

Testes padronizados na educação básica e seus efeitos sobre resultados educacionais

evidências das pesquisas acadêmicas

Setembro, 2022

EVE
Evidência Express

ENAP

Expediente

Presidente

Diogo Costa

Diretora-Executiva

Rebeca Loureiro de Brito

Diretora de Altos Estudos

Diana Coutinho

Diretor de Educação Executiva

Rodrigo Torres

Diretor de Desenvolvimento Profissional

Paulo Marques

Diretora de Inovação

Bruna Santos

Diretora de Gestão Interna

Alana Regina Biagi Silva Lisboa

Coordenador Geral de Ciência de Dados

Pedro Masson Sesconetto Souza

Capa e Diagramação

Samyra Lima

Equipe Evidência Express

Imagens

Unsplash

Autoria

Célio Belmiro

Doutorando em Economia (UFPE).

Breno Salomon Reis

Mestre em Políticas Públicas (Insper).

O Evidência Express (EvEx) é uma iniciativa da Diretoria de Altos Estudos da Escola Nacional de Administração Pública (Enap) em parceria com a Universidade Federal de Brasília (UnB). A missão do EvEx é melhorar a tomada de decisão do setor público. Para isso a equipe sintetiza, produz e dissemina evidências que possam servir de base para o desenho, monitoramento e avaliação de políticas públicas.

Avaliações completas de políticas públicas são intensivas em tempo e custos. A fim de agilizar esses processos, o EvEx produz relatórios ágeis de evidências para a consolidação do conhecimento disponível e introdução de novos pontos de vista.

Os resultados dos produtos EvEx apoiam tomadores de decisão do setor público federal, subsidiando avaliações Ex Ante, Ex Post ou Análises de Impacto Regulatório. Beneficiam também os gestores públicos subnacionais, pesquisadores, docentes, servidores e demais interessados na sociedade civil.

Os produtos EvEx analisam evidências qualitativas e quantitativas, podendo ser demandados de forma avulsa ou em pacotes, sobre:

- Evolução do problema no Brasil e no mundo;
- Público-alvo de uma política;
- Causas e consequências do problema ou política;
- Soluções existentes para o problema;
- Impactos de intervenções ou políticas públicas.

Para mais informações, consulte nossa página (www.enap.gov.br/pt/servicos/avaliacao-e-organizacao-de-evidencias) ou entre em contato: evidencia.express@enap.gov.br.



Sumário Executivo

- Este trabalho apresenta uma síntese rápida de evidências sobre o impacto da introdução de testes padronizados no sistema educacional.
- São analisados casos em que os testes são aplicados em conjunto com a introdução de um sistema de admissão centralizado, como a heterogeneidade de perfil dos estudantes afeta o desempenho nos testes e a relação entre testes padronizados e *accountability* no sistema escolar.
- Para elaborar a síntese foram analisados 17 artigos após um levantamento de 309 trabalhos por meio dos portais indexadores *Scopus* e *Web of Science* e da análise de conexões bibliográficas do trabalho de Machado e Szerman (2021) a partir da ferramenta *Connected Papers*.
- É identificado que a transição para um sistema de admissão centralizado gerou no Japão como consequências um efeito de migração geográfica (curto prazo) e impactos na carreira dos estudantes (longo prazo).
- O *background* familiar, características socioeconômicas e de acesso a infraestrutura e oportunidades educacionais têm impacto no desempenho dos alunos e consequente relação com a admissão no ensino superior.
- Embora a literatura aponte para uma reversão das diferenças de gênero nas notas dos testes padronizados, Saygin (2020) argumenta que a diferença de notas nos testes padronizados persistem mesmo que, em média, as mulheres alcancem notas relativamente melhores que as dos homens no primeiro ano de graduação e possuam desempenho mais elevado no ensino médio.
- Testes padronizados em alguns casos são utilizados em conjunto com medidas que visam promover a responsabilização dos resultados educacionais alcançados no sistema escolar. A avaliação de uma política na Coreia do Sul por Lee e Kang (2019) encontrou que uma maior responsabilização das escolas mensurada a partir dos resultados de estudantes em testes padronizados levou ao aumento do aprendizado em leitura e matemática.
- Os resultados encontrados nas pesquisas analisadas possuem validade interna, mas não necessariamente validade externa. Dessa forma, análises adicionais devem ser feitas para extrapolar suas conclusões para outros contextos, tendo em vista, especialmente, os mecanismos que explicam os achados.

Sumário

| | | |
|----------|--|-----------|
| 1 | Introdução | 5 |
| 2 | Metodologia | 7 |
| 3 | Resultados | 9 |
| 3.1 | Testes Padronizados em Sistemas de Admissão Centralizados | 9 |
| 3.2 | Testes Padronizados e Heterogeneidade no Acesso ao Ensino Superior ... | 15 |
| 3.3 | Testes Padronizados e Responsabilização Escolar | 20 |
| 4 | Considerações Finais | 25 |
| | Referências Bibliográficas | 26 |
| | Apêndice 1 | 28 |

1. Introdução

Este trabalho tem como objetivo apresentar, em uma síntese rápida de evidências, a relação entre testes padronizados e resultados educacionais, com foco em variáveis que mensuram o acesso ao ensino superior. São analisados estudos científicos que abrangem um conjunto de evidências que variam em relação às consequências dos testes para os participantes, ao tipo de sistema educacional em que estão inseridos e aos objetivos da política - ou programa - educacional dos quais fazem parte. Os principais conjuntos de resultados estão organizados em três blocos que se referem ao acesso ao ensino superior em sistemas de admissão centralizados, as heterogeneidades e a relação entre os testes e responsabilização escolar¹ (*school accountability*).

Um primeiro aspecto relevante da relação entre testes padronizados e o acesso ao ensino superior se refere às características do sistema educacional em que a avaliação está inserida. Como argumenta Szerman (2015), nos mercados educacionais em que os alunos têm preferências sobre as instituições e as instituições enfrentam limitações de capacidade há uma grande pesquisa teórica sobre mecanismos de escolha ótima de seleção de alunos, levando em consideração questões relacionadas à estabilidade, justiça e eficiência das alocações. Sem um consenso sobre o mecanismo mais eficiente, diferentes procedimentos de admissão são atualmente empregados em todo o mundo e isto afeta a forma como os testes padronizados impactam no acesso ao ensino superior (FRANCESCONI; SLONIMCZYK; YURKO, 2019; TANAKA; NARITA; MORIGUCHI, 2020; MACHADO; SZERMAN, 2021).

Assim, o primeiro conjunto de evidências apresentados, destaca como os testes padronizados associados aos sistemas centralizados, descentralizados ou parcialmente centralizados impactam um conjunto de variáveis relacionadas ao acesso ao ensino superior. Há, ainda, uma preocupação transversal a todos os sistemas de admissão, que se refere a questões de equidade e justiça na avaliação de desempenho por meio de testes padronizados. As evidências empíricas de diferenças persistentes nos resultados de testes padronizados entre grupos étnicos, gênero e situação socioeconômica, somada a uma sub-representação desses grupos em instituições de ensino superior ao redor do mundo levantam importantes debates sobre o papel de tais testes nas políticas de admissão (CORREA et al., 2019; SAYGIN, 2020). Nesse sentido, o segundo conjunto de evidências apresentado discute os efeitos heterogêneos dos testes padronizados, em diferentes contextos, sobre a performance de estudantes com base em seu *background* familiar (status socioeconômico), grupo étnico/racial e gênero. Finalmente, há também uma vertente da literatura que investiga o papel dos testes padronizados nas políticas de responsabilização escolar e na divulgação de informações sobre qualidade do ensino e outros indicadores educacionais que podem influenciar o comportamento dos agentes - pais, responsáveis, professores, gestores e toda a comunidade.

¹Decidimos adotar o termo responsabilização escolar para se referir aos diversos termos, em inglês, relacionados a *accountability*, como *school accountability system*.

Conforme argumentam Figlio e Loeb (2011) o termo responsabilização na educação é bastante amplo e a definição mais comum do termo se refere ao uso de mecanismos administrativos, baseados em dados, com o objetivo de aumentar os resultados dos estudantes. Boa parte dos sistemas atuais de responsabilização tomam a escola como unidade de análise - em detrimento de sistemas que exigem padrões mais elevados de alunos ou professores. Os incentivos (recompensas/punições) podem ser diretos - providos pelo governo - ou indiretos, através da disponibilização da informação para a comunidade. O problema do principal-agente prevê uma racionalização da questão: se os tomadores de decisão - pais, agentes locais, *policymakers* - têm dificuldade de monitorar as atividades da escola, então os educadores podem se comportar de maneira contrária aos interesses dos tomadores de decisão. Nesse casos, medidas mais efetivas de monitoramento dos educadores poderiam resultar em melhorias dos resultados escolares (FIGLIO; LOEB, 2011).

Avaliar as escolas em torno de uma métrica comum, por meio dos testes padronizados provê aos formuladores de política e o público em geral, informações sobre como as escolas, distritos escolares e, potencialmente, os professores, estão em comparação aos seus pares (FIGLIO; LOEB, 2011; CAMARGO et al., 2018; CUESTA; GONZÁLEZ; PHILIPPI, 2020). O terceiro conjunto de evidências apresentado trata, justamente, dos efeitos da aplicação de testes padronizados - e da divulgação de seus resultados - em associação à responsabilização escolar.

Além desta breve introdução, esta síntese de evidências encontra-se dividida em mais três seções. A segunda seção apresenta a metodologia de mapeamento e seleção dos trabalhos que compõem o núcleo de resultados. Esses, apresentados na terceira seção, dividem-se em três subseções que exploram, respectivamente, casos em que os testes são aplicados em conjunto com a introdução de um sistema de admissão centralizado, como a heterogeneidade de perfil dos estudantes afeta o desempenho nos testes e a relação entre testes padronizados e *accountability* no sistema escolar. Finalmente, a quarta e última seção, apresenta as considerações finais desta pesquisa.

2. Metodologia

Os estudos apresentados nesta síntese de evidências foram selecionados a partir de buscas nos bancos de dados de resumos e citações de artigos, *Scopus* e *Web of Science*, a partir das combinações de um conjunto de palavras-chave, apresentadas na tabela A1, no Apêndice. Adicionalmente, dada a relevância do trabalho de Machado e Szerman (2021) para a principal pergunta de pesquisa desta síntese, empreendemos uma análise da rede de conexões bibliográficas relevantes para o entendimento da questão. Essa análise foi feita por meio da ferramenta *Connected Papers*, e a figura A1, também no Apêndice, ilustra o procedimento de busca.

O processo de busca resultou em uma base de dados bibliográfica com 304 trabalhos científicos, entre artigos, livros, capítulos de livro, trabalhos apresentados em conferências, revisões de literatura e material editorial¹. O conjunto final de trabalhos selecionados para a leitura é resultado de um processo de filtragem composto por duas etapas.

A primeira etapa consistiu na estruturação de uma base única, contendo os trabalhos resultantes dos quatro conjuntos de buscas nos portais indexadores *Scopus* e *Web of Science*, para exclusão das duplicatas, das publicações não revisadas por pares e das publicações fora da área das Ciências Sociais. Nesta etapa, dos 304 trabalhos encontrados, 116 foram excluídos - 31 duplicados e 85 por não se enquadrarem na categoria de revisados por pares ou por pertencerem a áreas distintas das Ciências Sociais.

A segunda etapa consistiu na checagem, a partir da leitura dos *abstracts* (resumos) de quais dos 188 trabalhos restantes estavam adequados ao tema investigado por esta síntese. Os trabalhos classificados como adequados foram aqueles que, no resumo, descreviam a investigação de impactos - evidências ou efeitos - de testes padronizados nos estudantes em resultados educacionais, de maneira geral, em variáveis que mensuram o acesso ao ensino superior e em políticas - ou programas - relacionados à responsabilização escolar. Nesta etapa, dos 188 resumos lidos, 25 foram selecionados para leitura completa e consequente produção da síntese de evidências.

Contudo, observou-se no momento da leitura, que um pequeno conjunto dos trabalhos selecionados, embora adequados ao tema desta síntese, apresentam discussão qualitativa e/ou filosófica sem a consequente investigação - ou apresentação - das evidências existentes que corroborem com as discussões apresentadas. Por este motivo, optou-se por excluir este conjunto de 13 trabalhos da apresentação dos resultados. A Tabela 2.1, abaixo, resume os critérios de inclusão e exclusão apresentados.

¹Material editorial é um artigo que apresenta opiniões de uma pessoa, grupo ou organização. Inclui comentários, a depender do conteúdo, editoriais, entrevistas, discussões entre indivíduos, discussões pós-artigos, simpósios de mesa redonda, resumo da conferência e destaques da pesquisa.

Tabela 2.1: Critérios de Inclusão e Exclusão dos Estudos

| Inclusão | Exclusão |
|---|--|
| <ul style="list-style-type: none">▪ Estudos que avaliam os efeitos de testes padronizados em variáveis que mensuram acesso ao ensino superior e em outros resultados educacionais.▪ Estudos quantitativos (ou métodos mistos), já publicados e revisados por pares;▪ Estudos que avaliam os efeitos heterogêneos da performance nos testes padronizados com base no gênero e/ou raça/étnia;▪ Apenas textos em inglês | <ul style="list-style-type: none">▪ Documentos de literatura cinzenta, como relatório de organizações não-governamentais e materiais editoriais;▪ Artigos de áreas fora das Ciências Sociais e textos filosóficos e - ou - qualitativos |

Fonte: Elaborado pelos autores

Finalmente, também no momento da leitura, observou-se no já citado trabalho de Machado e Szerman (2021), uma vertente da literatura que trata sobre os impactos de sistemas de admissão centralizados - que irão requerer, necessariamente, algum processo padronizado para seleção dos candidatos - e realizou-se uma análise das conexões bibliográficas deste trabalho com outros relevantes para o tema. Neste processo, mais 5 trabalhos foram selecionados para leitura, totalizando 17 estudos que atenderam a todos os critérios de inclusão definidos para a produção desta síntese.

3. Resultados

No processo de busca, foram selecionados 17 artigos, que serão agrupados em três conjuntos para apresentação das evidências. Primeiro, são apresentados os resultados referentes aos testes padronizados e seus impactos - efeitos e evidências - em um conjunto de variáveis educacionais, principalmente no acesso ao ensino superior. Na sequência, uma nova seção discute a relação entre as características dos indivíduos, sua heterogeneidade, e como elas afetam os resultados obtidos em testes padronizados. Finalmente, são apresentadas as evidências referentes à relação entre testes padronizados e responsabilização escolar.

3.1 Testes Padronizados em Sistemas de Admissão Centralizados

Todo ano milhões de estudantes buscam uma vaga no sistema de ensino superior sob uma grande variedade de mecanismos de admissão (MACHADO; SZERMAN, 2021). Por meio de reformas que alteram de maneira significativa o panorama de acesso ao ensino superior, os mercados de educação estão gradualmente migrando para sistemas de admissão parcialmente (des)centralizados (FRANCESCONI; SLONIMCZYK; YURKO, 2019; KNIGHT; SCHIFF, 2022; MACHADO; SZERMAN, 2021). É inerente aos sistemas centralizados, ou parcialmente centralizados, a aplicação dos testes padronizados para avaliar as habilidades individuais dos estudantes. Neste primeira seção, são apresentados um conjunto de trabalhos que discutem as evidências encontradas - sobre acesso ao ensino superior, eficiência alocativa, aumento do bem-estar e integração geográfica dos mercados de educação - quando ocorre mudança no sistema de admissão para um nível maior de centralização e tais testes são aplicados.

Machado e Szerman (2021) buscam responder se, e como, a introdução de um sistema centralizado, que adota um teste padronizado para admissão, influencia a composição dos estudantes entrantes no contexto de admissão para o ensino superior. Para tanto, exploram uma única e ampla mudança de política, implementada pelo Ministério da Educação, o SISU, sistema centralizado que aloca os estudantes para as universidades públicas em todo o país, usando como base as notas do Exame Nacional do Ensino Médio (ENEM). Conforme argumentam, a introdução do SISU tornou o processo de inscrição para diversas instituições de ensino superior mais fácil, reduzindo os custos de procura e de obtenção de informação, dando mais agilidade ao processo de seleção e reduzindo, também, os custos financeiros que inscrições para múltiplas instituições acarretariam.

Para quantificar se, e como, o sistema centralizado afeta a composição dos estudantes, Machado e Szerman (2021) utilizam como especificação empírica um *event-study* e um modelo de diferenças em diferenças para explorar a variação no *timing* - momento - em que os programas passaram a adotar o SISU como método de seleção, comparando as características dos estudantes dos programas antes e após a centralização. Os resultados evidenciam que: (a) a participação das mulheres caiu em 1,2 p.p, o

que equivale a um declínio de 2,3% no número de estudantes; (b) houve um incremento de 0,9% na idade dos estudantes admitidos; e (c) não parece haver mudanças nos componentes raciais e de status socioeconômico, resultados que contrastam com os achados de Mello (2021), também discutidos nessa seção. As autoras também identificam um efeito de integração geográfica. No ano após a migração para o sistema centralizado houve um incremento de 2,9 p.p (20,4%) na probabilidade dos estudantes migrarem para outro estado, um padrão que apresenta tendência de crescimento ao longo dos anos.

Os autores também investigam se os efeitos variam mais a depender da seletividade das instituições. Se, por exemplo, os ganhos no incremento das notas ou mobilidade estiverem desproporcionalmente concentrados em instituições mais seletivas, a centralização estará aumentando a estratificação no mercado de educação. Enquanto os resultados apontam que as mudanças na composição de gênero não dependem da seletividade das instituições, o efeito de aumento da idade aparece no tercil mais baixo de classificação de qualidade das universidades, indicando que a repetição do exame está concentrada nestas instituições, de nível mais baixo, com potenciais notas mais baixas de admissão. Os resultados também indicam que pessoas mais velhas e que desempenharam abaixo da média no ENEM, passaram - com o SISU - a conseguir ingressar em um curso superior. As notas de entrada, por outro lado, crescem para todas as instituições de ensino superior, indicando um aumento da competição, mas são maiores para o tercil superior da distribuição de qualidade (MACHADO; SZERMAN, 2021).

Uma implicação dos resultados encontrados, segundo as autoras, é que as admissões centralizadas impactam a composição dos alunos admitidos, beneficiando relativamente mais os homens, mais velhos, de fora do estado e com altas habilidades, embora não necessariamente desloquem os alunos de baixo nível socioeconômico.

Os resultados sugerem que a centralização leva a uma maior competitividade entre os mercados integrados de educação. Este efeito seria explicado, principalmente, por um aumento das notas - consequência do aumento da competição - que irá requerer um esforço maior para a aprovação. Os efeitos da migração, contudo, estão concentrados, predominantemente, nas instituições de maior qualidade, levando a um aumento da estratificação entre as instituições de baixa e alta qualidade. As autoras sugerem que esses dois mecanismos podem superar os efeitos, para estudantes mais desfavorecidos, dos benefícios associados aos sistemas centralizados e outros mecanismos podem ser necessários para aumentar a representatividade desses estudantes, como as políticas afirmativas (MACHADO; SZERMAN, 2021).

O trabalho de Knight e Schiff (2022) estuda a introdução, no sistema de admissão americano, do CA (*Common Application platform*), na qual estudantes se inscrevem, no mesmo processo de seleção, para um amplo conjunto de instituições do ensino superior que fazem parte da plataforma. Historicamente, no processo de admissão descentralizado americano, os estudantes completam múltiplos processos de inscrição para as instituições de ensino. Isso requer um tempo substancial, envolve diversos tipos de custo e gera uma série de fricções que podem limitar as possibilidades de inscrição e,

em última instância, reduzir as opções de escolha (KNIGHT; SCHIFF, 2022).

Nesse sentido, Knight e Schiff (2022) estudam se a introdução do CA, um mecanismo de admissão centralizado, contribui para: (a) redução das fricções do processo de inscrição, resultando em maior número de aplicantes e aumentando o conjunto de escolha dos estudantes; (b) maior integração geográfica do mercado de educação; e (c) estratificação das instituições de ensino superior. Para responder a essas questões, os autores apresentam um modelo teórico simples do processo de admissão universitária, que inclui as decisões de inscrição, ofertas de admissão e a aceitação dos estudantes, e testam as previsões teóricas deste modelo utilizando dados da Pesquisa Anual de Faculdades do *College Board (College Board's Annual Survey of Colleges)* e do Sistema Integrado de Dados do Ensino Superior (*Integrated Postsecondary Education Data System - IPEDS*), que cobrem o período (1990 – 2016), combinados com o ano em que cada instituição tornou-se membro do CA.

As duas principais especificações empíricas são obtidas usando o estimador *two-way fixed effects* e um *event-study*. Os resultados indicam que, após a entrada no CA: (a) há um aumento no número de inscrições de 12%; (b) uma redução de 9% na taxa de aceite, que é condizente com a hipótese de maior oferta e oportunidades de escolha; (c) um incremento de 12% no número de estudantes admitidos; (d) um aumento de 1,4% do número de estudantes de fora do estado e de 0,3% do número de estudantes estrangeiros; (e) não há evidências conclusivas sobre estratificação, isto é, evidências de que a introdução do CA, contribuiu para que estudantes mais habilitados aplicassem para instituições mais seletivas aumentando as diferenças no perfil dos estudantes entre instituições; e (f) encontram evidências mistas sobre os efeitos nas notas do SAT, sem evidências de aumentos para o percentil 25 de distribuição das notas, mas evidências mais fortes e significativas de aumentos para o percentil 75 (9 pontos) (KNIGHT; SCHIFF, 2022).

Kapor, Karnani e Neilson (2022) estudam os impactos de bem-estar e de capital humano de uma reforma no sistema de admissão centralizada no Chile, em 2012, que promoveu uma expansão no número de vagas (aumento de cerca de 40%) e um grande programa de concessão de bolsas para cursos na área de licenciatura. Utilizando o método de *Gibbs sampler* para apresentar as estimativas empíricas do modelo teórico construído, os autores encontram: (a) aumento no número de matrículas de 8%; (b) redução das taxas de evasão ao final do primeiro ano de curso de 2 p.p (uma queda de 16%); e (c) um aumento nas taxas de graduação dentro do prazo de 1 p.p (um crescimento de 3%). Esses resultados são consistentes com o efeito esperado de um incremento das opções dentro de uma plataforma centralizada, que reduz consequências adversas das listas de espera, contribuindo para uma alocação eficiente com base nas principais escolhas dos indivíduos (KAPOR; KARNANI; NEILSON, 2022).

Já Tanaka, Narita e Moriguchi (2020) estudam os impactos de mudança no sistema de admissão do ensino superior japonês. No final do século XIX o Japão implementou mudanças no sistema de admissão que alteraram suas características de descentralizado para centralizado. Ao longo do século XX, contudo, o governo descentralizou e (re)centralizou o processo de admissão em momentos

distintos, gerando inúmeros experimentos naturais. Os autores exploram essas mudanças bidirecionais para identificar o impacto da chamada centralização meritocrática.

Os resultados indicam que a centralização teve efeitos de curto prazo no comportamento dos aplicantes e, conseqüentemente, no resultados das inscrições. Os incentivos estratégicos levaram, em primeiro lugar, os candidatos dos centros urbanos e rurais a escolher as escolas mais seletivas como primeira opção e, em segundo lugar, o sistema centralizado levou um número maior de candidatos habilitados de áreas urbanas a serem admitidos em instituições nas áreas rurais, muitas vezes depois de serem rejeitados por sua primeira escolha. No longo prazo, os autores investigam se houve impactos na carreira dos estudantes. Como o sistema centralizado foi desenhado para ser meritocrático e a distribuição dos alunos mais habilitados está nos centros urbanos, um possível resultado esperado é que a centralização deixe as áreas urbanas com acesso escolar desproporcionalmente maior em relação às áreas rurais (TANAKA; NARITA; MORIGUCHI, 2020).

Os resultados de curto prazo, estimados por meio de um modelo de efeitos fixos, evidenciam que, no sistema centralizado, há um aumento de 64% médio no número de submissões para a escola mais seletiva (no sistema descentralizado, essa média era de 25%); e uma redução do número de entrantes locais - nascidos na mesma região da escola - para todas as instituições, com resultados que variam, por instituição, de 17% a 49% (TANAKA; NARITA; MORIGUCHI, 2020).

Quanto aos impactos de longo prazo, estimados por meio de um modelo de diferenças em diferenças, os autores utilizam a distribuição geográfica dos estudantes para conjecturar que, sob o sistema centralizado, os estudantes da região de Tokyo devem experimentar um ganho mais alto do que os estudantes das demais regiões do país, uma vez que os estudantes mais habilitados estão localizados, desproporcionalmente, nas áreas urbanas. Para os coortes expostos ao sistema centralizado, o número de carreiras de elite para os nascidos dentro da área de Tokyo - em comparação aos nascidos fora - aumenta em 25% para os 0,05% mais ricos e 36% para os 0,01% mais ricos. Os efeitos são simétricos em relação à direção das mudanças - centralizado para descentralizado e descentralizado para centralizado (TANAKA; NARITA; MORIGUCHI, 2020).

Na mesma linha dos trabalhos anteriores, Francesconi, Slonimczyk e Yurko (2019) avaliam os efeitos na mobilidade dos estudantes e integração geográfica do mercado de educação, de uma reforma, em 2009, que exigiu que as universidades russas passassem a tomar as decisões de admissão com base nos resultados de um exame nacional do ensino médio. Antes da reforma, o acesso universitário era extremamente desigual, estudantes de baixo status socioeconômico eram considerados menos propensos a aplicar para vagas no ensino superior que seus pares de alto status socioeconômico. Uma importante dimensão da larga diferença de acesso ao ensino superior é que, antes da reforma, menos de 20% dos jovens russos eram nascidos nas maiores cidades do país, mas representavam cerca de 60% de todos os estudantes universitários. Apenas 1 em cada 10 estudantes universitários era nascido nas cidades/vilas do interior do país (FRANCESCONI; SLONIMCZYK; YURKO, 2019).

Os altos custos de admissão no sistema descentralizado se manifestavam de diversas formas, e

Francesconi, Slonimczyk e Yurko (2019) destacam a falta de transparência no processo de admissão, corrupção, necessidade de deslocamento para a área dos exames específicos de cada instituição, que eventualmente requeriam entrevistas presenciais e o custo de cursos preparatórios, como barreiras de acesso ao ensino superior para os moradores de regiões mais periféricas e com renda mais baixa. Para verificar a hipótese dos efeitos de mobilidade dos estudantes e de integração geográfica, os autores estimam um modelo de diferenças em diferenças que compara a probabilidade de mudança do local da residência de jovens que estão no último ano do ensino médio com a de jovens que estão em outras séries da educação secundária.

As estimativas base indicam que a reforma induziu uma taxa de saída do local de residência substancialmente maior, entre 12% e 16%, pós-reforma, o que representa um aumento de três vezes do observado antes da centralização, entre os graduados do ensino médio que vivem em pequenas vilas e cidades (FRANCESCONI; SLONIMCZYK; YURKO, 2019). Estes são, como argumentam, Francesconi, Slonimczyk e Yurko (2019) exatamente os lugares mais propensos a serem caracterizados por uma demanda reprimida por admissão no ensino superior, uma vez que as pequenas cidades e vilas não possuem, de maneira geral, grandes centros universitários mas são dotadas de escolas secundárias de alta qualidade.

Mello (2021) estuda a reforma implementada pelo governo brasileiro, que aprovou, em 2012, a lei de cotas, reservando 50% das vagas nas universidades federal do país para estudantes que completassem os três anos do ensino médio em escolas públicas, e estipulou subcotas - dentro dos 50% - baseadas em critérios raciais e situação econômica. A análise foca nos estudantes do 9º ano, uma vez que a entrada no ensino médio público precisa ocorrer nesta transição (primário para o secundário), das 50 microrregiões do Brasil com ao menos uma instituição federal de ensino superior. O estudo utiliza um estimador *two-way fixed effects* para comparar as mudanças nas taxas médias de transição entre escolas públicas e privadas. As unidades tratadas são as escolas localizadas em microrregiões que experimentaram uma variação no nível de exposição às cotas. O controle, em contraste, são as escolas localizadas em microrregiões em que os níveis não variaram (MELLO, 2021).

Os resultados mostram que a exposição a um incremento das cotas aumenta os movimentos dos alunos do 9º ano do ensino médio privado para o ensino médio público em 4,7 p.p (ou 31% considerando os anos posteriores de 2012 a 2016), esse efeito é mais forte para os alunos não-brancos (7,1 p.p) , que se beneficiam de uma subcota adicional e para as mulheres (5,4 p.p). Os efeitos também são muito maiores entre os alunos do 9º ano de baixo status socioeconômico em escolas particulares de baixa performance. Juntos, esses resultados sugerem que os indivíduos que optam pela transição do privado para o público são os que provavelmente menos se beneficiam da educação privada e enfrentam custos mais baixos de transição.

A Tabela 3.1 sintetiza as principais características dos estudos descritos anteriormente.

Tabela 3.1: Características dos Estudos sobre Testes Padronizados e Sistema de Admissão Centralizado

| Estudo | Objeto de estudo | Dados | Estratégia empírica | Principais resultados |
|--|--|---|---|---|
| Machado e Szerman (2021) | Implementação do SISU, um sistema de admissão centralizado, na composição dos estudantes (Brasil) | Censo do Ensino Superior; Base de dados do Enem; Fontes adicionais do Ministério da Educação | Diferenças em Diferenças; <i>Event Study</i> | 1. Participação das mulheres no ensino superior caiu em 2,3%; 2. incremento de 0,9% na idade dos estudantes admitidos; 3. incremento de 20,4% na probabilidade do estudante migrar para outro estado |
| Knight e Schiff (2022) | Introdução de plataforma de submissão de candidaturas centralizada (CA - <i>Common Application platform</i>) (EUA) | Pesquisa Anual de Faculdades do <i>College Board</i> ; Integrado Sistema Integrado de Dados do Ensino Superior (IPED); Ano de entrada no CA | <i>Two-way fixed effects</i> ; <i>Event Study</i> | 1. Aumento de 12% no número de inscrições nas instituições de ensino superior; 2. redução de 9% na taxa de aceite; 3. incremento de 12% no número de estudantes admitidos; 4. aumento de 1,4% de estudantes de fora do estado e de 0,3% de estudantes de fora do país |
| Kapor, Karnani e Neilson (2022) | Impactos de bem-estar e de capital humano de uma reforma no sistema de admissão centralizada (Chile) | Ministério da Educação do Chile (MINEDUC); Conselho Nacional de Educação (CNED); Prova de Seleção Universitária (PSU) | <i>Gibbs sampler</i> | 1. Aumento de 8% no número de matrículas; 2. redução das taxas de evasão ao final do primeiro ano de curso em 2 p.p (16%); 3. aumento nas taxas de graduação dentro do prazo em 1 p.p (3%) |
| Tanaka, Narita e Moriguchi (2020) | Mudanças no sistema de admissão do ensino superior japonês (Japão) | Dados sobre inscrições, matrículas e outros resultados através da digitalização de diversas fontes administrativas e não administrativas | Efeitos fixos e Diferenças em Diferenças | 1. Aumento de 64% no número de submissões para escolas mais seletivas; 2. redução do número de entrantes locais - nascidos na mesma região da escola - de 17% a 49%; 3. aumento do número de carreiras de elite para os nascidos dentro da área de Tokyo (25% para os 0,05% mais ricos e 36% para os 0,01% mais ricos) |
| Francesconi, Slonimczyk e Yurko (2019) | Efeitos na mobilidade dos estudantes após a introdução de sistema de admissão centralizado (Rússia) | Pesquisa Russa de Monitoramento Longitudinal (RLMS) | Diferenças em Diferenças | 1. Aumento da taxa de saída do local de residência - entre 12% e 16%, o que representou um aumento de três vezes a taxa observada antes da centralização |
| Mello (2021) | Mudanças nas taxas de transferência entre instituições privadas e públicas após a promulgação da lei de cotas (Brasil) | Censo da Educação Básica; Censo Demográfico (2010) | <i>Two-way fixed effects</i> | 1. Aumento de 4,7 p.p do movimento de alunos do 9º ano do ensino fundamental em instituições privadas para o ensino médio em escolas públicas (7,1 p.p para estudantes não-brancos, que se beneficiam de subcota adicional e 5,4 p.p para mulheres) |

Fonte: Elaborado pelos autores.

3.2 Testes Padronizados e Heterogeneidade no Acesso ao Ensino Superior

O próximo conjunto de trabalhos apresenta as evidências da aplicação de testes padronizados - principalmente relacionados ao acesso ao ensino superior - e sua relação com as características socioeconômicas e demográficas dos estudantes. Bastedo et al. (2022) destacam, no sistema educacional americano, o crescimento de tais testes. Segundo os autores, menos de 1000 estudantes realizaram, pela primeira vez, em 1901, o *College Board Exam*. Já na década de 50, cerca de meio milhão de estudantes realizavam o exame anualmente. Hoje, mais de 1.8 milhões de estudantes realizam o SAT (*Scholastic Aptitude Test*) ou ACT (*American College Testing*) a cada ano. Adicionalmente, 21 estados requerem que os estudantes realizem o SAT ou ACT como parte dos critérios estudiais de conclusão do ensino médio - e outros três adotam como opcional que os exames sejam feitos (BASTEDO et al., 2022).

Ao longo do século XX, diversas faculdades e universidades adotaram os testes padronizados como critérios de admissão. Apesar das aspirações iniciais de que tais testes poderiam ser utilizados como um censo das habilidades humanas, capaz de identificar talentos e habilidades em indivíduos de diferentes contextos, ainda permanecem fortes correlações entre as notas dos testes padronizados e fatores contextuais, exógenos ao ambiente escolar, como raça, gênero e status socioeconômico dos estudantes e das famílias (BASTEDO et al., 2022). Essas diferenças, contudo, não originadas com os testes de admissão, são um reflexo da variação sistêmica de longo prazo em recursos educacionais e oportunidades, que se manifestam através de diversos canais, dentre eles, nos resultados dos testes (BENNETT, 2022; DIXON-ROMÁN; EVERSON; MCARDLE, 2013).

Bastedo et al. (2022) conduzem um RCT (*Randomized-Controlled Trial*) para determinar se fornecer as pontuações dos testes dos alunos, comparando seu desempenho em relação aos demais da escola e da vizinhança, aumenta a probabilidade de admissão no ensino superior. Especificamente, os autores testam se os oficiais de admissão são mais propensos a admitir estudantes em situação de desvantagem se eles tiverem informações de qualidade sobre a média dos estudantes na escola e na vizinhança do estudante - ou seja, puderem comparar o desempenho relativo aos demais pares em situação semelhante (BASTEDO et al., 2022).

Os resultados da aplicação de um modelo logístico multinível indicam que as chances de recomendar a aceitação são 2,44 vezes maiores quando o contexto do SAT está presente do que quando não está. Os oficiais de admissão mostram maior propensão a recompensar estudantes de baixo status socioeconômico por uma performance maior que a média da vizinhança e da escola. Efeitos positivos também foram encontrados para estudantes de alto status socioeconômico, um indicativo, segundo os autores, de que esta política, sozinha, não seria suficiente para equilibrar o jogo entre os diferentes *backgrounds* socioeconômicos (BASTEDO et al., 2022).

Bennett (2022) examina um conjunto de cerca de 100 instituições privadas que adotaram uma política opcional de testes para entrada na graduação, entre 2005–2006 e 2015–2016. De maneira

geral, o autor justifica a importância do estudo nas discussões constantes por maior diversidade no ensino superior, em que se destacam a sub-representação de determinados grupos étnicos/raciais, e as diferenças de gênero e de status socioeconômico dos estudantes (BENNETT, 2022).

Usando CITS (*Comparative Interrupted Time Series*) e diferenças em diferenças como estratégias analíticas, o autor obtém os efeitos das políticas de teste opcional em métricas como admissão e composição do corpo estudantil. A principal pergunta de pesquisa é, se após a adoção da política de testes opcionais, as instituições experimentaram mudanças significativas na quantidade de inscritos, nas taxas de matrículas, e no perfil demográfico e socioeconômico dos estudantes, inclusive total de beneficiários dos *Pell Grant*¹, estudantes de baixa renda e mulheres (BENNETT, 2022).

O estudo encontra um aumento de 10,3% a 11,9% no número de estudantes de grupos sub-representados - mulheres, estudantes de baixa renda e de minorias raciais, com exceção de asiáticos - que se matriculam após a implementação da política. Ao mesmo tempo, como argumenta o autor, não houve mudanças significativas nas taxas de matrículas de estudantes brancos e asiáticos. Esses achados, contudo, correspondem a apenas 1% de aumento no número absoluto de estudantes sub-representados. Também oferece evidência sobre o aumento da diversidade socioeconômica, por meio do crescimento de 3,1% a 4,2% do número de bolsas *Pell Grant* que, mais uma vez, são modestos, representando apenas 1% da proporção de todos os alunos. Finalmente, encontra que a política também teve efeito no número de matrículas de mulheres, o que pode estar relacionado ao fato das mulheres se perceberem como melhores qualificadas para admissão sobre essa política. O efeito, de 4% no número total de alunos, é maior em magnitude que os efeitos somados para estudantes de grupos minoritários e estudantes beneficiários da *Pell Grant* (BENNETT, 2022).

Já o trabalho de Correa et al. (2019) investiga a importância do status socioeconômico das famílias e do *background* familiar nos resultados no ensino superior e no mercado de trabalho no Chile, e como esses estão relacionados a testes padronizados aplicados durante o ensino fundamental e testes padronizados para ingressar no ensino superior. Dessa forma, os autores contribuem para a literatura ao estudar se o *background* socioeconômico continua a influenciar os resultados durante o ensino terciário e nos estágios iniciais do mercado de trabalho. Para tanto, usam um rico conjunto de dados, contendo notas de dois testes padronizados (administrados no fim do ensino fundamental e para admissão universitária), variáveis socioeconômicas e salário inicial dos indivíduos. Usando um modelo econométrico, Correa et al. (2019) comparam os resultados acadêmicos e do mercado de trabalho, entre os estudantes que tiveram a mesma nota no teste padronizado do ensino fundamental, mas advém de contextos sociais diferentes.

Os resultados indicam que os estudantes com mais alto status socioeconômico tem notas 1,5 desvio padrão mais alto nos testes de admissão da faculdade que os alunos que tiveram os

¹Pell Grants são subsídios federais para estudantes de baixa renda que, ao contrário da maioria das ajudas federais, não precisa ser reembolsada. Faculdades e universidades devem atender a certas diretrizes federais para fazer parte do Programa Federal Pell Grant. A maioria das instituições de graduação atende a esses requisitos e cerca de 5.400 instituições se qualificam (GREELANE, 2021).

mesmos resultados acadêmicos nos testes do último ano do ensino fundamental, mas pertencem a um status socioeconômico mais baixo. Da mesma forma, entre os alunos que obtiveram notas mais altas nos testes de admissão para o ensino superior, aqueles com a situação socioeconômica mais alta estavam em famílias com salários mensais 0,7 desvios padrão maiores que aqueles com o SES mais baixo (CORREA et al., 2019). Portanto, os achados do estudo indicam que a diferença de nível socioeconômico persiste ao longo da vida dos jovens e explica diferenças de resultados de aprendizagem alcançados, mesmo em etapas educacionais mais avançadas, como a transição para entrada no ensino superior. De forma alternativa (ou complementar), também é possível que os testes padronizados SAT e ACT sejam especialmente sensíveis ao nível socioeconômico dos estudantes em função da forma como são estruturados.

Dixon-Román, Everson e McArdle (2013) contribuem com a discussão anterior ao estudar a associação entre renda familiar e desempenho no SAT. Parte dos trabalhos ao modelar a relação entre renda familiar e nota do SAT levam em consideração os seguintes pressupostos: (a) suposição de que a relação é linear (entre renda e nota do SAT); (b) suposição de que não há associação entre renda familiar e raça; e (c) não modelando explicitamente os efeitos diretos e indiretos da renda familiar após controlar pelos efeitos combinados da educação dos pais e desempenho no ensino médio. Na visão de Dixon-Román, Everson e McArdle (2013) autores, essas suposições subestimaram e simplificaram a relação existente.

Utilizando dados de estudantes do ensino médio que realizaram o exame do SAT em 2003 e estimando um modelo de equações estruturais (SEM), os autores identificam que: (a) em relação a renda familiar, a situação de pobreza trás efeitos adversos independente da raça; (b) os estudantes negros em situação de pobreza têm performance média 77 pontos abaixo no SAT enquanto os estudantes brancos, também em situação de pobreza, têm performance média de 44,2 pontos abaixo (em comparação com estudantes de famílias de renda média). Este efeito é não-linear e a diferença média se reduz a depender da faixa de renda que as famílias se encontram (DIXON-ROMÁN; EVERSON; MCARDLE, 2013).

O significativo efeito maior da renda para as famílias negras e as reduções expressivas nas diferenças de pontuação por grupo social na nota total do SAT sugerem uma influência relativa e substancial da renda familiar para viabilizar oportunidades sociais e educacionais. Essas oportunidades parecem ser especialmente limitadas para os estudantes pobres - brancos e negros - que prestaram o exame. Para os estudantes brancos, um aumento de 1 ponto no *GPA* está associado a um ganho de 180 pontos no SAT; para os estudantes negros, esse incremento é de 138 pontos. O efeito distinto para os estudante do ensino médio, juntamente com o efeito da pobreza, sugerem um efeito da escolaridade onde os candidatos negros, especialmente aqueles que vivem na pobreza, frequentam, em termos relativos, escolas de pior qualidade (DIXON-ROMÁN; EVERSON; MCARDLE, 2013).

Finalmente, Zwick (2019) sintetiza um conjunto de evidências sobre o papel dos testes admissionais nos EUA. Segundo a autora o amplo conjunto de políticas de admissão reflete o quão profunda

é a ambivalência nacional em torno dos testes de admissão. De um lado, recompensar a excelência acadêmica é considerado uma coisa boa. Nessa perspectiva, parece ser justo dar preferência às notas mais altas na alocação das vagas universitárias. Por outro, como já é sabido que os testes de admissão diferem entre grupos étnicos e socioeconômicos, os testes são vistos como antiéticos dado o princípio fundamental de igualdade de oportunidades.

A alegação de que os testes padronizados são injustos usualmente começa com uma descrição das diferenças de conceitos obtidas entre grupos distintos. Para comparar as diferenças socioeconômicas e étnicas, a autora utiliza dados do *Education Longitudinal Study* (ELS), um *survey* que inclui uma amostra nacional representativa de concluintes do ensino médio com mais de 13.000 estudantes. Os resultados apontam que há uma diferença substancial entre grupos socioeconômicos e étnicos: (i) a nota do teste é 1,1 desvio padrão mais alta para os estudantes de alto SES em comparação com os de baixo SES; (ii) a maior diferença de notas, entre asiáticos e negros, é de 1,3 desvio padrão (ZWICK, 2019).

Para a autora, o padrão generalizado de diferença entre os grupos étnicos e socioeconômicos aponta para as desigualdades sociais como a causa final. O desempenho acadêmico acaba refletindo a enorme variação nos recursos educacionais e oportunidades em todo o país. Em 2014, ano da disponibilidade dos dados, as taxas de pobreza para negros e hispânicos (30% e 36%, respectivamente), contrasta com as taxas de brancos e asiáticos (11% e 12%, respectivamente) (ZWICK, 2019).

Especificamente sobre as diferenças de performance de gênero em testes padronizados, Saygin (2020) e Saygin (2016) estuda se as diferenças de gênero existem em diferentes métodos de avaliação, utilizando dados do desempenho no ensino médio e as notas do teste padronizado para acesso ao ensino superior na Turquia. Embora a literatura aponte para uma reversão da tendência das diferenças de gênero na admissão estudantil, a diferença de notas nos testes padronizados persistem mesmo que, em média, as mulheres alcancem notas relativamente melhores que as dos homens no primeiro ano de graduação e possuam, também, desempenho mais elevado no ensino médio (SAYGIN, 2020).

Os resultados de um modelo de mínimos quadrados ordinários concebida pela autora revelam que a vantagem que as mulheres têm de melhor performance no ensino médio (3,2 pontos de diferença quando comparadas com os homens) se esvaí na comparação das notas dos testes padronizados, onde os homens possuem, em média, nota 4 pontos acima. Esses resultados se mantêm para os quantis de distribuição das notas, tanto para o GPA quanto o para o Teste de Seleção de Alunos (OSS na sigla Turca) (SAYGIN, 2020). A autora encontra, ainda, que as mulheres são menos propensas a repetir o exame e mais propensas a escolher faculdades de ranking mais baixo, revelando um comportamento de aversão ao risco. Diante destas constatações, ela argumenta que as diferenças de gênero na preferência por risco no momento de aplicar para vagas no ensino superior tem importantes implicações para as escolhas principais (SAYGIN, 2016).

Este segundo conjunto de evidências está sintetizado na tabela 3.2.

Tabela 3.2: Características dos Estudos Sobre Testes Padronizados e Heterogeneidade no Acesso ao ES

| Estudo | Objeto de estudo | Dados | Estratégia empírica | Principais resultados |
|------------------------------|--|--|--|---|
| Bastedo et al. (2022) | Relação entre relativização do desempenho dos alunos no SAT e probabilidade de admissão no ensino superior (EUA) | RCT com dados dos participantes (alunos) simulados e oficiais de admissão recrutados do Associação Nacional para Admissão na Faculdade e Aconselhamento (NACAC) | Logit | 1. Chances de recomendar a aceitação são 2,44 vezes maior quando a contextualização do SAT está presente do que quando não está. 2. Efeitos positivos para estudantes de baixo e alto status socioeconômico, respectivamente, 9% e 11% de incremento na probabilidade de recomendação de aceite |
| Bennett (2022) | Efeitos da adoção de políticas opcional de testes para entrada no ensino superior no comportamento e características demográficas dos estudantes (EUA) | Sistema Integrado de Dados do Ensino Superior (IPEDS); Secretaria de Educação Superior (Departamento de Educação Americano); Índice de Competitividade de Barron's | <i>Comparative Interrupted Time Series</i> (CITS); diferenças em diferenças | 1. Aumento de 10,9% a 11,9% no número de estudantes de grupos sub-representados que se matriculam após a implementação da política; 2. Crescimento de 3,1% a 4,2% do número de bolsas <i>Pell Grant</i> ; 3. Aumento de 4% no número total de mulheres matriculadas |
| Correa et al. (2019) | Importância do status socioeconômico das famílias e do <i>background</i> familiar nos resultados do ensino superior e dos anos iniciais no mercado de trabalho (Chile) | Sistema de Medição da Qualidade da Educação (SIMCE); Departamento de Avaliação, Mensuração e Cadastro Educacional (DEMRE); Sistema de Seguro Desemprego | Regressões de forma reduzida (<i>reduced-form regressions</i>) motivados pela construção de modelo empírico que define a forma funcional das funções estimadas | 1. estudantes com status socioeconômico mais alto tem notas 1,5 desvio padrão mais altas nos testes de admissão; 2. alunos com nota mais alta no teste de admissão recebiam salários mensais 0,7 desvio padrão maiores que aqueles com status socioeconômico mais baixo |
| Zwick (2019) | Síntese do conjunto de evidências sobre o papel dos testes admissionais na política e no mercado educacional (EUA) | Estudo Longitudinal de Educação (ELS) | Diferença de médias | 1. Diferenças substanciais entre grupos socioeconômicos e étnicos: a nota dos testes é 1,1 desvio padrão mais alta para estudantes de alto status socioeconômico, em comparação com os de baixo e, a maior diferença entre notas é entre asiáticos e negros, de 1,3 desvio padrão. |
| Saygin (2020), Saygin (2016) | Estudam a relação entre diferentes tipos de avaliação - desempenho no ensino médio e exames admissionais para acesso ao ensino superior - e as diferenças entre gênero (Turquia) | Levantamento do Conjunto de Dados de Candidatos ao OSS e Programas de Ensino Superior | Mínimos quadrados ordinários; Regressão quantílica | 1. Mulheres tem melhor performance no ensino médio que os homens (3,2 pontos de diferença) e pior desempenho nos testes padronizados (4 pontos de diferença); 2. As mulheres são menos propensas a repetir o exame e mais propensas a escolher faculdades de ranking mais baixo |

Fonte: Elaborado pelos autores.

3.3 Testes Padronizados e Responsabilização Escolar

Nos mercados de educação, os testes padronizados têm sido utilizados, há décadas, como uma prática comum para coletar informações sobre as escolas e estudantes, subsidiando decisões de políticas governamentais e um conjunto de escolhas do consumidor, como a decisão locacional (moradia) e da escola (CUESTA; GONZÁLEZ; PHILIPPI, 2020). Esse processo de avaliar a performance escolar com base no desempenho dos estudantes em sistemas padronizados de avaliação, atribuindo recompensas e punições com base nas conquistas alcançadas e no progresso empreendido afeta, diretamente, o comportamento dos agentes envolvidos - como pais, professores, alunos e gestores (CAMARGO et al., 2018; ANDRABI; DAS; KHWAJA, 2017).

Lee e Kang (2019) estudam como a introdução de testes padronizados de alto risco² impactou a responsabilização das escolas sobre os resultados educacionais. Essa introdução ocorreu por meio da política *No Student Below Basic* - NSBB (Nenhum Aluno Abaixo do Básico), que reformulou o sistema previamente existente de testes padronizados aplicados no ensino médio - *National Assessment Student Achievement* - NAEA (Avaliação Nacional de Desempenho do Aluno). A partir da mudança introduzida pela NSBB, os resultados nos testes passaram a ser utilizados em um sistema de recompensa para escolas e professores. Dessa forma, a política esperava que com esse sistema de recompensas as escolas aumentassem sua qualidade e, conseqüentemente, os estudantes alcançassem um nível de aprendizado superior. Para avaliar o efeito da mudança, os autores compararam os resultados dos testes padronizados aplicados no ensino médio com os resultados do (PISA) *Programme for International Student Assessment* como controle, o qual é um teste de baixo risco e não tem ligação com o sistema de recompensa recém-estabelecido, ou seja, não está sujeito aos incentivos introduzidos pela NSBB.

Os resultados dos testes nacionais pré-NSBB (ou seja, dentro do sistema de responsabilização escolar de baixo risco) mostram que havia pouco ou nenhum crescimento, em média, no desempenho dos alunos. Durante o período 2003–2009 houve um ganho de 0,01 desvio padrão por ano em leitura, 0,02 desvio padrão por ano em matemática no PISA, enquanto se observou um progresso também modesto nos resultados do NAEA, incremento de 0,06 desvio padrão por ano em leitura e 0,02 desvio padrão por ano em matemática. Em contraste, os resultados da NAEA mudaram substancialmente no período 2009–2012, época da vigência da política de responsabilização de alto risco, sem que isso se refletisse em ganhos de desempenho no PISA. Enquanto no NAEA houve um ganho, por ano, de 0,20 desvio padrão em leitura e 0,18 desvio padrão em matemática, no mesmo período houve uma perda de -0,03 desvio padrão em leitura e um ganho de 0,02 desvio padrão em matemática, no PISA (LEE; KANG, 2019).

As diferenças de desempenho dos alunos entre as escolas de melhor desempenho - quartil superior de distribuição - e as de pior desempenho - quartil inferior - tendem a seguir padrões

²Um teste de alto risco (*high-stakes test*) é um teste utilizado para tomar decisões importantes sobre alunos, educadores, escolas ou distritos. Em geral, alto risco significa que as pontuações dos teste são usadas para determinar punições ou recompensas. Por outro lado, um teste de baixo risco (*low-stakes test*) não trás conseqüências significativas ou públicas (EDGLOSSARY, 2021).

semelhantes de antes do NAEA e o PISA antes da reforma NSBB, em que não há redução significativa das lacunas de desempenho. Por outro lado, no período pós-reforma houve um aumento, medido pelo PISA, do *gap* entre as instituições, um aumento de 0,12 desvio padrão por ano em leitura e 0,09 em matemática. As diferenças de desempenho, contudo, caíram na avaliação interna, -0,23 desvio padrão por ano em leitura e menos 0,20 em matemática (LEE; KANG, 2019)

Assim, embora Lee e Kang (2019) não apresentem de forma clara o mecanismo, com base no desenho da política, para as discrepâncias observadas no desempenho dos estudantes no NAEA e PISA, argumentam que os incentivos positivos (recompensas) ou negativos (sanções), devem ser bem avaliados sob o risco de cair na “armadilha dos testes de alto risco”, uma vez que este tipo de política pode acarretar em consequências negativas não intencionais. Os autores concluem argumentando que quando os testes têm foco principal na habilidade de resolução de problemas reais, a escola é menos propensa a manipular o sistema e os resultados dos testes (LEE; KANG, 2019).

Leckie e Goldstein (2019) estudam a introdução, em 2016, de um novo sistema de responsabilização da escola secundária na Inglaterra. Essencialmente, uma pontuação total em oito assuntos foi introduzida como nova medida linear do desempenho do aluno ao final do ensino secundário (entre 15 e 16 anos), *Attainment 8*. Em adição, houve a introdução de uma nova medida que mensura o progresso dos alunos ao final da escolaridade primária - *Progress 8* (entre 10 e 11 anos). Os autores exploram uma decisão particularmente divisiva relevante não apenas para o *Progress 8*, mas para todas as medidas de valor acrescentado e sistema de responsabilização escolares. Para o *Progress 8*, o governo decidiu ignorar o *background* dos estudantes. Avaliam a importância prática desta decisão para responsabilidade escolar na Inglaterra, examinando até que ponto a pontuação das escolas, classificações e outros aspectos, mudam quando se leva em consideração o histórico do aluno, uma vez que na ausência de qualquer tipo de ajuste, diferentes grupos de alunos irão, tipicamente, apresentar diferentes progressos médios durante a fase escolar (LECKIE; GOLDSTEIN, 2019).

Dado o desempenho heterogêneo dos alunos, os autores levam em consideração a composição da escola com base em 7 características: (a) região de localização da escola; (b) tipo de instituição; (c) tipo de processo de admissão; (d) faixa de idade; (e) gênero; (f) religião; (g) nível de privação (pobreza). Os resultados mostram que, na consideração dessas características, 988 escolas - 32% do total, mudariam de faixa de classificação. O número de escolas considerados com desempenho abaixo do padrão mínimo pelo governo, cairia de 303 - 9,8% do total - para 196 - 6,3% do total. No outro extremo, o número de escolas designadas para a faixa acima da média diminuiria de 193 escolas - 6,2% de todas as escolas - para 148 - 4,8% do total. A intuição por trás, como argumentam os autores, é que ao serem consideradas as pontuações esperadas, menos alunos seriam considerados fazendo progressos irregulares e menos escolas seriam julgadas, conseqüentemente, como estando abaixo ou acima da média de desempenho de maneira imprecisa (LECKIE; GOLDSTEIN, 2019).

Outra vertente da literatura estuda como a disponibilidade e divulgação das informações sobre os testes padronizados tem impactos na tomada de decisão das famílias em relação a escolha escolar e

no comportamento dos agentes - professores, alunos, pais e gestores - dado os sinais enviados com a divulgação das métricas de desempenho. Nessa linha, Cuesta, González e Philippi (2020), Andrabi, Das e Khwaja (2017) e Camargo et al. (2018), apresentam importantes contribuições.

Cuesta, González e Philippi (2020) estudam as consequências que as distorções causadas pela presença heterogênea de estudantes de diferentes níveis de habilidade no dia das provas do SIMCE - *Sistema de Medición de la Calidad de la Educación* - (Sistema de Medição da Qualidade da Educação) têm na escolha da escola por parte dos agentes privados e na alocação de recursos educacionais por parte dos agentes governamentais. Os autores evidenciam que os estudantes de menor nível de habilidade comparecem com menor frequência no dia dos testes e isto acaba distorcendo a métrica utilizada para determinar a qualidade da escola e a consequente tomada de decisão por parte dos pais - e responsáveis - e do governo.

Os autores mostram que os estudantes de performance mais alta tem maior probabilidade de aparecer no dia dos testes - aumentam sua frequência em 3 p.p, enquanto os estudantes de baixa performance têm uma queda na frequência de 2 p.p. Ao estimar as distorções nas pontuações usando um método de imputação múltipla, para prever a nota dos estudantes que faltaram ao exame, os resultados encontram distorções de 0,10 desvio padrão da nota média dos testes, o que sugere que os padrões de absentismo no dia dos testes e as distorções subsequentes não são randômicas (CUESTA; GONZÁLEZ; PHILIPPI, 2020).

Para quantificar as implicações dessas distorções, os autores estimam um modelo de escolha do colégio e encontram que prover informações não distorcidas a respeito da qualidade da escola está associado a um aumento de 3% de mudanças de escola para os estudantes que fazem parte do coorte. Também mostram que dois programas públicos estão alocando recursos de maneira ineficiente devido às distorções. No primeiro programa, o governo remunera os professores das escolas que obtiveram desempenho médio suficientemente alto. Os autores realocam o bônus, baseados na remoção das distorções, e encontram que 13% dos recursos estão alocados de maneira ineficiente. No segundo programa, o governo usa a nota média das escolas para classificá-las em três grupos distintos e passar essa informação para que os pais (e responsáveis) possam tomar decisões de alocação. Usando o algoritmo de classificação encontram que 4% das escolas está classificadas de maneira incorreta e que isto leva a 2% dos estudantes estreatantes a escolher uma escola diferente (CUESTA; GONZÁLEZ; PHILIPPI, 2020).

Já Camargo et al. (2018) utilizam uma descontinuidade na regra de divulgação da pontuação do ENEM para testar se a divulgação afeta a performance dos estudantes em escolas públicas e privadas. Como argumentam, uma questão importante na literatura sobre responsabilização escolar é se a divulgação de informações é suficiente para influenciar o comportamentos dos professores e gestores ou se é necessário prover incentivos mais explícitos para esses agentes, como o pagamento performance da escola ou outros tipos de incentivos/punições. Os resultados empíricos encontrados sugerem que a divulgação dos resultados podem, por si, afetar o comportamento de professores e

gestores, desde que estes estejam sujeitos a incentivos de mercado, ou seja, o impacto da divulgação das notas difere para as instituições públicas e privadas.

Em conjunto, os resultados mostram que, em média, o efeito da divulgação das notas do ENEM das escolas foi de 0,10 a 0,17 desvio padrão maior para os alunos das escolas particulares em comparação aos alunos das escolas públicas. Ainda, o impacto da divulgação das notas é maior para os colégios particulares em mercados mais competitivos, os efeitos vão de 0,07 a 0,17 desvio padrão, o que fornece evidências de que os incentivos de mercado moldam a resposta das escolas privadas e destacam dois possíveis canais associados: (a) a divulgação das notas pode ajudar as escolas a aprender sobre sua qualidade, levando professores e gestores nas escolas tratadas a reagir às notícias sobre a qualidade das instituições; e (b) a divulgação das notas pode reduzir a seleção adversa em determinado mercado, reduzindo a assimetria de informação entre as escolas tratadas e outros participantes do mercado, o que pode levar a uma alteração do comportamento de professores e gestores (CAMARGO et al., 2018).

Na mesma linha, Andrabi, Das e Khwaja (2017) analisam os dados de um experimento que aumentou as informações, de maneira exógena, em 56 de 112 vilarejos do Paquistão (um país de baixa renda), por meio da disseminação de boletins que continham resultados dos testes das crianças e os resultados médios dos testes de todas as escolas na aldeia, em resultados de testes subsequentes, preço da educação privada e participação escolar.

Nas aldeias de tratamento, a pontuação média dos testes aumentou 0,11 desvio padrão, refletindo um ganho adicional de 42% sobre a pontuação dos testes nas aldeias de controle. Os custos da matrícula caíram nas escolas privadas em cerca de 17% em relação às escolas nas aldeias de controle. A matrícula geral entre crianças em idade primária aumentou 3 p.p nas vilas de tratamento e as escolas particulares com baixas pontuações nos testes tornaram-se mais propensas a fechar nas vilas de tratamento. Essa gama de impactos é substancial em relação a uma variedade de intervenções educacionais (geralmente mais caras) aplicadas em outros ambientes de baixa renda. Finalmente, os resultados também mostram que houve um ganho de 0,09 desvio padrão para os estudantes das escolas públicas nas vilas tratadas, em comparação com as de controle. Embora as escolas públicas enfrentem poucos mecanismos de mercado (ou administrativos), a pressão social da comunidade pode alterar o comportamento dos professores e gestores. Os autores provêm evidência consistente com esse canal, demonstrando um incremento significativo nas interações entre pais e escolas nas aldeias de tratamento, após a distribuição dos boletins (ANDRABI; DAS; KHWAJA, 2017).

A Tabela 3.3 sintetiza as principais características dos estudos descritos nesta seção.

Tabela 3.3: Características dos Estudos sobre Testes Padronizados e Responsabilização

| Estudo | Objeto de estudo | Dados | Estratégia empírica | Principais resultados |
|------------------------------------|--|--|--|--|
| Lee e Kang (2019) | Impacto dos testes de responsabilização de alto risco na melhora do desempenho dos estudantes e superar lacunas de aprendizado entre grupos distintos de estudantes | Avaliação Nacional de Desempenho do Aluno (NAEA); Programa de Avaliação Internacional de Estudantes (PISA) | Diferenças em Diferenças | 1. Durante a política de responsabilização de alto risco, houve um aumento da nota do NAEA (0,20 desvio padrão por ano em leitura e 0,18 desvio padrão por ano em matemática); 2. Durante este período, não houve, em contraste, incremento das notas do PISA; 3. No período de vigência da política houve um aumento da diferença entre as escolas de alto e baixo desempenho no PISA e uma redução no NAEA. |
| Leckie e Goldstein (2019) | Efeitos da consideração da características dos estudantes em uma métrica de responsabilização escolar (Inglaterra) | Tabelas de desempenho do ensino secundário; Banco de Dados Nacional de Alunos do Governo | Mínimos quadrados ordinários | 1. Quando ajustado pelas características dos estudantes: redução de 3,5 p.p nas escolas classificadas como abaixo da média; e redução de 1,4 p.p nas escolas classificadas como de melhor desempenho |
| Cuesta, González e Philippi (2020) | Consequência das distorções dos sinais de qualidade das escolas no mercado de educação (Chile) | Ministério da Educação - registros de escolas em funcionamento; registros estudantis; frequência escolar diária e desempenho no teste SIMCE (utilizado como medida de qualidade escolar) | Efeitos fixos; Modelo de escolha escolar | 1. Informações não distorcidas a respeito da qualidade da escola está associado a um aumento de 3% de mudanças da escola; 2. 13% dos recursos do programa de remuneração dos professores é alocado de forma errônea dada a distorção de sinal; 3. 4% das escolas estão classificadas de maneira incorreta, o que leva a 2% dos estudantes entrantes escolher uma escola diferente |
| Camargo et al. (2018) | Investigam como a divulgação da pontuação do ENEM afeta a performance dos estudantes em escolas públicas e privadas | Microdados do ENEM; Censo Escolar | Diferenças em Diferenças | 1. Efeito da divulgação das notas foi 0,10 – 0,17 desvio padrão para os alunos das escolas particulares em comparação aos alunos das escolas públicas; 2. Efeito da divulgação das notas foi maior para as escolas particulares em mercados mais competitivos (0,07 a 0,17 desvios padrão) |
| Andrabi, Das e Khwaja (2017) | Impactos da divulgação sobre o desempenho das escolas (e dos alunos) em resultados educacionais - desempenho nos testes subsequentes, preço da educação privada e participação escolar (Paquistão) | Projeto de Aprendizado e Realização Educacional nas Escolas de Punjab (SALTOS) | oi | 1. Nas aldeias de tratamento, a pontuação média dos testes aumentou 0,11 desvio padrão; 2. O custo da matrícula nas escolas particulares caiu em cerca de 17%; 3. a matrícula geral de crianças em idade primária aumentou em 3 p.p nas vilas de tratamento |

Fonte: Elaborado pelos autores.

4. Considerações Finais

Este trabalho apresentou evidências sobre a relação entre testes padronizados e um conjunto de resultados educacionais, principalmente o acesso ao ensino superior e a responsabilização escolar. Para tanto, empreenderam-se buscas na literatura acadêmica, por meio de consultas a bancos de dados de trabalhos científicos e investigação das referências bibliográficas de trabalhos-chave. Ao final do processo de busca, 17 artigos foram incluídos na revisão e agrupados em três conjuntos de resultados que tratam, respectivamente, de testes padronizados aplicados em sistemas centralizados de admissão; dos efeitos heterogêneos de testes padronizados dado o status socioeconômico, gênero e raça/etnia dos estudantes; e, finalmente, do papel de tais testes em políticas e/ou programas de responsabilização escolar.

As evidências sobre os sistemas de admissão centralizado, como é o caso brasileiro para parcela significativa das instituições federais e estaduais, destacam, de um lado, os benefícios da redução de uma série de fricções inerentes aos mercados descentralizados, como redução dos custos totais de aplicação para múltiplos centros, regras transparente de acesso e admissão, integração geográfica dos mercados de educação e democratização das oportunidades de acesso ao ensino superior a estudantes à despeito de sua condição socioeconômica (KNIGHT; SCHIFF, 2022; FRANCESCONI; SLONIMCZYK; YURKO, 2019). Por outro lado, a depender do desenho dos sistemas de admissão e das políticas educacionais em que estão inseridos, efeitos indesejáveis - e não intencionais - podem surgir, resultantes de tais sistemas, como a estratificação das instituições de ensino e deslocamento dos alunos de status socioeconômico mais baixo (CAMARGO et al., 2018).

De maneira geral, as evidências sobre os testes padronizados em sistemas de admissão centralizados destacam, principalmente, os efeitos positivos no aumento do número de inscrições, matrículas e integração geográficas das regiões de um mesmo território, além de efeitos ambíguos - ou inexistentes - referentes ao deslocamento de acesso a estudantes de baixo status socioeconômico. Contudo, vale destacar, o aumento médio da idade de admissão e uma redução do número de mulheres matriculadas.

As evidências sobre os efeitos heterogêneos da aplicação de testes padronizados e o desempenho divergente de grupos raciais/étnicos, gênero e de diferentes status socioeconômico, longe de ser um debate ultrapassado na literatura, parecem envolver um *trade-off* entre eficiência e equidade (ZWICK, 2019). De um lado, o crescimento expressivo da demanda pelo ensino superior, requer a existência de métricas que permitam comparar a habilidade e capacidade de estudantes advindos de diferentes contextos, para fins de seleção. Do outro, a baixa representação de determinados grupos da sociedade, como os já citados, em detrimento de uma sobre-representação - ou representação excessiva - de outros grupos vistos como privilegiados, aquece o debate sobre a forma como esses resultados são levados em consideração (BENNETT, 2022; ZWICK, 2019; SAYGIN, 2020).

Os resultados destacados nesta síntese, apontam para uma diferença persistente do desempenho e, conseqüentemente, na admissão entre homens e mulheres; grupos raciais/étnicos; e grupos de diferentes status socioeconômico e *background* familiar. Contudo, como destaca Zwick (2019), a diferença de notas no teste não indica, por si, que os testes são injustos, o ponto chave reside na forma como a informação será levada em consideração. No caso brasileiro, por exemplo, as ações afirmativas mitigam parte desses efeitos observados (MELLO, 2021).

Finalmente, o terceiro conjunto de evidências apresentados, destaca o papel dos testes padronizados nas políticas e programas de responsabilização escolar. Existe uma extensa literatura que relaciona as notas dos testes padronizados e o valor adicionado como medidas de qualidade para sistemas de responsabilização. Diferentes agentes - pais, professores, gestores - e as autoridades governamentais usam os resultados dos testes como indicadores de qualidade e as evidências de pesquisas sugerem que os resultados dos testes afetam, além das decisões de alocação de recursos e atenção dos governos, um conjunto de escolha das famílias, como local de moradia e escolha da escola (CUESTA; GONZÁLEZ; PHILIPPI, 2020; LEE; KANG, 2019).

As evidências destacadas sintetizam o papel que os testes possuem como métrica de performance do progresso e desempenho dos alunos - e escolas - e ressaltam o papel que a informação fornecida pelas avaliações, longe de servir apenas aos formuladores de políticas, é utilizada pelos pais, professores e gestores educacionais, alterando o comportamento desses agentes até mesmo sem a implementação de qualquer tipo de mecanismo de punição ou recompensa pelo desempenho (CAMARGO et al., 2018). Há também importante discussão sobre o formato que os incentivos são desenhados, como argumenta Lee e Kang (2019) e sobre como as métricas, ainda que padronizadas, de diferentes instituições, em distintos contextos, devem ser comparadas, como destacam Leckie e Goldstein (2019).

Concluimos essa síntese destacando que os impactos e efeitos observados de diferentes tipos de políticas são, antes de uma receita de sucesso - ou fracasso - de determinado tipo de intervenção, um resultado da combinação de elementos intrínsecos ao programa - política ou intervenção - com os elementos culturais, sociais e econômicos de cada contexto específico. Foge ao escopo desta síntese rápida julgar e avaliar a validade externa das intervenções destacadas, portanto, reiteramos que o aprimoramento na construção de políticas públicas internas voltadas para o contexto brasileiro deve seguir por mais análises qualitativas e quantitativas antes da replicação de um modelo.

Referências Bibliográficas

ANDRABI, T.; DAS, J.; KHWAJA, A. I. Report cards: The impact of providing school and child test scores on educational markets. *American Economic Review*, v. 107, n. 6, p. 1535–63, 2017. Citado 4 vezes nas páginas 20, 22, 23 e 24.

BASTEDO, M. N. et al. Contextualizing the sat: Experimental evidence on college admission recommendations for low-ses applicants. *Educational Policy*, SAGE Publications Sage CA: Los Angeles, CA, v. 36, n. 2, p. 282–311, 2022. Citado 2 vezes nas páginas 15 e 19.

BENNETT, C. T. Untested admissions: Examining changes in application behaviors and student demographics under test-optional policies. *American Educational Research Journal*, SAGE Publications Sage CA: Los Angeles, CA, v. 59, n. 1, p. 180–216, 2022. Citado 4 vezes nas páginas 15, 16, 19 e 25.

CAMARGO, B. et al. Information, market incentives, and student performance evidence from a regression discontinuity design in brazil. *Journal of Human Resources*, University of Wisconsin Press, v. 53, n. 2, p. 414–444, 2018. Citado 7 vezes nas páginas 6, 20, 22, 23, 24, 25 e 26.

CORREA, J. A. et al. The persistent effect of socioeconomic status on education and labor market outcomes: Evidence from chile’s administrative records. *Applied Economic Analysis*, Emerald Publishing Limited, 2019. Citado 4 vezes nas páginas 5, 16, 17 e 19.

CUESTA, J. I.; GONZÁLEZ, F.; PHILIPPI, C. L. Distorted quality signals in school markets. *Journal of Development Economics*, Elsevier, v. 147, p. 102532, 2020. Citado 5 vezes nas páginas 6, 20, 22, 24 e 26.

DIXON-ROMÁN, E. J.; EVERSON, H. T.; MCARDLE, J. J. Race, poverty and sat scores: Modeling the influences of family income on black and white high school students’ sat performance. *Teachers College Record*, SAGE Publications Sage CA: Los Angeles, CA, v. 115, n. 4, p. 1–33, 2013. Citado 2 vezes nas páginas 15 e 17.

EDGLOSSARY. *High-Stakes Test*. 2021. <<https://www.edglossary.org/high-stakes-testing/>>. Acessado em: 2022-09-09. Citado na página 20.

FIGLIO, D.; LOEB, S. School accountability. *Handbook of the Economics of Education*, Elsevier, v. 3, p. 383–421, 2011. Citado na página 6.

FRANCESCONI, M.; SLONIMCZYK, F.; YURKO, A. Democratizing access to higher education in russia: The consequences of the unified state exam reform. *European Economic Review*, Elsevier, v. 117, p. 56–82, 2019. Citado 6 vezes nas páginas 5, 9, 12, 13, 14 e 25.

GREELANE. *O que é um Federal Pell Grant?* 2021. <<https://www.greelane.com/pt/recursos/para-pais-e-estudantes/what-is-a-pell-grant-788453>>. Acessado em: 2022-04-09. Citado na página 16.

KAPOR, A.; KARNANI, M.; NEILSON, C. *Aftermarket frictions and the cost of off-platform options in centralized assignment mechanisms*. [S.l.], 2022. Citado 2 vezes nas páginas 11 e 14.

KNIGHT, B.; SCHIFF, N. Reducing frictions in college admissions: Evidence from the common application. *American Economic Journal: Economic Policy*, v. 14, n. 1, p. 179–206, 2022. Citado 5 vezes nas páginas 9, 10, 11, 14 e 25.

LECKIE, G.; GOLDSTEIN, H. The importance of adjusting for pupil background in school value-added models: A study of progress 8 and school accountability in england. *British Educational Research Journal*, Wiley Online Library, v. 45, n. 3, p. 518–537, 2019. Citado 3 vezes nas páginas 21, 24 e 26.

LEE, J.; KANG, C. A litmus test of school accountability policy effects in korea: cross-validating high-stakes test results for academic excellence and equity. *Asia Pacific Journal of Education*, Taylor & Francis, v. 39, n. 4, p. 517–531, 2019. Citado 5 vezes nas páginas 3, 20, 21, 24 e 26.

MACHADO, C.; SZERMAN, C. Centralized college admissions and student composition. *Economics of Education Review*, Elsevier, v. 85, p. 102–184, 2021. Citado 8 vezes nas páginas 3, 5, 7, 8, 9, 10, 14 e 30.

MELLO, U. Affirmative action and the choice of schools. 2021. Citado 4 vezes nas páginas 10, 13, 14 e 26.

SAYGIN, P. O. Gender differences in preferences for taking risk in college applications. *Economics of Education Review*, Elsevier, v. 52, p. 120–133, 2016. Citado 2 vezes nas páginas 18 e 19.

SAYGIN, P. O. Gender bias in standardized tests: evidence from a centralized college admissions system. *Empirical Economics*, Springer, v. 59, n. 2, p. 1037–1065, 2020. Citado 5 vezes nas páginas 3, 5, 18, 19 e 25.

SZERMAN, C. *The effects of a centralized college admission mechanism on migration and college enrollment: evidence from Brazil*. Tese (Doutorado), 2015. Citado na página 5.

TANAKA, M.; NARITA, Y.; MORIGUCHI, C. Meritocracy and its discontent: Long-run effects of repeated school admission reforms. *Research Institute of Economy, Trade and Industry (RIETI)*, 2020. Citado 4 vezes nas páginas 5, 11, 12 e 14.

ZWICK, R. Assessment in american higher education: The role of admissions tests. *The ANNALS of the American Academy of Political and Social Science*, SAGE Publications Sage CA: Los Angeles, CA, v. 683, n. 1, p. 130–148, 2019. Citado 5 vezes nas páginas 17, 18, 19, 25 e 26.

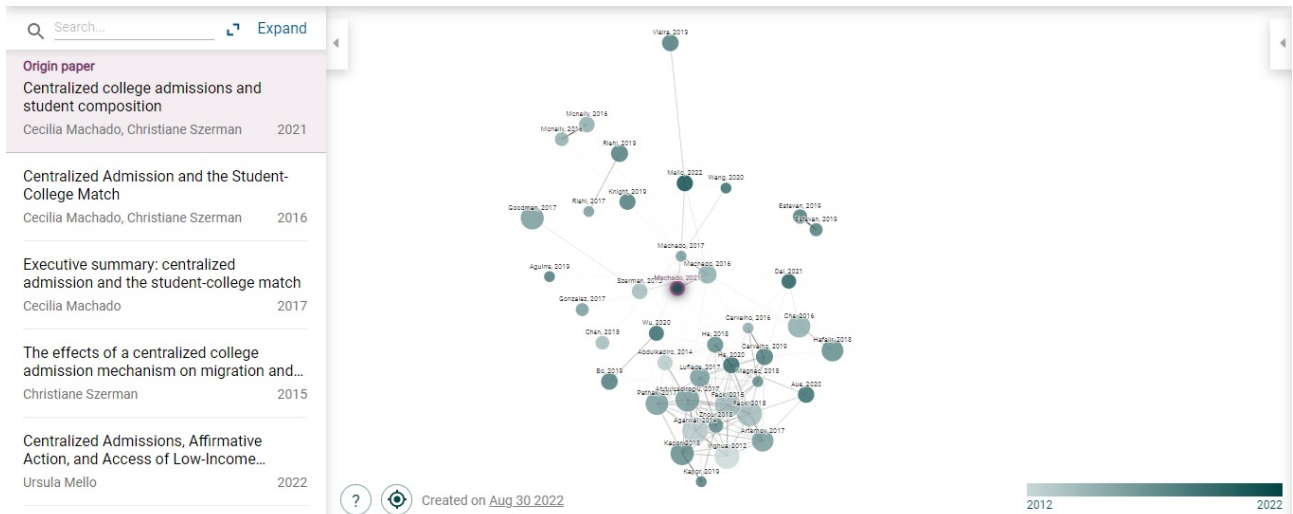
Apêndice

Tabela A1: Resultados da busca de literatura nas bases de dados

| Base de Dados | Expressões de Buscas | Resultados |
|-----------------------|--|------------|
| <i>Scopus</i> | ((standardized test) OR (standardized testing) OR (test score)) AND ((college access) OR (college admission)) AND (effect) OR (evidence) OR (impact)) | 135 |
| <i>Scopus</i> | ((standardized test) OR (standardized testing) OR (test score)) AND ((school accountability) OR (test-based accountability) OR (accountability performance) OR (educational market) OR (market incentive)) AND ((effect) OR (evidence) OR (impact)) | 73 |
| <i>Web of Science</i> | ((standardized test*) AND (college access)) OR ((standardized test*) AND (college admission)) OR ((test score) AND (college access)) OR ((test score) AND (college admission)) | 60 |
| <i>Web of Science</i> | ((standardized test*) AND (school accountability)) OR ((standardized test*) AND (test-based accountability)) OR ((standardized test*) AND (accountability performance)) OR ((standardized test*) AND (educational market)) OR ((standardized test*) AND (market incentive)) OR ((test score) AND (school accountability)) OR ((test score) AND (test-based accountability)) OR ((test score) AND (accountability performance)) OR ((test score) AND (educational market)) OR ((test score) AND (market incentive)) | 36 |

Fonte: Elaborada pelos autores.

Figura A1: Trabalhos relacionados a (MACHADO; SZERMAN, 2021)



Notas: Cada nó representa um trabalho acadêmico relacionado ao trabalho de origem. Os trabalhos são arranjados de acordo com sua similaridade, o tamanho nó representa o número de citações e a cor do nó o ano de publicação. Trabalhos mais próximos (similares) tem linhas de conexão mais fortes e estão mais próximos.

Evidência Express

EVEX **ENAP**