilizam o modelo probabilístico ordenado, como Lucas (2001) mostram que os efeitos das características individuais e familiares são semelhantes em todas as séries do ensino médio, porém existe uma manutenção da desigualdade de oportunidades entre estudantes de classes sociais diferentes — hipótese da desigualdade efetivamente mantida — isto é, a probabilidade de conclusão do ensino médio e o acesso ao Ensino Superior diferem entre classes sociais, e esta diferença se mantém mesmo que outras variáveis sejam consideradas, como zona residencial e grade curricular, se tornando um elemento marcante no processo de reprodução de desigualdades.

Nesse contexto, Holm e Jeager (2008, 2011) analisaram as probabilidades de transição escolar em alunos do ensino médio no Reino Unido, Estados Unidos, Dinamarca e Holanda com o intuito de identificar as diferenças entre as hipóteses de desigualdade constante e a de efeitos decrescentes a partir de técnicas bivariadas *probit*. Nesse sentido, constataram que da perspectiva familiar, o seu efeito se mantém constante ao longo de toda essa etapa de ensino, com exceção da escolaridade dos pais, que tem um efeito marginalmente superior na série final do que na inicial. Do ponto de vista do aluno, a partir do resultado de exames padronizados, notaram que boas notas são condicionadas por suas características cognitivas observáveis (presença em aula e atividades extracurriculares) e aspectos familiares, além de ser um bom preditor para a progressão entre séries. Todavia, relatam que tais resultados não têm nenhum efeito direto sobre a probabilidade em cursar níveis superiores de educação.

Ademais, De Witte et al (2013) realizou uma revisão de literatura com o intuito de identificar padrões e metodologias de análise de evasão relacionados à aspectos familiares (pais ausentes, problemas de saúde etc.), escolares (projetos pedagógicos, infraestrutura, perfis de professores e turmas etc.) e ambientais (mercado de trabalho, delinquência juvenil etc.). Nessa miríade de aspectos, estimar acuradamente tais correlações é uma tarefa complicada, porque o pesquisador comumente se depara com vários problemas nas bases de dados, ou não dispõe de informações suficientemente detalhadas. Apesar dessa situação, mostram que a literatura internacional tem consistentemente apontado que maiores probabilidades de evasão estão mais frequentemente correlacionadas com: reprovação e distorção idade-série; necessidade de iniciar no mercado de trabalho; estar em uma família desestruturada ou pouco escolarizada; características depreciativas

da escola ou da vizinhança, tais como infraestruturas precárias; e, gravidez precoce, no caso das meninas.

Em uma outra revisão da literatura promovida por Banerjee (2016), as pesquisas sobre os efeitos do ambiente escolar (público ou privado, estrutura pedagógica, tamanho e localização) e progressão nos estudos, indicam que as características da escola primária exercem uma influência mais significativa sobre as transições seriais e, conseqüente, diplomação no Ensino Médio do que os fatores contextuais da escola secundária. Também destaca evidências de que aqueles estudantes que frequentaram salas de tamanho médio ao longo do Ensino Fundamental são aqueles que apresentam a maior probabilidade de continuar nos estudos. Por fim, salienta que estudantes que frequentam escolas secundárias relativamente maiores e de regiões urbanas tendem a apresentar melhores rendimentos acadêmicos do que aqueles de escolas pequenas e rurais.

No Brasil, dentre outros, Leon e Menezes-filho (2002) descrevem a evolução do fluxo escolar para as séries-diplomas² entre 1984 e 1997, a partir de um painel de dados da Pesquisa Mensal de Emprego (PME), do IBGE. Com vistas a controlar possíveis problemas de viés de seleção, os autores utilizaram o método de Heckman para calcular as probabilidades de avanço e evasão escolar e concluíram que estudantes mais pobres, com mais idade, inseridos na PEA e meninos apresentam uma maior probabilidade de reprovação e de não continuar os estudos, especialmente alunos do 9º ano do ensino fundamental. Além disso, constataram não haver diferenças estatisticamente significativas para renda, gênero e diferenças regionais na explicação da evasão escolar condicional à reprovação, que evidencia a importância das taxas de repetência na justificativa para a não progressão nos estudos.

Outrossim, Menezes-Filho e Pazello (2007) corroboram com tais resultados. Os exercícios econométricos realizados por eles, com o objetivo de identificar quais são os determinantes do rendimento escolar na disciplina de Matemática, com base em dados do Sistema de Avaliação do Ensino Básico (SAEB) de 2003, mostraram que os fatores mais relevantes para o rendimento escolar são as características que tratam do interesse do aluno nos estudos: número de reprovações, distorção idade-série, idade de entrada no ciclo básico escolar; e as familiares: cor, nível de escolaridade da mãe, presença de computadores e livros em casa. Por outro lado, as características estruturais e pedagógicas das escolas, que englobam a experiência e salário dos professores e oferta de laboratórios, biblioteca e o número de alunos por sala, se mostraram pouco significativas para a melhora do rendimento do aluno. Todavia, o tempo que o aluno permanece na escola está positivamente correlacionado com o seu resultado escolar, o que levou os autores a concluir que uma política educacional, que busque capacitar melhor o estudante, deve priorizar o aumento da carga horária, mesmo que isto exija um maior número de alunos por turma.

Por sua vez, Souza et al (2012) também utilizaram os microdados da PME aplicados a modelos lineares de probabilidade e *probit*. Nesse contexto, estimaram os determinantes dos fluxos escolares do ensino fundamental para o médio e ao longo do último, em seis regiões metropolitanas, onde constataram que mulheres têm, em média, 6% a mais de chance de serem aprovadas do que homens, a defasagem escolar reduz a probabilidade de aprovação à medida que o aluno avança no ensino médio — a repetência não garante o aprendizado, porém tem efeitos negativos em termos de autoestima e motivação — enquanto que as características do mercado local de trabalho e de ambiente escolar não apresentam relevância para a probabilidade de aprovação. Destarte, ressaltam que as variáveis mais importantes para a promoção serial estão relacionadas às características do próprio estudante e de sua família.

Oliveira e Soares (2012) estimaram os determinantes da repetência a partir de uma análise em painel dos Censos Escolares dos anos de 2007 a 2010 para alunos do 1º ao 5º ano e do 6º ao 9º ano, que permite o acompanhamento do aluno ao longo do tempo. Então, com a técnica logit, mensuraram a probabilidade de repetência anual a partir de variáveis que retratam as características do aluno, da escola, dos professores e das turmas. Entre as diversas conclusões, destaca-se a maior probabilidade de repetência dos alunos do sexo masculino, portadores de necessidades especiais e

² 4ª e 8ª séries (ou 5º e 9º ano) do ensino fundamental e 3ª série do ensino médio.

que utilizam transporte público escolar, enquanto que alunos matriculados em escolas com melhor nível de infraestrutura e com maior carga horária de aula, possuem menores chances de repetência. Em relação à formação dos professores, verificou-se que professores mais especializados reprovam menos e que uma redução do número de alunos por sala não é significativa para o fluxo escolar.

Shirasu e Arraes (2015) buscaram identificar as causas da evasão e repetência escolar no ensino médio público, a partir de uma base de dados longitudinal, para o período de 2009 a 2011, obtida junto ao Sistema Permanente de Avaliação da Educação Básica do Ceará (SPAECE) utilizando o modelo *logit* multinível. Os autores concluíram que distorção idade-série, ser homem — mais predispostos a ingressar no mercado de trabalho — e quanto menor o grau de escolaridade dos pais, maior a probabilidade de repetência e evasão. Já em relação às características da escola, estas se mostraram pouco significativas, o que sugere que as políticas públicas educacionais devem ter como foco o aluno por meio do monitoramento da sua vida escolar.

Riani e Assis (2017) usaram dados da Pesquisa por Amostra de Domicílios de Minas Gerais (PAD-MG) para estimar as taxas de fluxo no ensino médio, entre os anos de 2012 e 2013, e os principais determinantes da promoção serial através de um modelo de regressão logística que controla fatores individuais e de *background* familiar. Destaca-se que o estudo corrobora com as evidências da literatura quanto ao melhor desempenho escolar das mulheres, ao estimar uma taxa de probabilidade de promoção, ao fim da 1ª série, 25% superior à taxa estimada de probabilidade de promoção dos meninos e que quanto maior a distorção idade-série, menores as chances de aprovação. Além disso, identificaram que com o avançar das séries, o efeito-diploma se sobrepõem aos demais controles, dado que estudantes que não trabalham têm 140% a mais de chances de serem aprovados do que aqueles que trabalham, porém essa situação não ocorre com a 3ª série. Ademais, identificaram também que estudantes residentes na zona rural tem um melhor rendimento escolar do que aqueles que vivem na área urbana.

Com o intuito de identificar os principais fatores relacionados à defasagem idade-série e o desempenho dos alunos, Portella, Bussmann e Oliveira (2017), por sua vez, com dados da PNAD de 2013, para estudantes entre 8 e 17 anos matriculados em escolas públicas e modelos de contagem, concluíram que ser mulher, morar com a mãe e em residências com maior poder aquisitivo estão associados a uma menor distorção idade-série. Quanto ao efeito-escola, notou-se que essa situação está relacionada à docentes mais escolarizados e uma carga horária maior, entretanto, o tamanho da sala revelou-se insignificante. Por fim, ressaltam que as variáveis referentes a trabalho, mostraram que aqueles estudantes que realizam alguma atividade profissional estão associados a um menor grau de distorção idade-série, que pode ser justificado por aspectos motivacionais de ter emprego ao final do ciclo básico de ensino.

Machado, Firpo e Gonzaga (2013), fazendo uso dos dados do SAEB de 2011 e do método de painel de efeitos fixos, notaram que quanto maior a dispersão em termos de idade em sala de aula, maiores as taxas de repetência, e, por consequência, maior a probabilidade de o jovem evadir. Corroborando com esses resultados, Junior e Stein (2017) verificaram que quanto maior a heterogeneidade etária, menor o aprendizado escolar e nem mesmo as qualidades técnicas dos docentes conseguem reverter este cenário.

Outro fator apontado na literatura como importante para o desempenho escolar são os programas de transferência de renda. Então, com o objetivo de identificar os efeitos desse tipo de política sobre a permanência escolar, Cacciamali, Tatei e Batista (2010) investigaram os efeitos do programa Bolsa Família sobre a probabilidade conjunta de trabalho e frequência escolar em crianças de 7 a 15 anos em situação socioeconômica adversa, a partir de dados da PNAD de 2004. Os testes econométricos foram realizados com modelos *probit* bivariados e identificaram que raça e gênero não tem poder preditivo sobre a frequência escolar, todavia ter pais alocados em trabalho informal e quanto maior o tamanho da família, maior a probabilidade de a criança trabalhar e não frequentar a escola, sobretudo no meio rural. Por fim, concluíram que apesar do programa de transferência de renda ser capaz de aumentar a frequência escolar não reduz a incidência de trabalho infantil. Nesta mesma linha de raciocínio, Ferro e Kassouf (2005) também investigaram os efeitos desse programa sobre o trabalho infantil, a partir de informações da PNAD de 2001. Para essa

análise, fizeram uso de um modelo probit ponderado e de mínimos quadrados ponderados e verificaram que devido à exigência de frequência escolar, isso por si só reduz a probabilidade de trabalho infantil, principalmente no meio rural, porém, os resultados são inconclusivos quanto à decisão dos pais em demandar trabalho de seus filhos.

Nos estudos que tratam do fluxo escolar, além da investigação sobre os fatores que o explicam, há um debate sobre a cultura da pedagogia da repetência e seus efeitos sobre a trajetória escolar do aluno e a potencial evasão. Ferrão, Beltrão e Santos (2002), a partir de dados do SAEB de 1999 aplicados a modelos de regressão hierárquicos, estimaram o impacto de políticas de não-repetência no desempenho escolar de alunos do 5º ano do ensino fundamental da região Sudeste. Entre os principais resultados, verifica-se que a interação entre nível socioeconômico com a política de promoção automática apresentou uma relação positiva com o rendimento escolar e que a perda de proficiência de alunos com defasagem idade-série se acentua com o número de reprovações. Por fim, concluíram que o rendimento escolar do aluno com maior probabilidade de repetir tende a melhorar caso seja promovido do que se fosse retido.

Sob essa visão, Riani, Silva e Soares (2012), através dos dados do PAD-MG, avaliam o *trade-off* entre repetir ou progredir no rendimento escolar de alunos do 4º ano do ensino fundamental entre 2008 e 2009. Para tal, também recorreram a técnicas de modelos hierárquicos e concluíram que aqueles os alunos reprovados apresentaram uma evolução com rendimento menor do que aqueles alunos que progredirem de série e que a baixa proficiência dos alunos reprovados é diretamente afetada pelos resultados da escola.

Em suma, os trabalhos apresentados nesta seção mostram que os problemas relacionados ao Ensino Médio são preocupantes e que a literatura da economia da educação enfrenta diversas dificuldades em suas pesquisas, como especificação de funções, a falta de qualidade dos dados e problemas teóricos e metodológicos.

3. A base de dados

As informações utilizadas nesta pesquisa provêm da Secretaria de Educação de Santa Catarina (SEDSC) e abrangem o período de 2008 a 2012, que trata exclusivamente do ensino médio regular³ da rede pública. A base de dados contém informações dos 428.622 alunos matriculados em 649 escolas em Santa Catarina, compostas majoritariamente por alunos do gênero feminino (53,7%), da raça branca (34,5%) e que moram no meio urbano (81,2%).

Na literatura de desempenho acadêmico, nota-se que um vetor de variáveis-chave é recorrentemente analisado, o qual é passível de observação na base SEDSC e representa os impactos dos segmentos familiar, individual e escolar no rendimento escolar. O ambiente familiar de tal vetor é representado por: (i) zona de residência; (ii) pela *proxy* de renda familiar⁴; e, (iii) bolsa - e.g., participam ou não de algum programa de transferência direta de renda (Bolsa Família). Por sua vez, o individual é composto por: (i) gênero; (ii) cor/raça; (iii) número de faltas cometidas no ano; (iv) transferência escolar; e, (v) repetência de série. E finalmente, o ambiente escolar: (i) número de funcionários (*proxy* para tamanho relativo de escola); (ii) idade média do professor (*proxy* para experiência profissional); (iii) existência de biblioteca; (iv) laboratório de informática; e, (v) quadra coberta. A Tabela (1), a seguir, traz as estatísticas descritivas para várias destas características:

³ Desconsidera-se os regimes de educação profissional, jovens e adultos e outros sistemas diferenciados.

⁴ Produto Interno Bruto Per capita Municipal obtido junto ao IBGE.

Tabela 1: Estatísticas descritivas

Características Individuais / Familiares			
Variável	Classificação	Total (mil)	%
Gênero	Feminino	230	53,7
	Masculino	198	46,3
Cor/ Raça	Branca	147	34,5
	Preta	3	0,8
	Parda	13	3,1
	Não declarado	264	61,6
Zona Residencial	Rural	83	18,9
	Urbana	345	81,1
Bolsa Família	Sim	5	1,3
	Não	423	98,8
Transferência Escolar*	Total	25	5,8
	Feminino	10	40
	Masculino	15	60
Alunos com pelo menos 1 reprovação	Total	27	6,3
	Feminino	10	37
	Masculino	17	63
Total de Matrículas		428	
Características Escolares			
Variável	Classificação	Total	%
Biblioteca	Sim	521	80,2
	Não	128	19,8
Laboratório Informática	Sim	578	89,1
	Não	71	10,9
Quantidade Média de Funcionários	45,89 (23,94)		
Quantidade Média Alunos/sala	26,05 (6,52)		
Idade Média Professor	40,32 (2,85)		
Total de Escolas	649		

Fonte: SEDSC. Formatação própria.

Em relação ao nível socioeconômico do aluno, a base de dados não possibilita extrair a renda familiar anual, mas permite, por exemplo, identificar se a família está inserida em programa de transferência de renda do governo federal (1,32%); além de identificar que as famílias moram majoritariamente em área urbana (80,73%). No que diz respeito às características estruturais das escolas, a grande maioria apresenta biblioteca, laboratório de informática e quadra coberta; e cerca de 80% das escolas catarinenses ofertam todo o ciclo básico.

Explorando as informações fornecidas pela base de dados, para aqueles alunos que permaneceram na rede pública de ensino, foram criadas duas variáveis⁵: uma de fluxo escolar, que trata da promoção da 1ª Série para 2ª Série e da 2ª Série para a 3ª Série (*promo*), e uma de permanência na escola pública (*per*). Contudo, a informação referente aos alunos que cursam a 3ª Série é dúbia, pois não há como inferir a condição do aluno ao final do ano (aprovação, reprovação, migração para sistema privado ou abandono, diplomação), o que impossibilita o seu uso. A Tabela (2), a seguir, ilustra esse procedimento ao apresentar indicadores de fluxo escolar e latência por série e por ano para o período analisado.

⁵ Tanto as taxas de transição quanto a taxa de latência são definidas a partir de uma interpretação *backward-looking* para cada aluno, similar a metodologia aplicada em LEON E MENEZES-FILHO (2002).

Tabela 2: Transição Escolar: Quantidade total de alunos. 2008 a 2011

Ano	Série	Situação			
		Repetência	Promoção	Latência	Total
2008	1	3.996	39.518	15.700	59.214
	2	1.484	38.267	7.941	47.692
	3	624		45.972	46.596
	Total	6.104	77.785	69.613	155.502
2009	1	5.489	42.105	12.936	60.530
	2	1.974	39.051	7.133	48.158
	3	565		44.844	45.409
	Total	8.028	81.156	64.913	154.097
2010	1	5.305	38.042	24.178	67.525
	2	2.094	34.045	16.484	52.623
	3	438		45.988	46.426
	Total	7.837	72.087	86.650	166.547
2011	1	2.937	36.451	11.291	50.679
	2	1.304	35.248	6.405	42.957
	3	321		36.480	36.801
	Total	4.562	71.699	54.176	130437

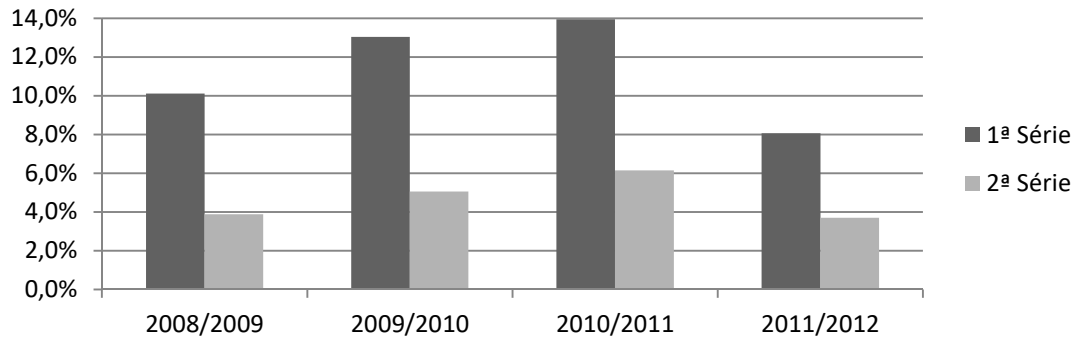
Fonte: SEDSC. Formatação própria.

A coluna latência, na Tabela (2) acima, nos indica a quantidade de alunos que por algum motivo qualquer não estão mais ativos na base de dados da SEDSC. Por exemplo, no ano de 2008, dos alunos que frequentaram a 1ª série, ocorreram 3996 repetências e 39518 aprovações, enquanto 15.700 estão em estado de latência, isto é, ao final do ano cursado, nada podemos inferir sobre reprovação, aprovação, evasão, migração para a rede privada ou mudança de Estado, apenas que não mais continuam os estudos no ano seguinte na rede pública.

Observe também, a evolução temporal da quantidade de alunos matriculados em cada série no Ensino Médio. Note, em especial, que em 2010 o número de matrículas no Ensino Médio regular sofreu um aumento de mais de 10%, porém, para todo o período em análise, todas as séries seguintes apresentaram redução no número de alunos, com ênfase para a 3ª série, que obteve uma queda de cerca de 25% no número de alunos matriculados no período. Observa-se ainda que, entre a 1ª e a 2ª série, o número de alunos que não continuam os estudos no sistema público foi de 17%, com ênfase em 2010, onde essa taxa atingiu 22%. Entretanto, na progressão da 2ª para a 3ª série, essa taxa se reduz à metade, o que pode indicar a sobreposição do efeito diploma sobre qualquer outra circunstância na decisão do aluno prosseguir na escola até a conclusão do ciclo escolar.

Em termos de fluxo escolar, a análise conjunta das Tabelas (1) e (2) e do Gráfico (1), abaixo, condicional à permanência do aluno na escola pública, indicam que ocorreu uma redução do número total de reprovados e que está concentrada na transição da 1ª para a 2ª Série, que representa 66,2% dos casos ocorridos no período analisado, enquanto que a repetência ao final da 2ª Série atinge um valor três vezes menor.

Gráfico 1: Evolução temporal da taxa de repetência



Fonte: SEDSC. Formatação própria.

A situação de risco para os alunos da 1ª Série fica ainda mais evidente ao se examinar os dados para latência, pois os seus números são, em média, quase 90% superiores aos verificados para a 2ª Série. Esses indicadores mostram que, qualquer política pública que venha a impactar no Ensino Médio, deve ter como foco o ano inicial do Ensino Médio, pois parece ser neste momento que a decisão daqueles alunos que apresentam algum tipo de carência em continuar os estudos no sistema público de ensino será tomada.

4. Metodologia Econométrica

Em sintonia com os modelos teóricos de rendimento e evasão escolar, esta seção descreve o modelo de regressão bivariado probabilístico (*biprobit*), que foi utilizado para analisar um conjunto de variáveis que condicionam simultaneamente a probabilidade de um aluno do Ensino Médio da rede pública de Santa Catarina em progredir de série e permanecer na rede de ensino. Os principais motivos que levam ao uso dessa abordagem são: a sua capacidade em modelar probabilidades sem a necessidade de impor qualquer especificação para a tomada de decisão e sua estrutura de estimação que controla para vies de seleção de escolha binária (CACCIAMALI, TATEI E BATISTA, 2010; HOLM E JAEGER, 2011; SASS ET AL, 2016).

Em termos econométricos, a técnica biprobit resulta da expansão do modelo univariados probit a um sistema de 2 equações, tal que (CAMERON E TRIVEDI, 2005):

$$y_{i1}^* = \mathbf{x}'_{i1}\beta_1 + \varepsilon_{i1}, \quad y_{i1} = 1 \text{ se } y_{i1}^* > 0, = 0 \text{ caso contrário;} \quad (1.1)$$

$$y_{i2}^* = \mathbf{x}'_{i2}\beta_2 + \varepsilon_{i2}, \quad y_{i2} = 1 \text{ se } y_{i2}^* > 0, = 0 \text{ caso contrário;} \quad (1.2)$$

em que os vetores \mathbf{x}_1 e \mathbf{x}_2 contêm as variáveis explicativas relacionadas à probabilidade de um agente i qualquer (aluno, firma, trabalhador) incorrer em algum evento, as variáveis latentes contínuas y_{i1}^* e y_{i2}^* são definidas pelos resultados observados y_{i1} e y_{i2} .

Para o cálculo das probabilidades conjuntas (aprovação e permanência), deve-se ter em mente o modelo probit univariado para cada uma das variáveis aleatórias dicotômicas y_{ij} que assumem valores 0 e 1 correspondentes a cada observação i e equação j . Por exemplo, para $Pr[y_{i1} = 1]$ (aluno ser aprovado), tem-se:

$$\begin{aligned} Pr[y_{i1} = 1] &= Pr[y_{i1}^* > 0] = Pr[\mathbf{x}'_{i1}\beta_1 + \varepsilon_{i1} > 0] \\ Pr[\varepsilon_{i1} > -\mathbf{x}'_{i1}\beta_1] &= Pr[\varepsilon_{i1} < \mathbf{x}'_{i1}\beta_1] = \Phi(\mathbf{x}'_{i1}\beta_1), \end{aligned} \quad (2)$$

onde $\Phi(\cdot)$ representa a função de distribuição normal padrão acumulada (CDF). Destarte, a distribuição conjunta de y_1 e y_2 condicional em \mathbf{x} , $Pr(Y_1 = y_{i1}, Y_2 = y_{i2} | \mathbf{x})$, a qual por conveniência notacional abreviamos para $Pr^{y_{i1}y_{i2}}$ poderá ter até quatro elementos:

$$\begin{aligned}
P11 &= \Pr(\varepsilon_{i1} < \mathbf{x}'_{i1}\beta_1, \varepsilon_{i2} < \mathbf{x}'_{i2}\beta_2) \\
P10 &= \Pr(\varepsilon_{i1} < \mathbf{x}'_{i1}\beta_1, \varepsilon_{i2} > \mathbf{x}'_{i2}\beta_2) \\
P01 &= \Pr(\varepsilon_{i1} > \mathbf{x}'_{i1}\beta_1, \varepsilon_{i2} < \mathbf{x}'_{i2}\beta_2) \\
P00 &= \Pr(\varepsilon_{i1} > \mathbf{x}'_{i1}\beta_1, \varepsilon_{i2} > \mathbf{x}'_{i2}\beta_2)
\end{aligned} \tag{3}$$

Aqui vale uma observação pertinente. Dada às características da base de dados, os cálculos possíveis nesta pesquisa se referem a $P11$ e $P01$, ou seja, as probabilidades conjuntas de um estudante aprovado permanecer na rede pública de ensino e de um estudante reprovado não permanecer (evadir).

Quanto aos termos de erro ε_1 e ε_2 , estes são descritos por:

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \Big| \mathbf{x}_1, \mathbf{x}_2 \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right], \tag{4}$$

que, por hipótese, seguem uma distribuição normal bivariada com média zero, variância unitária e coeficiente de correlação ρ (no caso desta pesquisa, a correlação entre aprovação e permanência escolar), sendo este o termo que conecta as equações 1.1 e 1.2.

As especificações (11), (1.2) e (3) juntas com a hipótese (4) é comumente conhecida como modelo bivariado probabilístico recursivo, o qual foi utilizado nesta pesquisa. Então, a função de distribuição conjunta de ε_1 e ε_2 que define esse modelo será $\Phi_2(\varepsilon_1, \varepsilon_2, \rho)$, que descreve a função de distribuição acumulada de uma distribuição normal padrão bivariada com coeficiente de correlação ρ . Neste caso, a probabilidade conjunta entre y_{i1} e y_{i2} pode ser expressa de forma compacta como:

$$\Pr^{y_{i1}y_{i2}}(\theta) = \Phi_2(q_1, q_2, \rho^*), \tag{5}$$

onde $q_1 = (2y_{i1} - 1)(\mathbf{x}'_{i1}\beta_1)$, $q_2 = (2y_{i2} - 1)(\mathbf{x}'_{i2}\beta_2)$, $\rho^* = q_1q_2\rho$, $\Pr^{y_{i1}y_{i2}}(\theta)$ descreve as probabilidades apresentadas em (3) estimadas no ponto $(y_{i1}, y_{i2}, \mathbf{x}'_{i1}, \mathbf{x}'_{i2})$ e o subscrito i , a unidade (aluno) observada na amostra. Os parâmetros para o modelo, assim como as probabilidades conjuntas serão estimadas a partir da função log-verossimilhança.

5. Modelo empírico

Em sintonia com a literatura de análise de fluxo escolar, esta seção descreve os modelos de regressão bivariados probabilísticos especificados, os quais utilizaram uma miríade de variáveis que condicionam, simultaneamente, a probabilidade de promoção serial⁶ de um aluno do ensino médio da rede pública de Santa Catarina/ e permanência na escola pública, para as transições do 1º para o 2º ano e do 2º para o 3º ano. A seguir, apresentam-se as variáveis utilizadas neste trabalho:

⁶ A base de dados fornecida pela SED-SC não nos permite ter certeza quanto à aprovação do aluno ao final do último ano do Ensino Médio. A base não possui nenhum indicador de graduação. Por esta razão, as informações referentes ao 3º ano não foram utilizadas nas estimações.

Tabela 4: Descrição das variáveis extraídas da SEDSC

Variáveis	Descrição
<i>Variáveis dependentes</i>	
aprov	Variável binária, na qual possui valor 1 se o estudante é aprovado ao final da 1º série e 0 caso contrário;
per	Variável binária, na qual possui valor 1 se estudante permanece na rede pública de ensino ao fim da 1º série e 0 caso contrário;
<i>Características individuais e familiares</i>	
menino	Variável binária, na qual possui valor 1 se o estudante é menino e 0 caso contrário;
branco	Variável binária, na qual possui valor 1 se o estudante se autodeclara de cor branca e 0 caso contrário;
Rural	Variável binária, na qual possui valor 1 se o estudante reside na zona rural e 0 caso contrário;
Bolsa	Variável binária, na qual possui valor 1 se a família do estudante recebe Bolsa Família (BF) e 0 caso contrário;
computador	Variável binária, na qual possui valor 1 se a família do estudante tem computador em casa e 0 caso contrário;
lnpib_pc	Logaritmo natural do PIB per capita por município;
<i>Características do ensino local</i>	
e_biblio	Variável binária, na qual possui valor 1 se a escola tem biblioteca e 0 caso contrário;
e_lab_info	Variável binária, na qual possui valor 1 se a escola tem laboratório de informática e 0 caso contrário;
i_funcionarios_razao	Razão entre número de estudantes e funcionários por escola;
prof_idade	Idade média dos professores por escola;
<i>Características de comportamento escolar individual</i>	
transfer	Variável binária, na qual possui valor 1 se a aluno mudou de escola ao longo do Ensino Médio e 0 caso contrário;
alunos_turma	Quantidade de alunos na turma do aluno;
faltas	Número de faltas anuais por aluno.

Fonte: Elaboração própria.

Para tanto, foram propostas três especificações para os modelos probabilísticos bivariados (*biprobit*) que têm as variáveis binárias *aprov* e *per* como dependentes. O primeiro modelo limita as variáveis explicativas as características individuais e familiares; o segundo adiciona aspectos estruturais e pedagógicos das escolas, que são diretamente afetados pela ação do Estado; e, por fim, o terceiro modelo inclui, além das características individuais e escolares, particularidades da vida acadêmica do jovem. Seguem as especificações *biprobit* em que o subscrito *i* refere-se ao aluno e *k* a escola:

Especificações *biprobit*:

M1

$$\begin{aligned}
 aprov_{ik} &= \alpha + \beta_1 menino_i + \beta_2 branco_i + \beta_3 rural_i + \beta_4 bolsa_i + \\
 &\quad \beta_5 computador_i + \beta_6 \ln pib_pc_i + \varepsilon_{ik} \\
 per_{ik} &= \alpha + \beta_1 menino_i + \beta_2 branco_i + \beta_3 rural_i + \beta_4 bolsa_i + \\
 &\quad \beta_5 computador_i + \beta_6 \ln pib_pc_i + \varepsilon_{ik}
 \end{aligned}$$

$$aprov_{ik} = \alpha + \beta_1 menino_i + \beta_2 branco_i + \beta_3 rural_i + \beta_4 bolsa_i + \beta_5 computador_i + \beta_6 \ln pib_pc_i + \beta_7 e_biblio_k + \beta_8 e_lab_info_k + \beta_9 i_funcionarios_razaok + \beta_{10} prof_idade_k + \varepsilon_{ik}$$

M2

$$per_{ik} = \alpha + \beta_1 menino_i + \beta_2 branco_i + \beta_3 rural_i + \beta_4 bolsa_i + \beta_5 computador_i + \beta_6 \ln pib_pc_i + \beta_7 e_biblio_k + \beta_8 e_lab_info_k + \beta_9 i_funcionarios_razaok + \beta_{10} prof_idade_k + \varepsilon_{ik}$$

$$aprov_{ik} = \alpha + \beta_1 menino_i + \beta_2 branco_i + \beta_3 rural_i + \beta_4 bolsa_i + \beta_5 computador_i + \beta_6 \ln pib_pc_i + \beta_7 e_biblio_k + \beta_8 e_lab_info_k + \beta_9 i_funcionarios_razaok + \beta_{10} prof_idade_k + \beta_{11} transfer_i + \beta_{12} faltas_i + \beta_{13} alunos_turma_{ik} + \varepsilon_{ik}$$

M3

$$per_{ik} = \alpha + \beta_1 menino_i + \beta_2 branco_i + \beta_3 rural_i + \beta_4 bolsa_i + \beta_5 computador_i + \beta_6 \ln pib_pc_i + \beta_7 e_biblio_k + \beta_8 e_lab_info_k + \beta_9 i_funcionarios_razaok + \beta_{10} prof_idade_k + \beta_{11} transfer_i + \beta_{12} faltas_i + \beta_{13} alunos_turma_{ik} + \varepsilon_{ik}$$

6. Resultados

As Tabelas (5), (6) e (7), a seguir, apresentam os resultados estimados para os três modelos propostos para os dois períodos de transição. Nota-se que, de modo geral, as características individuais, familiares, de estrutura do colégio e de comportamento escolar influenciam o desempenho acadêmico do jovem da rede pública, corroborando com os achados apresentados pela literatura da economia da educação.

Os resultados para o modelo M1, que trata exclusivamente do contexto familiar, estão dispostos na Tabela (5), demonstram que os estudantes do sexo masculino apresentam uma probabilidade menor de serem aprovados e de não evadir com relação às mulheres, porém essa discrepância é decrescente ao longo do ensino médio — os valores estimados para a 1ª Série são semelhantes aos obtidos por Riani e Assis (2017) nos três modelos. Um ano a mais de estudo está associado a uma queda de cerca de 50% em ambos os indicadores para os meninos. Quanto a localidade, os jovens residentes da zona rural aparentam ter um melhor rendimento escolar do que aqueles que moram nas cidades.

A atratividade do mercado de trabalho local, medida pelo PIB *per capita* do município, apresenta alta associação com a probabilidade de não aprovação e de evasão escolar ao longo de todo o ensino médio, isto é, quanto maior o PIB do município, maior serão as chances de repetência e evasão. Isso pode ser justificado pelo fato de cidades mais desenvolvidas economicamente terem maior oferta de trabalho e remunerações mais elevadas, além do próprio ambiente familiar que demanda trabalho por parte do jovem. Esse resultado vai de encontro ao apresentado em Souza et al (2012) e Banerjee (2016), porém corrobora com o obtido em Leon e Menezes-filho (2002), De Witte (2013), Neri et al (2015) e Shirasu e Arraes (2015).

Quanto ao recebimento de Bolsa Família (BF), os resultados se mostraram inconclusivos, apresentando divergência ao longo das transições escolares. Porém, essa análise deve ser vistas com ressalvas, pois o número de estudantes que declaram que a família está inscrita no programa é bastante reduzido, receber BF não implica em não trabalhar e as características do público-alvo do programa podem promover erros de estimação.

Tabela 5: Modelo M1: Resultados estimados

Variáveis	Transição 1 ^a - 2 ^a		Transição 2 ^a - 3 ^a	
	aprov	per	aprov	per
menino	-0,2256*** (-0,0046)	-0,1436*** -0,0047	-0,1175*** (0,0052)	-0,0713*** (0,0052)
branco	0,2316*** (-0,0109)	0,1826*** -0,0109	0,2463*** (0,0132)	0,2250*** (0,0132)
rural	0,2462*** (-0,006)	0,2113*** -0,0061	0,0480*** (0,0065)	0,0111* (0,0065)
bolsa	-0,0492** (-0,0197)	-0,0314 -0,0198	0,1369*** (0,0236)	0,1332*** (0,0238)
computador	0,1637*** (-0,0052)	0,1607*** -0,0053	0,0467*** (0,0057)	0,0475*** (0,0058)
lnpib_pc	-0,3018*** (-0,0058)	-0,3249*** -0,0058	-0,3690*** (0,0064)	-0,3875*** (0,0065)
Constant	2,9246*** (-0,0602)	3,3273*** -0,0606	3,7922*** (0,0669)	4,0614*** (0,0673)
ρ	0,9999		1	

Fonte: Elaboração própria.

Desvio-padrão: valores entre parênteses; *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

No contexto da estrutura física da escola, ao contrário de Banerjee (2016) e Portella, Bussmann e Oliveira (2017), as ofertas de biblioteca e de laboratório de informática, o tamanho relativo da escola e a experiência de professor apresentaram pouca significância na vida acadêmica de um jovem secundarista, e mais, esse efeito foi praticamente nulo na transição da 2^a para a 3^a série. A não relevância dessas variáveis pode ser explicada pela caracterização das escolas em análise — vide Tabela (1) — que têm a mesma dependência administrativa (Estado) e são semelhantes em termos de estrutura.

Tabela 6: Modelo M2: Resultados estimados

Variáveis	Transição 1 ^a - 2 ^a		Transição 2 ^a - 3 ^a	
	aprov	per	aprov	per
menino	-0,2274*** -0,0046	-0,1446*** -0,0047	-0,1183*** (0,0052)	-0,0725*** (0,0052)
branco	0,2335*** -0,0109	0,1840*** -0,0109	0,2459*** (0,0132)	0,2253*** (0,0132)
rural	0,2146*** -0,0062	0,1819*** -0,0062	0,0135** (0,0066)	-0,0177*** (0,0067)
bolsa	-0,0682*** -0,0197	-0,0457** -0,0199	0,1218*** (0,0236)	0,1224*** (0,0238)
computador	0,1760*** -0,0053	0,1696*** -0,0053	0,0594*** (0,0058)	0,0568*** (0,0058)
lnpib_pc	-0,2615*** -0,006	-0,2911*** -0,006	-0,3303*** (0,0066)	-0,3516*** (0,0066)
e_biblio	0,0509*** -0,0058	0,0486*** -0,0058	0,0760*** (0,0064)	0,0759*** (0,0064)
e_lab_info	0,0485*** -0,0071	0,0306*** -0,0071	0,0152* (0,0079)	0,0141* (0,0079)
i_funcionarios_razao	-0,0117*** -0,0005	-0,0089*** -0,0005	-0,0094*** (0,0006)	-0,0068*** (0,0006)
prof_idade	-0,0145*** -0,0009	-0,0141*** -0,0009	-0,0187*** (0,0010)	-0,0176*** (0,0010)
Constant	3,1115*** -0,0676	3,5614*** -0,0681	4,1614*** (0,0751)	4,3937*** (0,0755)
ρ	0,999		0,9998***	

Fonte: Elaboração própria.

Desvio-padrão: valores entre parênteses. *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Por fim, a Tabela (7) traz os resultados obtidos para o modelo M3, o qual inclui as características de comportamento escolar do estudante, ou seja, o interesse do aluno. Os resultados estimados indicam uma situação muito desfavorável para alunos que tiveram que se transferir de colégio, indicando que eles apresentam uma probabilidade de reprovação e de evasão, em média, 50% superior ao final da 1ª série em relação aqueles que cursaram todo o ensino médio no mesmo local, evidenciando a necessidade de um maior cuidado com esse grupo por parte dos agentes escolares. Em relação ao número de faltas, apesar desta estar relacionada a uma melhora do desempenho acadêmico e ser um indicativo de interesse do aluno, seus efeitos apesar de significantes foram pequenos.

No que se refere ao tamanho médio da sala, corroborando com Oliveira e Soares (2012) e Portella, Bussmann e Oliveira (2017), as estimativas indicam uma associação positiva entre o impacto da “enturmação” e as probabilidades de aprovação e continuidade dos estudos, porém de impacto reduzido.

Tabela 7: Modelo M3: Resultados estimados

Variáveis	Transição 1ª - 2ª		Transição 2ª - 3ª	
	aprov	per	aprov	per
menino	-0,2220*** -0,0046	-0,1387*** -0,0047	-0,1201*** (0,0052)	-0,0743*** (0,0052)
branco	0,2255*** -0,0109	0,1763*** -0,0109	0,2395*** (0,0132)	0,2188*** (0,0132)
rural	0,2153*** -0,0062	0,1827*** -0,0062	0,0137** (0,0067)	-0,0170** (0,0067)
bolsa	-0,0654*** -0,0197	-0,0432** -0,0199	0,1196*** (0,0236)	0,1201*** (0,0238)
computador	0,1698*** -0,0053	0,1634*** -0,0053	0,0595*** (0,0058)	0,0568*** (0,0058)
lnpib_pc	-0,2713*** -0,006	-0,3015*** -0,006	-0,3334*** (0,0066)	-0,3554*** (0,0067)
e_biblio	0,0510*** -0,0058	0,0486*** -0,0058	0,0762*** (0,0064)	0,0759*** (0,0064)
e_lab_info	0,0507*** -0,0071	0,0328*** -0,0071	0,0147* (0,0079)	0,0140* (0,0079)
i_funcionarios_razao	-0,0159*** -0,0006	-0,0131*** -0,0006	-0,0123*** (0,0006)	-0,0100*** (0,0006)
prof_idade	-0,0142*** -0,0009	-0,0139*** -0,0009	-0,0186*** (0,0010)	-0,0175*** (0,0010)
transfer	-0,5049*** -0,0255	-0,4923*** -0,0248	-0,2432*** (0,0119)	-0,2360*** (0,0119)
faltas_completas	-0,0236*** -0,0013	-0,0206*** -0,0012	-0,0211*** (0,0014)	-0,0199*** (0,0014)
alunos_turma	0,0127*** -0,0004	0,0128*** -0,0004	0,0076*** (0,0004)	0,0082*** (0,0004)
Constant	2,9178*** -0,0679	3,3738*** -0,0684	4,0478*** (0,0753)	4,2744*** (0,0758)
ρ	0,999		0,9999***	

Fonte: Elaboração própria.

Desvio-padrão: valores entre parênteses. *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$

Em relação ao parâmetro de correlação ρ entre aprovação e permanência, nota-se que o valor estimado é próximo de +1, o que significa que qualquer fator não-observável que estimule a continuidade dos estudos terá um efeito positivo sobre o rendimento acadêmico do estudante. Para verificar a sua significância estatística, realizou-se o teste LR no qual a hipótese nula de $\rho = 0$ foi rejeitada. Conclui-se, então, que aprovação e permanência são diretamente dependentes e as

equações devem ser estimadas conjuntamente em um modelo biprobit, ao invés de equações probit separadas.

No que se refere as probabilidades conjuntas preditas, a Tabela (8), abaixo, apresenta os resultados para duas situações exploradas neste trabalho: a primeira, $P01$, expressa a probabilidade do aluno reprovar e permanecer na rede pública; e a segunda, $P11$, indica a probabilidade do aluno ser aprovado e continuar os estudos na escola pública. Os resultados para o Modelo M1, que trata exclusivamente das características familiares, mostram que a probabilidade de um aluno reprovado permanecer na rede pública de ensino é bastante reduzida, atingindo um máximo de 5,4% para a transição da 2ª para a 3ª Série. Note que essa situação também ocorre nos Modelos M2 e M3, o que sinaliza que para aqueles alunos que reprovam, as características da escola e do comportamento individual são praticamente irrelevantes na decisão do aluno em evadir. Esses resultados corroboram com as evidências identificadas por De Witte (2013), Machado, Firpo e Gonzaga (2013) e Shirasu e Arraes (2015) ao trazer indícios de que aqueles alunos com maior probabilidade de evasão são os que reprovam e de que a decisão em não prosseguir nos estudos é predita majoritariamente pelos aspectos familiares.

Tabela 8: Probabilidades conjuntas preditas:

Transição: 1ª para 2ª Série				
Especificação	Cenário	Média	Mínimo	Máximo
M1	P01	0,0588	0,0192	0,1152
	P11	0,5254	0,2139	0,7915
M2	P01	0,0588	0,0118	0,1186
	P11	0,5254	0,1644	0,8033
M3	P01	0,0588	0,0110	0,1233
	P11	0,5203	0,0358	0,8367
Transição 2ª para 3ª Série				
Especificação	Cenário	Média	Mínimo	Máximo
M1	P01	0,0283	0,0000	0,05429
	P11	0,5949	0,2761	0,8137
M2	P01	0,0281	0,0001	0,0707
	P11	0,5948	0,2599	0,8245
M3	P01	0,0282	0,0000	0,07375
	P11	0,5981	0,0956	0,8435

Fonte: Formatação própria.

Os resultados apresentados também indicam que a aprovação é importante para explicar a continuidade dos estudos ao longo do ensino médio. Verifica-se um aumento na probabilidade média destes estudantes em concluir o ensino médio com o avançar das séries, que pode ser uma evidência da sobreposição do efeito-diploma sobre os demais controles, conforme expresso em Soares et al (2012) e Riani e Assis (2017). Nesse sentido, as políticas públicas voltadas para a manutenção do jovem na escola, devem em um primeiro momento tratar de questões relacionadas a repetência, motivadas pelos efeitos deletérios que esta tem sobre disposição, autoestima e ambiente de sala de aula, além de não solucionar problemas de aprendizado conforme explicitado em Neri et al (2015).

Esses resultados parecem indicar que as variáveis mais relevantes para explicar a probabilidade de aprovação ao longo do ensino médio estão diretamente relacionadas às características individuais e familiares e pouco relacionada ao efeito-escola, ratificando Menezes-Filho e Pazello (2007) e Souza et al (2012).

6.1 Os efeitos marginais para o modelo *biprobit*

Esta seção discorre sobre os efeitos marginais⁷ dos modelos bivariados relatados nas Tabelas (9), (10) e (11). Este tipo de análise oferece o benefício analítico de permitir considerar os efeitos de cada variável explicativa sobre as probabilidades conjuntas observadas — no caso desta pesquisa, *P01* e *P11*.

Para o caso de estudantes do sexo masculino, nota-se que eles apresentam uma maior probabilidade de não terem um rendimento escolar satisfatório do que as estudantes do sexo feminino, isto é, na média dos três modelos, possuem uma probabilidade 9% menor de serem aprovados e continuar os estudos. Quanto a ser branco e morar na zona rural, os resultados indicam que a chance do estudante ser aprovado e continuar os estudos aumentam em quase 10% em relação aos demais alunos. O fato da família do estudante estar inscrita no Programa Bolsa Família tem impacto bastante reduzido no comportamento estudantil do jovem quando comparado a jovens cujas famílias não recebem o auxílio público. Já o efeito da variável *computador*, mostrou-se de pouca relevância para o caso dos alunos reprovados que permaneceram na base de dados, entretanto, os alunos que tiveram acesso a computador possuíram uma probabilidade 6% maior de serem aprovados e permanecer na rede pública de ensino. Por fim, os efeitos marginais médios estimados para a variável *lnpib_pc* aparentam estar viesados, pois, imagina-se que, na média, o nível socioeconômico tenha um efeito positivo sobre a aprovação acadêmica.

Tabela 9: M1: Efeitos marginais (em %)

Variáveis	Média	<i>P01</i>	<i>P11</i>
Menino	0,4764	0,0338	-0,0898418
Branco	0,9522	-0,0210	0,0922355
Rural	0,1926	-0,0156	0,0980428
Bolsa	0,0142	0,0073	-0,01958
computador	0,2759	-0,0025	0,0652118
lnpib_pc	10,1753	-0,0066	-0,1201858

Fonte: Elaboração própria.

Os efeitos marginais estimados para os modelos M2 e M3 apresentaram resultados semelhantes para as características familiares apresentadas no modelo M1. Como esperado, as ofertas de biblioteca e laboratório de informática praticamente não influenciaram no desempenho dos alunos reprovados, mas promoveram uma ligeira melhora, de cerca de 2%, no desempenho de alunos aprovados em escolas que possuem esses dois equipamentos, em relação aos que cursam escolas que não têm.

Quanto ao tamanho relativo da escola, seu efeito marginal mostrou-se pouco relevante, porém, os valores estimados indicam que escolas maiores têm efeitos negativos no desempenho acadêmico daqueles alunos que não evadem, contrariando Banerjee (2016). No que diz respeito à idade média dos professores, mostrou-se não ser significativa para a melhora do rendimento escolar daqueles alunos que têm maior probabilidade de reprovar, corroborando com os resultados obtidos por Souza et al (2012) e Junior e Stein (2017).

⁷ Esses efeitos correspondem às variações na probabilidade de um por cento nas variáveis explicativas contínuas, e no caso das variáveis binárias (*dummies*), uma variação discreta de 0 para 1 (CAMERON E TRIVEDI, 2005).

Tabela 10: M2: Efeitos marginais (em %)

Variáveis	Média	P01	P11
menino	0,4764	0,0342	-0,0906
branco	0,9522	-0,0212	0,0930
rural	0,1926	-0,0145	0,0855
bolsa	0,0142	0,0093	-0,0272
computador	0,2759	-0,0039	0,0701
lnpib_pc	10,1753	-0,0094	-0,1042
e_biblio	0,7938	-0,0013	0,0203
e_lab_info	0,8734	-0,0074	0,0193
i_funcionarios_razao	7,5292	0,0012	-0,0046
prof_idade	40,7436	0,0003*	-0,0058

Fonte: Elaboração própria.

* Não é estatisticamente significativa.

Por fim, os efeitos marginais estimados para o Modelo M3, destacam que a transferência escolar tem efeitos expressivos no rendimento escolar do estudante. Os alunos que mudaram de colégio apresentaram uma probabilidade 20% menor de serem aprovados e permanecer na rede pública de ensino em relação àqueles que não tiveram que passar por essa situação ao longo do Ensino Médio. Quanto à quantidade de faltas, elas comprometem o rendimento escolar, por consequência, aumentam a probabilidade de reprovação dos estudantes, apesar da magnitude do efeito ser pequena. As estimativas para o efeito do número de alunos por turma também apresentaram uma magnitude reduzida, o que pode ser explicada pela homogeneidade no tamanho médio das turmas entre os colégios.

Tabela 11: M3: Efeitos marginais (em %)

Variáveis	Média	P01	P11
menino	0,4764	0,0343	-0,0884
branco	0,9522	-0,0210	0,0898
rural	0,1926	-0,0145	0,0857
bolsa	0,0142	0,0092	-0,0260
computador	0,2759	-0,0039	0,0676
lnpib_pc	10,1753	-0,0096	-0,1080
e_biblio	0,7938	-0,0014	0,0203
e_lab_info	0,8734	-0,0074	0,0202
i_funcionarios_razao	7,5292	0,0012	-0,0063
prof_idade	40,7436	0,0002	-0,0057
transfer	0,0091	0,0091	-0,2011
faltas	0,3865	0,0014	-0,0094
alunos_turma	26,2287	-0,0001	0,0050

Fonte: Elaboração própria.

* Não é estatisticamente significativa.

Os resultados estimados dos efeitos marginais para as três especificações indicam que as variáveis mais relevantes para explicar a probabilidade de aprovação ao longo do ensino médio do estudante catarinense estão diretamente relacionadas às características individuais e familiares e pouco relacionadas ao efeito-escola, como sugerem Menezes-Filho e Pazello (2007) e Souza et al (2012). Ressalta-se, novamente, que o formulador de política pública deve estar atento ao caso específico dos alunos que mudam de colégio, pois estes apresentaram uma probabilidade de desempenho escolar inferior aos alunos que cursaram todo o Ensino Médio na mesma instituição de ensino.

7. Conclusão

Este artigo buscou identificar as características individuais dos jovens e de suas famílias, de ambiente escolar e de comportamento em sala de aula que são decisivas para o fluxo escolar ao longo do ensino médio. Além disso, a pesquisa também procurou estimar as probabilidades conjuntas de um estudante aprovado permanecer na rede pública de ensino e de um estudante reprovado não permanecer (evadir), a partir de uma regressão probabilística bivariada. Para tanto, fez-se uso de uma base de microdados inédita e exclusiva administrada pela Secretaria de Estado da Educação de Santa Catarina (SEDSC), com a qual é possível acompanhar mais de 400 mil alunos ao longo de toda sua trajetória no Ensino Médio público catarinense entre os anos de 2008 a 2012.

Os resultados indicam que as características mais relevantes para explicar o fluxo escolar estão diretamente relacionadas às características individuais e familiares, porém tal efeito se reduz com o avançar das séries, corroborando com os resultados apresentados no trabalho seminal de Mare (1980), que pode ser uma evidência de que o efeito-diploma se sobrepõe a estes controles. De modo geral, as taxas de promoção serial são menores entre homens, não brancos e moradores de zona urbana com alto PIB *per capita*, enquanto que a família estar inscrita no programa bolsa família não apresentou efeitos significativos.

No que se refere aos fatores associados à escola, mostrou-se que a oferta de biblioteca e laboratório de informática pouco altera as chances de aprovação e permanência na escola, em todas as transições analisadas, enquanto que escolas maiores têm efeitos negativos no desempenho acadêmico daqueles alunos que não evadem. Esse é um indicativo de que a estrutura escolar ofertada nas escolas públicas pouco influencia no fluxo escolar de estudantes catarinenses, principalmente sobre aqueles alunos que têm pior desempenho acadêmico.

Em relação às características comportamentais dos jovens, essas indicam que o desinteresse pelos estudos já é inerente ao próprio indivíduo e que, provavelmente, as características da escola primária exercem uma influência mais significativa sobre as transições seriais e, conseqüente, diplomação no ensino médio do que os fatores contextuais da escola secundária. Isso pode justificar o fato de a transição da 1ª para a 2ª Série no ensino médio ser o momento mais crítico para situações de evasão, devido aos altos índices de reprovação e latência, indo ao encontro de evidências obtidas em outros trabalhos que evidenciam a importância das taxas de repetência na justificativa para a não progressão nos estudos. Ademais, destaca-se, ainda, o papel fundamental da aprovação na explicação da probabilidade de evoluir ao longo do ciclo educacional. Estimou-se que cerca de 80% dos alunos aprovados continuam os estudos na rede pública após a conclusão da 1ª série e que esta taxa ainda foi ligeiramente superior, em torno de 83%, para a transição da 2ª para a 3ª série.

À derradeira, tendo em vista a escassez de dados longitudinais para a condução de pesquisas que focam nos determinantes do fluxo escolar, esta pesquisa apresentou resultados que se adicionam à literatura nacional, sobretudo no que diz respeito à redução dos efeitos familiares ao longo do ensino médio, nas probabilidades de aprovação e permanência, e, ainda, a sobreposição do efeito-diploma sobre os demais controles. Ademais, os alunos que apresentam maior risco de não completar o estudo secundário são aqueles que apresentam um baixo rendimento acadêmico durante toda sua trajetória escolar. Destarte, a formulação de políticas públicas deve focar nas necessidades do aluno por meio do monitoramento de toda a sua vida escolar e não restrita a uma etapa, principalmente daqueles com maior risco de evadir e repetir.

Referências

- BANERJEE, P. A.. A systematic review of factors linked to poor academic performance of disadvantaged students in science and maths in schools. **Cogent Education**, 2016.
- CAMERON, A.C.; TRIVEDI, P.K. *Microeconometrics: methods and applications*. Cambridge University Press, 2005.
- CAMERON, S.; HECKMAN, J.. The dynamics of educational attainment for black, hispanic, and white males. **Journal of political Economy**, v. 109, n. 3, p. 455-499, 2001.

- DE WITTE et al. A critical review of the literature on school dropout. **Educational Research Review**, v. 10, p. 13-28, 2013.
- ECKSTEIN, Z.; WOLPIN, K. I. Why youths drop out of high school: The impact of preferences, opportunities, and abilities. **Econometrica**, v. 67, n. 6, p. 1295-1339, 1999.
- FERRÃO, M. E., BELTÃO, K. I., & SANTOS, D. P. Políticas de não-repetência e a qualidade da educação: evidências obtidas a partir da modelagem dos dados da 4ª série do SAEB-99. **Estudos em avaliação educacional**, (26), 47-74, 2002.
- FERRO, A. R.; KASSOUF, A. L. Avaliação do impacto dos programas de Bolsa Escola sobre o trabalho infantil no Brasil. **Pesquisa e Planejamento Econômico**, v. 35, p. 417, dez. 2005
- HOLM, A.; JAEGER, M. M. Selection bias in educational transition models: Theory and empirical evidence. *Social Policy and Welfare Services Working Paper* 11, 2008.
- HOLM, A.; JAEGER, M. M. Dealing with selection bias in educational transition models: The bivariate probit selection model. **Research in Social Stratification and Mobility**, v.29, n.3, p. 311-322, 2011.
- JUNIOR, M. V. W.; STEIN, G.. Heterogeneidade da turma e o aprendizado escolar: o papel das qualificações do professor. **Planejamento e Políticas Públicas**, Rio de Janeiro, v. 49, 2017.
- LEE, V E.. A Necessidade dos Dados Longitudinais na Identificação do Efeito-Escola. **Revista Brasileira de Estudos Pedagógicos**, Brasília, v. 91, n. 229, p.471-480, 2010.
- LEON, F. L. L.; MENEZES-FILHO, N. Reprovação, avanço e evasão escolar no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v.32, n.3, p.417-451, 2002.
- LUCAS, S.R. Effectively maintained inequality: Education transitions, track mobility, and social background effects. **American Journal of Sociology**, 106(6), pp.1642-1690, 2001.
- MARE, ROBERT D. Social background and school continuation decisions. **Journal of the American Statistical Association**, 75(370), p.295-305, 1980.
- MENEZES-FILHO, N.; PAZELLO, E. "Do teachers' wages matter for proficiency? Evidence from a funding reform in Brazil." **Economics of Education Review** 26.6. p. 660-672, 2007.
- NERI, M. et al. *Motivos da Evasão Escolar*. Fundação Getúlio Vargas, Rio de Janeiro, 2015.
- OLIVEIRA, L.F.B.D.; SOARES, S.S., 2012. Determinantes da repetência escolar no Brasil: uma análise de painel dos censos escolares entre 2007 e 2010. **Texto para discussão, IPEA**, n. 1706, 2012.
- PORTELLA, A. L.; BUSSMANN, T. B; OLIVEIRA, A. M. O.. A relação de fatores individuais, familiares e escolares com a distorção idade-série no ensino público brasileiro. **Nova Economia** v.27, n. 3, 2017.
- RIANI J.D.L.R.; ASSIS, M.A.Os determinantes do fluxo escolar no ensino médio em Minas Gerais (Brasil). *Anais do XX Encontro Nacional de Estudos populacionais*, pp.1-21, 2017.
- RUMBERGER, R.; LIMA, S. A. Why Students Drop Out: A Review of 25 Years of Research. **California Dropout Research Project**, Policy Brief 15, University of California, 2008.
- SASS, T. R.; ZIMMER, R. W.; GILL, B. P.; BOOKER, T. K.. Charter high schools' effects on long-term attainment and earnings. **Journal of Policy Analysis and Management**, 35(3), 683-706, 2016.
- SHIRASU, M.; ARRAES, R. Determinantes da evasão e repetência escolar no ensino médio do Ceará. **Revista Econômica do Nordeste**, v. 46, n. 4, p. 117-136, 2015.
- SOUZA, A. P. D.; PONCZEK, V. P.; OLIVA, B. T.; TAVARES, P. A. Fatores associados ao fluxo escolar no ingresso e ao longo do ensino médio no Brasil. **Pesquisa e Planejamento Econômico**, v. 42, n. 1, 2012.