

## ÁREA 12 - Economia Social e Demografia Econômica

### O impacto da Educação Infantil sobre o desempenho dos alunos brasileiros no Pisa<sup>1</sup>

Guilherme Irffi<sup>2</sup>

Márcio Aurélio Frota Pereira<sup>3</sup>

#### Resumo

Como o começo da história pode ter impacto ao longo da vida. Partindo dessa premissa, este artigo analisa o efeito da educação infantil sobre o desempenho dos estudantes brasileiros no PISA a partir dos métodos *Propensity Score Matching* (PSM) e *Propensity Score Generalized* (PSG). Quanto mais cedo o ingresso na educação infantil, comparado aos que entraram mais tardiamente, maior o desempenho no PISA. Assim, ingressar na creche, em comparação ao ingresso somente na pré-escola, alavanca a performance nessa prova. Por outro lado, ingressar na pré-escola, em relação a entrada diretamente no fundamental, aumenta a performance no PISA. Ao expor alunos a diferentes níveis de escolaridade, constatou-se, pelo método PSG, que existem efeitos heterogêneos sobre a performance. Com isso, o aumento de anos de escolaridade no ensino infantil, a partir de 2 até 5 anos, pode alavancar a performance, ao passo que, na travessia do 5 até os 6 anos a performance se reduz. Por sua vez, o retorno marginal do desempenho é crescente no acréscimo de escolaridade nos anos iniciais, isso até 2 anos em matemática e 3 anos em ciências e em leitura. Em síntese, verifica-se que a educação infantil pode contribuir para a melhora nos resultados educacionais.

**Palavras-Chave:** Educação Infantil, Efeito Creche, Efeito Pré-escola, Desempenho no PISA.

#### Abstract

How the beginning of the story can have an impact throughout life. Based on this premise, this article analyzes the effect of early childhood education on the performance of Brazilian students in PISA using the Propensity Score Matching (PSM) and Propensity Score Generalized (PSG) methods. The earlier the entry into early childhood education, compared to those who entered later, the higher the performance in PISA. Thus, entering day care, compared to entering preschool only, leverages performance in this test. On the other hand, entering preschool, in relation to entering elementary school directly, improves performance in PISA. By exposing students to different levels of education, it was found, by the PSG method, that there are heterogeneous effects on performance. As a result, increasing the number of years of schooling in early childhood education, from 2 to 5 years old, can boost performance, while performance drops from 5 to 6 years old. In turn, the marginal return to performance is increasing in the increase in schooling in the initial years, that is, up to 2 years in mathematics and 3 years in science and reading. In summary, it appears that early childhood education can contribute to improving educational outcomes.

**Keywords:** Education, Early Childhood, Impact Assessment, PISA Performance.

**Classificação JEL:** I21; I26

---

<sup>1</sup> Os autores agradecem aos comentários e sugestões da Janaína Feijó, Maitê Rimekká Shirasu, Marco Túlio Aniceto, Rafael Barros e Zilania Mariano. GErros e omissões são de nossa responsabilidade.

<sup>2</sup> Professor no Programa de Pós-graduação de Economia – CAEN/UFC. E-mail: [irffi@caen.ufc.br](mailto:irffi@caen.ufc.br)

<sup>3</sup> Doutor em Economia - UFC/CAEN. E-mail: [marcio.frota25@gmail.com](mailto:marcio.frota25@gmail.com)

## 1. Introdução

A primeira infância é uma fase primordial do processo educacional. Nesse período, o cérebro se desenvolve muito rapidamente, sendo sensível aos estímulos parentais e à influência externa (KNUDSEN, 2004). As primeiras experiências em casa ou em ambientes com cuidado e/ou educação moldam o desenvolvimento e a qualidade da arquitetura cerebral (YOSHIKAWA, 2013). Esses cuidados impulsionam uma série de habilidades, incluindo as cognitivas e as socioemocionais (FELFE; LALIVE, 2010). Além disso, impulsionam de forma cumulativa as habilidades necessárias para ter bons resultados futuros na escola (CARNEIRO; HECKMAN, 2003; CUNHA *et al.*, 2006). O que pode indicar que “aprender gera aprendizagem”, gerando benefícios no médio e no longo prazo (NORES; BARNETT, 2010).

Existe uma relação positiva da educação na primeira infância com os resultados educacionais futuros (BERLINSKI; GALIANI; GERTLER, 2009; BIETENBECK; ERICSSON; WAMALWA, 2019; DIMISKI, 2021; ESPING-ANDERSEN *et al.*, 2012; MCCOY *et al.*, 2017; SCHÜTZ, 2009; PHOLPHIRUL, 2017). Uma forma de mensurar os resultados educacionais acontece por meio de exames padronizadas, e o *Programme for International Student Assessment (PISA)* é uma delas.

Os resultados apresentam elevada desigualdade de performance acadêmica, sendo vários os fatores que afetam a performance no PISA; e, algum desses são, a origem, a qualidade dos professores e das instituições (FUCHS; WÖSSMANN, 2006), a condição socioeconômica (OCDE, 2019), a relação do aluno com o professor (YI; LEE, 2017), dentre outros.

A educação na primeira infância, de acordo com a literatura internacional tem se dedicado a explicar o papel da frequência escolar na primeira infância sobre a performance no PISA, isso tanto utilizando todos ou quase todos os países<sup>4</sup> que participam da avaliação (DEL BOCA; MONFARDINI; SEE, 2018); DIMISKI, 2021; SCHÜTZ, 2009) ou uma nação específica, como a Tailândia (PHOLPHIRUL, 2017) ou a Grécia (KARAKOLIDIS; PITSIA; EMVALOTIS, 2016). No entanto, esses estudos são limitados devido não explorarem as etapas da educação infantil (creche e pré-escola) e sim, “apenas”, se o estudante na avaliação do PISA, quando criança, frequentou um ano ou mais a educação infantil.

Os estudos nacionais corroboram com a literatura internacional, os quais constataram uma associação positiva entre educação na primeira infância e desempenho acadêmico. No entanto, em sua grande maioria, se limitam a bases nacionais como a do SAEB<sup>5</sup> (CURI; MENEZES-FILHO, 2009; MARQUES; CUNHA, 2020; PINTO; SANTOS; GUIMARÃES, 2017; REYNA; SILVA; LÍRIO, 2019), por vezes, a da Prova Brasil (SILVA JÚNIOR; GONÇALVES, 2016) ou as duas (FELÍCIO; VASCONCELLOS, 2007). São escassos os trabalhos que utilizam outras fontes de dados (BARROS *et al.*, 2020; CORREA; COMIN; TAI, 2014; MARKUS; OLIVEIRA; NISHIMURA, 2020).

A educação na primeira infância está em constante mudança, buscando desenvolver o físico, o psicológico, a inteligência e o social. Em outros termos, pode-se dizer que a educação é um fator de desenvolvimento de capital humano (CUNHA; HECKMAN, 2007). E, educação de qualidade, além de um direito, ainda, apresenta um elevado retorno em relação ao investimento (CUNHA *et al.* 2006; TEMPLE; REYNOLDS, 2007).

No Brasil, a educação infantil é constituída por duas etapas para as crianças de 0 a 5 anos. A creche (crianças de 0 a 3 anos de idade) é uma etapa não obrigatória, porém, conforme o PNE<sup>6</sup> (Plano Nacional de Educação) 2014-2024 deve-se alcançar 50% das crianças matriculadas até 2024. Por sua vez, a pré-escola (crianças de 4 a 5 anos de idade) se tornou obrigatória em 2013, com a implementação da Lei nº 12.796, a qual ajustou a Lei nº 9.394 de 1996 (LDB - Lei de Diretrizes e Bases da educação Nacional) à Emenda Constitucional de 2009. No entanto, essa fase educacional apresenta além da desigualdade de acesso, outros fatores relacionados a desigualdade de oferta em função da qualidade, o que pode comprometer o desenvolvimento da criança.

---

<sup>4</sup> Avaliação aplicada em países participantes da OCDE (Organização para a Cooperação e Desenvolvimento Econômico) e não participantes, que, atualmente, é o caso do Brasil.

<sup>5</sup> Sistema Nacional de Avaliação da Educação Básica.

<sup>6</sup> O PNE estabelece objetivos, diretrizes e metas para as políticas educacionais, por um período de dez anos, para as políticas educacionais, e em todos os níveis e modalidades de ensino.

A taxa de atendimento no Brasil, segundo o observatório da criança e do adolescente<sup>7</sup>, em 2020, alcançou 29,8% das crianças entre 0 e 3 anos de idade, isto é, menos de um terço estavam matriculadas em creches. Na pré-escola, segundo o observatório, a taxa de matrícula foi de 83,4% para crianças entre 4 e 5 anos de idade. Há uma evolução em relação aos anos anteriores, todavia, essas taxas devem crescer e, ainda, promover a equidade educacional por forma de acesso aos cuidados e educação nessa fase, contribuindo para o desenvolvimento e formação desses.

Diante do exposto, o artigo contribui com o debate sobre a educação na primeira infância, com ênfase nos impactos gerados sobre o desempenho dos alunos brasileiros no PISA. De forma mais específica, são analisados os ingressos em cada etapa da educação infantil, isto é, se foram atendidos desde a creche vis a vis se frequentaram apenas pré-escola, ou se ingressaram mais tardiamente – diretamente no fundamental. Além da etapa, busca-se averiguar se o nível de exposição à educação na primeira infância apresenta diferentes retornos educacionais.

Para atingir esses objetivos, utiliza-se a base de dados do Brasil no PISA de 2018. Para captar a diferença de performance entre os grupos foram utilizados os métodos *Propensity Score Matching* (PSM) para captar o efeito sobre a média, e o *Propensity Score Generalized* (PSG) para testar a hipótese de heterogeneidade no efeito educacional. Por fim, faz-se uso da análise de sensibilidade de Rosenbaum (2002) para testar a robustez dos resultados.

Além desta introdução, este artigo é composto por mais quatro seções. A próxima apresenta uma revisão da literatura e trata da evolução da educação infantil no Brasil. A descrição da base de dados e da estratégia de identificação, bem como modelagem empírica são expostas na terceira sessão. Em seguida, são apresentados e discutidos os resultados. E, por fim, são tecidas as considerações finais.

## **2. Revisão de Literatura**

### **2.1 Primeira Infância e Performance Acadêmica**

A primeira infância é um período crucial e que necessita de acompanhamento para um pleno desenvolvimento da criança junto à sociedade. Devido a suma importância dessa fase, muitos estudos voltaram sua atenção para os cuidados e/ou educação nos primeiros anos. Um desses é o de Burchinal *et al.* (1997), o qual relatam que cuidados intensivos na primeira infância estão associados a melhores desempenhos cognitivos. Em contrapartida, a falta de cuidado nessa etapa impacta de forma desastrosa sobre resultados educacionais futuros (SCHADY; 2006). De acordo com Cunha e Heckman (2007), as lacunas de habilidades podem ser reduzidas a partir de intervenções voltadas aos primeiros anos. Essas intervenções têm um grande retorno educacional, ao passo que, intervenções mais tardias no ciclo de vida (corretivas ou compensatórias) têm menor retorno (Carneiro e Heckman, 2003).

Nota-se que a educação infantil pode contribuir para resultados educacionais futuros. Segundo Berlinski, Galiani e Gertler (2009), um ano de cuidado infantil aumenta em 23% desvio padrão o desempenho dos alunos. Já Felfe e Lalive (2010) constatam que ter sido atendido em creches alavancam as habilidades cognitivas em crianças de 2 ou 3 anos ou de 5 a 10 anos. Por sua vez, Nores e Barnett (2010) indicam que as crianças que tiveram alguma assistência na primeira infância apresentaram ganhos cognitivos e de escolaridade os quais são mantidos ao longo do tempo. E, segundo Cortázar *et al.* (2020), a educação nessa etapa impacta de forma positiva não apenas sobre a performance acadêmica, mas também sobre as chances de repetência ou abandono escolar, reduzindo-as no longo prazo.

Voltando-se para esse efeito, só que direcionado a performance no PISA, tem-se o estudo de Schütz (2009), que em seus achados, observou que há uma relação positiva entre a frequência na educação infantil com os resultados no PISA, sendo maior em países que possuem uma porcentagem mais elevada de crianças que frequentam a primeira infância em instituições privadas. Por sua vez, Pholphirul (2017) também relata a mesma tendência, mas considera apenas os dados da Tailândia. Em contrapartida, de acordo com Dimiski (2021), a não frequência à educação infantil é um determinante robusto com impacto negativo no desempenho dos alunos em ciências.

---

<sup>7</sup> Os dados foram obtidos por meio de consulta *online* no *website* do Observatório da Criança e do Adolescente mantido pela Fundação Abrinq.

Em relação à literatura brasileira da área, há uma gama de estudos que comprovaram o impacto positivo da educação infantil sobre a performance (BARROS *et al.*, 2020; CORREA; COMIN; TAI, 2014; CURI; MENEZES-FILHO, 2009; FELÍCIO; VASCONCELLOS, 2007; MARQUES; CUNHA, 2020; REYNA; SILVA; LÍRIO; 2019; SANTOS, 2015; SILVA JÚNIOR; GONÇALVES, 2016). Um desse é o de Felício e Vasconcellos (2007), o qual constataram que ingressar na educação infantil alavanca a performance em matemática. Por outro lado, Silva Júnior e Gonçalves (2016) relatam impacto semelhante, mas na proficiência da língua portuguesa. Segundo Felício e Vasconcellos (2007), se caso 100% das crianças frequentam a educação infantil, seria possível elevar em mais de 11% a proficiência. Segundo Curi e Menezes-Filho (2009), o atendimento na pré-escola alavanca as chances de conclusão do ensino fundamental I e II, do médio e do universitário, o que, em seus achados, aumentou em um ano e meio o grau de escolaridade. Já, segundo os autores, o atendimento em creche, por sua vez, tem influência menor, contribuindo na conclusão “apenas” do ensino médio e do universitário.

No entanto, segundo Barros *et al.* (2020), esse desenvolvimento cognitivo está relacionado à qualidade da creche a qual a criança frequentou. Além disso, à condição socioeconômica (CORREA; COMIN; TAI, 2014). Sendo que as crianças vulneráveis são as que mais se beneficiam da passagem pela pré-escola (SANTOS; 2015). Contudo, Pinto, Santos e Guimarães (2017) destacam que a direção do impacto da educação infantil sobre a performance está relacionado ao nível educacional da mãe. Em síntese, nota-se que a educação na pré-escola pode influenciar as performances acadêmicas futuras.

## 2.2 Evolução da Educação Infantil no Brasil

O primeiro grande marco da educação infantil brasileira se deu com a Constituição Federal de 1988, que tornou o atendimento de crianças de 0 a 6 anos em creches e escolas dever do estado. A LDB de 1996 integrou a educação infantil à educação básica, equiparando essa etapa ao ensino fundamental e médio. Os Art 29º e 30º da LDB estabelecem que a educação infantil é a primeira etapa da educação básica, a qual deve ter como finalidade o desenvolvimento integral da criança até os cinco anos, nos seus aspectos físico, psicológico, intelectual e social. Por sua vez, os Art 11º, 12º e 13º, estabelecem que compete aos municípios ofertar educação infantil, em creches e em pré-escolas, e que cada centro educacional fica responsável pela elaboração e execução das propostas pedagógicas. Além disso, no Art. 62º, é estabelecido que o professor, para lecionar na educação básica, deve possuir nível superior, sendo este em curso de licenciatura plena.

Em 1998, o Ministério da Educação e Cultura (MEC), em conjunto com especialistas e representantes de todos os estados organizaram o documento Subsídios para o Credenciamento e Funcionamento das Instituições de Educação Infantil, o qual contribuiu para a formação de normas e de diretrizes relacionadas à Educação Infantil. Simultaneamente, foi executada, pelo MEC, uma pesquisa para conhecer as propostas pedagógicas-curriculares adotadas em todo o país e os princípios que norteavam essas práticas. Com os resultados da pesquisa, o MEC elaborou o Referencial Curricular Nacional para Educação Infantil (RCNEI), buscando oferecer uma base nacional comum para os currículos, apesar de não ser obrigatória.

A Lei nº 10.172 de 2001, implementou o PNE 2001-2010, e trouxe a preocupação de assegurar, em todos os municípios, recursos para a educação infantil. Desse modo, a meta 18 do PNE indica que, de forma prioritária para educação infantil, deve-se destinar 10% dos recursos de desenvolvimento e manutenção do ensino não vinculados ao FUNDEF<sup>8</sup> e outros disponíveis a cada ente municipal. Em 2006 houve mudança na Emenda Constitucional nº 53, que antecipou o acesso ao ensino fundamental para os 6 anos de idade, passando a educação infantil a contemplar a faixa etária de 0 a 5 anos.

A educação infantil, embora reconhecida como direito de todas as crianças e dever do estado, passou a ser obrigatória para as crianças de 4 e 5 anos apenas com a Emenda Constitucional nº 59 de 2009, que estabeleceu a obrigatoriedade da educação básica dos 4 aos 17 anos. Além disso, estabeleceu-se as Diretrizes Curriculares Nacionais da Educação Infantil (DCNEI) que são um conjunto de definições doutrinárias sobre princípios, fundamentos e procedimentos na educação básica que orientam as escolas na organização, articulação, desenvolvimento e avaliação de suas propostas pedagógicas. No entanto, apenas

---

<sup>8</sup> Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério.

em 2013, com a Lei nº 12.796, houve a implementação da obrigatoriedade na LDB, no qual, indica que crianças de 4 e 5 anos devem estar obrigatoriamente matriculadas em instituições de educação infantil.

Em 2014, há o PNE 2014-2024, instituído pela Lei nº 13.005, o qual estabeleceu como meta universalizar, até 2016, a educação infantil na pré-escola para as crianças de 4 (quatro) a 5 (cinco) anos de idade e ampliar a oferta de educação infantil em creches de forma a atender, no mínimo, 50% (cinquenta por cento) das crianças de até 3 (três) anos até o final da vigência deste PNE.

Outro marco de grande relevância adveio da inclusão da educação infantil na Base Nacional Comum Curricular (BNCC), sendo que sua última versão foi publicada em 2017. Essa direciona o que deve ser ensinado nas escolas, englobando todas as fases da educação básica, desde a educação infantil até a fundamental. Assim, no documento, há o reconhecimento das creches e pré-escolas como espaços primordiais para o desenvolvimento da criança visto que, muitas vezes, é onde ocorre a primeira separação dos pequenos com os seus vínculos familiares. Desse modo, as instituições de ensino de educação infantil têm, como principal objetivo, ampliar o universo de experiências, conhecimentos e habilidades das crianças, diversificando e consolidando novas aprendizagens, atuando de maneira complementar à educação familiar. Além disso, o documento aborda a importância do brincar e as interações na primeira infância, estabelecendo seis direitos de desenvolvimento e aprendizagem no ensino infantil: conviver, brincar, participar, explorar, expressar e autoconhecimento.

### 3. Notas Metodológicas

#### 3.1 Fonte e descrição dos Dados

Para estimar o efeito da educação infantil sobre o desempenho futuro, utilizam-se os dados do exame do PISA de 2018. Como indicadores de impacto são consideradas as notas de proficiência em matemática, em leitura e em ciências, as quais são contínuas, variando de 0 a 1.000.

A resposta do aluno sobre o momento de ingresso na creche ou na pré-escola é utilizada como estratégia de identificação. Após retirar os valores faltantes, a base final é composta por 3.310 estudantes, podendo variar devido ao momento de ingresso na educação infantil.

O período de duração de cuidados e/ou educação infantil na primeira infância considerado foi de 6 anos, sendo constituído pela creche de 0 a 3 anos e pela pré-escola de 4 a 6 anos. Destaca-se que embora a Lei nº 11.274/2006 tenha estabelecido que a partir de 2010 a educação infantil fosse composta por 5 anos de duração, considerou-se a duração da educação infantil antes dessa lei. Haja vista que os alunos que participaram do PISA 2018, possuíam, em média, de 15 anos e dois meses a 16 e três meses, esses nasceram entre 2002 e 2003, tendo o período de educação infantil de 6 anos.

Para atender em parte aos objetivos, cruzou-se os seguintes grupos de alunos: 1) os que ingressaram na creche *versus* os que ingressaram na pré-escola; 2) os que ingressaram na pré-escola *versus* os que ingressaram diretamente no ensino fundamental; 3) os que tiveram educação durante toda a primeira infância (creche e pré-escola) *versus* os que ingressaram diretamente no ensino fundamental.

Para construir os grupos de comparação fez-se uso das respostas dos alunos sobre a idade de ingresso na educação da primeira infância, como se observa no Quadro 1. Além disso, a idade de ingresso é utilizada para captar o número de anos de educação nessa fase. Para os alunos que ingressaram entre 0 e 1 ano de idade tiveram 6 anos de atendimento na primeira infância, enquanto que os que ingressaram com 1 ano tiveram 5 anos e, assim, sucessivamente.

Quadro 1: Variáveis dependentes e grupos de tratamento e controle

Dependentes	Descrição	Valor da variável			
<i>PVIMATH</i>	Nota de matemática	0 a 1000			
<i>PVIREAD</i>	Nota de leitura	0 a 1000			
<i>PVISCIE</i>	Nota de ciências	0 a 1000			
Tratamentos binários					
Grupo de Tratamento			Grupo de Controle		
Grupos	Idade que ingressou na escola	0 a 3 anos	4 a 6 anos ou mais	Sem E.I.	
Creche x Pré Escola	0 a 3 anos	Não	Sim	Não	
Pré-Escola x sem E.I.	4 a 6 anos	Não	Não	Sim	

E. I. x sem E.I.	0 a 3 + 4, 5 e 6	Não	Não	Sim
Tratamento contínuo				
Variável				Valor da variável
nanos	Número de anos de cuidados e/ou educação na primeira infância			1 a 6 anos ou mais

Fonte: Elaboração própria. Nota: E.I. refere-se à educação infantil completa.

O Quadro 2 apresenta os controles utilizados para suprir o viés de variável omitida e satisfazer a condição de independência condicional nas metodologias empregadas. Dentre os controles, estão as características dos alunos, o ano do ensino que o aluno cursava, a educação do pai e da mãe, a repetência. Além disso, índices como o de *status* ocupacional parental, o de posses em casas, o de posses culturais, o de grau econômico, social e cultural e, por fim, o de riqueza. Ainda fazem parte dos controles questões direcionadas ao responsável como a de leituras *online*<sup>9</sup> e o motivo do aluno tutelado ter sido atendido na primeira infância. Por fim, tem os controles a nível escolar, que captam a região, o total do número de professores e a razão do número de alunos pela quantidade de professores na escola.

Os índices supracitados são formados por transformações aritméticas ou recodificação de um ou mais itens. Podendo ser variáveis derivadas com base na escala de TRI. Um desses é a variável HISEI, formado a partir das respostas sobre as ocupações dos pais, codificadas conforme a classificação de quatro dígitos da *International Standard Classification of Occupations (ISCO)*. Os valores são mapeados em uma escala de intervalo de acordo com a *International Socio-Economic Index (ISEI)* do *status* ocupacional que é desenvolvida como uma escala que reflete o *status* socioeducacional.

A *ESCS* é uma pontuação composta de três componentes principais: educação dos pais, maior *status* educacional dos pais e posses em casa. Por sua vez, *HOMEPOS* é um índice construído a partir de respostas dos alunos sobre a disponibilidade de 16 itens domésticos em casa, incluindo três itens específicos do país que foram vistos como medidas adequadas de riqueza familiar dentro do contexto do Brasil. Sendo um índice que resume todos os itens domésticos e de posse.

A *CULTPOSS* segue uma linha parecida a *HOMEPOS*, só que voltado para a cultura, sendo um índice composto pelos seguintes questionamentos: 1) se tem livros de literatura; 2) se tem livros de poesia; 3) se tem trabalhos de arte; 4) se tem livros de arte; e, 5) se tem instrumentos musicais.

O índice *WEALTH* é uma pontuação baseada no dimensionamento da TRI, que tem por base os seguintes questionamentos voltados para o aluno: se há uma sala só sua para estudos, se há um *link* de *internet*, se há televisões, se há carros, se há quartos com banheira ou chuveiro, se há computadores e se há tablets. Por fim, tem-se a variável *STRATIO*, que é a média aritmética do número de alunos na escola, ponderada pelo número de professores, sendo considerados os pesos de 0,5 quando o professor trabalha em regime de tempo parcial e de 1 quando trabalha em regime de tempo integral.

## Quadro 2: Descrição dos controles

	<i>gen</i>	1 se for menina, 0 caso contrário.
	<i>idad</i>	Idade em anos do aluno.
	<i>ano1</i>	1 se estiver cursando o primeiro ano do ensino médio, 0 caso contrário.
	<i>ano2</i>	1 se estiver cursando o segundo ano do ensino médio, 0 caso contrário.
	<i>ano3</i>	1 se estiver cursando o terceiro ano do ensino médio, 0 caso contrário.
Controles ao nível do aluno	<i>edm1</i>	1 se a mãe possui o ensino fundamental completo, 0 caso contrário.
	<i>edm2</i>	1 se a mãe possui o ensino médio completo, 0 caso contrário.
	<i>edm3</i>	1 se a mãe possui o ensino superior completo, 0 caso contrário.
	<i>edp1</i>	1 se a pai possui o ensino fundamental completo, 0 caso contrário.
	<i>edp2</i>	1 se a pai possui o ensino médio completo, 0 caso contrário.
	<i>edp3</i>	1 se a pai possui o ensino superior completo, 0 caso contrário.
	<i>HISEI</i>	Índice do mais alto <i>status</i> ocupacional parental.
	<i>REPEAT</i>	1 se repetiu algum ano, 0 caso contrário.
	<i>ESCS</i>	Índice do mais alto grau econômico, social e cultural.

<sup>9</sup> No questionário direcionado ao responsável nas questões relacionadas à leitura, pode-se marcar quatro alternativas: 1ª eu não conheço o que é isso; 2ª nunca ou quase nunca; 3ª diversas vezes em um mês e 4ª diversas vezes em um dia. A Partir dessas, criou-se binárias, assim, se o responsável marcou a partir da segunda alternativa recebeu 1, caso contrário 0.

	<i>HOMEPOS</i>	Índice de posses em casa.
	<i>CULTOS</i>	Índice de posses culturais.
	<i>WEALTH</i>	Índice de riqueza familiar.
Controles ao nível parental	<i>ler1</i>	1 se ler <i>e-mails</i> , 0 caso contrário
	<i>ler2</i>	1 se participa de <i>chats online</i> ( <i>WhatsApp</i> , <i>Messenger</i> ), 0 caso contrário.
	<i>ler3</i>	1 se ler notícias <i>online</i> , 0 caso contrário.
	<i>ler4</i>	1 se pesquisa por informações <i>online</i> para aprender sobre um tópico particular, 0 caso contrário.
	<i>ler5</i>	1 se participa de discussões em grupo ou fóruns <i>online</i> .
	<i>ricep2</i>	1 se o filho teve cuidados e/ou educação na primeira infância devido eu/nós não podia/podíamos cuidar da criança, 0 caso contrário.
	<i>ricep3</i>	1 se o filho teve cuidados e/ou educação na primeira infância devido eu/nós queria/queríamos um estímulo de aprendizagem adicional para a criança, caso contrário.
	<i>ricep4</i>	1 se o filho teve cuidados e/ou educação na primeira infância devido a maioria das outras crianças frequentarem na época, 0 caso contrário.
Controles ao nível escolar	<i>centro_oeste</i>	1 se a escola localiza-se na região centro-oeste, 0 caso contrário.
	<i>nordeste</i>	1 se a escola localiza-se na região nordeste, 0 caso contrário.
	<i>norte</i>	1 se a escola localiza-se na região norte, 0 caso contrário.
	<i>sul</i>	1 se a escola localiza-se na região sul, 0 caso contrário.
	<i>sudeste</i>	1 se a escola localiza-se na região sudeste, 0 caso contrário.
	<i>escp</i>	1 se a escola é pública, 0 caso contrário.
	<i>TOTAT</i>	Total do número de professores.
	<i>STRATIO</i>	Proporção do número de alunos na escola em relação ao número de professores.

Fonte: Elaborado pelos autores.

## 3.2 Estratégias Empíricas

### 3.2.1 Propensity Score Matching (PSM)

A estratégia de identificação proposta é comparar os grupos de tratados e controles apresentados no Quadro 1 a fim de estimar o impacto de ter ingressado na creche, na pré-escola ou de ter tido toda a educação infantil sobre a performance nas provas do PISA. Para tanto, faz-se uso do método PSM, desenvolvido por Rosenbaum e Rubin (1983) buscando equiparar a probabilidade de um grupo de receber o tratamento, considerando as diversas características observáveis,  $X$ , em comum entre os dois grupos. Assim, presume-se que cada membro do grupo de tratados possui um equivalente no grupo de controle. Para evitar o problema de dimensionalidade, substitui o vetor  $X$  de características observáveis pela sua probabilidade  $p(X)$ , que é denominado de *escore de propensão*:

$$p(X) = Pr(T = 1|X) \quad (1)$$

Considerando a hipótese de seleção nos observáveis, a independência entre o resultado potencial de não ter sido tratado e a decisão de participação do programa serão válidos. Assim,

$$Y_i(0) \perp T_i|X \Rightarrow Y_i(0) \perp T_i|p(X) \quad (2)$$

onde  $Y_i$  é resultado potencial da variável de impacto (desempenho nas provas do PISA de 2018),  $T_i$  é o tratamento (algum grau de educação na primeira infância) e  $X_i$  é vetor de variáveis observáveis.

Essa técnica capta o efeito médio de tratamento nos tratados fazendo o pareamento entre indivíduos que têm certo grau de educação infantil com outro que tem menor ou nenhum, utilizando o *escore de propensão*. No entanto, como esse é desconhecido, torna-se necessário aplicar um modelo de estimativa binário, como o *logit* ou o *probit*. No presente estudo, utilizou-se o *logit*, que pode ser visualizado logo abaixo:

$$Pr(T = 1|X = x) = \frac{\exp(x\beta)}{1 + \exp(x\beta)} \quad (3)$$

Para comparar os tratados com os não tratados, utilizou-se o pareamento do vizinho mais próximo. Esse estimador compara os indivíduos que foram tratados com os que não foram, através do *escore de propensão*, isto é, equipara os indivíduos com *escore de propensão* mais próximo. Para tanto, foi utilizado o pareamento com reposição, assim, o grupo de controle pode ser comparado com diferentes observações dos tratados.

Supondo independência condicional, o efeito médio de tratamento nos tratados é dado por  $E[Y_{1i} - Y_{0i}|X_i]$ , com essa hipótese, o efeito nos tratados resultante do pareamento, aplicando a lei de expectativas iteradas sobre  $X_i$ , é:

$$E[Y_{1i} - Y_{0i}|T_i=1] = E\{E[Y_i|P(X_i), T_i = 1] - E[P(X_i, T_i), T_i = 0]\} |T_i= 1] \quad (4)$$

sendo o estimador do pareamento por escore de propensão:

$$E[Y_{1i} - Y_{0i}|T_i=1] = \frac{1}{N_T} [\sum_{i \in D} Y_{1,i} - \sum_{j \in C} w(i,j) Y_{1,i}] \quad (5)$$

onde,  $N_T$  representa o número de indivíduos pertencentes a região de suporte comum e  $w(i,j)$  são os pesos utilizados na agregação dos indivíduos pertencentes ao contrafactual, no qual, depende do escore de propensão do aluno tratado  $i$ ,  $P(X_i)$ , e do escore de propensão dos não foram  $j$ ,  $P(X_j)$ .

### 3.2.3 Propensity Score Generalized (PSG)

A metodologia do PSM considera o tratamento como uma intervenção homogênea entre os tratados, ou seja, como uma variável binária. Essas são viáveis quando busca-se averiguar os impactos dos grupos expostos no Quadro 1. No entanto, a variável que gerou os grupos de comparação de respostas dos alunos é uma variável contínua, tornando necessário uma abordagem que capte o impacto do grau de exposição ao tratamento (anos de escolaridade na primeira infância) sobre a variável de resposta (desempenho nas provas do PISA).

O desafio é comparar alunos com características similares suficientes, mas com intensidades diferentes de exposição ao tratamento, com o objetivo de construir um cenário de quase experimento. Nesse intuito, Imbens (2000) e Hirano e Imbens (2004) propuseram a estimação de dose resposta. Essa, baseia-se na estimação da probabilidade de recebimento de cada nível de tratamento  $T_i = t$ , dado por  $r(t, x)$ . Os autores demonstram que, condicional à probabilidade de exposição ao nível de tratamento  $t$ , o resultado potencial na ausência de intervenção independe da *status* de tratamento. Essa premissa é denominada de ignorabilidade fraca, que pode ser expressa como:

$$Y(t) \perp 1 || [T = t] | r(t, x), \forall t \quad (8)$$

Para estimar a probabilidade de cada nível de tratamento,  $r(t, x)$ , pressupõe-se que o tratamento tenha uma distribuição normal, condicional as variáveis de controle, ou seja,  $T_i | X \sim N(X\beta, \sigma^2)$ . Após obter as estimativas do vetor  $\beta$ , é possível captar o valor da função de densidade de probabilidade associada ao valor observado do tratamento, pela seguinte equação:

$$\hat{r}_i = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left(-\frac{1}{2\hat{\sigma}^2} (T_i - f(\hat{\beta}, X_i))^2\right) \quad (9)$$

A relação condicional entre o resultado potencial e a  $r(t, x)$ , dada por  $Y | r(t, x)$ , é estimada para obter o resultado sobre o tratamento para definir o PSG da seguinte forma:

$$\hat{E}[Y(\hat{T})] = \sum_{i=1}^n \frac{\hat{\alpha}_0}{n} + \alpha_1 f(\hat{T}_i) + \alpha_2 f(\hat{T}_i)^2 + \alpha_3 \hat{r}_i(\hat{T}_i) + \alpha_4 \hat{r}_i^2(\hat{T}_i) + \alpha_5 \hat{r}_i(\hat{T}_i) \quad (10)$$

Para cada nível de tratamento  $\hat{T}$ , a expressão acima, apresenta um resultado potencial médio  $\hat{E}[Y(\hat{T})]$ . A função dose resposta é expressa devida a captação de vários níveis de exposição ao tratamento. Em seguida, obtém-se o impacto médio do incremento marginal ( $\epsilon > 0$ ) de anos de escolaridade na primeira infância sobre o desempenho dado por:

$$\theta(\hat{T}) = E[Y_i(\hat{T} + \epsilon)] - E[Y_i(\hat{T})] \quad (11)$$

Cabe ressaltar que as abordagens de pareamento pelo escore de propensão utilizam-se da suposição de independência condicional para estabelecer as relações casuais.

### 3.2.4 Análise de Sensibilidade de Rosenbaum (2002)

Os fatores não observáveis podem enviesar as estimativas dos efeitos médios de tratamento nos tratados. Para captar o possível impacto de viés de seleção advinda de variáveis não observáveis, empregou-se a técnica de limites de Rosenbaum (Rosenbaum, 2002). Essa, busca inferir qual a influência de uma eventual variável omitida sobre o viés de seleção existente na probabilidade de participação no tratamento e que prejudicaria as estimativas dos efeitos causais.



A técnica verifica a robustez dos resultados a variáveis omitidas. Os limites de Rosenbaum podem ser definidos da seguinte forma:

$$\frac{1}{e^y} \leq \frac{\theta_i(1-\theta_j)}{\theta_j(1-\theta_i)} \leq e^y \quad (12)$$

onde  $i$  e  $j$  são dos indivíduos com características observáveis em uma distribuição logística e  $y$  expressa a medida do grau de ruptura do viés de seleção. Desse modo, quando  $y = 0$ , o grau de associação será igual a 1, o que consequentemente indica que os resultados não são influenciados por variáveis omitidas.

## 4 Análise e Discussão dos Resultados

### 4.1 Análise Descritiva

As estatísticas descritivas das notas na prova do PISA 2018 para todos os grupos de tratados e não tratados estão expressas na Tabela 1. As estatísticas dos controles estão explícitas nas Tabelas A.2 e A.3, no apêndice A. Observa-se que a média do desempenho em matemática, em leitura e em ciências de alunos que foram atendidos na creche é maior do que os que não foram. No entanto, esse grupo de tratados apresenta maior dispersão em torno da média. As notas dos estudantes que foram atendidos na creche são maiores, em média, em torno de 6,33% por cento (24,8 pontos) em matemática, de 5,48% (23,4 pontos) em leitura e de 5,73% (23,8 pontos) em ciências quando comparado aos que ingressaram diretamente na pré-escola.

Ao comparar alunos que fizeram a pré-escola com os que não tiveram educação na primeira infância (ingressaram diretamente no ensino fundamental) apresentaram, em média, desempenho inferior, em torno de 3,78% (-15.4 pontos) em matemática, de 2,68% (-11,8 pontos) em leitura e de 4,13% (-17.9 pontos) em ciências.

Tabela 1: Descrição estatística dos grupos

Variável dependente	Variáveis respondidas pelos alunos											
	Creche x Pré-Escola				Pré-Escola x Sem E.I.				E.I. x sem E.I.			
	Creche		Pré-Escola		Pré-Escola		sem E.I.		E.I.		sem E.I.	
	Média	D.p.	Média	D.p.	Média	D.p.	Média	D.p.	Média	D.p.	Média	D.p.
Nota de Matemática	416,6	91,9	391,8	80,9	391,8	80,9	407,2	82,5	416,6	91,9	407,2	82,5
Nota de Leitura	450,9	101,1	427,5	93,0	427,5	93,0	439,3	89,9	450,9	101,1	439,3	89,9
Nota de Ciências	439,3	94,8	415,5	86,9	415,5	86,9	433,4	78,5	439,3	94,8	433,4	78,5
Nº de obs.	1.788		1.428		1.498		94		1.788		94	

Fonte: Elaborado pelos autores. Nota: EI é referente à educação infantil, Pré-E. a Pré-Escola e D.p. a desvio padrão.

Os alunos que tiveram toda a educação na primeira infância (isto é, frequentaram creche e pré-escola), vis a vis os que não tiveram, apresentaram, em média, notas superiores. Desse modo, o diferencial de notas é maior em 2,30% (9,4 pontos), em 2,65% (11,6 pontos) e em 1,37% (5,9 pontos), respectivamente, em matemática, em leitura e em ciências a favor dos alunos que tiveram educação durante toda a primeira infância.

Portanto, constata-se que os grupos de tratamentos são superiores aos grupos de controle, com exceção da comparação entre os que fizeram pré-escola com os que não tiveram EI. De forma complementar, tem-se os testes de médias no apêndice C, os quais são significativos a 5% no cruzamento dos grupos creche x pré-escola, os demais cruzamentos são insignificantes a esse nível. Essa média de comparação entre os grupos é uma forma “ingênua” de comparação, para captar de forma correta os diferenciais de notas, utiliza-se nos próximos tópicos, técnicas que isolam o efeito da escolaridade na primeira infância, o que esclarecerá, se realmente essas diferenças se mantêm.

### 4.2 Resultado da estimação do *Propensity Score Matching* (PSM)

Inicialmente, seguindo Becker e Ichino (2002), estima-se os *logits*, para cada grupo de tratamento, que expressam as chances de alunos pertencerem ao grupo considerado. Essas estimativas podem ser encontradas no apêndice D. Além disso, por meio do apêndice E é possível observar a densidade das distribuições antes e após o pareamento pelos escores de propensão para cada tratamento. Observa-se que

as distribuições são bastantes semelhantes após o procedimento, de forma que se verifica a sobreposição entre as distribuições dos tratados e não tratados.

No apêndice F, tem-se a robustez do pareamento, o qual indica que em todos os níveis de escolaridade considerado como tratamento, tem-se a redução dos Pseudo- $R^2$ , dos vieses médio e mediano. Além disso, os testes de razão de verossimilhança (LR) e os p-valores indicam que a amostra é balanceada ao utilizar o vetor de variáveis observáveis. No entanto, isso não foi constatado no balanceamento do grupo de tratamento dos que tiveram educação completa na primeira infância.

A estimativa dos diferenciais de notas dos grupos de tratados e de não tratados por PSM podem ser encontrados na Tabela 2, os quais, em grande parte, são significativos ao nível de 5%, todavia, a comparação dos alunos atendidos em toda a EI *versus* os que ingressaram diretamente no ensino fundamental não apresentou significância estatística.

Tabela 2: Efeito médio de tratamento nos tratados

Efeito	Nota de matemática	Nota de leitura	Nota de ciências	Nº de obs.
Variáveis advindas das respostas dos alunos				
Creche x pré-escola	15,565* (3,288)	10,615* (2,880)	11,457* (3,200)	3.216
Pré-escola x sem E.I.	19,314* (5,743)	17,908* (7,125)	12,647* (5,890)	1.522
Educação infantil x sem E.I.	15,491 (9,942)	-1,062 (15,065)	2,323 (10,204)	1.882

Fonte: elaborado pelos autores. Nota: \* p-valor < 0,05.

Em média, constata-se que um aluno atendido na creche apresentou desempenho superior em matemática, em leitura e em ciências, respectivamente, de 15,56, de 10,61 e de 11,45 pontos, em relação aos que fizeram apenas a pré-escola. Já os alunos que ingressaram na pré-escola, comparados aos que ingressaram somente no ensino fundamental, têm, em média, notas maiores de 19,31 pontos em matemática, de 17,90 pontos em leitura e de 12, 647 pontos em ciências.

Assim, pode-se inferir que ter frequentado educação infantil impacta positivamente o desempenho acadêmico futuro. Essa prerrogativa está de acordo com estudos internacionais (BIETENBECK; ERICSSON; WAMALWA, 2019; CORTÁZAR, 2020; DIMISKI, 2021; ESPING-ANDERSEN, 2012; FELFE e LAVIVE, 2010; HOLLA *et al.*, 2021; SHÜTZ, 2009; e PHOLPHIRUL, 2017) e nacionais (CURI; MENEZES-FILHO, 2009; FELÍCIO; VASCONCELLOS, 2007; REYNA; SILVA; LÍRIO, 2019; PINTO; SANTOS; GUIMARÃES, 2017).

#### 4.4 Estimativa do *Propensity Score Generalized* (PSG)

Para inferir se o desempenho dos alunos no PISA é influenciado pela quantidade de anos de estudo (cuidados) na primeira infância, utiliza-se o PSG. Antes de dar início a estimação, é necessário que os resíduos sejam normalmente distribuídos. Para tanto, fez-se uso do teste de Kolmogorov-Smirnov de normalidade dos resíduos, no qual, é satisfeito a 5% de significância para as três áreas de conhecimento.

A dose resposta e o efeito marginal do impacto dos números de anos de estudo na primeira infância sobre as áreas de conhecimentos, podem ser visualizados nos Gráficos 1, 2 e 3. Para obter os intervalos de confiança aplicou-se a técnica de *bootstrap* com 500 repetições. Cabe ressaltar, que nas três disciplinas alguns pontos do efeito marginal são insignificantes, isso acontece quando os intervalos de confiança tocam o zero no eixo horizontal (que representa o nível zero de tratamento).

O aumento dos anos iniciais de escolaridade na primeira infância, de 1 para 2 anos, aumenta, em média, a performance em matemática e reduz em ciências e leitura. Por outro lado, o aumento da escolaridade a partir de 2 anos até 5 anos, em média, aumenta o desempenho nas três disciplinas. Entretanto, quando a mudança de exposição passa do 5º para o 6º ano a performance cai nas três áreas de conhecimento.

O efeito marginal indica que, em média, o acréscimo do aumento de escolaridade inicial na primeira infância traz retornos ainda maiores na performance acadêmica. Assim, em matemática, em média, alunos que tiveram 1 ano de escolaridade, caso tivessem ingressado um ano mais cedo, tendo 2 anos de escolaridade, teriam acréscimos maiores de desempenho (retornos marginais crescentes). Já em leitura e ciências, o acréscimo de anos, até 3 anos de escolaridade, segue a mesma tendência. Todavia, o aumento

de escolaridade em matemática a partir de 2 anos e em leitura e ciências a partir de 3 anos, traz retornos marginais decrescente sobre a performance. Essa tendência se mantém até 5º ano, pois quando a exposição passa do 5º para o 6º ano de escolaridade, há maiores retornos nas performances.

Os resultados corroboram os achados de Marques e Cunha (2020), que constataram que a educação infantil no Brasil tem efeitos heterogêneos sobre o desempenho. Além disso, em parte, com o estudo internacional de Berlinski, Galian e Getler (2009), que mostra que o aumento da escolaridade nos primeiros anos aumenta a performance acadêmica.

Gráfico 1: Dose Resposta e efeito marginal do tratamento em matemática

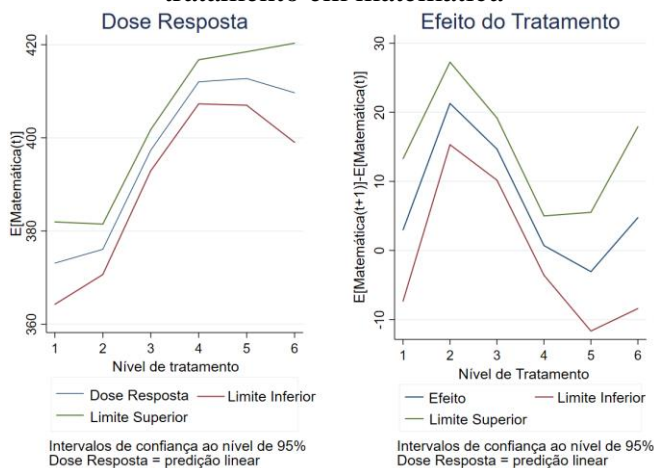


Gráfico 2: Dose Resposta e efeito marginal do tratamento em leitura

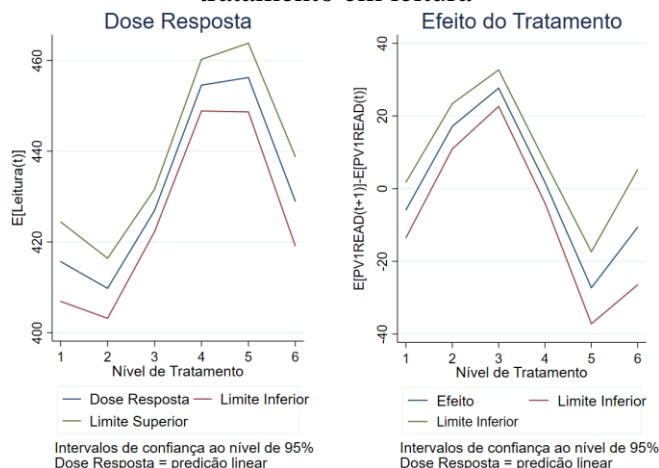
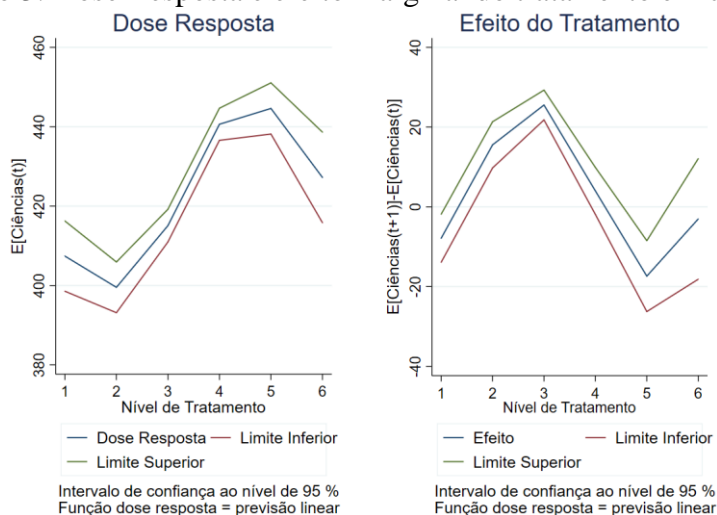


Gráfico 3: Dose Resposta e efeito marginal do tratamento em ciências



Fonte: Elaborado pelos autores com base nos dados do PISA 2018. Nota: a sigla D.R. corresponde à dose resposta.

#### 4.5 Análise de Sensibilidade

Para observar se os resultados acima são robustos a variáveis omitidas, aplicou-se o teste de sensibilidade de Rosenbaum, o qual são apresentados na Tabela 3. O teste estabelece a partir de que ponto crítico de viés em características não observáveis rejeita-se a hipótese nula do tratamento ser aleatório após o pareamento. Desse modo, considera-se um parâmetro de sensibilidade  $\Gamma$ , no qual, representa o nível de distanciamento da aleatoriedade do tratamento, onde  $\Gamma = 1$ , o impacto do tratamento é livre de viés. Os resultados são satisfatórios, uma vez que reforçam a não existência de viés de possíveis características não observáveis que afetam o resultado para um  $\Gamma = 1,05$ .

Tabela 3: Análise de sensibilidade de Rosenbaum

Creche x pré-escola	Pré-escola x sem E.I	Educação infantil x sem E.I.
---------------------	----------------------	------------------------------

Gamma ( $\Gamma$ )	Mate. sig+	Leit. sig+	Ciên. sig+	Mate. sig+	Leit. sig+	Ciên. sig+	Mate. sig+	Leit. sig+	Ciên. sig+
1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01
1,05	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02

Fonte: Elaborado pelos autores. Nota: Mate., Leit., Ciên. se referem, respectivamente, à matemática, à leitura e à ciências.

## 5. Considerações Finais

Visando contribuir para o debate sobre a educação infantil, com ênfase nos impactos sobre o desempenho acadêmico dos estudantes, optou-se por comparar alunos que ingressam mais cedo na educação infantil com os que entraram tardiamente. Além disso, entender como o tempo de exposição na educação infantil pode afetar o desempenho futuro.

Para atender ao objetivo proposto fez-se uso de dados do Brasil no PISA de 2018. A partir destes, tendo como medidores de impacto as notas de matemática, de leitura e de ciências dos alunos no PISA, tentou-se captar o efeito da educação infantil comparando os seguintes grupos: 1) creche *versus* pré-escola; 2) pré-escola *versus* sem educação infantil; 3) educação infantil completa (creche e pré-escola) *versus* sem educação infantil. Para captar o impacto entre os grupos fez-se uso do PSM. Já para inferir o impacto da quantidade de anos de estudo na primeira infância utilizou-se a técnica do PSG. Por fim, aplicou-se os limites de Rosenbaum (2002) para verificar a robustez dos efeitos.

A partir do PSM, observou-se que, em média, os estudantes que ingressaram na creche têm melhor performance em comparação aos que entraram diretamente na pré-escola em todas as competências avaliadas. Já os alunos que ingressaram na pré-escola têm, em média, performances maiores nas três competências em comparação aos que não tiveram EI.

O PSG, medida de dose resposta, indica que, o aumento de anos de escolaridade a partir de 2 até 5 anos, em média, aumenta a performance. Em termos marginais, pode-se inferir que o acréscimo de anos de escolaridade nos anos iniciais traz retornos maiores de desempenho, isso até 2 anos em matemática e 3 anos em ciências e em leitura. No entanto, à medida que se ultrapassa esses anos o retorno no desempenho é decrescente. Por fim, ressalta-se que o teste de Rosenbaum (2002) indicou que os efeitos não são influenciados por variáveis omitidas.

Diante do exposto, a educação infantil mostra-se um impulsionador do desenvolvimento cognitivo. Além disso, o grau de exposição nesse nível educacional influencia o desempenho. Para estudos futuros cabe testar se essa relação se mantém para diferentes níveis socioeconômicos.

## Referências

- BARROS, R. P.; CARVALHO, M.; FRANCO, S.; MENDONÇA, R.; ROSALÉM, A. (2011). Uma avaliação do impacto da qualidade da creche no desenvolvimento infantil. *Pesquisa e planejamento econômico*, v. 41, p. 213-232.
- BECKER, S. O.; ICHINO, A. (2002). Estimation of average treatment effects based on propensity scores. *The stata journal*, v. 2, n. 4, p. 358-377.
- BERLINSKI, S.; GALIANI, S.; GERTLER, P. (2009). The effect of pre-primary education on primary school performance. *Journal of Public Economics*, v. 93, n. 1-2, p. 219-234.
- BIETENBECK, J.; ERICSSON, S.; WAMALWA, F. M. (2019). Preschool attendance, schooling, and cognitive skills in East Africa. *Economics of Education Review*, v. 73, p. 101909.
- BRASIL. *Constituição da República Federativa do Brasil* (1988).
- BRASIL. *Emenda Constitucional nº 53, de 19 de dezembro de 2006*. Dá nova redação aos arts. 7º, 23º, 30º, 206º, 208º, 211º e 212º da Constituição Federal e ao art. 60º do Ato das Disposições Constitucionais Transitórias.
- BRASIL. *Lei n. 9.394 de 20 de dezembro de 1996*. Estabelece as Diretrizes e Bases da Educação Nacional.
- BRASIL. *Lei nº 10.172, de 09 de janeiro de 2001*. Aprova o Plano Nacional de Educação.

- BRASIL. *Lei nº 11.274 de fevereiro de 2006*. Altera a redação dos artigos 29, 30, 32 e 87 da Lei nº 9.394/1996, que estabelece as diretrizes e bases da educação nacional, dispondo sobre a duração de nove anos para o ensino fundamental, com matrícula obrigatória a partir dos 6 seis anos de idade.
- BRASIL. *Lei nº 12.796 de 4 de abril de 2013*. Altera a Lei nº 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional, para dispor sobre a formação dos profissionais da educação e dar outras providências.
- BRASIL. *Lei nº 13.005 de junho de 2014*. Aprova o Plano Nacional de Educação e dá outras providências.
- BRASIL. Ministério da Educação. *Base Nacional Comum Curricular*. Brasília: MEC, 2017.
- BRASIL. Ministério da Educação. Secretaria de Educação Básica. *Diretrizes Curriculares Nacionais para a Educação Infantil*.
- BRASIL. Ministério da Educação e do Desporto. Secretaria de Educação Fundamental. *Subsídios para o Credenciamento e Funcionamento das Instituições de Educação Infantil*.
- BRASIL. Ministério da Educação e do Desporto. Secretaria de Educação Fundamental. *Referencial Curricular Nacional para a Educação Infantil*.
- BURCHINAL, M. R.; CAMPBELL F. A.; BRYANT D. A.; WASIK B. H.; RAMEY C. T. (1997). Early intervention and mediating processes in cognitive performance of children of low-income African American families. *Child development*, p. 935-954.
- CARNEIRO, P.; HECKMAN, J. (2003). *Human Capital Policy* (No. 821). Institute of Labor Economics (IZA).
- CORREA, E.; COMIM, F.; TAI, S. H. T. (2014). Impactos da Creche na Primeira Infância: efeitos dependendo das características da família e do grau de exposição ao centro de cuidado. *Anais do XLII Encontro Nacional de Economia*. ANPEC.
- CORTÁZAR, A.; MOLINA, M. D. L. Á.; SÉLMAN, J.; MANOSALVA, A. (2020). Early childhood education effects on school outcomes: Academic achievement, grade retention and school drop Out. *Early Education and Development*, v. 31, n. 3, p. 376-394.
- CUNHA, F.; HECKMAN, J. J. (2007). The technology of skill formation. *American Economic Review*, v. 97, n. 2, p. 31-47.
- CUNHA, F.; HECKMAN, J. J.; LOCHNER, L.; MASTEROV, D. V. (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, v. 1, p. 697-812.
- CURI, A. Z.; MENEZES-FILHO, N. A. (2009). A relação entre educação pré-primária, salários, escolaridade e proficiência escolar no Brasil. *Estudos Econômicos* (São Paulo), v. 39, p. 811-850.
- DEL BOCA, D.; MONFARDINI, C.; SEE, S. G. (2018), "Government Education Expenditures, Pre-Primary Education and School Performance: A Cross-Country Analysis", *IZA Discussion Paper*, No. 11375.
- DIMISKI, A. (2021). Factors that affect Students' performance in Science: An application using Gini-BMA methodology in PISA 2015 dataset. *Review of Economic Analysis*, v. 13, n. 2, p. 157-211.
- ESPING-ANDERSEN, G.; GARFINKEL, I.; HAN, W. J.; MAGNUSON, K.; WAGNER, S.; WALDFOGEL, J. (2012). Child care and school performance in Denmark and the United States. *Children and youth services review*, v. 34, n. 3, p. 576-589.
- FELFE, C.; LALIVE, R. (2010). *How does early child care affect child development? Learning from the children of German unification*.
- FELÍCIO, F.; VASCONCELLOS, L. (2007). O efeito da educação infantil sobre o desempenho escolar medido em exames padronizados. *Anais do XXXV Encontro Nacional de Economia*.
- FUCHS, T.; WÖSSMANN, L. (2008). What accounts for international differences in student performance? A re-examination using PISA data. In: *The economics of education and training* (pp. 209-240). Physica-Verlag HD.
- FUNDAÇÃO ABRINQ. *Taxa bruta de matrícula em Creches*. 2021. Disponível em: <https://observatoriocrianca.org.br/cenario-infancia/temas/educacao-infantil/1081-taxa-bruta-de-matricula-em-creches?filters=1,77>
- FUNDAÇÃO ABRINQ. *Taxa bruta de matrículas em pré-escolas*. 2021. Disponível em: <https://observatoriocrianca.org.br/cenario-infancia/temas/educacao-infantil/543-taxa-bruta-de-matricula-em-pre-escolas?filters=1,82>
- HIRANO, K.; IMBENS, G. W. (2004). The propensity score with continuous treatments. *Applied Bayesian modeling and causal inference from incomplete-data perspectives*, v. 226164, p. 73-84.

- HOLLA, A.; BENDINI, M. M.; DINARTE DIAZ, L. I.; TRAKO, I. (2021). *Is Investment in Preprimary Education Too Low? Lessons from (Quasi) Experimental Evidence across Countries*, No. 9723. The World Bank.
- IMBENS, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika*, v. 87 n. 3, p. 706-710.
- KARAKOLIDIS, A.; PITSIA, V.; EMVALOTIS, A. (2016). Examining students' achievement in mathematics: A multilevel analysis of the Programme for International Student Assessment (PISA) 2012 data for Greece. *International Journal of Educational Research*, v. 79, p. 106-115.
- KNUDSEN, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of cognitive neuroscience*, 16, n. 8, p. 1412-1425, 2004.
- MARKUS, N. M.; OLIVEIRA, F. R.; NISHIMURA, F. N. (2020). Efeito da Educação Pré-primária no Desempenho Escolar. *Anais do XXV Encontro Regional de Economia*.
- MARQUES F. C.; CUNHA M. S. (2020). ANÁLISE DOS IMPACTOS HETEROGÊNEOS DA EDUCAÇÃO INFANTIL NO PAÍS. *Anais do XLVIII Encontro Nacional de Economia. ANPEC*.
- MCCOY, D. C.; YOSHIKAWA, H.; ZIOL-GUEST, K. M.; DUNCAN, G. J.; SCHINDLER, H. S.; MAGNUSON, K.; SHONKOFF, J. P. (2017). Impacts of early childhood education on medium-and long-term educational outcomes. *Educational Researcher*, v. 46, n. 8, p. 474-487.
- NORES, M.; BARNETT, W. S. (2010). Benefits of early childhood interventions across the world: (Under) Investing in the very young. *Economics of education review*, v. 29, n. 2, p. 271-282.
- PHOLPHIRUL, P. (2017). Pre-primary education and long-term education performance: Evidence from Programme for International Student Assessment (PISA) Thailand. *Journal of Early Childhood Research*, v.15, n. 4, p. 410-432.
- PINTO C. C. X.; SANTOS, D.; GUIMARÃES, C. (2017). The Impact of Daycare Attendance on Math Test Scores for a Cohort of Fourth Graders in Brazil. *The Journal of Development Studies*, v. 53, n. 9, p. 1335-1357.
- REYNA, E. F.; SILVA M. M. C.; E LÍRIO V. S. (2019). Novas evidências sobre o efeito da pré-escola nos testes de proficiência do Saeb. *XVII Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos*.
- ROSENBAUM, P. R.; RUBIN, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, v. 70, n. 1, p. 41-55.
- ROSENBAUM, P. R. (2002). Sensitivity to hidden bias. In *Observational studies* (p. 105-170). Springer, New York, NY.
- SANTOS, D. D. (2015). *Impactos do ensino infantil sobre o aprendizado: benefícios positivos, mas desiguais*. 260 f. Tese (Livre-Docência) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, Ribeirão Preto.
- SCHADY, N.; GALIANI, S.; SOUZA, A. P. (2006). Early childhood development in Latin America and the Caribbean [with comments]. *Economía*, v. 6, n. 2, p. 185-225.
- SCHÜTZ, G. (2009). *Does the quality of pre-primary education pay off in secondary school? An international comparison using PISA 2003* (No. 68). Ifo Working Paper.
- SILVA JUNIOR, W. S. D.; GONÇALVES, F. D. O. (2016). Evidências da relação entre a frequência no ensino infantil e o desempenho dos alunos do ensino fundamental público no Brasil. *Revista Brasileira de Estudos de População*, v. 33, p. 283-301.
- TEMPLE, J. A.; REYNOLDS, A. J. (2007). Benefits and costs of investments in preschool education: Evidence from the Child-Parent Centers and related programs. *Economics of Education Review*, v. 26, n. 1, p. 126-144.
- YI, H. S.; LEE, Y. (2017). A latent profile analysis and structural equation modeling of the instructional quality of mathematics classrooms based on the PISA 2012 results of Korea and Singapore. *Asia Pacific Education Review*, 18(1), 23-39.
- YOSHIKAWA, H.; WEILAND, C.; BROOKS-GUNN, J.; BURCHINAL, M. R., ESPINOSA, L. M.; GORMLEY, W. T.; ZASLOW, M. J. (2013). Investing in Our Future: The Evidence Base on Preschool Education. *Society for Research in Child Development*.

Tabela A.1: Análise descritiva da idade de ingresso na primeira infância (resposta dos alunos)

Variável	Matemática					Leitura					
	Valores das variáveis	Média	D. P.	Mínimo	Máximo	Obs.	Média	D. P.	Mínimo	Máximo	Obs.
Idades em que o aluno iniciou a primeira infância	Não teve educação	407,20	82,54	245,10	593,90	94	439,30	89,94	230,10	608,40	94
	idade = 1 ou menos	419,10	102,20	224,20	797,60	318	447,50	108,70	197,90	766,30	318
	idade = 2	430,20	92,04	127,80	725,70	537	464,10	102,70	197,80	744,40	537
	idade = 3	408,00	87,16	179,60	708,20	933	444,50	96,82	167,10	732,50	933
	idade = 4	398,90	83,46	177,40	705,80	728	437,80	93,99	158,40	704,30	728
	idade = 5	392,30	78,87	204,80	671,90	451	424,70	91,22	209,50	770,00	451
	idade = 6 ou mais	370,10	73,27	157,00	594,40	249	402,20	88,17	144,70	659,90	249
	Ciências										
	Valores das variáveis	Média	D. P.	Mínimo	Máximo	Obs.					
	Não teve educação	433,40	78,52	256,40	603,10	94					
	idade = 1 ou menos	443,00	102,20	219,60	747,50	318					
	idade = 2	451,10	94,09	192,30	685,00	537					
	idade = 3	431,30	91,86	148,80	693,00	933					
	idade = 4	425,20	88,25	178,50	661,40	728					
	idade = 5	413,70	84,28	224,90	693,40	451					
	idade = 6 ou mais	390,20	82,55	179,80	618,40	249					

Fonte: Elaboração própria. Obs. significa o número de observações.

Tabela A.2: Análise descritiva dos controles

Variáveis	Creche x Pré-escola				Pré-escola x sem E.I.				EI x sem EI			
	Creche		Pré-escola		Pré-escola		Sem E.I.		EI		Sem EI	
	Média	D. P.	Média	D. P.	Média	D. P.	Média	D. P.	Média	D. P.	Média	D. P.
<i>Gen</i>	0,56	0,50	0,57	0,50	0,57	0,50	0,36	0,48	0,56	0,50	0,36	0,48
<i>Idad</i>	15,91	0,28	15,91	0,29	15,91	0,29	15,92	0,29	15,91	0,28	15,92	0,29
<i>ano 1</i>	0,43	0,49	0,54	0,50	0,54	0,50	0,48	0,50	0,48	0,50	0,48	0,50
<i>ano 2</i>	0,54	0,50	0,44	0,50	0,44	0,50	0,46	0,50	0,50	0,50	0,46	0,50
<i>ano 3</i>	0,03	0,18	0,02	0,13	0,02	0,13	0,06	0,25	0,03	0,16	0,06	0,25
<i>edm1</i>	0,17	0,38	0,19	0,39	0,19	0,39	0,21	0,41	0,18	0,38	0,21	0,41
<i>edm2</i>	0,65	0,48	0,51	0,50	0,51	0,50	0,49	0,50	0,59	0,49	0,49	0,50
<i>edm3</i>	0,09	0,28	0,06	0,24	0,06	0,24	0,07	0,26	0,08	0,27	0,07	0,26
<i>edp1</i>	0,18	0,38	0,19	0,39	0,19	0,39	0,18	0,39	0,18	0,39	0,18	0,39
<i>edp2</i>	0,57	0,50	0,45	0,50	0,45	0,50	0,49	0,50	0,52	0,50	0,49	0,50
<i>edp3</i>	0,09	0,28	0,07	0,25	0,07	0,25	0,06	0,25	0,08	0,27	0,06	0,25
<i>HISEI</i>	46,77	23,51	39,08	22,10	39,08	22,10	41,85	22,42	43,35	23,21	41,85	22,42
<i>REPEAT</i>	0,21	0,41	0,26	0,44	0,26	0,44	0,26	0,44	0,23	0,42	0,26	0,44
<i>ESCS</i>	-0,81	1,25	-1,25	1,21	-1,25	1,21	-1,18	1,31	-1,01	1,25	-1,18	1,31
<i>HOMEPOS</i>	-1,25	1,13	-1,51	1,01	-1,51	1,01	-1,49	0,98	-1,37	1,08	-1,49	0,98
<i>CULTPOSS</i>	-0,49	0,81	-0,63	0,76	-0,63	0,76	-0,56	0,80	-0,55	0,79	-0,56	0,80
<i>WEALTH</i>	-1,26	1,09	-1,47	0,97	-1,47	0,97	-1,44	0,92	-1,36	1,04	-1,44	0,92
<i>ler1</i>	0,62	0,49	0,50	0,50	0,50	0,50	0,60	0,49	0,57	0,50	0,60	0,49
<i>ler2</i>	0,87	0,34	0,82	0,38	0,82	0,38	0,79	0,41	0,85	0,36	0,79	0,41
<i>ler3</i>	0,81	0,39	0,74	0,44	0,74	0,44	0,70	0,46	0,78	0,41	0,70	0,46
<i>ler4</i>	0,83	0,38	0,76	0,43	0,76	0,43	0,75	0,44	0,80	0,40	0,75	0,44
<i>ler5</i>	0,74	0,44	0,68	0,47	0,68	0,47	0,72	0,45	0,71	0,45	0,72	0,45
<i>centro_oeste</i>	0,08	0,27	0,10	0,31	0,10	0,31	0,14	0,35	0,09	0,28	0,14	0,35
<i>Nordeste</i>	0,07	0,25	0,10	0,31	0,10	0,31	0,10	0,30	0,08	0,28	0,10	0,30
<i>Norte</i>	0,44	0,50	0,26	0,44	0,26	0,44	0,18	0,39	0,36	0,48	0,18	0,39
<i>Sul</i>	0,10	0,30	0,16	0,37	0,16	0,37	0,32	0,47	0,13	0,33	0,32	0,47
<i>Sudeste</i>	0,32	0,47	0,38	0,49	0,38	0,49	0,27	0,44	0,35	0,48	0,27	0,44
<i>Escp</i>	0,83	0,79	0,99	0,82	0,99	0,82	0,88	0,32	0,90	0,81	0,88	0,32
<i>TOTAT</i>	37,74	39,96	35,61	33,12	35,61	33,12	45,78	37,84	36,80	37,09	45,78	37,84
<i>STRATIO</i>	27,39	15,73	28,94	17,62	28,94	17,62	27,15	16,09	28,08	16,61	27,15	16,09
Nº de obs.	1.788		1.428		1.428		94		3216		94	

Fonte: Elaboração própria. Nota: EI é referente à Educação Infantil.



## Apêndice C - Testes de Médias

Tabela C: Teste de médias das variáveis advindas do aluno

Disciplina		Média	D. Padrão	Intervalo de C.		Estatística t
Creche x Pré-escola						
Matemática	Creche	416,6	91,9	412,4	420,9	8,026
	Pré-Escola	391,8	80,9	387,6	396,0	
	Diferença	24,84				
Leitura	Creche	450,9	101,1	446,2	455,6	6,770
	Pré-Escola	427,5	92,98	422,7	432,3	
	Diferença	23,44				
Ciências	Creche	439,3	94,8	434,9	443,7	7,355
	Pré-Escola	415,5	86,9	410,9	420,0	
	Diferença	23,85				
Pré-escola x sem Educação Infantil						
Matemática	Pré-escola	391,8	80,9	387,6	396,0	-1,784
	Sem EI	407,2	82,5	390,3	424,1	
	Diferença	-15,39				
Leitura	Pré-escola	427,5	93,0	422,7	432,3	-1,199
	Sem EI	439,3	89,9	420,9	457,7	
	Diferença	-11,84				
Ciências	Pré-escola	415,5	86,9	410,9	420,0	-2,106
	Sem EI	433,4	78,5	417,3	449,5	
	Diferença	-19,05				
Educação Infantil x sem Educação Infantil						
Matemática	EI	416,6	91,9	412,4	420,9	0,975
	Sem EI	407,2	82,5	390,3	424,1	
	Diferença	9,446				
Leitura	EI	450,9	101,1	446,2	455,6	1,089
	Sem EI	439,3	89,9	420,9	457,7	
	Diferença	11,59				
Ciências	EI	439,3	94,81	434,9	443,7	0,590
	Sem EI	433,4	78,5	417,3	449,5	
	Diferença	5,880				

Fonte: Elaboração própria.

Apêndice D - Estimação dos *logits*

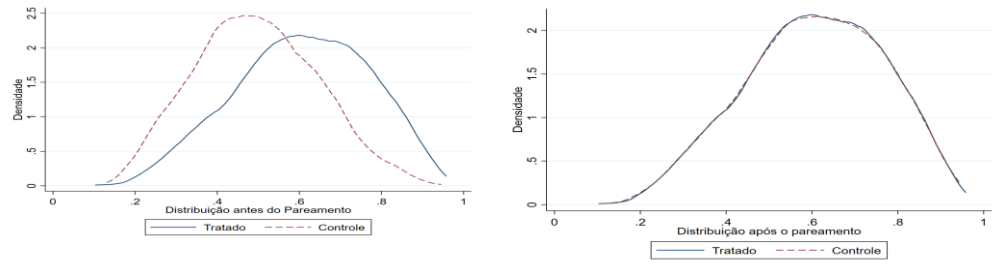
Tabela D: Resultados do *logit*

Variáveis	Variáveis advindas de respostas dada pelo aluno					
	Creche x Pré-Escola		Pré-escola x sem E.I		E. I. x sem E.I	
	coef.	E.P.	coef.	E.P.	coef.	E.P.
<i>Gen</i>	-0,05	0,08	0,96*	0,23	0,90*	0,23
<i>Idad</i>	-0,38*	0,15	0,23*	0,07	0,25*	0,06
<i>ano2</i>	0,46*	0,10	-0,48*	0,28	-0,28	0,27
<i>ano3</i>	0,99*	0,28	-2,09*	0,56	-1,51	0,50
<i>edm1</i>	0,25	0,14	-0,11	0,38	-0,06	0,37
<i>edm2</i>	0,35*	0,14	0,19	0,40	0,32	0,39
<i>edm3</i>	0,19	0,17	-0,42	0,48	-0,31	0,48
<i>edp1</i>	0,15	0,12	-0,17	0,37	-0,06	0,36
<i>edp2</i>	0,21	0,12	-0,18	0,36	-0,14	0,36
<i>edp3</i>	-0,10	0,17	0,31	0,51	0,29	0,52
<i>HISEI</i>	0,01*	0,00	-0,01	0,01	-0,01	0,01
<i>REPEAT</i>	0,14	0,11	-0,28	0,32	-0,21	0,31
<i>ESCS</i>	0,00	0,11	0,23	0,32	0,17	0,32
<i>HOMEPOS</i>	-0,11	0,15	0,55	0,45	0,51	0,42
<i>CULTPOSS</i>	0,19*	0,07	-0,40	0,22	-0,27	0,20
<i>WEALTH</i>	0,20	0,13	-0,27	0,39	-0,19	0,36
<i>ler1</i>	0,17	0,09	-0,72*	0,30	-0,53	0,29
<i>ler2</i>	0,02	0,14	0,34	0,38	0,40	0,38
<i>ler3</i>	0,04	0,14	0,71	0,39	0,59	0,37
<i>ler4</i>	0,09	0,14	0,25	0,41	0,29	0,39
<i>ler5</i>	-0,01	0,11	-0,53	0,34	-0,48	0,33
<i>centro_oeste</i>	-0,19	0,18	-0,58	0,48	-0,66	0,47
<i>Nordeste</i>	1,11*	0,15	0,36	0,44	0,92*	0,43
<i>Sul</i>	-0,34	0,18	-0,81	0,44	-1,04*	0,44
<i>Sudeste</i>	0,05	0,15	0,24	0,43	0,17	0,42
<i>SC013Q01TA</i>	-0,19*	0,05	0,31	0,34	0,03	0,19
<i>TOTAT</i>	0,00	0,00	-0,01*	0,00	0,01	0,00
<i>STRATIO</i>	-0,01*	0,00	0,00	0,01	0,00	0,01
<i>const.</i>	5,36*	2,42	-	-	-	-

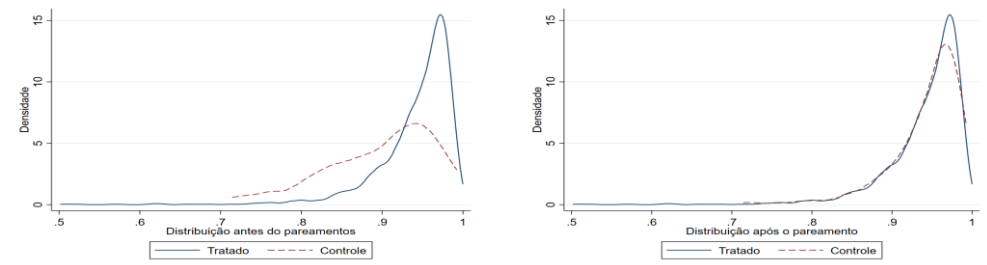
Fonte: Elaboração própria. Nota: \* p-valor < 0,05.

## Apêndice E - Distribuições Antes e Depois do Pareamento

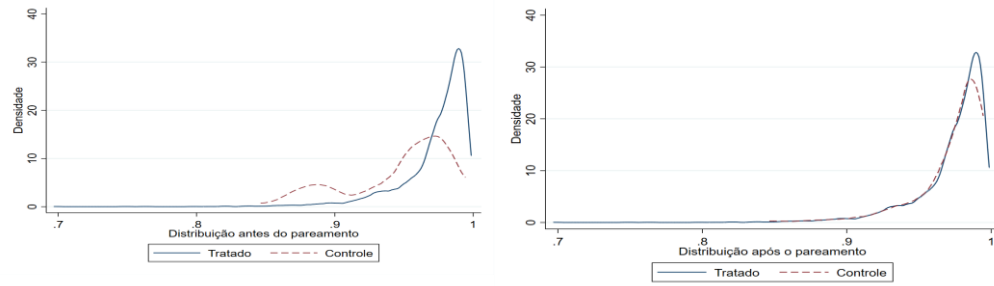
Figura E: Distribuições antes (a esquerda) e depois (a direita) do pareamento  
Creche x Pré-escola



Pré-escola x sem E.I



Com E.I x Sem E.I



Fonte: Elaboração própria.

Apêndice F - Balanceamento Pré e Pós Balanceamento

Tabela F: Balanceamento pré e pós balanceamento - Variáveis advindas de respostas dada pelo aluno

	Pseudo- $R^2$	LR $\chi^2$	P-valor	Viés Médio	Viés Mediano
Creche x Pré-Escola					
Não pareados	0,088	386,77	0,000	16,2	15,6
Pareados	0,005	24,51	0,654	2,4	2,4
Pré-escola x sem E.I					
Não pareados	0,100	70,82	0,000	11,6	8,5
Pareados	0,058	31,36	0,240	6,2	5,5
E.I. x sem E.I					
Não pareados	0,095	81,17	0,000	13,0	8,6
Pareados	0,152	1356,14	0,000	14,2	11,5

Fonte: Elaboração própria.