

JULIANA MARA DE FÁTIMA VIANA GANDRA

**O IMPACTO DA EDUCAÇÃO EM TEMPO INTEGRAL NO
DESEMPENHO ESCOLAR: UMA AVALIAÇÃO DO
PROGRAMA MAIS EDUCAÇÃO**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Economia, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2017

Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa

T

G153i
2017

Gandra, Juliana Mara de Fátima Viana, 1988-

O impacto da educação em tempo integral no desempenho
escolar : uma avaliação do Programa Mais Educação / Juliana
Mara de Fátima Viana Gandra. – Viçosa, MG, 2017.
xiii, 84f. : il. (algumas color.) ; 29 cm.

Inclui anexos.

Orientador: Cristiana Tristão Rodrigues.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f.70-74.

1. Economia. 2. Educação integral - Avaliação. 3. Política
públicas. 4. Capital humano . I. Universidade Federal de Viçosa.
Departamento de Economia. Programa de Pós-graduação em
Economia. II. Título.

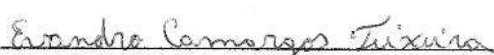
CDD 22. ed. 330

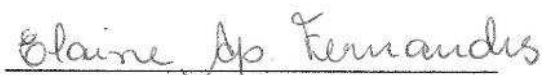
JULIANA MARA DE FÁTIMA VIANA GANDRA

**O IMPACTO DA EDUCAÇÃO EM TEMPO INTEGRAL NO
DESEMPENHO ESCOLAR: UMA AVALIAÇÃO DO
PROGRAMA MAIS EDUCAÇÃO**

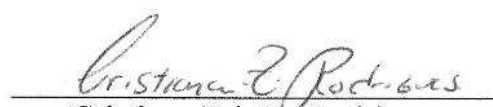
Dissertação apresentada à Universidade
Federal de Viçosa, como parte das
exigências do Programa de Pós-
Graduação em Economia, para obtenção
do título de *Magister Scientiae*.

APROVADA: 23 de fevereiro de 2017.


Evandro Camargos Teixeira


Elaine Aparecida Fernandes


Roni Barbosa Moreira


Cristiana Tristão Rodrigues
(Orientadora)

“Aos meus queridos pais, Paulo e Sueli. ”

AGRADECIMENTOS

Agradeço primeiramente à Deus que em sua infinita bondade, me permitiu retomar o caminho acadêmico em um momento muito especial em minha vida. Aos meus pais, Paulo e Sueli, agradeço imensamente pelo incansável apoio e preocupação. Espero um dia poder retribuir a vocês tudo aquilo que já fizeram por mim. Ao meu irmão, Renan, agradeço a torcida e as energias positivas.

Ao meu marido, Nagai, agradeço pelo companheirismo, pela grande amizade, pelo carinho e pela paciência. Você foi meu grande incentivador neste projeto e sempre tem um papel fundamental em todas as minhas conquistas.

Agradeço à UFV e ao Departamento de Economia que mais uma vez me acolheu me proporcionando toda infraestrutura necessária para a realização deste trabalho. Aos meus professores agradeço pelos profundos ensinamentos e inspiração. Especialmente à prof. ^a Cristiana, agradeço pela disponibilidade, atenção e paciência que sempre me recebeu em sua sala e pelas preciosas contribuições dadas ao trabalho. À prof. ^a Elaine pelo apoio profissional e pessoal. Obrigada pelas boas conversas e pela confiança em meu trabalho! E aos demais membros da banca, professores Evandro e Roni, agradeço profundamente cada sugestão.

Por fim, agradeço à CAPES pelo imprescindível apoio financeiro que recebi durante os 24 meses de mestrado.

SUMÁRIO

LISTA DE TABELAS	vi
LISTA DE FIGURAS	viii
LISTA DE QUADROS	ix
RESUMO	x
ABSTRACT	xii
1 INTRODUÇÃO	1
1.1 Considerações Iniciais	1
1.2 O Problema e sua Importância	4
1.3 Hipótese	9
1.4 Objetivos	9
1.4.1 Objetivo Geral	9
1.4.2 Objetivos Específicos	9
1.5 Estrutura do Trabalho	9
2 A EDUCAÇÃO INTEGRAL E O PROGRAMA MAIS EDUCAÇÃO	10
2.1 Os principais pontos no debate da Educação Integral no Brasil	10
2.2 O modelo de Funcionamento do Mais Educação	12
3 REFERENCIAL TEÓRICO	16
3.1 A Função de Produção da Educação	16
3.2 O impacto de políticas públicas educacionais mensurados pela função de produção da educação	19
4 METODOLOGIA	24
4.1 A Construção do Grupo de Controle por Meio do Propensity Score Matching	24
4.2 Aplicação do Método Diferenças em Diferenças	30
4.3 Variáveis e definição das amostras	33
4.4 Fonte de Dados	38
5 RESULTADOS E DISCUSSÕES	40
5.1 Perfil das escolas analisadas	40
5.2 A Construção dos grupos de controle	47
5.2.1 O grupo de controle para escolas de 5º ano da amostra 1	47
5.2.2 O grupo de controle para escolas de 9º ano da amostra 1	52
5.2.3 Os grupos de controle para as demais amostras	55

5.3 Impacto do Mais Educação para escolas com cinco, três e um ano de programa.....	56
5.3.1 Impacto para escolas com cinco anos de programa – amostra 1	56
5.3.2 Impacto para escolas com três e um ano de programa – amostra 2 e 3	62
5.4 Considerações Finais	67
REFERÊNCIAS BIBLIOGRÁFICAS.....	70
ANEXOS	75

LISTA DE TABELAS

Tabela 1 - Índice de Desenvolvimento da Educação Básica (IDEB) – 2005 a 2013.	2
Tabela 2 - Amostras para 5º ano e 9º ano	35
Tabela 3 - Variáveis de resultado	36
Tabela 4 - Variáveis de Controle (usadas nos modelos Probit)	37
Tabela 5 - Notas médias de escolas participantes e não participantes (amostra 1)..	40
Tabela 6 - Notas médias de escolas participantes e não participantes (amostra 2)..	41
Tabela 7 - Notas médias de escolas participantes e não participantes (amostra 3)..	42
Tabela 8 - Características médias das escolas participantes e não participantes em 2007 (amostra 1)	43
Tabela 9 - Características médias das escolas participantes e não participantes em 2009 (amostra 2)	45
Tabela 10 - Características médias das escolas participantes e não participantes em 2011 (amostra 3)	46
Tabela 11 - Resultados do Probit para participação no Mais Educação, 5º ano, usando dados pré tratamento de 2007 (amostra 1)	48
Tabela 12 - Testes de Média entre escolas do Mais Educação e escolas não participantes, antes e depois do pareamento, 5º ano	50
Tabela 13 - Resultados do Probit para participação no Mais Educação, 9º ano, usando dados pré tratamento de 2007 (amostra 1)	53
Tabela 14 - Testes de Média entre escolas do Mais Educação e escolas não participantes, antes e depois do pareamento, 9º ano	54
Tabela 15 - Impacto do Mais Educação nas notas da Prova Brasil entre 2007 e 2013 (amostra 1)	57
Tabela 16 - Impacto do Mais Educação nas notas da Prova Brasil entre 2007 e 2009 (amostra 1)	59
Tabela 17 - Impacto do Mais Educação nas notas da Prova Brasil entre 2009 e 2013 (amostra 2)	62
Tabela 18 - Impacto do Mais Educação nas notas da Prova Brasil entre 2011 e 2013 (amostra 3)	64
 Tabela A1 - Resultados do Probit para participação no Mais Educação, 5º ano, usando dados pré-tratamento de 2009 (amostra 2).....	 75

Tabela A2 - Resultados do Probit para participação no Mais Educação, 9º ano, usando dados pré-tratamento de 2009 (amostra 2).....	76
Tabela A3 - Resultados do Probit para participação no Mais Educação, 5º ano, usando dados pré-tratamento de 2011 (amostra 3).....	77
Tabela A4 - Resultados do Probit para participação no Mais Educação, 9º ano, usando dados pré-tratamento de 2011 (amostra 3).....	78
Tabela A5 - Testes de Média entre escolas do Mais Educação e escolas não participantes, antes e depois do pareamento, 5º ano (amostra 2).....	79
Tabela A6 - Testes de Média entre escolas do Mais Educação e escolas não participantes, antes e depois do pareamento, 9º ano (amostra 2).....	80
Tabela A7 - Testes de Média entre escolas do Mais Educação e escolas não participantes, antes e depois do pareamento, 5º ano (amostra 3).....	81
Tabela A8 - Testes de Média entre escolas do Mais Educação e escolas não participantes, antes e depois do pareamento, 9º ano (amostra 3).....	82

LISTA DE FIGURAS

Figura 1 - Distribuição dos escores de propensão estimados para as escolas tratadas e controle de 5º ano, antes e depois do pareamento, amostra 1	52
Figura 2 - Distribuição dos escores de propensão estimados para as escolas tratadas e controle de 9º ano, antes e depois do pareamento, amostra 1	55
Figura A1 - Distribuição dos escores de propensão estimados para as escolas tratadas e controle de 5º ano, antes e depois do pareamento, amostra 2	83
Figura A2 - Distribuição dos escores de propensão estimados para as escolas tratadas e controle de 9º ano, antes e depois do pareamento, amostra 2	83
Figura A3 - Distribuição dos escores de propensão estimados para as escolas tratadas e controle de 5º ano, antes e depois do pareamento, amostra 3	84
Figura A4 - Distribuição dos escores de propensão estimados para as escolas tratadas e controle de 9º ano, antes e depois do pareamento, amostra 3	84

LISTA DE QUADROS

Quadro 1 - Resumo dos Macrocampos do Mais Educação a partir de 2013	13
Quadro 2 - Escolas participantes do Programa Mais Educação entre 2008 e 2013.	34

RESUMO

GANDRA, Juliana Mara de Fátima Viana, M.Sc., Universidade Federal de Viçosa, fevereiro de 2017. **O impacto da educação em tempo integral no desempenho escolar: Uma avaliação do Mais Educação.** Orientadora: Cristiana Tristão Rodrigues.

Diante da necessidade urgente de iniciativas que promovam a qualidade da educação brasileira, o governo federal criou em 2007 o Programa Mais Educação. Sua finalidade é levar ao ensino básico os moldes da escola em tempo integral, modelo já previsto na Lei de Diretrizes e Bases da Educação (LDB), mas que permaneceu, por muito tempo, relegado a um segundo plano devido a prioridade de ações que levassem à universalização do acesso ao ensino fundamental. A estratégia do programa é ampliar a jornada escolar de crianças e adolescentes matriculados nas escolas da rede pública brasileira para, no mínimo, 7 horas diárias, contemplando atividades optativas em diversos campos, como o acompanhamento pedagógico, a educação ambiental, prática de esportes, dentre outros. Entre os objetivos da política estão a melhoria no desempenho dos alunos, redução do abandono, da reprovação e da distorção idade/série, além de sua característica inclusiva, que busca proporcionar mais possibilidades socioeducativas aos alunos mais vulneráveis. Após mais de cinco anos de programa, é natural que surjam questionamentos sobre sua efetividade, se os esforços empreendidos nessa perspectiva estão sendo efetivos em melhorar a qualidade da educação básica. Diante disso, o objetivo proposto nesse trabalho foi determinar o impacto do Programa Mais Educação no desempenho escolar de alunos do 5º e 9º ano, em testes da Prova Brasil. Para obter o efeito médio de tratamento do programa, duas técnicas foram utilizadas: o *Propensity Score Matching* e o método de Diferenças em Diferenças. O primeiro método possibilitou a construção de um grupo de controle adequado. O segundo foi empregado para comparar os resultados do grupo de tratamento e de controle, antes e depois da intervenção. Com dados do Censo Escolar e da Prova Brasil para os anos pré-tratamento – 2007, 2009 e 2011 e pós-tratamento – 2013, foram construídas três amostras para escolas com diferentes anos de exposição ao programa: cinco, três e um ano de adesão. Os resultados encontrados não mostraram-se satisfatórios. O efeito do programa foi nulo para o teste de português das turmas de 9º ano em escolas com cinco anos de programa e nos testes de matemática das turmas de 9º ano em escolas com um ano de programa.

Em todas as outras amostras analisadas foram encontrados efeitos negativos sobre as notas de português e matemática das turmas de 5º e 9º ano. Os resultados também mostraram que o efeito negativo é maior para aquelas escolas expostas a mais tempo no programa, contrariando a ideia de que o amadurecimento da política pudesse beneficiar os resultados. Em termos do que foi encontrado no estudo, a ampliação da jornada escolar promovida pelo Programa Mais Educação não mostrou capacidade em melhorar o desempenho escolar de alunos do ensino básico da rede pública, não permitindo comprovar sua importância em prol de uma educação de qualidade. Diante da importância do tema para o desenvolvimento da educação pública brasileira, os resultados apresentados neste estudo fornecem uma valiosa ferramenta para a avaliação do programa, possibilitando elevar o debate à sociedade, que junto aos formuladores de políticas públicas podem promover o aperfeiçoamento do Mais Educação.

ABSTRACT

GANDRA, Juliana Mara de Fátima Viana, M.Sc., Universidade Federal de Viçosa, February, 2017. **The impact of Full-Time Education on School Performance: An Evaluation of Mais Educação.** Adviser: Cristiana Tristão Rodrigues.

Given the urgent need for initiatives to promote the quality of Brazilian education, the federal government created in 2007 the Mais Educação Program. Its purpose is to bring basic education into the mold of a full-time school, a model already envisaged in the Law on Guidelines and Bases of Education (LDB), but which remained for a long time relegated to the background due to the priority of actions that led Universal access to basic education. The program's strategy is to extend the school day of children and teenagers enrolled in Brazilian public schools for at least 7 hours a day, contemplating optional activities in various fields, such as pedagogical accompaniment, environmental education, sports practice, among others. Among the objectives of the policy are the improvement in student performance, reduction of dropout rates, failure rates, age/grade distortion, and its inclusive characteristic, which seeks to provide more socio-educational possibilities for the most vulnerable students. After more than five years of the program, it is natural that questions arise about its effectiveness, if the efforts undertaken in this perspective are being effective in improving the quality of basic education. Therefore, the objective of this study was to determine the impact of the Mais Educação Program on the school performance of students in the 5th and 9th grade, in Prova Brasil. To obtain the average treatment effect of the program, two techniques were used: The Propensity Score Matching and the Differences in Differences method. The first method allowed the construction of an adequate control group. The second was used to compare the results of the treatment and control groups, before and after the intervention. With data from the School Census and the Prova Brasil for the pre-treatment years – 2007, 2009 or 2011 and post-treatment – 2013, three samples were built for schools with different years of exposure to the program: five, three and one year of adherence. The results were not satisfactory. The effect of the program was null for the Portuguese test of the 9th grade classes in schools with five years of program and in the math tests of the 9th grade classes in schools with a program year. In all other analyzed samples, negative effects were found on the Portuguese and mathematics grades of the 5th and 9th grade classes. The results also showed that the negative effect is greater for those schools

exposed to more time in the program, contrary to the idea that the maturing of the policy could benefit the results. In terms of what was found in the study, the expansion of the school day promoted by the Mais Educação Program did not show the capacity to improve the school performance of students of the basic education, not allowing to prove its importance in favor of a quality education. Given the importance of the theme for the development of Brazilian public education, the results presented in this study provide a valuable tool for the evaluation of the program, allowing to raise the debate to the society, which together with the public policy makers can promote the improvement of Mais Educação.

1 INTRODUÇÃO

1.1 Considerações Iniciais

Estudos com foco na questão educacional são frequentes na literatura justificados, principalmente, pelos benefícios que o investimento em capital humano desempenha na promoção do desenvolvimento econômico e social, ajudando tanto no aumento da renda absoluta (de indivíduos e do país como um todo) quanto na diminuição das desigualdades.

Além do crescimento da renda, os retornos sociais da educação causam externalidades positivas, ou seja, proporcionam indiretamente maiores cuidados com a saúde, reduz a criminalidade, facilita a comunicação e promove a consciência política e social além de outros benefícios para a sociedade (BARBOSA FILHO e PESSÔA, 2010).

No Brasil, a Constituição de 1988 abriu caminho para um período de redemocratização, proporcionando mudanças importantes nas políticas educacionais que passaram a ter dois focos: o atendimento escolar e a qualidade do ensino. De fato, o país avançou na cobertura educacional. Segundo dados do IBGE (2014), nos últimos dez anos, houve um crescimento substantivo no acesso à educação infantil. De 2004 a 2013, as taxas de escolarização das crianças de 0 a 3 anos e de 4 e 5 anos de idade subiram de 13,4% e 61,5% para 23,2% e 81,4%, respectivamente. A taxa de frequência escolar bruta das pessoas de 6 a 14 anos de idade permaneceu próxima da universalização, alcançando 98,4% em 2013. Em ritmo mais lento, a proporção de jovens de 15 a 17 anos de idade que frequentava escola cresceu 2,5 pontos percentuais, passando de 81,8% em 2004 para 84,3% em 2013¹.

Contudo, observa-se que o progresso alcançado no acesso à educação não tem sido acompanhado pela melhoria na qualidade. A qualidade da educação brasileira ainda é baixa quando analisados os níveis de aprendizagem adequados a cada série e também em relação aos outros países (VELOSO, 2009). Este cenário pode ser observado pelo Índice de Desenvolvimento da Educação Básica (IDEB) – um importante indicador de qualidade da educação criado pelo INEP/MEC em 2007². Ele

¹ É importante salientar que embora a maior parte de crianças e jovens entre 6 a 17 anos esteja na escola, isto não significa que os alunos estão na série correta. Ver dados de distorção idade-série em IBGE (2014).

² O Ideb foi criado pelo Inep em 2007, em uma escala de zero a dez. O indicador é calculado a partir dos dados sobre aprovação escolar, obtidos no Censo Escolar, e médias de desempenho nas avaliações do Inep, o Saeb e a Prova Brasil. A série histórica de resultados do Ideb se inicia em 2005.

sintetiza dois conceitos igualmente importantes para a qualidade da educação: aprovação e média de desempenho dos estudantes em língua portuguesa e matemática. A evolução dos resultados está exposta na Tabela 1 abaixo.

Tabela 1 - Índice de Desenvolvimento da Educação Básica (IDEB) – 2005 a 2013

Unidade da Federação	Ensino Fundamental – Anos Iniciais		Ensino Fundamental – Anos Finais		Ensino Médio	
	2005	2013	2005	2013	2005	2013
Brasil	3,8	5,2	3,5	4,2	3,4	3,7
Região Norte	3,0	4,3	3,2	3,8	2,9	3,1
Região Nordeste	2,9	4,3	2,9	3,7	3,0	3,3
Região Sudeste	4,6	5,9	3,9	4,6	3,6	3,9
Região Sul	4,4	5,8	3,8	4,3	3,7	3,9
Região Centro-Oeste	4,0	5,5	3,4	4,5	3,3	4,1

Adaptado do Anuário da Educação Básica 2015. Fonte: MEC/INEP/SAEB e Censo Escolar

Os números mostram que o Brasil ainda encontra-se em uma posição bastante desafiadora visto que a meta é atingir o patamar educacional da média dos países da OCDE, que seria um Ideb igual a 6,0. A evolução positiva registrada nos anos iniciais do Ensino Fundamental tem ritmo menor nos anos finais do Ensino Fundamental e se transforma em estagnação no Ensino Médio (TODOS PELA EDUCAÇÃO, 2015). Além disso, tem-se que lidar com as desigualdades regionais, que colocam as regiões Norte e Nordeste em posições inferiores às outras regiões e à média brasileira.

A extensão de resultados desfavoráveis também é encontrada em avaliações internacionais. O Brasil participa do Programa Internacional de Avaliação de Estudantes³ (Pisa, sigla em inglês) desde 2000, sua primeira edição. A partir de então, houve um significativo progresso no desempenho dos estudantes brasileiros, mas o país ainda ocupa as últimas colocações nos *rankings* do exame, apresentando resultados bem preocupantes: em 2012, quase metade dos estudantes do país (49,2%) apresentaram saber apenas o básico em leitura, ficando abaixo do nível de proficiência considerado pelo exame. Em matemática, a situação é ainda pior. Ao todo, 67,1% dos estudantes estão abaixo do nível de proficiência⁴.

³ Promovido pela Organização para a Cooperação e o Desenvolvimento Econômico (OCDE), o Pisa é uma avaliação trienal aplicada a estudantes de 15 anos em cerca de 70 países. Cada edição tem foco em uma área. Em 2000 e em 2009, o Pisa mirou a proficiência em leitura. A matemática foi o tema central em 2003 e novamente em 2012. Em 2006, a prioridade foi ciências, e esse tema se repete em 2015.

⁴ Resultados sintetizados pelo Anuário da Educação Básica 2015 (Todos pela Educação, 2015). O Pisa resume o que os estudantes conseguem fazer em cada disciplina em uma escala com seis níveis. São considerados proficientes alunos que alcançam o nível 2 em cada avaliação.

Os resultados negativos anteriormente apresentados são apenas alguns exemplos do cenário desfavorável da educação brasileira. Situação que instiga discussões que busquem compreender quais seriam os principais determinantes do desempenho escolar e conseqüentemente da qualidade da educação. Os estudos têm apontado para três fatores: os associados à estrutura familiar, os relacionados ao próprio aluno e aqueles atribuídos ao nível da escola (chamados na literatura de fator-escola ou efeito-escola). O efeito escola pode ser atribuído a várias dimensões da realidade escolar, tais como tamanho da escola, turma, recursos, tipo de gestão, qualificação do professor, tempo, entre outros.

Coleman *et al* (1966) e Hanushek (1986) já destacavam que o impacto da escola tem relevância pequena, especialmente quando comparado à parcela explicada pelo *background* familiar. Análises para dados brasileiros foram feitas por Menezes Filho (2007), Paes de Barros *et al* (2001) e Esposito *et al* (2000) e os resultados seguiram esta mesma direção. No entanto, não se pode descartar os fatores escolares na trajetória de aprendizado do aluno que, em alguns casos, podem amenizar a importância do ambiente familiar e social no resultado educacional⁵. Ao pensar em estratégias que incidam sobre o fator-escola, permite-se que políticas educacionais entrem na agenda pública e busquem ações efetivas que promovam a qualidade do ensino.

Nesse sentido, nos últimos anos, o país vivenciou a criação e a implementação de diversas políticas públicas e programas, iniciando uma nova etapa na reforma educacional. A Lei nº 11.274 de 2006 instituiu o Ensino Fundamental de Nove Anos que permitiu a matrícula da criança aos seis anos de idade, iniciando o processo de alfabetização mais cedo e incluindo um ano a mais no Ensino Fundamental. Em 2007, entrou em vigor o Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação⁶ (Fundeb). Esta medida proporcionou uma nova forma de financiamento da educação, aumentando os aportes a Educação Básica e garantindo maiores recursos por aluno na rede pública.

⁵ Sobre estudos relativos ao efeito-escola ver Barros e Lam (1996), Felício e Fernandes (2005), Biondi e Felício (2007), Riani e Rios-Neto (2008), Barbosa (2009), entre outros.

⁶ O Fundeb, criado em 2006, substituiu o Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério (Fundef) e se tornou um dos mais importantes mecanismos de distribuição dos recursos destinados à Educação Básica no Brasil. Enquanto o Fundef se restringia ao financiamento do Ensino Fundamental, o Fundeb contempla toda a Educação Básica. O mecanismo de financiamento é composto pela arrecadação estadual de impostos, complementada pelo governo federal. A distribuição dos recursos é feita com base no número total de alunos matriculados, o que contribui para a redução das desigualdades (Todos pela Educação, 2015).

Também entre os programas criados com o objetivo de melhorar a qualidade da educação, estão as políticas de ampliação da jornada escolar ou de educação integral. Nessa perspectiva, o governo federal criou o Programa Mais Educação. Instituído pela Portaria Interministerial nº 17 de 24 de abril de 2007 e regulamentado pelo Decreto 7.083 de 27 de janeiro de 2010, constitui-se como estratégia do Ministério da Educação para induzir a ampliação da jornada escolar e a organização curricular na perspectiva da Educação Integral⁷.

1.2 O Problema e sua Importância

A educação integral começou a ser definida legalmente por meio da Lei de Diretrizes e Bases da Educação Nacional (LDB) de 1996, que prevê a ampliação progressiva da jornada escolar em direção a uma educação integral (artigo 34 e § 5º do artigo 87º da Lei nº 9.394 de 20 de dezembro de 1996). Mais recentemente, o novo Plano Nacional da Educação – PNE (instituído pela Lei nº 13.005 de 25 de junho de 2014), reafirmou o direcionamento dos esforços para uma educação brasileira em tempo integral por meio da meta nº 6, que impõe que até 2024, no mínimo, 50% (cinquenta por cento) das escolas públicas devem oferecer educação em tempo integral, de forma a atender, pelo menos, 25% (vinte e cinco por cento) dos alunos da educação básica (BRASIL, 2014).

A consolidação da educação integral enquanto política pública fundamenta-se na relação positiva entre o número de horas que o aluno passa na escola e seus resultados de aprendizado. Mas, no geral, os benefícios da ampliação da jornada escolar ainda não estão consolidados na literatura. Pesquisas como as de Gomes (2005), Menezes Filho (2007) e Unicef (2010) apontam práticas e resultados que demonstram relações positivas entre tempo e desempenho escolar. Contudo, Cavaliere (2007) observa que não há uma associação automática entre mais tempo e melhor desempenho e que esta relação deve ser vista com cautela dadas as particularidades de cada caso.

As práticas que ampliam o tempo na escola buscam priorizar a formação de crianças, adolescentes e jovens mais vulneráveis, tendo também como propósito

⁷ Existe na literatura uma divisão entre educação integral e educação em tempo integral. A primeira defende a ideia de uma formação “mais completa possível” do aluno, priorizando não só a aprendizagem de conteúdo, mas também as demais dimensões do desenvolvimento humano. Já a segunda trata da ampliação da carga horária escolar. Estes termos serão melhor apresentados em seção própria. Mas por muitas vezes neste trabalho serão tratados como sinônimos uma vez que o tempo é condição necessária para a educação integral.

promover a equidade e inclusão social.⁸ Espera-se que o reforço cognitivo por meio da oferta de múltiplas atividades de aprendizagem amplie o repertório cultural e social de crianças e adolescentes, favorecendo maior inclusão e promovendo melhores resultados escolares, contribuindo consequentemente para aumentar a qualidade da educação (UNICEF, 2011).

No Brasil, as experiências de ampliação da jornada escolar tomaram fôlego somente a partir de 2008, embora esta iniciativa já estivesse presente em algumas cidades brasileiras como Rio de Janeiro nas décadas 1980-1990, Belo Horizonte em 2006, Apucarana (PR) em 2001, Cuiabá (MT) e Sorocaba (SP) em 2007, com a coordenação de suas respectivas secretarias municipais de educação. Estes programas se caracterizavam principalmente por serem práticas isoladas, restritas (poucos alunos atendidos) e implementadas de forma diferenciada, sem uma padronização das metodologias de trabalho. A implementação do Programa Mais Educação, em 2008, concretizou as ações do governo federal no âmbito da educação em tempo integral, contribuindo para o adensamento dessa perspectiva na educação pública brasileira.

A estratégia do programa é induzir a construção da agenda de educação integral nas redes estaduais e municipais de ensino por meio da ampliação da jornada escolar nas escolas públicas para, no mínimo, 7 horas diárias⁹, contemplando atividades optativas em diversos macrocampos, como o acompanhamento pedagógico, a educação ambiental, prática de esportes, dentre outros¹⁰. E, de fato, o programa tem contribuído para o avanço na educação em tempo integral. De acordo com dados do Censo Escolar (MEC, 2014), o número de matrículas em tempo integral nas escolas públicas brasileiras foi superior a três milhões de alunos matriculados no ensino fundamental em 2013, apresentando uma elevação de 45,2% em relação a 2012. Isto corresponde a 12,5% dos alunos da rede pública, frente aos 8,3% registrados em 2012 e 4,7% em 2010.

⁸ Como destaca MEC (2011), esse ideal está presente na legislação educacional brasileira e pode ser apreendido na Constituição Federal, nos artigos 205, 206 e 227; no Estatuto da Criança e do Adolescente (Lei n.º 9089/1990); na Lei de Diretrizes e Bases da Educação Nacional (Lei n.º 9394/1996); no Plano Nacional de Educação (Lei n.º 10.179/2001), dentre outros dispositivos legais.

⁹ A jornada de sete horas diárias é a jornada mínima considerada de tempo integral pelo governo federal. É interessante ressaltar que a média diária de horas-aula do ensino fundamental brasileiro é de 4,5 horas (MEC, 2015a).

¹⁰ As atividades fomentadas bem como o modelo de funcionamento do programa são descritas mais detalhadamente na próxima seção.

Com estratégias bem definidas e com seu alcance aumentando a cada ano, o Programa Mais Educação vem-se consolidando na agenda governamental, alicerçado em aportes de recursos cada vez maiores disponíveis às políticas educacionais¹¹.

Diante do não consenso sobre os benefícios da educação integral e a necessidade urgente de ações que promovam a educação brasileira, é necessário que sejam feitas avaliações da efetividade do programa, principalmente sobre seu impacto na qualidade da educação, por meio da melhoria do desempenho escolar dos estudantes. De acordo com Rua (2009), uma avaliação de efetividade permite que outros efeitos de espectro mais amplos, chamados de secundários ou imprevistos, sejam mensurados à luz dos objetivos principais conferindo valor à política. Portanto, apesar de não ser um objetivo imediato da política, é razoável esperar que as mudanças provocadas pelo programa nas escolas participantes beneficiem o desenvolvimento cognitivo dos alunos e consequentemente o desempenho escolar.

Como já foi discutido até aqui, a qualidade destaca-se como a principal carência educacional no país e, neste contexto, a gestão adequada de políticas educacionais torna-se um elemento fundamental para que este cenário sofra mudanças mais rápidas e significativas. Estudos com o foco em avaliações de políticas governamentais tem como objetivo assegurar o aperfeiçoamento dos programas, permitindo aos agentes públicos, avaliar seus resultados e decidir pela manutenção, interrupção ou revisão das ações implementadas. Funcionam também como mecanismos poderosos de gestão capazes de dar maior transparência às ações públicas, contribuindo para a participação e controle social.

A avaliação de políticas que ampliam a jornada escolar já foi tema de trabalhos na literatura internacional. Estes trabalhos investigaram os efeitos de um dia letivo mais extenso nos resultados cognitivos dos alunos. De Cicca (2007) estimou o impacto do dia escolar estendido no desempenho de crianças matriculadas no Jardim de Infância nos Estados Unidos. O autor encontrou resultados positivos sobre testes padronizados de Matemática e Leitura.

Cerdan-Infantes and Vermeersch (2007) avaliaram o Programa de Escola em Tempo Integral do Uruguai, em testes padronizados de alunos da 6ª série. Usando a técnica de pareamento, os autores constroem um grupo comparável de escolas e

¹¹ De 2008 a 2013 o investimento direto em educação básica passou de 3,9% do PIB para 4,4%. O gasto médio por aluno aumentou 58% neste mesmo período. Dados disponíveis em <http://www.todospelaeducacao.org.br/indicadores-da-educacao>. Acesso em 21/12/2016.

mostram que os alunos em escolas atendidas pelo programa melhoraram seus resultados em 0,07 de desvio padrão em Matemática e 0,04 em Linguagem, por ano de participação no programa de tempo integral. Ainda na América Latina, os trabalhos de Marín (2006) e Bellei (2009) encontraram impactos positivos no desempenho de alunos participantes do Programa chileno de extensão da jornada escolar (*Jornada Escolar Completa*).

Seguindo esta linha, alguns trabalhos já se preocuparam em avaliar os efeitos de programas de educação integral no desempenho dos estudantes brasileiros. Em âmbito regional, dois estudos foram realizados com enfoque em programas estaduais. Aquino e Kassouf (2011) examinaram o impacto do programa Escola de Tempo Integral da rede pública do estado de São Paulo sobre o desempenho escolar nos anos de 2007 e 2008. Os resultados não mostraram melhoria significativa no desempenho acadêmico dos estudantes que participaram do programa, quando comparados com aqueles que frequentaram o ensino regular. Já Soares *et al* (2014) mostram que o programa mineiro Projeto Escola de Tempo Integral apresentou resultados positivos nas avaliações em larga escala de Língua Portuguesa e Matemática no período de 2006 a 2009.

Especificamente ao Programa Mais Educação, destacam-se os trabalhos de Pereira (2011), Mendes (2011) e Xerxenevsky (2012) que encontraram diferentes impactos do programa em testes de Língua Portuguesa e Matemática. Pereira (2011) e Xerxenevsky (2012) preocuparam-se com o impacto a nível estadual. O primeiro analisou o impacto do Mais Educação no desempenho dos alunos em Minas Gerais. O estudo não encontrou nenhum efeito significativo da extensão do dia escolar em testes de Língua Portuguesa e Matemática. Já o segundo, estimou o impacto do Mais Educação no Rio Grande do Sul. Para o 5º ano, houve um efeito positivo e estatisticamente significativo sobre os resultados dos testes de Português e negativo nos testes de Matemática. Não houve resultados significativos sobre os resultados dos testes para os alunos do 9º ano. Estendendo os resultados ao âmbito nacional, Mendes (2011) avaliou os efeitos causais do programa no desempenho dos alunos de 5º e 9º anos de escolas primárias públicas do Brasil, medido pela Prova Brasil 2009. O estudo não encontra nenhum efeito estatisticamente significativo do Mais Educação no desempenho acadêmico dos alunos em Português e uma diminuição significativa nos resultados dos testes em Matemática.

Os estudos citados anteriormente não apontam uma única direção para os efeitos do programa e se caracterizam por serem análises de curto prazo, visto que seus efeitos foram avaliados em resultados nos anos de 2008 e, principalmente 2009. É importante ressaltar que os resultados de curto prazo podem ser afetados pelas dificuldades e desafios encontrados na implementação e gestão do programa. Neste sentido, o tempo torna-se essencial para a adequação das ações e sua resposta no desempenho dos alunos, justificando a realização de análises de impacto a médio prazo.

Somente um estudo explorou o impacto do programa em um período de tempo maior. Recentemente, Almeida *et al* (2015) apresentou os resultados de uma pesquisa que considerou o período de análise até 2011. O estudo encontrou impacto negativo do programa sobre o desempenho médio das escolas em matemática tanto no curto quanto no médio prazo¹². Porém, os resultados indicaram que os efeitos negativos em matemática reduzem-se com o passar do tempo. Segundo Almeida *et al* (2015), a consolidação do programa e a melhoria em sua execução obtidas com o tempo podem melhorar os impactos e reduzir os efeitos negativos do programa, sugerindo que a avaliação do programa em períodos mais longos pode levar a resultados diferentes.

Nessa perspectiva, esta pesquisa se distingue dos demais trabalhos encontrados na literatura por apresentar duas contribuições. A primeira consiste em ampliar o período analisado, estendendo a avaliação a escolas com cinco anos de exposição ao programa. Uma avaliação mais extensa é importante para verificar se os resultados encontrados até agora podem estar sendo influenciados por problemas de gestão e/ou implementação. A segunda é comparar os impactos do programa Mais Educação em amostras com tempo de exposição distintos (escolas com três e um ano de exposição ao programa), o que permitiu verificar se as escolas que aderiram ao programa mais recentemente se beneficiaram de uma melhoria da política.

Portanto, pretende-se com este estudo ampliar o conhecimento sobre os impactos que os programas governamentais de educação implementados recentemente, sobretudo os de educação integral, provocam no desempenho escolar dos alunos e na qualidade da educação pública brasileira. Além disso, procura-se promover uma discussão em torno da ampliação do tempo escolar como estratégia de

¹² O estudo definiu como curto prazo a exposição ao programa no período de um ano. O médio prazo, abrange uma exposição de três anos. Para efeito deste trabalho, o termo médio prazo corresponderá a um período de exposição de cinco anos.

melhoria da educação pública brasileira. Esta pesquisa é importante para fomentar uma avaliação de efetividade do programa, contribuindo para determinar se o programa está atingindo seus objetivos ou se eles precisam ser revistos em função dos resultados encontrados.

1.3 Hipótese

A ampliação da jornada escolar promovida pelo Programa Mais Educação, tem efeito positivo sobre o desempenho escolar de alunos do ensino fundamental da rede pública de ensino.

1.4 Objetivos

1.4.1 Objetivo Geral

Avaliar o impacto do Programa Mais Educação no desempenho médio das escolas da rede pública de ensino brasileira no período entre 2007 e 2013.

1.4.2 Objetivos Específicos

- a) Determinar o impacto da ampliação da jornada escolar sobre as notas de Português e Matemática da Prova Brasil, aplicada aos alunos do 5º e 9º ano das escolas participantes.
- b) Comparar o impacto em escolas com diferentes tempos de exposição ao programa.

1.5 Estrutura do Trabalho

Este trabalho está estruturado em seis seções. Além da seção introdutória, na seção 2 faz-se uma discussão em cima dos principais pontos da Educação Integral como política pública educacional além de uma apresentação do programa Mais Educação. Na seção 3, apresenta-se o referencial teórico que norteia este estudo, relacionando o desempenho escolar aos seus fatores determinantes. A metodologia usada para testar a hipótese levantada e os objetivos propostos é descrita na seção 4. Os resultados obtidos são apresentados e discutidos na seção 5 e, por fim, a seção 6 conclui este trabalho.

2 A EDUCAÇÃO INTEGRAL E O PROGRAMA MAIS EDUCAÇÃO

2.1 Os principais pontos no debate da Educação Integral no Brasil

A Lei de Diretrizes e Bases da Educação Nacional (LDB) deu um grande passo no debate da educação integral ao prever o aumento progressivo da jornada escolar para o regime de tempo integral (art. 34 e 87) e favorecer projetos de educação integral. Antes disso, porém, duas experiências já antecipavam esse desafio e foram pioneiras na experiência brasileira de escola pública em tempo integral. O Centro Educacional Carneiro Ribeiro – Escola-Parque –, na Bahia dos anos de 1950, e os Centros Integrados de Educação Pública (Cieps), no Rio de Janeiro dos anos 1980 tinham por objetivo complementar o currículo formal do ensino fundamental com atividades diversificadas, permitindo ao aluno a oportunidade de participar de um conjunto de experiências esportivas, artísticas e recreativas, além de incluir atividades relacionadas à saúde e de garantir atendimento médico-odontológico (ERNICA, 2006).

Apesar do conhecimento destas experiências e da previsão legal, a proposição da educação integral enquanto política pública permaneceu, por muito tempo, relegado a um segundo plano provavelmente devido a prioridade de ações que levassem à universalização do acesso ao ensino fundamental. Após os esforços empenhados na conquista desta meta prioritária, a educação integral consolida-se enquanto política pública, confirmada por iniciativas presentes em todas as esferas governamentais – municipal, estadual e federal –, difundidas nas diferentes regiões do país. Trata-se em sua grande maioria, de iniciativas recentes, com seu início entre os anos 2005 e 2009. A política é então desenhada buscando não só o acesso, mas principalmente garantir a permanência de crianças e adolescentes na escola, assegurando-as um ensino público de qualidade.

Mesmo confiantes do potencial da educação integral para alavancar a aprendizagem e reduzir as desigualdades do país, há na literatura um grande debate sobre a qualidade desta oferta e em sua capacidade de atingir os propósitos desenhados pela política. Muitas vezes, o ceticismo quanto aos efeitos da jornada escolar estendida encontra-se explicada na dicotomia “educação integral” e “educação em tempo integral”, conceitos que podem definir os alcances da política a ser implementada.

O termo educação em tempo integral limita-se à ampliação do tempo escolar, o tempo diário de permanência dos alunos na escola, com a incorporação de atividades diferenciadas de acordo com os objetivos de cada escola (ou programa). O segundo

termo é definido de forma mais abrangente. De acordo com MEC (2009a), a Educação Integral se caracteriza pela ideia de uma formação “mais completa possível” para o ser humano, embora não haja consenso sobre quais pressupostos e metodologias a constituiriam.

Segundo Cavaliere (2007), na concepção de educação integral, uma escola precisa envolver as múltiplas dimensões da vida das crianças e adolescentes, promovendo a eles vivência reflexivas. Para Maurício (2009), a educação integral permite ao aluno desenvolver seus aspectos afetivo, cognitivo, físico e social conjuntamente. O autor afirma ainda que as atividades nas quais os alunos são expostos devem envolver multiplicidade de aspectos para benefício de seu desenvolvimento.

O tempo integral é visto como essencial para que estes aspectos sejam incorporados à vida escolar. Porém, segundo Cavaliere (2007), maior quantidade de tempo não determina por si só práticas escolares qualitativamente diferentes. Há ainda dentro desta discussão, o paralelismo turno e contraturno, modelo adotado pela maioria das escolas brasileiras em tempo integral. Neste modelo, as escolas têm suas atividades organizadas em turno, período em que os estudantes têm aulas de conteúdos presentes no “núcleo comum” (disciplinas obrigatórias) e contraturno, quando as atividades desenvolvidas na ampliação do tempo na escola assumem o caráter “complementar” no formato artístico, cultural e esportivo. Diante da perspectiva da educação integral, alguns estudos como Unicef (2011) mostram que este formato pode trazer limitações ao aprendizado e tem consigo o desafio de aproximar os tempos, os turnos e a metodologia de trabalho com propostas de atividades integradas que possam despertar o interesse dos estudantes (UNICEF, 2011).

Nas publicações do Ministério da Educação que fundamentam o debate nacional¹³, há uma clara preocupação para que a proposta de educação integral se encaixe no contexto de vulnerabilidade social vivenciado por muitas crianças e adolescentes no país. Desta forma, segundo MEC (2009a), a proposta de Educação Integral deve estar conjugada à ideia de proteção social, pressupondo políticas integradas (intersetoriais) que considerem, além da educação, outras demandas dos alunos, articuladas entre os campos da educação, do desenvolvimento social, da saúde, do esporte, da inclusão digital e da cultura.

¹³ Ver MEC, 2009a e MEC, 2009b.

Nessa perspectiva, a qualificação do tempo ampliado deve considerar atividades educativas diferenciadas que sejam capazes de promover o rendimento escolar tanto pelo desenvolvimento de habilidades cognitivas e intelectuais quanto pela superação das implicações trazidas pela situação de vulnerabilidade e risco social. Resta saber se os programas que promovem a educação integral, como o Mais Educação, estão estruturados de forma necessária a apoiar as escolas brasileiras neste que parece ser um desafio quando observada a precariedade da educação básica ofertada hoje.

2.2 O modelo de Funcionamento do Mais Educação

O Programa Mais Educação operacionalizado pela Secretaria de Educação Básica – SEB, promove a extensão da jornada escolar financiando atividades extras realizadas nas escolas através de recursos do Programa Dinheiro Direto na Escola – PDDE do Fundo Nacional de Desenvolvimento da Educação – FNDE para as escolas prioritárias.

O governo federal repassa recursos para ressarcimento de monitores, aquisição dos kits de materiais, contratação de pequenos serviços e obtenção de materiais de consumo e permanentes para a realização de atividades optativas organizadas por temas, chamados de macrocampos. Até 2012, o MEC propôs para as escolas urbanas, atividades organizadas em 10 (dez) macrocampos¹⁴, sendo eles:

- Acompanhamento pedagógico;
- Educação ambiental;
- Esporte e lazer;
- Direitos humanos em educação;
- Cultura e artes;
- Cultura digital;
- Promoção da saúde;
- Comunicação e uso de mídias;
- Investigação no campo das ciências da natureza e
- Educação econômica.

As escolas inscritas no programa escolhiam seis atividades dentro das possibilidades ofertadas, de acordo com seu projeto educacional em curso. Uma destas

¹⁴ O detalhamento dos macrocampos bem como suas atividades pode ser encontrado em MEC (2013).

atividades obrigatoriamente deve compor o macrocampo acompanhamento pedagógico. Este, por sua vez, contempla as diferentes áreas do conhecimento, sendo elas: Alfabetização; Ciências; História e Geografia; Línguas Estrangeiras e Matemática. A atividade escolhida dentro deste macrocampo deve favorecer a articulação entre o currículo estabelecido da escola e as atividades pedagógicas propostas pelo Programa Mais Educação. As atividades podem ser realizadas dentro ou fora da área da escola, e, neste caso, através de parcerias com ONGs ou entidades públicas/privadas (MEC, 2011).

Em 2012, as escolas do campo foram incluídas no programa e, para essas unidades de ensino, foi criado um desenho de organização das atividades em 07 (sete) macrocampos, que foi mantido para a adesão ao Mais Educação em 2013. São eles:

- Acompanhamento Pedagógico;
- Agroecologia;
- Cultura, Artes e Educação patrimonial;
- Educação em Direitos Humanos;
- Esporte e Lazer;
- Iniciação Científica;
- Memória e História das Comunidades Tradicionais.

Em 2013, foi pensado um agrupamento dos macrocampos disponíveis às escolas urbanas, e a partir de então, quatro atividades deveriam ser escolhidas dentro do universo exposto no Quadro 1:

Quadro 1 - Resumo dos Macrocampos do Mais Educação a partir de 2013

Escolas urbanas que já desenvolviam o Mais Educação	Escolas urbanas com ingresso no Mais Educação em 2013
Acompanhamento Pedagógico Comunicação, Uso de Mídias e Cultura Digital e Tecnológica Cultura, Artes e Educação Patrimonial Educação Ambiental e Sociedade Sustentável Educação em Direitos Humanos Promoção da Saúde Esporte e Lazer	Acompanhamento Pedagógico Comunicação, Uso de Mídias e Cultura Digital e Tecnológica Cultura, Artes e Educação Patrimonial Educação Ambiental, Desenvolvimento Sustentável e Economia Solidária e Criativa/Educação Econômica Esporte e Lazer

Fonte: Elaborado pela autora com a partir de MEC (2013).

A adesão das escolas é feita de forma voluntária, mas seguindo critérios de priorização ao atendimento, definidos anualmente pelo MEC. Em 2008, primeiro ano de implementação do programa, as escolas atendidas apresentavam baixo IDEB e eram situadas em capitais ou cidades de regiões metropolitanas com mais de 200 mil habitantes, totalizando 1409 escolas atendidas e 386 mil estudantes (MEC, 2009c). Em 2010, foram adicionados como critérios, escolas contempladas com PDDE/Integral no ano de 2008 e 2009 e escolas situadas nas capitais e nas cidades das nove regiões metropolitanas, bem como aquelas com mais de 90 mil habitantes. Cerca de 10 mil escolas e 2,3 milhões de alunos foram beneficiados neste ano. Já em 2011, eram aptas ao programa escolas estaduais ou municipais de baixo IDEB que foram contempladas com o PDDE/Escola 2009, localizadas em territórios de vulnerabilidade social e situadas em cidades com população igual ou superior a 18.844 habitantes (MEC, 2015b).

Seguindo o plano de expansão de 2012, as escolas passariam a ser selecionadas pelos seguintes critérios prioritários: serem do campo ou urbanas, sendo as urbanas com índices igual ou superior a 50% de estudantes beneficiários do Programa Bolsa Família. Dessa forma, participam da adesão ao Mais Educação 32.074 escolas, com a perspectiva de chegar a 45.000 em 2013. A previsão era que até 2014, o Programa Mais Educação estivesse em todo território nacional, chegando a 60.000 escolas (MEC, 2013).

Apesar de cada escola poder estabelecer seu critério de seleção dos estudantes atendidos, o MEC fornece algumas orientações gerais para a definição destes critérios, sendo as principais: i) estudantes com defasagem idade/série; ii) estudantes dos anos finais da 1ª fase do ensino fundamental (4ªsérie/5ºano) e da 2ª fase do ensino fundamental (8ªsérie/9ºano), entre os quais há maior saída extemporânea; e iii) estudantes de séries/anos nos quais são detectados índices de saída extemporânea e/ou repetência (MEC, 2011).

O Programa Mais Educação estabelece o mínimo de cem estudantes para o início das atividades (exceto nas escolas em que a quantidade de estudantes registrados no Censo Escolar do ano anterior seja inferior a este número), contudo, não estabelece número máximo de estudantes significando que cada escola pode então selecionar quais e quantos alunos participarão do programa em um dado ano, não necessariamente chegando à totalidade deles (MEC, 2011).

A partir de 2016, o programa sofreu uma reformulação passando a se chamar Programa Novo Mais Educação. Instituído pela Portaria nº 1.144, de 10 de outubro de 2016, o novo formato do programa direciona suas ações ao desenvolvimento da capacidade de aprender, tendo como meios básicos o pleno domínio da leitura, da escrita e do cálculo, passando a ter como um de seus objetivos principais a melhoria do rendimento e desempenho escolar (MEC, 2017).

3 REFERENCIAL TEÓRICO

3.1 A Função de Produção da Educação

Os benefícios da educação começaram a ser sistematizados pela Teoria do Capital Humano, que demonstra a estreita relação da habilidade produtiva e o estoque de conhecimento adquirido pelo homem com o crescimento da renda individual e agregada. Em trabalhos pioneiros como os de Schultz (1960; 1961), Becker (1962) e Mincer (1958; 1974), o gasto acumulado em educação passou a ser considerado investimento em capital humano, indicando que a renda futura do indivíduo depende de uma escolha individual e racional quanto ao seu tempo de formação. O crescimento da escolaridade média da população economicamente ativa proporcionaria aumentos na produtividade, podendo, portanto, indicar que uma parcela significativa da expansão da renda agregada seria consequência da expansão do capital humano.

De maneira geral, pode-se observar que a medida mais frequentemente utilizada para captar a educação tem sido anos de estudo concluídos. Nos trabalhos citados anteriormente, a escolaridade é comumente utilizada como uma medida aproximada da habilidade individual. Entretanto, esta medida de resultado educacional apresenta uma limitação. Nestes casos, assume-se que um ano de escolaridade produz o mesmo efeito sobre o desempenho do aluno ou em suas habilidades ao longo do tempo e em todos os países. Esta medida simplesmente conta o tempo gasto nas escolas sem julgar o que acontece nas escolas - portanto, não fornecem um quadro completo ou preciso de resultados (HANUSHEK, 1979). Da mesma forma, Bowles (1970) destaca que os efeitos da educação na produtividade do trabalho e nos rendimentos vão além dos anos de escolaridade adquiridos e precisa, portanto, considerar o desenvolvimento de habilidades e atitudes cognitivas em consequência do serviço escolar.

Não se pode afirmar que todas as escolas dentro de um país ou entre países são as mesmas em termos de conhecimento transmitido e qualidade. A incorporação dos efeitos escolares na análise permite entender a habilidade cognitiva adquirida na escolarização como reflexo da qualidade da educação oferecida¹⁵. Então, o que se propõe é ir além dos benefícios proporcionados pela quantidade de educação adquirida, e entender como o desempenho de habilidades individuais pode variar em

¹⁵ De acordo com Hanushek e Luque (2003), utiliza-se o termo qualidade para se referir à base de conhecimento e habilidades analíticas transmitidas pelas escolas.

resposta à oferta educacional, dado as diferenças encontradas nas políticas e programas escolares.

Estudos recentes encontraram evidências apontando uma relação de causalidade entre qualidade da educação e renda (individual e agregada). Uma ampla variedade de trabalhos com foco no mercado de trabalho aponta que as diferenças individuais em capacidades cognitivas podem afetar os rendimentos¹⁶. Hanushek e Kimko (2000) demonstram que as diferenças de qualidade em escolas também têm um impacto importante na produtividade e nas taxas de crescimento nacional. Desta forma, a literatura também se concentrou em investigar os determinantes da qualidade da educação, desenvolvendo um modelo que se convencionou chamar de função de produção da educação (ou função de produção educacional).

O Relatório Coleman (COLEMAN *et al*, 1966) é comumente citado na literatura como o primeiro documento que se baseia na estrutura conceitual da função de produção da educação, preocupando-se em entender o relacionamento entre insumos e produtos escolares e os determinantes dos resultados acadêmicos dos estudantes. De acordo com Bowles (1970), a função de produção educacional pode ser definida como a relação existente entre os insumos escolares e dos estudantes com uma medida de produto escolar, podendo ser usada tanto em estudos que discutem a formação de capital humano quanto naqueles que investigam a alocação ótima de recursos educacionais.

Segundo Hanushek (1986), o conceito de função de produção é uma poderosa ferramenta que em sua forma básica pode ser aplicada em uma ampla variedade de setores, atingindo até mesmo a educação. Apesar disso, Hanushek (1986) ressalta as diferenças entre a função de produção padrão e a educacional. O autor observa que na teoria microeconômica, a função de produção é geralmente conhecida pelos tomadores de decisão e envolve poucos insumos, que são mensurados perfeitamente e podem variar livremente. Porém, na realidade educacional há diferenças - a função não é conhecida e alguns insumos não podem ser alterados a livre escolha das políticas educacionais.

De acordo com Bowles (1970), inicialmente a função de produção da educação pode ser definida por:

¹⁶ Esta tendência pode ser encontrada nos trabalhos de Barbosa Filho e Pessôa (2010) e Hanushek e Luque (2003).

$$A = f(X_1, \dots, X_m; X_n \dots, X_v; X_w, \dots, X_z) \quad (1)$$

em que A é alguma medida de produto escolar; $X_1, \dots, X_m$ são variáveis correspondentes ao ambiente escolar que podem envolver características de infraestrutura da escola, qualificação dos professores, o período de tempo em que o aluno é exposto aos insumos escolares, etc; $X_n \dots, X_v$ refere-se a influência no aprendizado adquirida fora da escola como a escolaridade dos pais; e $X_w, \dots, X_z$ são variáveis que representam habilidades inerentes ao estudante, já presentes antes da entrada na escola.

Nesta função, a qualidade da educação é representada pelo desempenho dos estudantes em testes que medem conhecimento e a habilidade analítica, geralmente medidas pelo resultado dos alunos em testes padronizados de matemática, ciências ou língua. Hanushek e Luque (2003) afirmam que embora a qualidade da educação seja difícil de se definir com precisão e a utilização de testes padronizados como medida desta variável seja controverso na literatura, ainda assim são os melhores indicadores disponíveis de qualidade para a relação causa e efeito que se deseja estabelecer.

Como já previsto em Hanushek (1986), também é bastante comum o uso de medidas de quantidade como produto educacional, das quais tem-se as taxas de evasão, aprovação, frequência escolar e distorção idade-série. Contudo, as habilidades cognitivas medidas por testes padronizados ainda são considerados como a dimensão mais importante da qualidade educacional por decisores políticos, em parte relacionado ao esforço de promover *accountability* (HANUSHEK, 2006).

A maior parte dos estudos envolvendo a produção educacional relaciona os resultados alcançados pelos estudantes, seja o nível de proficiência em testes padronizados ou medidas de quantidade, com as suas características individuais, familiares, com outros estudantes na escola e variáveis da própria escola. Segundo Hanushek (2008), o modelo subjacente que tem evoluído como resultado desta linha de pesquisa é muito simples: o produto do processo educacional representado pelo desempenho de cada aluno é diretamente relacionado aos insumos que podem ser controlados pelos formuladores de políticas (por exemplo, as características das escolas, professores, currículos, e assim por diante) e também aqueles não tão controlados, como famílias e amigos e os dons inatos ou capacidades de aprendizagem dos alunos.

O *background* familiar é construído por características socioeconômicas como a educação dos pais, a renda e o tamanho da família. Os insumos escolares normalmente incluem as características dos professores (nível de escolaridade, experiência, sexo, raça, etc.), da organização escolar (turmas, infraestrutura, despesas administrativas, políticas educacionais, etc.) e as características inerentes à comunidade na qual a escola está inserida. As características dos pares, quando inseridas, refletem as características socioeconômicas agregadas ou o desempenho medido para uma escola ou sala de aula.

Hanushek (1979) destaca que apesar do resultado educacional poder ser medido em pontos discretos no tempo, o processo educacional é cumulativo. Insumos aplicados em algum momento do passado podem afetar os níveis atuais de desempenho dos alunos. Desse modo, considerando o efeito do tempo, Hanushek (1979) apresenta seu modelo para a função de produção educacional que pode ser definida da seguinte forma:

$$A_{it} = f(B_i^{(t)}, P_i^{(t)}, S_i^{(t)}, L_i) \quad (2)$$

em que A_{it} corresponde ao rendimento do estudante i no ano t ; $B_i^{(t)}$ é o vetor com variáveis referentes ao *background* familiar cumulativo ao tempo t ; $P_i^{(t)}$ o vetor com características dos pares também cumulativo ao tempo t ; $S_i^{(t)}$ é o vetor com os insumos escolares que foram acumulados até o tempo t ; e, L_i corresponde às habilidades inatas dos estudantes.

Se a estimação da equação (2) pode ser feita em diferentes pontos no tempo, ao considerar t^* como momento passado, Hanushek (1979) sugere uma versão alternativa deste modelo que poderá ser aplicada para estimar uma mudança no rendimento do aluno entre os tempos t e t^* , expressa pela seguinte equação:

$$A_{it} = f^*(B_i^{(t-t^*)}, P_i^{(t-t^*)}, S_i^{(t-t^*)}, L_i, A_{it^*}) \quad (3)$$

3.2 O impacto de políticas públicas educacionais mensurados pela função de produção da educação

Dentro de uma perspectiva de políticas públicas educacionais, a abordagem da função de produção da educação oferece dois tipos de subsídios. Ao estimar os parâmetros estruturais de f pode-se determinar quais dos vários insumos do processo educacional são mais importantes ao desenvolvimento das habilidades cognitivas dos

estudantes - seu peso sobre o desempenho do estudante nos testes avaliados. Além disso, como destaca Bowles (1970), o entendimento do resultado estimado para a função de produção educacional permite a formação de expectativas mais precisas quanto aos custos e benefícios das ações de políticas educacionais implementadas.

Em uma abordagem mais recente, Glewwe e Kremer (2006) discutem questões relativas à provisão da educação: como programas e sistemas educacionais afetam a quantidade (taxas de matrícula, anos de escolaridade e alfabetização) e a qualidade (conhecimento e habilidades adquiridas medidas por meio de testes) de educação obtida pelas crianças nos países em desenvolvimento.

De acordo com Glewwe e Kremer (2006), para entender o impacto de políticas educacionais sobre os anos de escolaridade e habilidades adquiridas é preciso supor que cada família, representada pelos pais da criança, maximizam uma função de utilidade sujeita a algumas restrições. Os principais argumentos dessa função são o consumo de bens e serviços (inclusive lazer) em diferentes pontos no tempo, e a escolarização de cada criança.

A função de produção de aprendizagem é considerada no modelo como uma das restrições enfrentadas e pode ser descrita como (GLEWWE e KREMER 2006, p. 965):

$$A = a(S, Q, C, H, I) \quad (4)$$

em que A é o desempenho escolar, S são os anos de escolaridade, Q é o vetor de características da escola (fator-escola), C é um vetor de características das crianças, H é um vetor de características da família e I é um vetor de insumos educacionais sob controle dos pais como frequência na escola e materiais escolar.

As variáveis incluídas em C e H são exógenas, podendo representar respectivamente a habilidade inata da criança e a educação dos pais. Algumas características das crianças capazes de afetar o resultado A como sua saúde, podem ser endógenas ao modelo e assim serem incluídas no vetor I , em que todas as variáveis são endógenas.

Outro importante conjunto de variáveis que deve ser considerado nesse modelo são os preços relacionados à escolarização, representado pelo vetor P . Este vetor inclui taxas (mensalidades) escolares, preços dos materiais escolares até o custo de oportunidade do trabalho infantil. Segundo Glewwe e Kremer (2006), embora P não

esteja explícito na equação (4), por não ter efeito direto na aprendizagem, seu efeito aparece através das decisões tomadas para as variáveis S e I.

Considerando um cenário simples para análise, assume-se que apenas uma escola está disponível e que os pais não têm controle sobre as características dessa escola. Assim, todas as variáveis em P e Q são exógenas à família. Para maximizar a função de utilidade da família, os pais escolhem S e I, o que implica que os anos de escolaridade (S) e os insumos educacionais (I) podem ser expressos como funções dos quatro vetores de variáveis exógenas, sendo assim escritas:

$$S = f(Q, C, H, P) \quad (5)$$

$$I = g(Q, C, H, P) \quad (6)$$

Em (5) e (6), H inclui não somente as variáveis familiares em (4) mas também todas as variáveis com efeitos indiretos, tal como as restrições de crédito e o desejo dos pais quanto à escolaridade dos filhos.

Substituindo (5) e (6) em (4), tem-se a forma reduzida da equação para A:

$$A = h(Q, C, H, P) \quad (7)$$

A equação (7) reflete uma relação causal que incorpora os efeitos diretos (Q, C, H) e os indiretos (P) sobre o desempenho escolar.

Em uma situação mais realista em que as famílias poderiam escolher entre mais de uma escola, Q e P seriam endógenos, mesmo estado fixados para cada escola. Neste cenário, as famílias maximizariam sua utilidade em relação à cada escola e depois escolheriam aquela que leve à utilidade mais alta. Condicionados à escolha dessa escola, S e I seriam então definidos como no primeiro caso.

Os formuladores de políticas estão preocupados, principalmente, com os impactos das políticas educacionais sobre os anos de escolaridade S e o desempenho acadêmico dos alunos A. Uma relação causal é então estabelecida uma vez que uma mudança Q_i qualquer no vetor Q (ou em P) como uma redução no tamanho da classe ou a ampliação da carga horária escolar pode causar alterações em S e A, tendo seu impacto medido pelas equações (5) e (7).

Na função de produção educacional representada pela equação (4) uma mudança em Q_i mantidas todas as outras variáveis constantes fornece a derivada parcial no produto educacional A em relação a Q_i . Em contraste, a equação (7) fornece a derivada total de A em relação à Q_i , pois permite mudanças em S e I em resposta à mudança em Q_i . Os pais podem responder a um melhor ensino aumentando sua oferta

de insumos educacionais ou fazendo o contrário, se entender que melhor ensino é um substituto para estes insumos.

Glewwe e Kremer (2006) destacam que a escolha da equação pode levar a diferentes resultados, sendo então necessário esclarecer a relação que está sendo estimada. Os autores acreditam que por esse motivo encontra-se diferentes resultados entre os diversos estudos. Ao analisar o impacto de políticas nos resultados acadêmicos, a forma reduzida representada por (7) proporcionaria não só o impacto do programa sobre a variável resultado A mas também permitiria aos formuladores de políticas conhecer a resposta de otimização dos agentes (dada por S e I) a partir das mudanças em Q e P . Assumindo que o custo destas mudanças não são difíceis de calcular, os benefícios em termos de aumento de S e A poderão ser comparados aos custos.

Este quadro também pode ser utilizado para examinar as políticas educacionais que não alteram Q e P diretamente, mas, em vez disso, mudam a forma como a escola está organizada como a descentralização ou um programa de incentivo ao professor. Estes tipos de política podem mudar o que acontece em sala de aula, favorecendo os resultados escolares.

Formalmente, as políticas educacionais definidas como PE podem ser incluídas no vetor Q e interagir com as características da comunidade local, denotada por L , para determinar a qualidade da escola e até os preços de insumos educacionais (como nos casos de políticas que permitam que a comunidade estabeleça taxas escolares). Assim tem-se:

$$Q = q(L, PE) \quad (8)$$

$$P = p(L, PE) \quad (9)$$

Substituindo (8) e (9) nas equações (5) e (7) obtêm-se a forma reduzida da relação causal entre as políticas educacionais e a escolaridade bem como os resultados acadêmicos dos alunos:

$$S = j(C, H, L, PE) \quad (10)$$

$$A = k(C, H, L, PE) \quad (11)$$

Cabe salientar que o modelo pode apresentar limitações. Os resultados obtidos pela forma reduzida apesar de considerar a reação das famílias quando há uma mudança na qualidade da escola, ignora os efeitos sobre o bem-estar familiar

proporcionado por um aumento ou redução na compra de bens de consumo. Intuitivamente, se os pais respondem a um aumento em Q_i reduzindo seus gastos em insumos I , o bem-estar familiar aumentaria comprando mais de outros bens de consumo. Da mesma maneira, o aumento na qualidade da escola poderia reduzir o tempo que os pais gastam ajudando suas crianças com o dever de casa.

O conhecimento dessas funções relaciona diretamente as políticas de educação com os principais resultados de interesse para os formuladores de políticas. Tendo em vista a relação que deseja-se estabelecer neste trabalho, entre ampliação da jornada escolar e resultado acadêmico, essa abordagem mostra-se adequada para avaliar o Mais Educação que, em seu caso, promove uma mudança direta em um insumo educacional, alterando Q (horas-aula) e gerando impactos no desempenho escolar medido em A (testes de Português e Matemática da Prova Brasil).

4 METODOLOGIA

Nessa seção apresenta-se a metodologia utilizada para a avaliação de impacto do programa Mais Educação. O objetivo é utilizar uma estratégia metodológica que permita comparar a performance das escolas participantes antes e depois da implementação do programa. Seguindo a linha dos trabalhos que avaliaram o Programa Mais Educação (PEREIRA, 2011; XERXENEVSKY, 2012; ALMEIDA *et al*, 2015), o *Propensity Score Matching* e o método de Diferenças em Diferenças foram as duas técnicas utilizadas para estimação do efeito médio de tratamento do programa.

Na primeira etapa da implementação da estratégia metodológica, busca-se construir um grupo de controle adequado com base nas características anteriores ao programa, selecionando escolas que sejam similares às aquelas participantes (grupo de tratamento). Para tanto, utilizou-se uma técnica de pareamento, o *Propensity Score Matching* (PSM). Em um segundo momento, o método de Diferenças em Diferenças (ou Dupla Diferença) foi empregado para comparar os resultados do grupo de tratamento e de controle, antes e depois da intervenção. O uso deste método possibilitou a inclusão de características não observáveis que se não consideradas podem levar a estimativas enganosas quanto ao efeito do programa. A combinação do PSM com o método Diferenças em Diferenças permitiu controlar as unidades avaliadas por suas condições iniciais, minimizando a existência de viés de seleção.

A seguir, estas duas ferramentas são formalizadas e por último as bases de dados utilizadas são apresentadas.

4.1 A Construção do Grupo de Controle por Meio do *Propensity Score Matching*

O principal interesse deste trabalho é comparar o desempenho de escolas participantes do Mais Educação antes e depois da implementação do programa. Mais especificamente, quer saber o impacto da ampliação da jornada escolar promovida pelo Mais Educação na performance dos alunos de 5º e 9º anos nos testes de Português e Matemática da Prova Brasil. Como em outros trabalhos de avaliação de impacto, o primeiro desafio enfrentado foi construir um grupo de controle que permitisse que esta estimação fosse feita sem a presença de viés.

O impacto do Mais Educação sobre o desempenho médio das escolas seria mais precisamente estimado pela diferença entre o resultado das escolas participantes na presença e na ausência do programa. No entanto, não se pode observar ambos os resultados para a mesma escola ao mesmo tempo. Trata-se do problema de

contrafactual não observado (CALIENDO; KOPEINIG, 2008). A inferência do impacto de um tratamento sobre o resultado de uma escola envolve especulações sobre como seria a trajetória dessa escola caso não tivesse recebido o tratamento.

Formalmente, este problema pode ser representado da seguinte maneira: no caso de um tratamento binário como a participação no programa Mais Educação, o indicador tratamento D_i é igual a 1, se a escola i recebe tratamento e 0 em caso contrário. Os resultados possíveis são, então, definidos como $Y_i(D)$ e o efeito do tratamento para uma escola i pode ser escrito como:

$$\tau_i = Y_i(1) - Y_i(0) \quad (12)$$

Sendo assim, o efeito médio do tratamento seria dado por:

$$E(\tau|D = 1) = E[Y(1)|D = 1] - E[Y(0)|D = 1] \quad (13)$$

O problema de avaliação fundamental surge porque apenas um dos resultados potencial é observado para cada escola. Como a média contrafactual para aquelas que estão sendo tratadas - $E[Y(0)|D = 1]$ - não é observada, tem-se que escolher um substituto adequado para ela a fim de estimar o efeito causal. Considerar o resultado médio de escolas não participantes do programa como uma aproximação - $E[Y(0)|D = 0]$ na equação (14), torna-se uma alternativa, porém limitada, uma vez que é muito provável que os componentes que determinam a decisão de tratamento também determinam a variável resposta de interesse. Os resultados das escolas tratadas e do grupo de comparação irão diferir, mesmo na ausência de tratamento, o que conduz ao viés de seleção amostral.

$$B = E[Y(1)|D = 1] - E[Y(0)|D = 0] \quad (14)$$

A presença do viés pode ser observada ao se fazer a diferença do efeito médio de tratamento obtido pelas duas situações. O viés de seleção só é eliminado quando se garante que:

$$E[Y(0)|D = 1] - E[Y(0)|D = 0] = 0 \quad (15)$$

No caso de programas em que a atribuição de tratamento é aleatória e isto é assegurado, o efeito médio do tratamento é identificado (a equação 15 se verifica). Especialmente, no caso do programa Mais Educação, as escolas aptas a participarem do programa são pré-selecionadas pelo MEC a partir de critérios de seleção definidos previamente, sendo sua adesão voluntária. Uma vez que o Mais Educação foi direcionado à escolas com vulnerabilidade social e baixo IDEB, não se pode pressupor

a aleatoriedade desta seleção. Além disso, as características iniciais destas escolas são importantes fatores que determinam seus resultados. Assim, é provável que as escolas e seus alunos se diferenciem quanto as suas características dentro dos critérios de seleção ao programa. Portanto, a escolha das escolas que constituirão o grupo de controle deve considerar a presença de viés de seleção.

Para contornar o problema do viés de seleção, foi utilizado uma técnica muito comum na literatura: o *Matching* (Pareamento). Sua ideia básica é encontrar em um grande grupo de não participantes, unidades que são semelhantes aos participantes em todas as características (observáveis) relevantes pré-tratamento - um vetor de variáveis X , e, assim, construir um grupo de controle que represente adequadamente o que aconteceria com as escolas participantes se elas não estivessem inscritas no programa.

A inclusão de um grande número de variáveis em X , sejam elas contínuas ou bidimensionais, dificulta a realização do *Matching*. Para solucionar o problema, Rosenbaum e Rubin (1983) sugerem que ao invés de parear as unidades com base em todo o vetor X , pode-se parear as unidades usando uma função de X que resume toda a informação contida neste vetor. Trata-se de sumarizar as características observáveis de cada unidade em uma única variável, o *Propensity Score* (escore de propensão). Esta técnica de pareamento é chamada de *Propensity Score Matching*.

Rosenbaum e Rubin (1983) definem o *Propensity Score* como a probabilidade condicional de se receber o tratamento dado um vetor de variáveis observáveis. Formalmente, tem-se:

$$P(D = 1|X) = P(X) \quad (16)$$

em que D é uma variável binária que assume valor igual a 1, se a escola participa do programa Mais Educação, e 0 caso contrário, e X refere-se ao vetor de características observáveis que afetam a seleção para o programa.

As escolas participantes foram então combinadas em função dessa probabilidade, o *Propensity Score*, com aquelas não participantes. Tal procedimento, segundo Rosenbaum e Rubin (1983), gera estimadores muito próximos daqueles obtidos em experimentos aleatórios quando atendidas as seguintes hipóteses de identificação: a primeira chamada de Independência Condicional – também conhecida por ignorabilidade, ou ainda não-confundimento, diz que dado um conjunto de covariáveis observáveis X que não são afetadas pelo tratamento, os resultados potenciais são independentes da atribuição do tratamento. Formalmente:

$$Y(1), Y(0) \perp D|P(X), \quad \forall X \quad (17)$$

em que $\perp$ indica independência.

Na prática, a hipótese de independência condicional significa que não pode-se ter fatores não observáveis que afetam a participação no Mais Educação, como auto seleção de escolas ou características da comunidade onde a escola pertence que não são observáveis. A implicação desta hipótese é que o resultado de uma unidade no grupo de controle é um bom previsor do resultado potencial na ausência de tratamento de uma unidade no grupo de tratamento, que possui o mesmo vetor de variáveis observáveis (X). Para Caliendo e Kopeinig (2008), esta é uma hipótese forte e tem que ser justificada pela qualidade dos dados em questão.

A segunda hipótese é o suporte comum ou a condição de sobreposição. A hipótese de suporte comum exige que na distribuição do *Propensity Score* haja uma região onde as probabilidades estimadas de participação no Mais Educação sejam semelhantes entre as escolas participantes e as escolas não participantes. Formalmente, tem-se:

$$0 < P(D = 1|X) < 1 \quad (18)$$

De acordo com Khandker, Koolwal e Samad (2010), a eficácia do PSM também depende de um número grande e aproximadamente igual de observações participantes e não participantes de modo que uma região substancial de suporte comum possa ser encontrada.

Para proceder com o pareamento a partir das suposições acima e então encontrar um grupo de comparação mais próximo aos participantes do programa, o primeiro passo é a realização da estimação do modelo de participação do programa. Usualmente o *Propensity Score* é estimado por modelos de escolha binária, tais como os modelos *Logit* e *Probit*. O resultado do modelo representará a probabilidade estimada de participação, de modo que toda observação da amostra selecionada, seja ela participante ou não participante do programa, terá seu *Propensity Score* estimado $\hat{P}(D = 1|X) = \hat{P}(X)$. Para este trabalho, a função $P(X)$ foi estimada a partir de um modelo *Probit*¹⁷, em que a variável dependente foi um indicador binário de participação do Mais Educação e as variáveis de controle foram as características das escolas e dos municípios em 2007, antes da implementação da política.

¹⁷ O modelo estimado de participação no programa foi obtido pelo comando `psmatch2` do programa Stata.

Segundo Khandker, Koolwal e Samad (2010), as variáveis observáveis constituintes do vetor X devem ser selecionadas de forma que todas as características que provavelmente determinam a participação no programa sejam incluídas. Além disso, é importante que estas variáveis não sejam afetadas pelo tratamento, garantido que a hipótese de independência condicional seja válida (CALIENDO e KOPEINIG, 2008). Portanto, a seleção do vetor de variáveis exógenas de cada modelo estimado contou com observância de fatores ligados à participação no programa, sendo eles relacionados à infraestrutura da escola, seus indicadores escolares e também as características das cidades em que situam, todos referentes ao ano de 2007, antes da existência do programa.

Khandker, Koolwal e Samad (2010) alertam que incluir muitas variáveis em X na equação de participação também deve ser evitada; a superespecificação do modelo pode resultar em erros padrão mais elevados para o *Propensity Score* estimado. A predição perfeita da participação das escolas também será prejudicial, levando à falha do suporte comum. Se isso acontece, o pareamento não poderá ser executado, pois não haverá sobreposição entre as probabilidades estimadas dos dois grupos, dentro das características observadas.

A qualidade dos modelos de participação estimados foi então verificada a fim de garantir que as hipóteses fossem atendidas. A escolha das variáveis baseou-se também em exercícios de inclusão e exclusão de variáveis (método "*hit or miss*") e na análise de significância estatística para obter a predição correta do modelo. A região de suporte comum foi definida para todos os pareamentos realizados, o que pode levar a exclusão de algumas escolas, a fim de assegurar a compatibilidade entre os dois grupos.

Testes de balanceamento também foram realizados para verificar se unidades com valores idênticos do *Propensity Score* tinham a mesma distribuição quanto às características observáveis analisadas, independentes se tratadas ou não. Para que o PSM funcione, os grupos de tratamento e de comparação devem ser equilibrados, na medida em que os escores de propensão semelhantes são baseados na distribuição semelhante das variáveis observadas em X. Khandker, Koolwal e Samad (2010) salientam que embora uma unidade tratada e seu correspondente não tratado possam ter os mesmos escores de propensão, eles não são necessariamente semelhantes em suas características observáveis se existirem erros de especificação na equação de participação.

Portanto, após a estimação do modelo *Probit*, foram feitos testes de diferenças de médias, antes e depois do pareamento, para verificar se o conjunto de variáveis X estavam balanceadas entre os grupos. Também aplicou-se o teste de viés, que permitiu analisar se a presença do viés foi reduzida após o pareamento¹⁸. Por fim, gráficos de densidade para a distribuição do *Propensity Score* foram obtidos para verificar se eram semelhantes entre escolas tratadas e controle após o pareamento. A aplicação destes testes¹⁹ resultou em grupos de controle e tratamento balanceados e dentro da região de suporte comum.

O pareamento baseado no *Propensity Score* irá depender de uma métrica pré-determinada, que definirá a proximidade do escore de propensão das unidades tratadas em relação ao escore de propensão das unidades não tratadas. Diferentes procedimentos podem ser adotados para a realização do pareamento. As técnicas mais utilizadas na literatura para este propósito são: Pareamento ao Vizinho mais próximo (*Nearest-Neighbor Matching*), Pareamento Radial (*Radius Matching*), Pareamento Estratificado (*Stratification Matching*) ou Pareamento de Kernel (*Kernel Matching*) (BECKER e ICHINO, 2002; CALIENDO e KOPEINIG, 2008; KHANDKER; KOOLWAL e SAMAD, 2010).

O pareamento ao vizinho mais próximo é o mais utilizado entre os algoritmos de pareamento. Consiste em parear cada unidade do grupo de tratamento a uma unidade do grupo de controle que tenha o *Propensity Score* mais próximo. O pareamento pode ser realizado com ou sem reposição. A escolha por reposição indica que uma mesma unidade do grupo de controle pode ser pareada com mais de uma unidade participante (KHANDKER, KOOLWAL e SAMAD, 2010).

No pareamento radial, cada unidade tratada é comparável apenas com as unidades de controle cujos escores de propensão estejam contidos em uma vizinhança do escore de propensão de unidades tratadas definida a priori. Se a dimensão da vizinhança (ou seja, o raio) é muito pequena, é possível que algumas unidades tratadas não sejam pareadas devido à falta de unidades de controle na vizinhança (BECKER e ICHINO, 2002).

O método de pareamento por estratificação, por sua vez, particiona o suporte comum do *Propensity Score* em um conjunto de intervalos (estratos) de tal forma que

¹⁸ Tanto o teste de médias quanto o teste de viés são realizados pelo comando *pstest* do Stata.

¹⁹ A estrutura de todos os testes realizados (além da exposição de outros) estão disponíveis em Caliendo e Kopeinig (2008).

as unidades dentro de cada intervalo, tratadas e controle, em média possuem o mesmo *Propensity Score*. Por fim, o pareamento de Kernel usa uma média ponderada de todos os não participantes para construir o pareamento para cada participante. Becker e Ichino (2002) explicam que neste método, todas as unidades tratadas são pareadas com uma média ponderada de todas as unidades de controle, com pesos que são inversamente proporcionais à distância entre os escores de propensão de controles e tratadas.

Apesar de cada método apresentar diferenças dentro do *trade-off* qualidade e quantidade de observações para pareamento, Becker e Ichino (2002) ressaltam que nenhum deles é a priori melhor que os outros. Assintoticamente, todos os estimadores de PSM devem produzir os mesmos resultados, porque com o tamanho crescente da amostra, as comparações se aproximam apenas de correspondências exatas (SMITH, 2000).

Em pequenas amostras a escolha do algoritmo de correspondência pode ser importante. Por outro lado, se houver uma grande quantidade de observações não-tratadas comparáveis, como é o caso das amostras selecionadas neste trabalho, pode valer a pena usar mais de um vizinho mais próximo (ou Kernel) para obter maior precisão nas estimativas (CALIENDO e KOPEINIG, 2008).

Deste modo, este trabalho seguiu esta orientação e realizou o pareamento com os cinco vizinhos mais próximos, com reposição. A utilização de cinco vizinhos permitiu que uma maior parte da amostra fosse explorada, facilitando a construção de um grupo de comparação que melhor represente o contrafactual.

Por fim, é importante ressaltar que apesar do *Propensity Score Matching* fornecer de forma satisfatória o efeito médio do programa após a realização do pareamento, ele não foi usado neste trabalho com esta finalidade. Com seu uso, buscou-se somente construir um grupo de controle adequado e para que isto fosse feito, as suposições subjacentes ao método foram observadas, garantidas pelos testes realizados. A relação causal válida entre tratamento e resultado na qual este trabalho busca revelar, não se restringe à satisfação dessas hipóteses, sendo os impactos de interesse estimados por meio de diferenças em diferenças (modelo discutido a seguir).

4.2 Aplicação do Método Diferenças em Diferenças

Após o procedimento exposto na seção anterior, o método de Diferenças em Diferenças (DD) foi aplicado na amostra pareada a fim de comparar escolas do grupo

de tratamento e controle em termos de mudanças no desempenho em testes de Português e Matemática da Prova Brasil ao longo do tempo, frente aos resultados observados antes da implementação do programa (em 2007). A aplicação desta metodologia foi realizada em dados em painel, o que permite a inclusão de fatores não observados que também afetam o resultado, isto é, considera a influência da heterogeneidade não observada entre os grupos, mas que são invariantes ao longo do tempo.

De maneira geral, para a especificação do modelo, considera-se dois períodos no tempo, em que $t = 0$ é o período antes do programa e $t = 1$ é o período após a implementação do programa; Y_t^T e Y_t^C representam, respectivamente, o resultado para os participantes e não participantes do programa, no tempo t . O método DD estima o impacto médio do programa da seguinte forma:

$$DD = E(Y_1^T - Y_0^T | D_1 = 1) - E(Y_1^C - Y_0^C | D_1 = 0) \quad (19)$$

em que $D_1 = 1$ representa a presença do programa (grupo de tratados) e $D_1 = 0$ o grupo de não tratados. Com dados de participantes e não participantes antes da implementação do programa e também depois da intervenção, a diferença é calculada entre os resultados médios observados para os grupos de tratamento e de controle antes e depois da intervenção do programa (KHANDKER, KOOLWAL e SAMAD, 2010).

O método da dupla diferença pode ser aplicado por meio da especificação de uma equação, estimada por Mínimos Quadrados Ordinários. A equação básica do modelo assume então a seguinte forma:

$$Y_{it} = \alpha + \beta T_{it}t + \rho T_{it} + \gamma t + \varepsilon_{it} \quad (20)$$

em que T é uma variável binária que assume valor unitário se a escola é tratada, e valor nulo caso contrário; t também é uma variável binária, assumindo valor 1 no período pós programa e valor 0 caso contrário. O termo de erro é dado por ε_{it} . As variáveis correspondentes ao período de tempo e ao *status* de tratamento são incluídas separadamente para captar os efeitos médios do tempo sobre a medida de resultado assim como o efeito da unidade fazer parte do grupo de tratados ou não.

A introdução do termo de interação na equação procura captar o que ocorreu especificamente com o grupo de tratamento no período pós-programa, ou seja, se a média da variável de resultado para esse grupo em particular tornou-se diferente após a intervenção. Vê-se claramente que a função do termo de interação é captar o impacto do programa que é medido por β .

A estimação do impacto por meio do método DD também pode ser feita a partir de um modelo de regressão com efeitos fixos. A inserção de características fixas não observáveis no tempo é a principal vantagem do método, pois muitas vezes, estas influenciam também a participação no programa (MENEZES FILHO, 2012). As características inerentes à administração de cada escola, sua forma de organização e a possível participação da comunidade no dia a dia escolar são exemplos de características não observáveis que podem afetar tanto os resultados avaliados quanto a decisão de participação no programa e precisam ser controladas para evitar a presença de viés de auto-seleção.

Especificamente para este trabalho, a equação principal foi especificada como:

$$Y_{it} = \beta_0 + \beta_1 MaisEducação_i + \beta_2 D2013_t + \beta_3 (MaisEducação_i \times D2013_t) + S_i + \varepsilon_{it} \quad (21)$$

em que Y_{it} refere-se ao resultado de interesse (pontuação nos testes de Português e Matemática da Prova Brasil) da escola i no ano t ; $MaisEducação$ é uma variável que assume valor igual a 1 se a escola participa do programa; $D2013$ é uma *dummy* que assume valor 1 para o ano de 2013; a variável $(MaisEducação \times D2013)$ é uma interação das variáveis anteriores que assume valor 1 apenas para as escolas do grupo de tratamento no período depois da implementação do programa; S_i é o efeito fixo em nível da escola e ε_{it} é o erro aleatório (não observado)²⁰. O principal coeficiente de interesse é β_3 que representa o efeito médio do programa nos resultados das escolas.

Nessa regressão, pode-se mostrar que o coeficiente β_3 , associado à interação, é exatamente o impacto do programa sobre o resultado Y na equação (19). As equações seguintes são capazes de demonstrar essa relação. Para o grupo de tratamento antes e depois da implementação do programa tem-se:

$$Y_0^T = \beta_0 + \beta_1 + S + \varepsilon \quad (22)$$

$$Y_1^T = \beta_0 + \beta_1 + \beta_2 + \beta_3 + S + \varepsilon \quad (23)$$

Para o grupo de controle as equações são dadas por:

$$Y_0^C = \beta_0 + S + \varepsilon \quad (24)$$

$$Y_1^C = \beta_0 + \beta_2 + S + \varepsilon \quad (25)$$

Como se queria demonstrar o impacto do programa será:

²⁰ O método permite que variáveis de controle, variantes no tempo, sejam incluídas na equação 21, possibilitando reduzir o viés ocasionado pela omissão de variáveis.

$$(Y_1^T - Y_0^T) - (Y_1^C - Y_0^C) = \beta_3 = DD \quad (26)$$

Ao assumir que a heterogeneidade não observada é invariante no tempo, os efeitos fixos individuais são eliminados na equação (26), de modo que o viés é cancelado por meio da diferenciação.

A estimativa do efeito do tratamento não será viesada, sob as seguintes hipóteses: a) a equação (equação 21) for corretamente especificada; b) a variância dos termos de erro é constante e c) os termos de erros não são correlacionados com as outras variáveis da equação²¹ (KHANDKER; KOOLWAL; SAMAD, 2010).

Ainda de acordo com Khandker, Koolwal e Samad (2010), o último desses pressupostos, também conhecido como pressuposto de tendência paralela, é o mais crítico. Essa hipótese propõe que a trajetória temporal da variável de resultado para o grupo de controle represente o que ocorreria com o grupo tratado caso não houvesse a intervenção. De outra maneira, isso significa que as características não observadas que afetam a participação no programa não variam ao longo do tempo com o *status* do tratamento.

Segundo Menezes Filho (2012), essa hipótese não pode ser testada diretamente nos dados, mas uma indicação de sua validade aparece quando as trajetórias dos dois grupos são parecidas pré-programa. Portanto, quando existem dados para vários períodos de tempo pré-programa, tipicamente testa-se se as séries da variável de resultado para os dois grupos possuem a mesma tendência temporal²².

Em um último esforço para que a presença de viés nos resultados estimados fosse completamente reduzida, as regressões estimadas foram ponderadas com os pesos obtidos da aplicação do *Propensity Score Matching*. Para Hirano, Imbens e Ridder (2003), ponderar as observações de controle de acordo com seu escore de propensão produz um estimador totalmente eficiente.

4.3 Variáveis e definição das amostras

Para a aplicação da estratégia empírica escolhida, a unidade de análise foi as escolas estaduais e municipais de todo o território brasileiro. Como o objetivo do trabalho foi estimar o impacto do programa Mais Educação no médio prazo, a amostra principal (amostra 1) contou com escolas que aderiram ao programa em seu primeiro ano de funcionamento (2008) e permaneceram ininterruptamente até o ano de 2013,

²¹ $Cov(\varepsilon_{it}, MaisEducação_i) = 0$; $Cov(\varepsilon_{it}, D2013_t) = 0$; $Cov(\varepsilon_{it}, MaisEducação_i \times D2013_t) = 0$

²² Não foi possível obter os dados da Prova Brasil de 2005 para a realização deste teste.

último ano com dados para as variáveis resultado utilizadas. Dessa forma, foi possível comparar o resultado obtido com outros trabalhos já presentes na literatura e observar de que maneira o tempo de participação foi importante para determinar o impacto nas notas médias de português e matemática da Prova Brasil das primeiras escolas participantes.

Indo ao encontro do segundo objetivo do trabalho, mais duas amostras foram criadas a fim de avaliar o impacto do programa em escolas com diferentes tempos de exposição. Uma delas (amostra 2) contou com escolas que aderiram ao Mais Educação em 2010, ou seja, após dois anos de programa. A outra amostra (amostra 3) foi construída com escolas que só aderiram ao programa pela primeira vez em 2012. Em ambas, só foram consideradas escolas que permaneceram no programa ininterruptamente até 2013, a partir do seu ano de adesão. Com isso, foi possível comparar os resultados para escolas com cinco, três e um ano de exposição ao programa.

O potencial grupo de controle foi constituído por escolas que nunca participaram do programa Mais Educação e aplicado nas três amostras construídas. Assim, escolas que foram expostas ao programa em algum momento entre 2008 e 2013, mas não se encaixam nos critérios já definidos anteriormente foram excluídas da análise, evitando que o tempo de exposição ao programa, mesmo que mínimo, influenciasse nos resultados observados. O Quadro 2 resume o número de escolas brasileiras que já foram expostas ao programa durante o período analisado.

Quadro 2 - Escolas participantes do Programa Mais Educação entre 2008 e 2013

Escolas	Total
Que participaram do programa ME neste período	52263
Que aderiram em 2008 e permaneceram até 2013	1185
Que aderiram em 2010 e permaneceram até 2013	4170
Que aderiram em 2012 e permaneceram até 2013	16328

Fonte: elaboração própria a partir dos dados fornecidos pelo MEC

A construção das amostras partiu das informações contidas nos bancos de dados da Prova Brasil e do Censo Escolar. O elo de ligação entre os diferentes bancos de dados foi o código INEP, um número de oito dígitos que identifica todas as escolas brasileiras. Foram consideradas somente as escolas que estavam com o *status* de “ativa” nos anos utilizados para compor a amostra. Assim, excluídas aquelas que não constavam com informações para a Prova Brasil, foram adicionadas a cada escola

informações sobre os indicadores educacionais produzidos pelo INEP (taxas de abandono e aprovação), informações sobre a cidade em que a escola se situa (População, PIB *per capita* e IDHM), nos anos pré-tratamento – 2007 (amostra 1), 2009 (amostra 2) ou 2011 (amostra 3) e pós-tratamento – 2013 e também a participação no programa.

Para cada amostra foi construído um painel²³, contendo informações pré e pós-tratamento de acordo com os anos analisados. As amostras também foram divididas para as séries analisadas, portanto para o 5º e 9º anos²⁴. A composição de cada amostra é descrita na Tabela 2 abaixo:

Tabela 2 - Amostras para 5º ano e 9º ano

5º ano					
	Grupo de Tratamento		Grupo Controle		Total
	Descrição	Nº de escolas	Descrição	Nº de escolas	
Amostra 1	Escolas com turmas de 5º ano que aderiram em 2008 e permaneceram até 2013	711	Escolas com turmas de 5º ano que nunca participaram do programa	10922	11633
Amostra 2	Escolas com turmas de 5º ano que aderiram em 2010 e permaneceram até 2013	2577	Escolas com turmas de 5º ano que nunca participaram do programa	12058	14635
Amostra 3	Escolas com turmas de 5º ano que aderiram em 2012 e permaneceram até 2013	6629	Escolas com turmas de 5º ano que nunca participaram do programa	12075	18704
9º ano					
	Grupo de Tratamento		Grupo Controle		Total
	Descrição	Nº de escolas	Descrição	Nº de escolas	
Amostra 1	Escolas com turmas de 9º ano que aderiram em 2008 e permaneceram até 2013	715	Escolas com turmas de 9º ano que nunca participaram do programa	8339	9054
Amostra 2	Escolas com turmas de 9º ano que aderiram em 2010 e permaneceram até 2013	2241	Escolas com turmas de 9º ano que nunca participaram do programa	9338	11579
Amostra 3	Escolas com turmas de 9º ano que aderiram em 2012 e permaneceram até 2013	5011	Escolas com turmas de 9º ano que nunca participaram do programa	9619	14630

Fonte: elaboração própria

²³ Para a construção de um painel balanceado foram excluídas escolas que não constavam nos dois anos que compõe o painel. Além disso, procedeu-se com as exclusões de *missing*, ou seja, foram excluídas escolas que possuíam dados em branco em qualquer uma das variáveis selecionadas.

²⁴ Turmas remanescentes do antigo sistema de séries escolares (4ª e 8ª séries) também estão incluídas na análise, respectivamente como 5º e 9º ano.

Apesar do banco de dados utilizado para a identificação de escolas tratadas ser o mesmo para o 5º e 9º anos dentro de cada amostra, o número de escolas tratadas em cada ano escolar é diferente devido a informações faltantes para a Prova Brasil e o Censo Escolar em cada série. Além disso, não são todas as escolas com turmas de 5º ano que também contam com turmas de 9º ano, acarretando um número de escolas menor para este último. O mesmo ocorre na definição dos grupos de escolas controle.

Definidas as amostras pôde-se então aplicar a metodologia escolhida para estimar o impacto do Mais Educação no desempenho médio de escolas nos testes de Português e Matemática da Prova Brasil. Os resultados das regressões para estimação de impacto foram apresentados separadamente para as análises realizadas para o 5º e 9º anos. Portanto, para cada ano escolar de cada amostra, dois modelos de efeitos fixos foram estimados, um deles medindo o impacto do programa sobre as notas de Português e o outro sobre as notas de Matemática. Ao todo, doze modelos de impacto foram estimados. As variáveis indicadoras de resultado selecionadas para a pesquisa são descritas na Tabela 3 abaixo.

Tabela 3 - Variáveis de resultado

Variável	Descrição
Media_5º_LP	Nota Média da escola em Português (5º ano)
Media_5º_MT	Nota Média da escola em Matemática (5º ano)
Media_9º_LP	Nota Média da escola em Português (9º ano)
Media_9º_MT	Nota Média da escola em Matemática (9º ano)
Fonte dos Dados e Anos Utilizados	
Amostra 1: Informação da Prova Brasil para 2007 e 2013	
Amostra 2: Informação da Prova Brasil para 2009 e 2013	
Amostra 3: Informação da Prova Brasil para 2011 e 2013	

Fonte: elaboração própria

O modelo *Probit* utilizado para estimar o *Propensity Score* e posteriormente facilitar o pareamento, foi aplicado para cada ano escolar (5º e 9º anos) em cada amostra. As variáveis de controle selecionadas para a estimação do modelo foram aquelas que tinham potencial para afetar a participação no programa. Além das características de infraestrutura da escola, foram consideradas informações de indicadores educacionais e características da cidade da escola – sendo estas últimas, critérios utilizados pelo MEC para indicar as escolas que eram aptas a participar do programa (Tabela 4).

Tabela 4 - Variáveis de Controle (usadas nos modelos Probit)

Variável	Descrição
Informações do Censo Escolar (2007, 2009, 2011 e 2013)	
Lab_Informática	Variável Dummy, igual a 1 se a escola possui laboratório de informática
Lab_Ciências	Variável Dummy, igual a 1 se a escola possui laboratório de ciências
Biblioteca	Variável Dummy, igual a 1 se a escola possui biblioteca
Nu_Salas	Número de salas de aula da escola
Nu_Computadores	Número de computadores disponíveis na escola
Alimentação	Variável Dummy, igual a 1 se a escola oferece refeição aos alunos
Nu_Funcionários	Número de empregados da escola
Água	Variável Dummy, igual a 1 se a escola possui abastecimento de água por rede ou poço
Esgoto	Variável Dummy, igual a 1 se a escola possui escoamento de esgoto por rede ou fossa
Saneamento	Variável de interação (Água_Adeq x Esgoto_Adeq)
Quadra	Variável Dummy, igual a 1 se a escola possui quadra de esportes
Ens_Médio	Variável Dummy, igual a 1 se a escola oferece ensino médio
Nu_Turmas	Número total de turmas (classes) que a escola possui
Nu_Matrículas	Número total de matrículas que a escola possui
Duração_Média	Duração média da aula em minutos
Indicadores do INEP (2007, 2009, 2011 e 2013)	
IDEB_5	Índice de qualidade da Educação para o 5º ano
IDEB_9	Índice de qualidade da Educação para o 9º ano
Tx_Aprovação_5	% de alunos aprovados nas turmas de 5º ano
Tx_Aprovação_9	% de alunos aprovados nas turmas de 9º ano
Tx_Abandono_5	% de alunos que abandonaram a escola no 5º ano
Tx_Abandono_9	% de alunos que abandonaram a escola no 9º ano
IBGE e PNUD	
População	População no município da escola
PIB_pc	PIB <i>per capita</i> no município da escola
IDHM_2000	Índice de Desenvolvimento Humano Municipal (2000)
IDHM_2010	Índice de Desenvolvimento Humano Municipal (2010)

Fonte: elaboração própria. Nota: Valores monetários a preços de 2013

É importante ressaltar que nem todas as variáveis descritas na Tabela 4 foram utilizadas em cada modelo estimado. Conforme explicado anteriormente, as variáveis selecionadas também levaram em conta a especificação do modelo e o pressuposto de balanceamento.

4.4 Fonte de Dados

As variáveis de controle que foram utilizadas para compor o modelo *Probit* pertencem a diferentes fontes de dados. As características das escolas e de seus alunos foram obtidas através dos microdados do Censo Escolar de 2007, 2009, 2011 e 2013. O Censo Escolar é um levantamento de dados estatísticos educacionais de âmbito nacional realizado todos os anos e coordenado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira - INEP/MEC. Trata-se do principal instrumento de coleta de informações da educação básica, que abrange as suas diferentes etapas e modalidades e contempla dados sobre os estabelecimentos de ensino, turmas, alunos, profissionais escolares em sala de aula e rendimento escolar. Cada escola recebe um número de identificação único, que se repete em outras bases de dados INEP, possibilitando identificar e relacionar suas informações.

Adicionalmente, também foram utilizadas informações obtidas junto ao Instituto Brasileiro de Geografia e Estatística – IBGE e do Atlas de Desenvolvimento Humano no Brasil²⁵ para explorar características relativas aos municípios das escolas (População, PIB *per Capita* e IDHM – Índice de Desenvolvimento Humano Municipal). É importante ressaltar que as variáveis de controle selecionadas foram aquelas com potencial para influenciar os resultados nos testes da Prova Brasil além daquelas que desenham os critérios estabelecidos pelo MEC para a participação no programa.

A informação de participação no programa foi obtida por meio de solicitação direta ao Ministério da Educação – MEC. Após contato por e-mail com a Coordenação-Geral de Educação Integral²⁶, foi recebida uma tabela com o nome e o código INEP de todas as escolas que participavam do programa Mais Educação entre os anos de 2008 e 2013. A partir da repetição dos códigos durante os anos analisados foi possível determinar as escolas que comporiam as amostras analisadas.

Os resultados educacionais, variáveis para as quais foram investigados os impactos do programa, consistiram no desempenho médio das escolas em testes de Português e Matemática da Prova Brasil. A Prova Brasil consiste em testes padronizados aplicados na quarta e oitava séries (quinto e nono anos) do ensino

²⁵Construído a partir de dados do Censo Demográfico em uma parceria do Programa das Nações Unidas para o Desenvolvimento - Pnud, Fundação João Pinheiro e do Instituto de Pesquisa Econômica Aplicada - IPEA.

²⁶ educacaointegral@mec.gov.br

fundamental em que os estudantes respondem a questões de Língua Portuguesa (com foco em leitura) e Matemática (com foco na resolução de problemas). A aplicação do exame é de caráter censitário, bianual, atingindo todas as escolas que possuem, no mínimo, 20 alunos matriculados nas séries/anos analisados.

Os anos analisados nesta pesquisa foram 2007, 2009, 2011 e 2013. Este último é o ano mais recente com disponibilidade de dados, ainda não explorado em outros trabalhos, permitindo assim que os impactos do programa sejam mensurados em um tempo maior de implementação.

Destaca-se por fim que a escolha por avaliar o impacto do programa sobre o desempenho médio da escola e não do aluno justifica-se pela impossibilidade de obter informações sobre a participação individual no programa. Apesar dos critérios sugeridos pelo MEC para a escolha dos alunos participantes, não há meios de identificar dentro das bases de dados disponíveis, quais estudantes são atingidos e suas características.

5 RESULTADOS E DISCUSSÕES

Esta seção inicia-se com uma análise das estatísticas descritivas das três amostras analisadas. Após a apresentação sintética dos dados, serão expostos os resultados referente a criação dos grupos de controle por meio do *Propensity Score Matching*. Por fim, os resultados das avaliações de impacto realizadas são analisados a partir da subseção 5.3.

5.1 Perfil das escolas analisadas

Nesta subseção, são apresentados os perfis das escolas participantes e não participantes do programa Mais Educação em cada amostra construída. As Tabelas 5, 6 e 7 reportam a diferença entre as escolas que aderiram ao Mais Educação e todas as escolas não participantes para as duas variáveis de interesse, sendo elas as notas de Português e Matemática da Prova Brasil. As médias foram obtidas para os anos pré e pós-tratamento, o que permite observar a evolução das notas durante o período analisado. A Tabela 5, expõe os resultados para a amostra 1.

Tabela 5 - Notas médias de escolas participantes e não participantes (amostra 1)

	Escolas Participantes			Escolas Não Participantes			Diferença de média entre os grupos	
	2007	2013	%	2007	2013	%	2007	2013
Média para escolas com 5º ano								
Notas Português	164,43	181,53	10,20	183,24	206,79	12,85	-18,81***	-25,25***
Notas Matemática	180,19	194,41	7,89	201,77	224,63	11,33	-21,57***	-30,21***
Nº de observações	711	711		10922	10922			
Média para escolas com 9º ano								
Notas Português	221,46	232,05	7,78	236,46	248,17	4,95	-14,99***	-16,11***
Notas Matemática	229,53	233,45	1,71	250,29	254,25	1,58	-20,75***	-20,80***
Nº de observações	715	715		8339	8339			

Fonte: resultados da pesquisa. *significativo a 10%; **significativo a 5%; ***significativo a 1%

Os resultados da Tabela 5 evidenciam as diferenças expressivas nas notas de Português e Matemática quando compara-se escolas participantes e não participantes. Pôde-se observar que para ambos os anos, as escolas que contam com a presença do programa obtiveram médias mais baixas que seus pares do grupo de controle, tanto em

português quanto matemática, nas duas séries analisadas. Para o 5º ano, além de persistir entre os anos, a diferença entre escolas tratadas e controle ainda aumentou. Para as escolas com turmas de 9º ano, tem-se uma situação um pouco melhor – a diferença entre os grupos se elevou ligeiramente em Português e manteve-se praticamente constante para Matemática.

Como já informado, o resultado da Prova Brasil é divulgado por escolas, o que permite avaliar o sistema. De maneira geral os resultados encontrados expõem a situação crítica da educação brasileira. Em média, pode-se constatar que de 2007 a 2013, as notas evoluíram, mas ainda encontram-se em um nível muito baixo em uma escala de proficiência que vai de 0 a 500 pontos.

Os mesmos testes são realizados para as amostras 2 e 3. Os resultados são apresentados nas Tabelas 6 e 7 a seguir.

Tabela 6 - Notas médias de escolas participantes e não participantes (amostra 2)

	Escolas Participantes			Escolas Não Participantes			Diferença de média entre os grupos	
	2009	2013	%	2009	2013	%	2009	2013
Média para escolas com 5º ano								
Notas Português	178,07	187,70	5,41	192,38	204,79	6,45	-14,33***	-17,09***
Notas Matemática	196,17	202,20	3,07	215,27	222,55	3,38	-19,10***	-20,34***
Nº de observações	2577	2577		12058	12058			
Média para escolas com 9º ano								
Notas Português	237,33	238,12	0,33	246,12	246,81	0,28	-8,78***	-8,69***
Notas Matemática	238,63	241,05	1,01	251,43	253,04	0,64	-12,80***	-11,99***
Nº de observações	2241	2241		9338	9338			

Fonte: resultados da pesquisa. *significativo a 10%; **significativo a 5%; ***significativo a 1%

Tabela 7 - Notas médias de escolas participantes e não participantes (amostra 3)

	Escolas Participantes			Escolas Não Participantes			Diferença de média entre os grupos	
	2011	2013	%	2011	2013	%	2011	2013
Média para escolas com 5º ano								
Notas Português	175,35	177,69	1,33	198,55	204,26	2,87	-23,19***	-26,56***
Notas Matemática	192,65	191,90	-0,39	219,93	221,94	0,91	-27,29***	-30,03***
Nº de observações	6629	6629		12075	12075			
Média para escolas com 9º ano								
Notas Português	228,40	230,74	1,03	246,77	246,87	0,04	-18,37***	-16,14***
Notas Matemática	235,49	235,69	0,09	255,25	253,03	-0,87	-19,76***	-17,34***
Nº de observações	5011	5011		9619	9619			

Fonte: resultados da pesquisa. *significativo a 10%; **significativo a 5%; ***significativo a 1%

A mesma tendência de comportamento das notas foi observada para as escolas com turmas de 5º ano nas Tabelas 6 e 7. As notas médias de Português e Matemática evoluem ao longo dos anos analisados, mas a diferença aumenta entre escolas com o programa e aquelas que não participam. No 9º ano, essa diferença parece estável na amostra 2 e até cai na amostra 3, neste caso, reflexo da estagnação da nota de Português e queda da nota de Matemática em escolas não participantes.

Apesar de significativas, as diferenças encontradas entre as notas de escolas participantes e não participantes ainda não podem substantiar uma análise de causalidade. Essa análise precisa ser parcimoniosa, uma vez que é bem provável que as escolas que aderiram ao Mais Educação sejam bem distintas daquelas que não fazem parte do programa. É importante lembrar que o cenário encontrado pode ter reflexos da própria seleção ao programa que priorizou escolas com baixo IDEB e situadas em territórios de vulnerabilidade social²⁷. Além disso, a adesão ao programa é feita de forma voluntária.

Foi possível verificar a diferença existente entre as escolas tratadas e as escolas presentes no grupo de controle a partir de uma análise de diferenças de médias para as

²⁷ Como caracterizado no referencial teórico, as características sociais dos estudantes são grandes fatores que afetam o desempenho escolar.

características observáveis. As Tabelas 8, 9 e 10 apresentam os resultados de testes de médias para as variáveis observadas no período pré-tratamento.

Tabela 8 - Características médias das escolas participantes e não participantes em 2007 (amostra 1)

Variáveis	Escolas com 5º ano			Escolas com 9º ano		
	Escolas tratadas	Escolas controle	Diferença	Escolas tratadas	Escolas controle	Diferença
Lab_Informática	0,50	0,48	0,02	0,58	0,74	-0,16***
Lab_Ciências	0,14	0,11	0,03**	0,27	0,33	-0,06***
Biblioteca	0,75	0,71	0,04**	0,87	0,88	-0,01
Nu_Salas	12,81	11,21	1,60***	14,57	13,35	1,21***
Nu_Computadores	9,29	8,93	0,36***	10,94	13,84	-2,90***
Alimentação	0,92	0,96	-0,04***	0,91	0,96	-0,05***
Nu_Funcionários	65,40	44,0	21,40***	77,64	62,46	15,19***
Água	0,98	0,99	-0,01***	0,98	0,99	-0,01***
Esgoto	0,98	1,0	-0,02***	0,98	1,00	-0,02***
Saneamento	0,97	0,99	-0,02***	0,97	0,99	-0,02***
Quadra	0,54	0,61	-0,07***	0,70	0,82	-0,12***
Ens_Médio	0,08	0,13	-0,05***	0,42	0,57	-0,15***
Nu_Turmas	30,67	21,51	9,16***	35,50	28,44	7,06***
Nu_Matrículas	902,92	580,23	322,69***	1184,57	887,56	297,01***
Duração_Média	252,61	265,02	-12,41***	253,20	269,36	-16,16***
IDEB	3,55	4,71	-1,16***	2,82	4,04	-1,22***
Tx_Aprovação	82,57	91,71	-9,13***	73,99	84,17	-10,18***
Tx_Abandono	5,16	1,04	4,12***	11,69	4,04	7,65***
População	1.122.093	784.174	337.918***	1.359.456	979.929	379.526***
IDHM2000	0,65	0,64	0,01***	0,65	0,64	0,01***
IDHM2010	0,75	0,74	0,01***	0,75	0,74	0,01***
PIB_pc	17.185	23.035	-5.849***	18.127	23.319	-5.192***

Fonte: resultados da pesquisa. *significativo a 10%; **significativo a 5%; ***significativo a 1%

As diferenças observadas para a amostra 1 estão muito relacionadas aos critérios de seleção definidos pelo MEC. No início do programa, as escolas atendidas seriam aquelas com baixo IDEB, situadas em territórios de vulnerabilidade social e em cidades com mais de 200 mil habitantes. Dessa forma, antes do início do programa, as escolas tratadas apresentavam na média, indicadores educacionais mais baixos, maior número de matrícula e estavam situadas em cidades maiores. Especialmente quanto ao IDEB, a média para escolas tratadas foi menor que a meta traçada pelo MEC para

aquele ano que era 3.9 para os anos iniciais do ensino fundamental e 3.5 para os anos finais.

Os indicadores educacionais analisados também mostraram números inferiores para as escolas com turmas de 9º ano. De fato, este dado reflete uma característica do sistema educacional brasileiro. Geralmente, os problemas se agravam nos anos finais do Ensino Fundamental devido ao acúmulo de defasagens trazidas pelo estudante ao longo da sua vida escolar²⁸, que pode impulsionar a repetência e as taxas de evasão e defasagem idade-série.

Em relação as características de infraestrutura, as escolas com turmas de 5º ano, participantes e não participantes apresentaram, na média, pequenas diferenças quanto a presença de laboratórios de informática, ciências, biblioteca e saneamento básico adequado. Para as escolas com turmas de 9º ano, essa diferença torna-se um pouco maior para a presença de laboratório de informática e número de computadores na escola.

Na amostra 2 (Tabela 9), o teste de médias realizado para as variáveis observáveis antes do tratamento, isto é, para o ano de 2009, não apontou diferenças entre escolas participantes e não participantes quanto a presença de saneamento adequado. As escolas tratadas, tanto para o 5º quanto para o 9º ano, continuam apresentando médias mais baixas para os indicadores educacionais. Porém, de forma positiva, estes indicadores evoluíram em relação à sua média de 2007.

²⁸ Esta tendência fica clara no estudo de Todos pela Educação, 2015.

Tabela 9 - Características médias das escolas participantes e não participantes em 2009 (amostra 2)

Variáveis	Escolas com 5º ano			Escolas com 9º ano		
	Escolas tratadas	Escolas controle	Diferença	Escolas tratadas	Escolas controle	Diferença
Lab_Informática	0,55	0,62	-0,07***	0,80	0,86	-0,06***
Lab_Ciências	0,13	0,11	0,02***	0,30	0,34	-0,04***
Biblioteca	0,60	0,55	0,05***	0,78	0,59	0,19***
Nu_Salas	11,46	11,09	0,37***	13,89	13,00	0,89***
Nu_Computadoes	9,99	12,44	-2,45***	16,06	18,55	-2,50***
Alimentação	0,99	0,99	0,00	0,99	0,99	0,00
Nu_Funcionários	52,68	44,20	8,48***	70,25	60,50	9,75***
Água	0,97	0,98	-0,01	0,98	0,98	0,00
Esgoto	1,00	1,00	0,00	1,00	1,00	0,00
Saneamento	0,97	0,98	-0,01	0,98	0,98	0,00
Quadra	0,54	0,63	-0,09***	0,73	0,81	-0,08***
Ens_Médio	0,08	0,12	-0,04***	0,42	0,55	-0,13***
Nu_Turmas	23,70	20,96	2,75***	30,06	26,67	3,39***
Nu_Matrículas	647,30	545,25	102,06***	938,65	811,83	126,82***
Duração_Média	255,73	264,51	-8,78***	259,25	268,48	-9,23***
IDEB	4,26	5,16	-0,90***	3,46	4,26	-0,80***
Tx_Aprovação	86,91	92,54	-5,63***	79,72	86,70	-6,99***
Tx_Abandono	2,47	0,91	1,56***	5,98	3,07	2,91***
População	807.554	749.085	58.469	873.783	952.403	-78.619
IDHM2000	0,62	0,63	-0,01***	0,63	0,63	0,00**
IDHM2010	0,73	0,73	0,00**	0,73	0,73	0,00
PIB_pc	20.997	23.858	-2.860***	22.610	23.939	1.328***

Fonte: resultados da pesquisa. *significativo a 10%; **significativo a 5%; ***significativo a 1%

A última amostra analisada também mostrou diferenças significativas entre as escolas participantes e não participantes em suas características observáveis. Entretanto, nesta amostra, observou-se uma mudança na média das variáveis que espelham os critérios de seleção do programa. A partir de 2010, o critério populacional foi relaxado, permitindo a participação no programa de escolas situadas em cidades menores. Consequentemente, a média populacional caiu de forma acentuada em relação aos anos anteriores.

Esta mudança também refletiu nas características das escolas. Curiosamente as escolas desta amostra apresentaram médias piores em saneamento básico adequado e IDHM em relação as amostras anteriores. O número de matrículas também se alterou. Agora, escolas tratadas possuem médias menores que aquelas do grupo de controle.

Adicionalmente, todas as variáveis de infraestrutura apresentaram médias mais baixas para as escolas tratadas. Estes resultados permitem pensar que a abertura dada ao critério populacional permitiu que escolas de menor porte e/ou situadas em cidades de infraestrutura mais precária participassem do programa, mudando o perfil das escolas tratadas.

Tabela 10 - Características médias das escolas participantes e não participantes em 2011 (amostra 3)

Variáveis	Escolas com 5º ano			Escolas com 9º ano		
	Escolas tratadas	Escolas controle	Diferença	Escolas tratadas	Escolas controle	Diferença
Lab_Informática	0,67	0,79	-0,12***	0,81	0,92	-0,11***
Lab_Ciências	0,06	0,11	-0,05***	0,15	0,35	-0,20***
Biblioteca	0,47	0,54	-0,07***	0,63	0,60	0,03***
Nu_Salas	9,58	11,57	-1,99***	11,18	13,45	-2,27***
Nu_Computadores	13,53	19,36	-5,83***	16,51	24,22	-7,71***
Alimentação	1,00	1,00	0,00	1,00	1,00	0,00
Nu_Funcionários	42,34	45,46	-3,12***	50,74	61,24	-10,50***
Água	0,92	0,98	-0,06***	0,93	0,98	-0,05***
Esgoto	0,99	1,00	-0,01***	0,99	1,00	-0,01***
Saneamento	0,92	0,98	-0,06***	0,93	0,98	-0,05***
Quadra	0,34	0,64	-0,30***	0,46	0,81	-0,34***
Ens_Médio	0,06	0,12	-0,06***	0,28	0,55	-0,26***
Nu_Turmas	19,00	20,72	-1,72***	22,23	26,14	-6,91***
Nu_Matrículas	475,90	525,38	-49,47***	610,10	771,75	-161,65
Duração_Média	251,98	267,24	-15,26***	254,35	271,67	-17,32***
IDEB	4,27	5,42	-1,15***	3,57	4,39	-0,82***
Tx_Aprovação	86,68	93,72	-7,04***	86,79	87,43	-0,63***
Tx_Abandono	2,46	0,81	1,64***	5,37	2,95	2,42***
População	243.496	779.768	-536.271***	298.661	970.156	-671.495***
IDHM2000	0,50	0,63	-0,13***	0,50	0,63	-0,13***
IDHM2010	0,65	0,73	-0,08***	0,65	0,73	-0,08***
PIB_pc	14.535	28.557	-14.015***	15.103	28.778	-13.674***

Fonte: resultados da pesquisa. *significativo a 10%; **significativo a 5%; ***significativo a 1%

Quanto aos indicadores educacionais, as escolas participantes ainda apresentaram números pré-tratamento inferiores quando comparados aqueles de escolas não participantes. Entretanto, a evolução constatada nas médias educacionais de 2007 para 2009 (amostras 1 e 2) tornou-se mais tímida em 2011 (amostra 3), com o único destaque positivo sendo a crescente observada na taxa de aprovação das

escolas tratadas com turmas de 9º ano. Os outros valores, no entanto, mostraram-se estagnados ao longo de dois anos.

De maneira geral, em todas as amostras foi possível observar que há diferenças na maioria das covariáveis e elas se mostraram estatisticamente significativas. Diante da presença de heterogeneidade entre os grupos, foi preciso a aplicação de um método que lide com essas diferenças, permitindo a construção de um grupo de controle adequado para cada amostra, que represente adequadamente a trajetória das escolas tratadas na ausência do programa. Como já exposto na metodologia, partiu-se então para a aplicação do *Propensity Score Matching*. O pareamento selecionou as escolas do grupo de controle mais similares às tratadas de forma que o impacto do programa fosse estimado sem a presença de viés. O resultado deste procedimento é apresentado na próxima subseção.

5.2 A Construção dos grupos de controle

Esta seção apresenta os resultados obtidos da primeira etapa da estratégia empírica. A aplicação de modelos *Probit* em cada amostra possibilitou construir grupos de controle adequados para a estimação dos impactos do Mais Educação. Usando dados referente ao ano pré-tratamento, dois modelos foram estimados para cada amostra, um para o 5º ano e outro para escolas de 9º ano.

Foi atribuído a cada unidade da amostra um *Propensity Score* e a partir dele foi possível a realização do pareamento aos 5 vizinhos mais próximos, métrica que selecionou para o grupo de controle somente aquelas unidades mais similares às tratadas, excluindo as demais.

5.2.1 O grupo de controle para escolas de 5º ano da amostra 1

O resultado da estimação do primeiro modelo *Probit* é apresentado na Tabela 11. Os valores foram obtidos para dados de 2007, período anterior ao tratamento das escolas presentes na amostra 1. Esta amostra reúne como tratadas escolas com turmas de 5º ano que aderiram ao Mais Educação em 2008 e permaneceram até 2013.

Tabela 11 - Resultados do Probit para participação no Mais Educação, 5º ano, usando dados pré tratamento de 2007 (amostra 1)

Variáveis	Coeficientes
População	0,000*** (0,000)
IDEB	-1,319*** (0,080)
IDHM2000	8,828*** (0,444)
Saneamento	-0,523*** (0,198)
Lab_Ciências	0,203** (0,084)
Biblioteca	0,253*** (0,063)
Quadra	-0,170*** (0,057)
Nu_Salas	-0,045*** (0,007)
Nu_Funcionários	0,010*** (0,001)
Ens_Médio	-0,861*** (0,095)
Duração_Média	-0,003*** (0,001)
Nu_Matrículas	0,001*** (0,000)
Media_5º_LP	0,018*** (0,004)
Media_5º_MT	-0,019*** (0,004)
Tx_Aprovação	0,014*** (0,003)
Tx_Abandono	0,032*** (0,005)
Número de Observações	11.633
LR chi2	2447,83
Prob>chi2	0,00
Pseudo R2	0,4574

Fonte: Resultados da Pesquisa.

Nota: Erros padrão em parênteses. *p<0,10, **p<0,05, ***p<0,01

O modelo foi estimado para 11633 observações, sendo 711 escolas tratadas e 10922 escolas do potencial grupo de controle. Como já pontuado anteriormente, o modelo foi controlado por características das escolas e de suas respectivas cidades que acredita-se serem fatores que influenciam a decisão de participação. O conjunto de

variáveis encontrado na Tabela 11 expressa o resultado do exercício “*hit or miss*”, utilizado para melhorar a qualidade do modelo²⁹.

A estatística LR, que testa se todos os coeficientes da regressão são simultaneamente iguais a zero, foi significativa a 1%, indicando que pode-se descartar a hipótese de insignificância conjunta dos coeficientes da regressão. Pelos resultados encontrados, vê-se que escolas com menores índices IDEB e maior número de matrículas tem maior probabilidade de participar no Mais Educação. Curiosamente, a população municipal reduz a probabilidade de participação no programa, apesar de ser critério de adesão, cidades com mais de 200 mil habitantes.

A probabilidade estimada de adesão ao Mais Educação é negativamente associada à presença de saneamento básico adequado, quadra de esportes e turmas de ensino médio. Por outro lado, afetam positivamente essa probabilidade laboratório de ciências, biblioteca e o número de funcionários. Além disso, a pontuação média das escolas em Português, a taxa de aprovação e a taxa de abandono elevam a probabilidade de participação. Uma maior nota média de Matemática, no entanto, diminuem as chances de participação.

Com *Propensity Score* estimado para toda unidade da amostra, realizou-se o pareamento das escolas tratadas com as cinco escolas do grupo de controle mais semelhantes em características observáveis, ou seja, com escores de propensão mais próximos (uma aplicação do algoritmo dos cinco vizinhos mais próximos com reposição). O pareamento foi realizado dentro do suporte comum, ou seja, em uma área de intercessão nas distribuições dos escores de propensão entre tratados e controle. Com isso, duas unidades tratadas foram excluídas do pareamento por possuírem escore de propensão fora dessa área. Das 10922 escolas não tratadas possíveis de pareamento, foram excluídas 9614, resultando em uma amostra pareada com 709 escolas tratadas e 1308 escolas controle.

A Tabela 12 reporta o resultado do teste de médias para as unidades tratadas e controle antes e depois do pareamento. Com o teste foi possível saber se a condição de balanceamento entre as variáveis dos dois grupos foi atendida. A aplicação do teste t permite testar a hipótese de que o valor médio de cada variável é igual para escolas

²⁹ Iniciou-se a estimação dos modelos *Probit* com um número amplo de variáveis que acreditava-se ser importante para determinar a participação de cada escola. Considerando a dificuldade de atender a hipótese de equilíbrio e o suporte comum, reduziu-se o número de variáveis até que o modelo fosse melhor especificado.

tratadas e controle. Este teste é feito antes e depois do pareamento. Para p maior que 0,10, a hipótese nula não pode ser rejeitada no nível de significância de 10%. Portanto, se a qualidade do pareamento for boa, espera-se que não haja diferenças estatisticamente significativas entre as médias dos dois grupos após o pareamento.

O teste ainda calculou o viés presente nas médias das variáveis antes e depois do pareamento e indica a variação. Este "viés" é definido como a diferença entre os valores médios do grupo de tratamento e do grupo não tratado, dividido pela raiz quadrada da variância média das duas amostras. Assim, foi possível acompanhar o percentual de redução no valor absoluto desse viés.

Tabela 12 - Testes de Média entre escolas do Mais Educação e escolas não participantes, antes e depois do pareamento, 5º ano

(Continua)

Variável	Amostra	Média		% viés	% viés reduzido	Teste t	
		Trat.	Cont.			t	$p > t $
População	Não pareada	1,1 e ⁰⁶	7,8 e ⁰⁵	20,2	82,3	4,03	0,000
	Pareada	1,1 e ⁰⁶	1,1 e ⁰⁶	3,6		0,67	0,501
IDEB	Não pareada	3,55	4,71	-157,1	96,8	-37,46	0,000
	Pareada	3,56	3,52	5,1		0,87	0,387
IDHM2000	Não pareada	0,651	0,637	21,0	70,5	4,48	0,000
	Pareada	0,651	0,647	6,2		1,08	0,280
Saneamento	Não pareada	0,969	0,991	-16,2	69,8	-5,87	0,000
	Pareada	0,969	0,976	-4,9		-0,78	0,437
Lab_Ciências	Não pareada	0,142	0,111	9,2	98,2	2,50	0,013
	Pareada	0,141	0,140	0,2		0,03	0,976
Biblioteca	Não pareada	0,752	0,710	9,5	96,0	2,40	0,016
	Pareada	0,753	0,751	0,4		0,07	0,941
Quadra	Não pareada	0,544	0,612	-13,7	41,5	-3,57	0,000
	Pareada	0,546	0,506	8,0		1,49	0,137
Nu_Salas	Não pareada	12,813	11,21	28,5	90,1	7,56	0,000
	Pareada	12,815	12,947	-2,8		-0,40	0,689
Nu_Funcionários	Não pareada	65,405	43,997	77,1	92,4	21,75	0,000
	Pareada	65,439	67,072	-5,9		-0,72	0,471
Ens_Médio	Não pareada	0,081	0,128	-15,1	56,7	-3,61	0,000
	Pareada	0,082	0,102	-6,6		-1,31	0,192
Duração_Média	Não pareada	252,61	265,02	-33,3	88,1	-8,39	0,000
	Pareada	252,68	254,14	-4,0		-0,92	0,360
Nu_Matrículas	Não pareada	902,92	580,23	75,7	96,9	22,05	0,000
	Pareada	902,71	912,71	-2,3		-0,32	0,752

Tabela 12 - Testes de Média entre escolas do Mais Educação e escolas não participantes, antes e depois do pareamento, 5º ano

(Conclusão)

Variável	Amostra	Média		% viés	% viés reduzido	Teste t	
		Trat.	Cont.			t	p> t
Media_5º_LP	Não pareada	164,44	183,25	-124,2	96,6	-28,86	0,000
	Pareada	164,45	165,09	-4,5		-0,84	0,400
Media_5º_MT	Não pareada	180,19	201,77	-132,2	96,6	-29,79	0,000
	Pareada	180,21	180,94	-4,5		-0,96	0,338
Tx_Aprovação	Não pareada	82,573	91,707	-83,4	89,2	-25,46	0,000
	Pareada	82,693	81,708	9,0		1,25	0,202
Tx_Abandono	Não pareada	5,162	1,044	81,3	96,2	29,69	0,000
	Pareada	5,070	4,916	3,1		0,38	0,705
R de Rubin	B	21,8					
	R	0,70					

Fonte: Resultados da Pesquisa

Na tabela é possível observar que antes do pareamento, todas as variáveis possuíam médias estatisticamente diferentes entre as escolas participantes e não participantes. Com a aplicação do pareamento com os cinco vizinhos mais próximos, as diferenças estatísticas foram extintas e assim, para todas as variáveis a hipótese de médias iguais não pode ser rejeitada. Consequentemente, observou-se uma importante redução no viés.

As notas médias de Português e Matemática – variáveis resultado para as quais o impacto do programa foi estimado, obtiveram um dos maiores percentuais de redução de viés (96,6%), reduzindo significativamente a diferença de médias para escolas tratadas e controle após o pareamento. Em relação as notas de Português, por exemplo, antes do pareamento as escolas tratadas possuíam uma média de 164,44 pontos. Já para as escolas do grupo de controle a média era 183,25. Com o pareamento e a exclusão das observações fora do suporte comum, as escolas tratadas obtiveram média de 164,45 e as escolas controle pareadas apresentaram média de 165,09 pontos.

Os resultados obtidos no teste de médias parecem demonstrar que o pareamento realizado com os cinco vizinhos mais próximos foi eficiente ao gerar um grupo controle semelhante o suficiente ao grupo de tratamento para que o impacto do Mais Educação possa ser estimado. Um teste adicional pode ser feito por meio das estatísticas B de Rubin e R de Rubin. O primeiro calcula a diferença absoluta padronizada nas médias do escore de propensão dos tratados e do grupo de controle pareados enquanto o segundo calcula a razão entre as variâncias do escore de

propensão dos dois grupos. Rubin (2001) recomenda que B seja menor que 25 e que R esteja entre 0,5 e 2 para que as amostras sejam consideradas suficientemente balanceadas. Para o modelo estimado obteve-se B e R dentro dos valores adequados, sendo 21,8 e 0,70 respectivamente.

Por fim, gráficos de densidade de Kernel que mostram a distribuição dos escores de propensão estimados foram obtidos para certificar mais uma vez sobre a qualidade do pareamento realizado. A Figura 1 mostra a distribuição dos escores de propensão estimados para escolas participantes e não participantes antes e depois do pareamento.

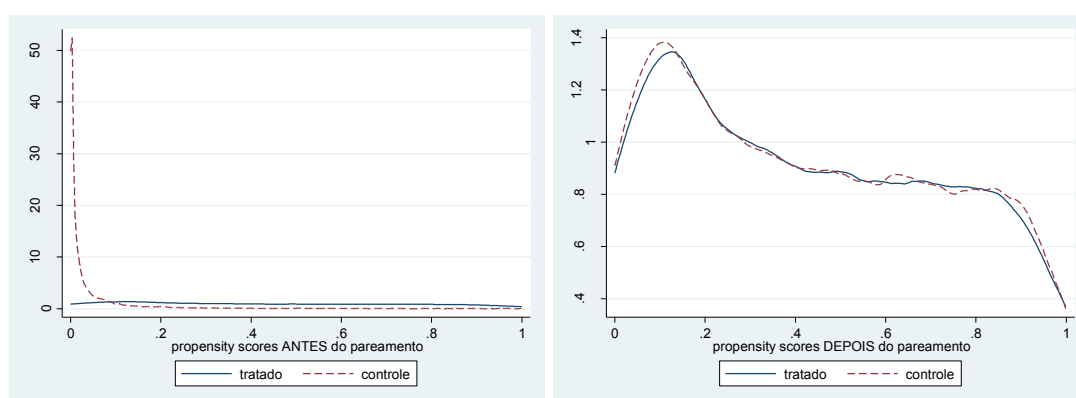


Figura 1 - Distribuição dos escores de propensão estimados para as escolas tratadas e controle de 5º ano, antes e depois do pareamento, amostra 1

Fonte: Resultados da Pesquisa.

No primeiro gráfico é possível observar a diferença na distribuição dos escores de propensão entre escolas tratadas e controle. Antes do pareamento, a maioria das unidades do grupo de controle apresentavam valor estimado no primeiro décimo do escore de propensão enquanto as unidades tratadas eram mais uniformemente distribuídas. Após o pareamento, a distribuição entre escolas tratadas e escolas selecionadas do grupo de controle tornou-se muito similar dentro da região de suporte comum.

5.2.2 O grupo de controle para escolas de 9º ano da amostra 1

O mesmo exercício foi realizado para escolas com turmas de 9º ano presentes na amostra 1 (escolas que iniciaram no programa em 2008 e permaneceram até 2013). Neste modelo *Probit*, 9054 observações foram utilizadas sendo 715 escolas tratadas e 8339 escolas no potencial grupo de controle. Os resultados são apresentados na Tabela 13.

Tabela 13 - Resultados do Probit para participação no Mais Educação, 9º ano, usando dados pré tratamento de 2007 (amostra 1)

Variáveis	Coeficientes
IDEB	-1,932*** (0,081)
IDHM2000	8,245*** (0,433)
Quadra	-0,277*** (0,069)
Nu_Computadores	-0,016*** (0,003)
Nu_Funcionários	0,007*** (0,001)
Duração_Média	-0,008*** (0,001)
Nu_Matrículas	0,000*** (0,000)
Media_9º_LP	0,046*** (0,004)
Media_9º_MT	-0,039*** (0,004)
Tx_Aprovação	0,048*** (0,003)
Tx_Abandono	0,026*** (0,004)
Número de Observações	9.054
LR chi2	2568,95
Prob>chi2	0,00
Pseudo R2	0,5136

Fonte: Resultados da Pesquisa

Nota: Erros padrão em parênteses. *p<0,10, **p<0,05, ***p<0,01

Neste modelo a estatística LR também foi significativa a 1%. Individualmente os coeficientes estimados também apresentaram-se estatisticamente significativos. Nota-se que a influência das variáveis do modelo na probabilidade de participação no programa Mais Educação foi semelhante a encontrada no modelo para o 5º ano. Dentre as variáveis que representam indicadores educacionais, maior IDEB e notas médias de matemática reduzem a probabilidade de adesão ao programa enquanto a nota média de português e as taxas de aprovação e abandono aumentam as chances.

A aplicação da métrica dos cinco vizinhos mais próximos também foi responsável pelo pareamento realizado nesta amostra. Com a imposição do suporte comum, 22 unidades do grupo de tratamento foram excluídas do pareamento. A amostra pareada contou com 693 unidades do grupo de escolas tratadas e 1069 do grupo de controle, excluindo, portanto, 7260 escolas não participantes.

O teste de médias também foi realizado para mostrar a adequabilidade do grupo de controle criado. A maioria das variáveis obtiveram redução de viés acima de 90%, não apresentando diferenças de média estatisticamente significativas após o pareamento entre escolas tratadas e controle. As notas médias de Português e Matemática da Prova Brasil não apresentaram-se estatisticamente iguais após o pareamento, contudo, os resultados mostraram uma redução significativa no viés contido nestas duas variáveis.

Tabela 14 - Testes de Média entre escolas do Mais Educação e escolas não participantes, antes e depois do pareamento, 9º ano

Variável	Amostra	Média		% viés	% viés reduzido	Teste t	
		Trat.	Cont.			t	p> t
IDEB	Não pareada	2,825	4,046	-178,2	98,0	-43,75	0,000
	Pareada	2,857	2,832	3,6		0,67	0,504
IDHM2000	Não pareada	0,649	0,639	13,9	90,8	2,92	0,004
	Pareada	0,649	0,649	1,3		0,24	0,809
Quadra	Não pareada	0,700	0,825	-29,5	99,8	-8,24	0,000
	Pareada	0,706	0,705	0,1		0,01	0,991
Nu_Computadores	Não pareada	10,943	13,844	-26,7	82,5	-6,60	0,000
	Pareada	11,063	10,556	4,7		0,90	0,370
Nu_Funcionários	Não pareada	77,645	62,457	50,0	85,3	12,87	0,000
	Pareada	77,351	79,582	-7,4		-0,85	0,397
Duração_Média	Não pareada	253,2	269,36	-52,8	95,1	-12,46	0,000
	Pareada	253,49	254,28	-2,6		-0,60	0,548
Nu_Matrículas	Não pareada	1184,6	887,56	59,6	84,6	15,92	0,000
	Pareada	1174,6	1129	9,2		1,44	0,151
Media_9º_LP	Não pareada	221,47	236,46	-97,3	82,8	-23,94	0,000
	Pareada	221,79	224,38	-16,8		-2,87	0,004
Media_9º_MT	Não pareada	229,54	250,29	-127,0	86,2	-29,57	0,000
	Pareada	229,92	232,79	-17,6		-3,30	0,001
Tx_Aprovação	Não pareada	73,997	84,175	-76,7	93,2	-21,47	0,000
	Pareada	74,257	73,57	5,2		0,81	0,417
Tx_Abandono	Não pareada	11,691	4,043	88,0	90,0	29,93	0,000
	Pareada	11,293	10,529	8,8		1,21	0,227
R de Rubin	B	29,4					
	R	0,54					

Fonte: Resultados da Pesquisa

As estatísticas B e R de Rubin mostraram-se um pouco inconclusivas. O resultado encontrado para B de Rubin é maior que 25, enquanto o R de Rubin está dentro do valor adequado para uma amostra balanceada.

Para concluir a análise da qualidade do pareamento foram obtidos gráficos da distribuição dos escores de propensão para a amostra de 9º ano (figura 2).

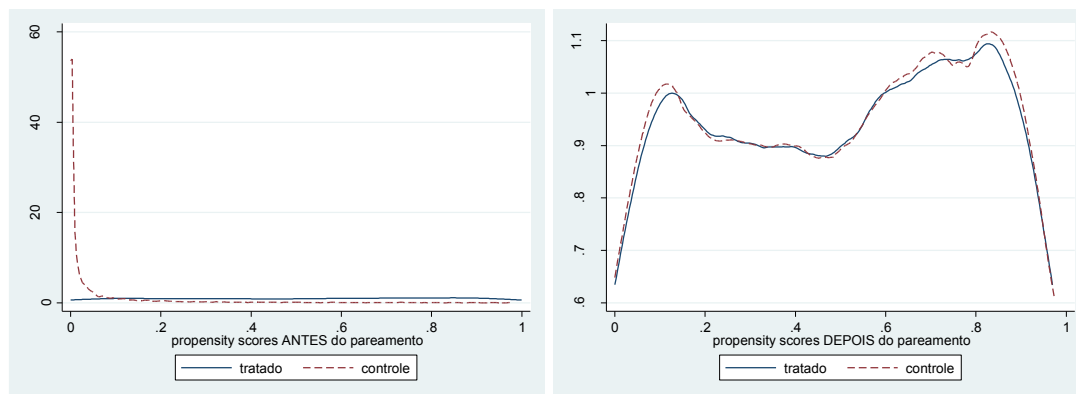


Figura 2 - Distribuição dos escores de propensão estimados para as escolas tratadas e controle de 9º ano, antes e depois do pareamento, amostra 1

Fonte: Resultados da Pesquisa.

Observa-se que antes do pareamento a distribuição das probabilidades de participação eram bem distintas entre escolas tratadas e escolas do potencial grupo de controle. Após a realização do pareamento, essa distribuição torna-se bem semelhante dentro da região de suporte comum.

Em comparação com as diferenças presentes antes do procedimento, os dados da Tabela 14 sugere um bom balanceamento das características observáveis entre escolas de tratamento e controle após o pareamento. Somente as variáveis representativas das médias de Português e Matemática ainda permaneceram com diferenças. Apesar disso, a estatística R de Rubin e a distribuição semelhante dos escores de propensão pôde assegurar a existência do balanceamento. Ademais, não pode-se desprezar o fato de que as regressões estimadas para o impacto do programa foram ponderadas pelos pesos que cada unidade do grupo de controle recebe ao ser pareada. A aplicação deste peso pode amenizar a presença de viés que ainda possa persistir após o pareamento realizado.

5.2.3 Os grupos de controle para as demais amostras

O mesmo exercício foi realizado para as amostras 2 e 3. Em cada uma delas foi estimado um modelo de probabilidade de participação para escolas com turmas de 5º

ano e outro modelo para escolas com turmas de 9º ano (da mesma forma como foi feita na amostra 1). Os resultados são apresentados nas Tabelas A1 a A4 do Anexo.

Os testes para verificar a qualidade do pareamento realizado e assim a adequabilidade dos grupos de controle criados também foram realizados em cada caso. Na amostra 2 para o 5º ano (Tabela A5), as notas de Português e Matemática ainda persistiram estatisticamente diferentes após o pareamento. Contudo, a diferença entre as médias mostrou-se muito pequena e a qualidade do pareamento pôde ser confirmada pelas estatísticas B e R de Rubin.

O mesmo acontece na amostra 3 para as escolas de 9º ano (Tabela A8). A nota média de Matemática permaneceu estatisticamente diferente entre escolas tratadas e as escolas escolhidas para o grupo de controle. Entretanto o pareamento realizado permitiu uma redução de 95.1% no viés para esta variável, permitindo que o balanceamento da amostra fosse confirmado pelas estatísticas B e R de Rubin.

Os gráficos de densidade de Kernel que apresentam a distribuição dos escores de propensão entre tratados e controle, antes e depois do pareamento, estão dispostos nas Figuras A1 a A3 do anexo. Em todos os casos é possível observar que após o pareamento a distribuição dos escores torna-se bastante semelhante entre as escolas tratadas e as unidades do grupo de controle, garantindo assim mais uma evidência da qualidade dos pareamentos realizados.

Os resultados encontrados nesta etapa possibilitaram a construção de grupos de controle similares às escolas tratadas, que fossem representativos da trajetória dessas escolas na ausência do programa. Com isso, foi possível controlar as unidades de cada amostra por suas características observáveis, minimizando a existência de viés de seleção. Na próxima subseção são apresentados os resultados dos modelos de estimação de impactos do programa Mais Educação, obtidos por meio da aplicação do método de Diferenças em Diferenças.

5.3 Impacto do Mais Educação para escolas com cinco, três e um ano de programa

5.3.1 Impacto para escolas com cinco anos de programa – amostra 1

Definido o grupo de controle adequado para cada ano em cada amostra, o impacto do programa Mais Educação sobre as notas médias de português e matemática da Prova Brasil puderam ser estimados. Os resultados foram obtidos a partir da

estimação da equação especificada em (21), considerando em todos os modelos o efeito fixo ao nível da escola.

A Tabela 15 apresenta os impactos médios da participação no Mais Educação sobre os resultados do 5º e 9º ano, para a amostra 1 que consiste em escolas que aderiram ao Mais Educação em 2008 e permaneceram até 2013.

Tabela 15 - Impacto do Mais Educação nas notas da Prova Brasil entre 2007 e 2013 (amostra 1)

Variável Dependente	Notas de Português	Notas de Matemática
Impacto no 5º ano		
Dummy ME*Ano2013	-8,485*** (1,195)	-11,527*** (1,258)
Ano (2013)	25,581*** (1,022)	25,764*** (1,068)
Constante	164,770*** (0,299)	180,574*** (0,315)
Efeitos Fixos	Sim	Sim
R2	0,6105	0,5669
Nº de Observações	4034	4034
Impacto 9º ano		
Dummy ME*Ano2013	-2,669 (2,040)	-5,982*** (2,230)
Ano (2013)	13,077*** (1,930)	9,662*** (2,151)
Constante	223,084*** (0,510)	231,352*** (0,557)
Efeitos Fixos	Sim	Sim
R2	0,2668	0,1233
Nº de Observações	3524	3524

Fonte: Fonte: Resultados da Pesquisa.

Nota: Erros padrão em parênteses. *p<0,10, **p<0,05, ***p<0,01. Todas as regressões ponderam as observações de controle segundo os pesos definidos com os 5 vizinhos mais próximos.

Os resultados obtidos mostram que para as escolas com turmas de 5º ano, o programa Mais Educação tem impacto negativo e estatisticamente significativo sobre as notas de Português e Matemática. A presença do programa provoca uma queda de 8,49 pontos nas notas de português e 11,53 pontos nas notas médias de Matemática, ao nível de 1% de significância. Comparados as notas médias de 2007, estes valores representam uma queda de 5,18% e 6,41% respectivamente.

Já para as turmas de 9º ano de amostra não foi encontrado impacto estatisticamente significativo para a nota de Português. Contudo, para o teste de matemática encontra-se o mesmo padrão das turmas de 5º ano, com ressalva feita a

magnitude do impacto, que desta vez foi menor: uma redução de 5,98 na nota ou uma queda de 2,60% em relação à média de 2007.

Os primeiros resultados encontrados não foram bons visto que estar-se a analisar uma política pública. O Mais Educação foi formulado dentro da necessidade de melhoria nos resultados educacionais brasileiros. Portanto, sempre ao analisar uma política implementada, principalmente a de educação em tempo integral que tem seu alcance aumentado a cada ano, espera-se que os recursos investidos nessa ação promovam resultados positivos.

Resultados de impacto negativos já foram encontrados em estudos anteriores. Mendes (2011) ao avaliar o impacto sobre turmas de 4^a série/5^o ano chega a um impacto negativo sobre a nota de Matemática. O mesmo acontece em Xerxenevsky (2012) e Almeida *et al* (2015). Porém, em todos, os resultados negativos foram menores. É importante destacar que nos trabalhos anteriores os impactos foram estimados para escolas com tempo menor de exposição ao programa. Uma amostra com exposição de cinco anos, como é feita aqui, é inédita entre os estudos já realizados. Isso torna o resultado mais intrigante, uma vez que é plausível supor que com um tempo maior de exposição ao programa, os possíveis problemas de implementação que poderiam ter atrapalhado a execução do programa já teriam sido sanados.

Para então verificar se o cenário encontrado é reflexo de alguma característica das escolas que compõem a amostra, uma nova estimação para a amostra 1 foi feita utilizando os resultados da Prova Brasil para 2009. Com este exercício, buscou-se verificar qual foi o impacto médio do programa no curto prazo, quando estas escolas tinham somente um ano de exposição ao programa, comparando a performance destas escolas entre 2007 e 2009. As unidades dos grupos de escolas tratadas e controle foram as mesmas da análise anterior, obtidas pelo pareamento já realizado para esta amostra. Os resultados estão descritos na Tabela 16.

Tabela 16 - Impacto do Mais Educação nas notas da Prova Brasil entre 2007 e 2009 (amostra 1)

Variável Dependente	Notas de Português	Notas de Matemática
Impacto no 5º ano		
Dummy ME*Ano2009	-6,268*** (1,214)	-9,828*** (1,231)
Ano (2009)	15,102*** (1,094)	19,355*** (1,095)
Constante	125,026*** (0,294)	180,800 (0,298)
Efeitos Fixos	Sim	Sim
R2	0,3847	0,4626
Nº de Observações	3976	3976
Impacto 9º ano		
Dummy ME*Ano2009	-5,406** (2,221)	-6,541*** (2,076)
Ano (2009)	15,199*** (2,136)	8,663*** (2,019)
Constante	223,203*** (0,540)	231,476*** (0,505)
Efeitos Fixos	Sim	Sim
R2	0,3191	0,1248
Nº de Observações	3461	3461

Fonte: Fonte: Resultados da Pesquisa.

Nota: Erros padrão em parênteses. *p<0,10, **p<0,05, ***p<0,01. Todas as regressões ponderam as observações de controle segundo os pesos definidos com os 5 vizinhos mais próximos.

A avaliação do impacto de curto prazo do Mais Educação mostra que após um ano de exposição ao programa, as escolas participantes também tiveram uma redução nas notas de Português e Matemática. A nota dos alunos das escolas tratadas com turmas de 5º ano teve uma queda de 6,29 pontos no teste de português e 9,83 pontos no teste de matemática, ambos a nível de significância de 1%.

Para o 9º ano o impacto foi um pouco menor, mas ainda sim negativo. As notas de Português tiveram, em média, uma redução de 5,41 pontos (ao nível de significância de 5%) e para matemática a redução foi de 6,54 pontos (ao nível de significância de 1%).

Nesta estimação obteve-se o impacto do programa para um tempo de exposição semelhante ao encontrado em trabalhos anteriores. O que se vê são impactos negativos também no curto prazo. Os resultados encontrados mostraram-se um pouco menor que aqueles encontrados na primeira estimação para as turmas de 5º ano. Neste caso, um maior tempo de exposição ao programa, ao contrário do que normalmente se espera,

resultou em impactos negativos maiores sobre as escolas participantes. Para as escolas com turmas de 9º ano esta comparação pode ser feita para as notas de matemática e a conclusão segue na mesma direção.

Inicialmente, é preciso pensar nos fatores que levaram ao insucesso do programa. Seguindo a linha adotada por outros trabalhos, poder-se-ia argumentar que desafios na implementação dada a nova abordagem sobre os estudantes, professores e hábitos escolares poderiam gerar dificuldades na condução do programa pelas escolas e refletir no desempenho de seus alunos. Mas isso geralmente acontece quando há pouco tempo de programa, situação em que as escolas contam com pouca experiência para gerir os problemas encontrados no dia-a-dia, que não parece ser o caso da amostra avaliada.

O que parece ser plausível, no entanto, é que os resultados negativos encontrados tenham relação com outros indicadores educacionais, a taxa de reprovação e a taxa de abandono. Estudos constataram que a presença do Mais Educação reduz as taxas de reprovação e abandono³⁰. É natural intuir que estes alunos que contaram com a ajuda do programa para permanecerem na escola ou progredir de ano apresentam um rendimento inferior aos outros. Dessa forma, a inclusão desses alunos com piores desempenhos nas avaliações da Prova Brasil pode ter reduzido o desempenho médio da escola ao longo dos anos.

Um outro fator importante que pode estar relacionado com os resultados encontrados pode ser a pouca conexão entre as atividades desenvolvidas pelos alunos do programa e as habilidades cobradas nos testes da Prova Brasil. Ao longo dos anos, as orientações quanto às atividades que as escolas deveriam desenvolver com seus alunos foram atualizadas pelos editais do programa, mas uma característica mantida até 2013 foi a obrigatoriedade de adesão à uma atividade do macrocampo acompanhamento pedagógico. No entanto, a escola não precisa necessariamente escolher entre matemática ou letramento. Há diversas outras temáticas como Ciências, História e Geografia. Portanto, as escolas participantes do programa não necessariamente estão trabalhando no contraturno assuntos relacionados aos conteúdos cobrados na Prova Brasil.

Esse parece ter sido de fato um aspecto importante no alcance dos objetivos do programa. Recentemente, o Mais Educação sofreu uma grande reformulação, mas

³⁰ Mendes (2011) e Pereira (2011)

manteve entre seus principais objetivos a melhoria nos resultados de aprendizagem do ensino fundamental. Dentre as mudanças, a escola que aderir ao programa agora precisa realizar duas atividades de acompanhamento pedagógico, sendo elas Língua Portuguesa e matemática. Também ficam determinadas as horas desprendidas nessas atividades, que pode ser de duas horas e meia ou quatro horas cada, dependendo do tipo de programa selecionado pela escola (MEC, 2017).

Uma outra hipótese que pode ser considerada para entender o resultado negativo do programa está ligada, principalmente às discussões pedagógicas. É evidente nos documentos que norteiam o programa e até em estudos que já avaliaram a implementação do Mais Educação *in loco*, a ênfase assistencialista dada na construção e na execução do programa, de forma que a ampliação da jornada escolar seja construída numa perspectiva compensatória, apoiando os estudantes em suas carências.

O aluno atendido pelo programa passa a maior parte do tempo envolvido em atividades que não necessariamente versam sobre o conteúdo do ensino regular, como culturais e esportivas, muitas vezes desarticuladas³¹ ao currículo básico obrigatório. A escola torna-se um atrativo para aqueles alunos em que as dificuldades socioeconômicas pesam no dia-a-dia escolar e talvez isso possa até justificar o impacto positivo do programa sobre a taxa de evasão.

Segundo Mosna (2014), a desintegração do currículo transforma as atividades do contraturno em uma espécie de “apoio”, com regras mais flexíveis, caracterizando cada tempo da escola de forma diferente. Na visão da autora, essa falta de diálogo reflete no alto índice de infrequência de estudantes nas atividades do programa, que muitas vezes comparecem no turno inverso somente em atividades de seu interesse.

O relato feito por Mosna (2014) refere-se à sua avaliação em escolas do Rio Grande do Sul, mas pode estar ocorrendo em todo Brasil. Pesa ainda a falta de apoio técnico e institucional dado às escolas participantes após a adesão ao programa, que carecem de monitoramento como coloca Santos (2014). Essas questões podem demonstrar que talvez possa haver problemas no desenho do programa, e desenvolvidas da forma que estão, as atividades propostas aos estudantes podem ser ineficientes para promover as habilidades que são cobradas em testes padronizados

³¹ A desarticulação de conteúdo e a necessidade de reestruturação curricular é uma crítica constante dos trabalhos da área educacional/pedagógica quanto ao alcance dos programas que estimulam a ampliação do tempo escolar. Ver Cavaliere (2007; 2009).

como a Prova Brasil. Isto pode então estar refletindo nos resultados negativos alcançados pelo programa.

Dando continuidade à avaliação do programa, os próximos resultados expõem o impacto do Mais Educação em escolas há três anos no programa e logo em seguida para a amostra com um ano de programa.

5.3.2 Impacto para escolas com três e um ano de programa – amostra 2 e 3

A Tabela 17 apresenta os resultados encontrados para a amostra 2 – que reúne as escolas que aderiram ao Mais Educação em 2010 e permaneceram até 2013.

Tabela 17 - Impacto do Mais Educação nas notas da Prova Brasil entre 2009 e 2013 (amostra 2)

Variável Dependente	Notas de Português	Notas de Matemática
Impacto no 5º ano		
Dummy ME*Ano2013	-3,439*** (0,411)	-4,579*** (0,455)
Ano (2013)	13,071*** (0,279)	10,624*** (0,314)
Constante	177,639*** (0,103)	195,685*** (0,114)
Efeitos Fixos	Sim	Sim
R2	0,3458	0,1968
Nº de Observações	16432	16432
Impacto 9º ano		
Dummy ME*Ano2013	-1,346** (0,566)	-2,423*** (0,502)
Ano (2013)	2,129*** (0,439)	4,909*** (0,377)
Constante	237,306*** (0,142)	238,698*** (0,125)
Efeitos Fixos	Sim	Sim
R2	0,0090	0,0598
Nº de Observações	12340	12340

Fonte: Fonte: Resultados da Pesquisa.

Nota: Erros padrão em parênteses. *p<0,10, **p<0,05, ***p<0,01. Todas as regressões ponderam as observações de controle segundo os pesos definidos com os 5 vizinhos mais próximos.

Como pode-se observar, o impacto do Mais Educação continua sendo negativo para as escolas analisadas. Os resultados demonstraram que as escolas de 5º ano com a presença do programa de 2010 a 2013, tiveram uma redução de 3,44 pontos na nota de Português e 4,58 pontos sobre a nota de Matemática, ambos ao nível de 1% de

significância. Em relação as notas médias de 2009 (antes da exposição), houve uma queda de 1.93% e 2.34%, respectivamente.

Em relação às escolas com turmas de 9º ano, o Mais Educação resultou em um impacto negativo de 1,35 pontos para as notas de Português (ao nível de 5% de significância) e 2,42 pontos sobre as notas de Matemática (ao nível de significância de 1%), uma queda de 0,57% e 1,01% em relação as notas médias de Português e Matemática de 2009.

Apesar de os impactos negativos permanecerem, vê-se que a tamanho da queda nas notas diminuiu quando comparado aos resultados obtidos para a amostra 1. De forma peculiar, tem-se que para a amostra com menos tempo de exposição ao programa os impactos negativos do Mais Educação sobre as notas de Português e Matemática da Prova Brasil são menores.

Na Tabela 18 observam-se os impactos encontrados para uma amostra que continha escolas com pouco tempo de participação no programa, mais especificadamente, um ano (2012-2013). Com estes resultados, pode-se verificar se a tendência de redução dos impactos negativos do programa se confirma para escolas com tempo de exposição menor.

Tabela 18 - Impacto do Mais Educação nas notas da Prova Brasil entre 2011 e 2013 (amostra 3)

Variável Dependente	Notas de Português	Notas de Matemática
Impacto no 5º ano		
Dummy ME*Ano2013	-1,787*** (0,357)	-2,727*** (0,411)
Ano (2013)	4,119*** (0,302)	1,980*** (0,340)
Constante	175,503*** (0,089)	192,873*** (0,103)
Efeitos Fixos	Sim	Sim
R2	0,0444	0,0071
Nº de Observações	27628	27628
Impacto 9º ano		
Dummy ME*Ano2013	-0,778** (0,385)	-0,388 (0,384)
Ano (2013)	3,116*** (0,321)	0,582* (0,316)
Constante	228,598*** (0,096)	253,979*** (0,096)
Efeitos Fixos	Sim	Sim
R2	0,0344	0,0009
Nº de Observações	20714	20714

Fonte: Fonte: Resultados da Pesquisa.

Nota: Erros padrão em parênteses. *p<0,10, **p<0,05, ***p<0,01. Todas as regressões ponderam as observações de controle segundo os pesos definidos com os 5 vizinhos mais próximos.

Para as escolas com turmas de 5º ano, o Mais Educação contribuiu com uma queda de 1,78 pontos na nota de Matemática e 2,75 pontos na nota do teste de Português, ambos estatisticamente significativos a 1%. Para escolas com turmas de 9º ano, a nota de Português reduziu-se 0,78 pontos ao nível de significância de 5% enquanto para as notas de matemáticas não houve impacto estatisticamente significativo.

Como os resultados indicam, escolas com exposição de um ano ao programa tiveram impactos negativos reduzidos, comparados às amostras anteriores. Ao que parece aquelas que aderiram ao programa recentemente se beneficiaram da evolução que a política alcançou ao longo dos anos. É natural pensar que as escolas que aderiram ao programa no início da política estiveram suscetíveis a problemas de implementação, tal como organização da escola, infraestrutura insuficiente às novas atividades e, principalmente, recursos. Com o passar do tempo, os pontos negativos e positivos da política deram lugar à diversas experiências, que puderam ser compartilhadas entre

escolas e os órgãos responsáveis, facilitando o trabalho de organização das atividades daquelas escolas que aderiram ao programa em anos mais recentes.

As escolas que aderiram ao programa recentemente também podem ter se beneficiado dos estudos já realizados sobre o programa, principalmente daqueles que puderam acompanhar *in loco*, relatando as experiências de sucesso e insucesso de escolas que já iniciaram a reorganização de seu currículo para o caminho da educação integral.

Por outro lado, os resultados encontrados neste estudo ligam o alerta principalmente para o conjunto de escolas da primeira amostra, as pioneiras no programa. Os resultados não favoráveis mostraram-se constantes já que não é a primeira vez que avaliações negativas do programa foram feitas para as escolas que começaram em 2008 (tal como o estudo de Almeida *et al*, 2015). Com a tendência de crescimento do programa e a indicação de permanência da política na composição destas escolas, é interessante que elas sejam identificadas e que sejam feitos estudos mais detalhados, de forma individual, para que os principais problemas sejam levantados.

Por fim, faz-se necessário destacar as principais limitações encontradas para este estudo. A primeira delas refere-se à unidade de análise utilizada. Uma avaliação mais precisa seria feita se contasse com os principais afetados pela política – os alunos. Mas, infelizmente, as bases de dados disponíveis não identificam os alunos participantes do programa e nem quantos são. Sabe-se apenas que são no mínimo cem em cada escola participante, de acordo com as instruções contidas nos manuais do programa.

Ainda sobre os alunos do Mais Educação, não é possível obter nas bases de dados qual série estão cursando. Isso pode representar um impacto significativo nas variáveis de impacto, uma vez que os testes são aplicados a alunos ao final do 5º e 9º anos. No entanto, há a orientação de que a preferência sobre a participação no programa seja dada a alunos em séries finais dos dois ciclos do ensino fundamental, o que pode ser um alento aos resultados encontrados.

Outra observação importante é quanto as atividades realizadas em cada escola adepta ao programa. Como já fora discutido, as escolas têm liberdade de escolha sobre as atividades desempenhadas no contraturno, o que as tornam diferentes para qualquer análise de impacto. Uma sugestão a próximos trabalhos é que essa diferenciação seja

feita, o que poderia levar a análises mais apuradas sobre quais conteúdos traz realmente efeito sobre o desempenho dos alunos.

Não espera-se com este trabalho esgotar todas as nuances da educação em tempo integral. Sabe-se perfeitamente que o assunto é extenso e as experiências de sucesso podem acontecer sob diversas variações dado a heterogeneidade do sistema educacional brasileiro. Contudo, espera-se oferecer mais instrumental a esta discussão, visto que as políticas educacionais estão cada vez mais consolidadas nos moldes da educação integral.

5.4 Considerações Finais

Pensar propostas para a educação é um grande desafio diante do quadro em que se encontra o Brasil. Há algum tempo, governos, instituições e pesquisadores vêm se engajando em pesquisas que tornem real um avanço na qualidade educacional brasileira. Os atrasos já foram enormes e a sociedade sente até hoje, os reflexos de um quadro crítico que perdurou até o final da década de 1980, mas que mudou de direção com a bem-sucedida política de universalização.

Aliada aos avanços, a política de ensino fundamental de 9 anos mostrou o comprometimento das novas políticas educacionais com a inclusão e a equidade. O próximo passo foi aumentar o tempo que o aluno brasileiro passa na escola. Nesse sentido, foi criado o Programa Mais Educação, que pretende levar a escolas públicas brasileiras a ampliação da jornada escolar para no mínimo 7 horas diárias. Entre os objetivos do programa estão a melhoria no desempenho dos alunos, redução do abandono, da reprovação e da distorção idade/série, além de proporcionar mais possibilidades socioeducativas aos alunos mais vulneráveis.

Sabe-se que após uma política pública ser implementada há a necessidade de acompanhá-la permitindo aos governantes, os beneficiados e até mesmo a sociedade, saber se os esforços empenhados, sejam eles recursos humanos, técnico e financeiros estão gerando o retorno esperado. Nesse aspecto, entram os estudos de avaliação de políticas públicas, cada vez mais importantes para o país, fundamentais por dar mais transparência às ações públicas.

Diante disso, este trabalho analisou os efeitos do Mais Educação sobre um importante indicador educacional, a Prova Brasil, que mede a proficiência dos alunos do ensino público em testes de Português e Matemática. O objetivo principal deste trabalho foi estender a análise já realizada na literatura, avaliando o impacto do programa sobre as escolas pioneiras na política, que aderiram ao Mais Educação em 2008 e permaneceram até 2013, em um período de cinco anos de exposição ao programa. Pretendeu-se também analisar o impacto em escolas que aderiram mais recentemente e que tinham três e um ano de participação no Mais Educação. Assim seria possível comparar os efeitos estimados e discutir as principais causas da diferença nos resultados encontrados.

Para alcançar os objetivos do estudo utilizou-se conjuntamente duas técnicas comuns de avaliação de impacto de políticas públicas: o *Propensity Score Matching* e o Diferenças em Diferenças. A aplicação da primeira permitiu a construção de um

contrafactual que espelhasse o desempenho das escolas que aderiram ao Mais Educação na ausência do programa. A utilização dessa estratégia permite que possíveis vieses de seleção sejam minimizados e assim não influenciem os resultados encontrados. Logo após a construção do grupo de controle, aplicou-se o modelo de diferenças em diferenças que procedeu com a comparação dos resultados entre o grupo de escolas tratadas e controle antes e depois do tratamento.

Em todas as amostras analisadas o Mais Educação contribuiu negativamente para o desempenho médio nas avaliações de português e matemática da Prova Brasil (exceto em alguns casos em que o efeito foi inexistente). Uma peculiaridade encontrada na análise foi impactos negativos maiores para escolas que estão a mais tempo no programa. A amostra de escolas com três anos de programa obteve uma queda menor que a encontrada para aquelas com cinco anos de programa. Da mesma forma, escolas com um ano de exposição ao programa tiveram resultados negativos de menor impacto que aquelas que estavam há três anos.

Primeiramente, buscou-se entender os resultados negativos. Sabe-se que as limitações na base de dados, por não permitir que a análise seja feita sobre o aluno participante e sim sobre a escola, não permite que muitas questões sejam levantadas. Uma hipótese levantada foi sobre a redução das taxas de evasão. Alguns estudos já mostraram que a presença do Mais Educação tem reduzido o abandono escolar. Se alunos com piores desempenhos agora permanecem na escola, é possível que o desempenho médio nos testes avaliados tenha diminuído ao longo do tempo.

Outra questão levantada é a falta de conexão entre as atividades desenvolvidas no contraturno e o conteúdo cobrado na Prova Brasil. Dentre as atividades escolhidas pela escola para serem financiadas pelo programa, somente uma delas obrigatoriamente deveria ser de acompanhamento pedagógico. Além disso, a escola poderia escolher entre diversos conteúdos, não sendo necessariamente matemática e letramento. Assim sendo, pouco ou nenhum impacto teriam sobre as notas dos testes de português e matemática da Prova Brasil.

A falta de conexão entre as atividades executadas no período estendido e o currículo básico obrigatório também é uma grande crítica ao programa. Alguns estudos na literatura já sinalizaram que a falta de monitoramento e apoio técnico as escolas participantes também podem dificultar o sucesso do programa.

A educação em tempo integral continua sendo um dos pilares na estrutura de políticas que buscam a qualidade do ensino brasileiro. Recentemente, essa proposta

foi estendida ao Ensino Médio, com a Medida Provisória nº746 de 2016, responsável pela reforma do Ensino Médio. Mas será que mais tempo na escola indica melhor qualidade? Os resultados têm sido controversos. Como estratégia de inclusão, os resultados positivos são vistos de maneira rápida. Contudo, parece que para realmente gerar impacto no desempenho dos alunos, os moldes de implementação precisam ser revistos.

Como toda boa experiência em educação, é preciso que mais estudos individualizados sejam feitos naquelas escolas que apresentam sucesso no programa. Boas práticas precisam ser externalizadas. De outra forma, não menos importante, a construção de um sistema de acompanhamento aos estudantes beneficiados pela política possibilitaria bases de dados mais completas aos estudos, promovendo um ganho de qualidade nas análises realizadas. Sabe-se que educação é um tema amplo, muitas vezes complexo e dificilmente esgotado. Espera-se que as evidências encontradas neste estudo contribuam para o enriquecimento do debate e sirva de subsídio para muitas propostas que ainda virão.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, R.; BRESOLIN, A.; BORGES, B.; MENDES, K.; MENEZES, N. A. F. Assessing the Impacts of Mais Educação on Educational Outcomes: Evidence from 2008 to 2011. In: 12º Seminário Itaú Internacional de Avaliação Econômica de Projetos Sociais – Fundação Itaú Social 2015. Anais...São Paulo: Fundação Itaú Social, 2015.
- AQUINO, J. M.; KASSOUF, A. L. A ampliação da jornada escolar melhora o desempenho acadêmico dos estudantes? Uma avaliação do programa escola de tempo integral da rede pública do estado de São Paulo. In: Encontro Nacional de Economia – ANPEC 2011, XXXIX, 2011, Foz do Iguaçu. Anais... Foz do Iguaçu: ANPEC, 2011.
- BARBOSA, M. L. O. Desigualdade e desempenho: uma introdução à sociologia da escola brasileira. Belo Horizonte: Argvmentvm, v. 272, 2009.
- BARBOSA FILHO, F. H.; PESSÔA, S. A. Educação e crescimento: O que a evidência empírica e teórica mostra? Revista EconomiA, v. 11, n. 2, p. 265-303, 2010.
- BARROS, R.; LAM, D. Income Inequality in Education, and Children's Schooling Attainment in Brazil. In: BIRDSALL, Nancy; SABOT, Richard (orgs.). Opportunity Foregone: Education in Brazil. Washington, D.C: Inter-American Development Bank, 1996, cap. 12, pp. 337-366.
- BECKER, G. Investment in human capital: A theoretical analysis. Journal of Political Economy Part 2: Investment in Human Beings, v. 70, n. 5, p. 9–49, 1962.
- BECKER, S. O.; ICHINO, A. Estimation of Average Treatment Effects Based on Propensity Score. The Stata Journal, v. 2, n. 4, p. 358-377, 2002.
- BELLEI, C. Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile. Economics of Education Review, v. 28, p. 629–640, 2009.
- BIONDI, R. L.; FELÍCIO, F. D. Atributos escolares e o desempenho dos estudantes: uma análise em painel dos dados do SAEB. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Brasília-DF, 2007
- BOWLES, S. Toward an Educational Production Function. In Education, income and human capital. Ed.: W. LEE HANSEN. New York: NBER, pp. 11-60, 1970.
- BRAZIL. Lei Ordinária nº 13.005 de 25 de junho de 2014. Diário Oficial da República do Brasil, Poder Executivo, Brasília, DF, 26 jun. 2014. Edição Extra, p.1.
- CALIENDO, M.; KOPEINIG, S. Some practical guidance for the implementation of propensity score matching, Journal of Economic Surveys, v. 22, n.1, p. 31-72, 2008.
- CAVALIERE, A. M. Tempo de escola e qualidade na educação pública. Educação & Sociedade, vol, 28, nº 100 – Especial. p. 1015-1035, Campinas: out, 2007.
- CAVALIERE, A. M. Escolas de tempo integral versus estudantes em tempo integral. Em Aberto, v. 22, n. 80, p. 51-63, 2009.

CERDAN-INFANTES, P.; VERMEERSCH, C. “More Time Is Better: An Evaluation of the Full-Time School Program in Uruguay”, Working Paper, World Bank. 2007.

COLEMAN, J. S. *et al.* Equality of Educational Opportunity. Washington: US Government Printing Office, 1966.

DE CICCIA, P. Does full-day kindergarten matter? Evidence from the first two years of schooling. *Economics of Education Review*, v. 26, n. 1, p. 67-82, 2007.

ERNICA, M. Percursos da educação integral no Brasil. In: Seminário Nacional Tecendo Redes para a educação integral. São Paulo: CENPEC, 2006.

ESPOSITO, Y. L.; DAVIS, C.; NUNES, M. M. Sistema de avaliação do rendimento escolar: o modelo adotado pelo Estado de São Paulo. *Revista brasileira de educação*, v. 13, p. 25-53, 2000.

FELICIO, F.; FERNANDES, R. O efeito da qualidade da escola sobre o desempenho escolar: uma avaliação do ensino fundamental no estado de São Paulo. In: Encontro Nacional de Economia, 33, Natal, 2005.

GOMES, C. A. A escola de qualidade para todos: abrindo as camadas da cebola. *Revista Ensaio: Avaliação e Políticas Públicas em Educação*, v. 13, n. 48, 2005.

GLEWWE, P.; KREMER, M. Schools, teachers and educational outcomes in developing countries. In *Handbook of the economics of education*, Amsterdam, v. 2, 2006.

HANUSHEK, E. Conceptual and Empirical Issues in the Estimation of Educational Production Functions. *The Journal of Human Resources*, v. 14, n. 3, p. 351-388, Summer, 1979.

HANUSHEK, E. The economics of schooling: production and efficiency in public schools. *Journal of Economic Literature*, v. 24, 1986.

HANUSHEK, E. Alternative school policies and the benefits of general cognitive skills. *Economics of Education Review*, v. 25, n. 4, p. 447-462, 2006.

HANUSHEK, E. Education Production Functions. *The New Palgrave Dictionary of Economics*. Hoover Institution, Stanford University, 2008.

HANUSHEK, E.; KIMKO, D. Schooling, labor force quality, and the growth of nations. *American Economic Review* v. 90, n. 5, p. 1184-1208, 2000.

HANUSHEK, E.; LUQUE, J. Efficiency and equity in schools around the world. *Economics of Education Review*, v. 22, p. 481-502, 2003.

HIRANO, K.; IMBENS, G.; RIDDER, G. Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score. *Econometrica*, v. 71, n.4, p. 1161-1189, Jul, 2003.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Síntese de indicadores sociais: uma análise das condições de vida da população brasileira, 2014. Rio de Janeiro, 2014.

KHANDKER, S. R.; KOOLWAL, G. B.; SAMAD, H. A. Handbook on impact evaluation – quantitative methods and practices. [S.I]: The World Bank, 2010.

MARÍN, A.G. Evaluación del Impacto de la Jornada Escolar Completa. Dissertação de Mestrado, Universidade do Chile. 2006.

MAURICIO, L. V. Escritos, representações e pressupostos da escola pública de horário integral. Revista Em Aberto, v. 22, n. 80, p. 1-165, 2009.

MENDES, K. M. O Impacto do Programa Mais Educação no Desempenho dos Alunos da Rede Pública Brasileira. Monografia da Universidade de São Paulo. São Paulo, 2011.

MENEZES FILHO, N. A. Os Determinantes do Desempenho Escolar do Brasil. Instituto Futuro Brasil, Ibmec-SP, FEA-USP. 2007.

MENEZES FILHO, N. A. (org). Avaliação econômica de projetos sociais. São Paulo: Fundação Itaú Social, 2012.

MINCER, J. Investment in human capital and personal income distribution. Journal of Political Economy, v. 66, n. 4, p. 281–302, 1958.

MINCER, J. Schooling, Experience and Earnings. National Bureau of Economic Research, distributed by Columbia University Press, New York., 1974.

MINISTÉRIO DA EDUCAÇÃO. Educação integral: texto referência para o debate nacional. Brasília: MEC, Secad, 2009a.

MINISTÉRIO DA EDUCAÇÃO. Programa Mais Educação: gestão intersetorial no território. Brasília: MEC, Secad, 2009b.

MINISTÉRIO DA EDUCAÇÃO. Manual de educação integral para obtenção e apoio financeiro através do programa dinheiro direto na escola – PDDE, no exercício de 2009. Brasília: MEC, Secad, 2009c.

MINISTÉRIO DA EDUCAÇÃO. Programa Mais Educação. Passo a Passo. Brasília: SEB/MEC, 2011.

MINISTÉRIO DA EDUCAÇÃO. Programa Mais Educação. Passo a Passo. Brasília: SEB/MEC, 2013.

MINISTÉRIO DA EDUCAÇÃO. Censo Escolar da Educação Básica 2013: resumo. Brasília: INEP/MEC, 2014.

MINISTÉRIO DA EDUCAÇÃO. Indicadores Educacionais. Brasília: INEP/MEC, 2015a. Disponível em <http://portal.inep.gov.br/indicadores-educacionais>. Acesso em 1º de novembro de 2015.

MINISTÉRIO DA EDUCAÇÃO. Programa Mais Educação. Brasília: INEP/MEC, 2015b. Disponível em <http://portal.mec.gov.br/programa-mais-educacao/apresentacao>. Acesso em 1º de novembro de 2015.

MINISTÉRIO DA EDUCAÇÃO. Programa Novo Mais Educação, 2017. Disponível em <http://portal.mec.gov.br/programa-mais-educacao/apresentacao>. Acesso em 30 de janeiro de 2017.

MOSNA, R. M. P. Avaliação da política pública “Programa Mais Educação” em escolas de ensino fundamental da rede estadual de ensino do Rio Grande do Sul: os impactos na qualidade da educação e no financiamento do ensino fundamental. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade Federal do Rio Grande do Sul. Porto Alegre, 2014.

PAES DE BARROS, R. *et al.* Determinantes do desempenho educacional no Brasil. Pesquisa e Planejamento Econômico. v. 31, n. 1, p. 1-42, abr, 2001.

PEREIRA, G. C. Uma avaliação de impacto do programa Mais Educação no Ensino Fundamental. Dissertação (Mestrado em Políticas Públicas, Estratégia e Desenvolvimento). Universidade Federal do Rio de Janeiro, 2011.

RIANI, J. D. L. R.; RIOS-NETO, E. L. G. Background familiar versus perfil escolar do município: qual possui maior impacto no resultado educacional dos alunos brasileiros. Revista Brasileira de Estudos de População, v. 25, n. 2, p. 251-269, 2008.

ROSENBAUM, P.; RUBIN, D. The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika, v. 70, n. 1, p. 41-55, 1983.

RUA, M. G. Políticas Públicas. Florianópolis: CAPES – UAB, 2009.

RUBIN, D. B. Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation. Health Services & Outcomes Research Methodology, v. 2, p. 169-188, 2001.

SANTOS, F. M. Políticas públicas de ampliação da jornada escolar na perspectiva da educação integral fazem diferença? Um estudo do Programa Mais Educação. Tese (Doutorado em Educação) – Universidade Católica de Brasília. Brasília, 2014.

SCHULTZ, T. Capital formation by education. The Journal of Political Economy, v. 68, n. 6, p. 571–583, 1960.

SCHULTZ, T. Investment in Human Capital. American Economic Review, v.51, n. 1, p. 1-17, 1961.

SMITH, J. A. Critical Survey of Empirical Methods for Evaluating Active Labor Market Policies, Schweizerische Zeitschrift fr Volkswirtschaft und Statistik, v. 36, n. 3, p. 1-22, 2000.

SOARES, T. M. *et al.* Escola de Tempo Integral: resultados do projeto na proficiência dos alunos do Ensino Fundamental das escolas públicas da rede estadual de Minas Gerais. Revista Ensaio: Avaliação e Políticas Públicas em Educação, v. 22, n. 82, p. 111-130, 2014.

TODOS PELA EDUCAÇÃO. Anuário brasileiro da educação básica. São Paulo, SP: Moderna, 2015.

UNICEF – FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA. Caminhos do direito de aprender. Boas práticas de 26 municípios que melhoraram a qualidade da educação. Brasília, 1º ed, Vol. 66, 2010.

UNICEF. Tendências para a educação integral. São Paulo: Fundação Itaú Social – CENPEC, 2011.

VELOSO, F. 15 Anos de Avanços na Educação no Brasil: Onde Estamos? In: VELOSO, F. et al (Org.). Educação Básica no Brasil: construindo o país do futuro. Rio de Janeiro: Editora Elsevier, 2009.

XERXENEVSKY, L. L. Programa Mais Educação: avaliação do impacto da educação integral no desempenho de alunos no Rio Grande do Sul. Dissertação (Mestrado em Economia do Desenvolvimento). Pontifícia Universidade Católica do Rio Grande do Sul. Porto Alegre, 2012.

ANEXOS

Tabela A1 - Resultados do Probit para participação no Mais Educação, 5º ano, usando dados pré-tratamento de 2009 (amostra 2)

Variáveis	Coefficientes
População	0,000*** (0,000)
Saneamento	0,268*** (0,083)
Lab_Ciências	0,115*** (0,044)
Biblioteca	0,267*** (0,030)
Quadra	-0,072** (0,030)
Nu_Salas	-0,008** (0,004)
Nu_Computadores	-0,013*** (0,002)
Duração_Média	0,000 (0,000)
Nu_Turmas	0,021*** (0,002)
Media_5º_LP	0,010*** (0,010)
Media_5º_MT	-0,027*** (0,002)
Tx_Aprovação	-0,013*** (0,001)
Tx_Abandono	0,012*** (0,005)
Número de Observações	14635
LR chi2	2080,40
Prob>chi2	0,00
Pseudo R2	0,1527

Fonte: Resultados da Pesquisa

Nota: Erros padrão em parênteses. *p<0,10, **p<0,05, ***p<0,01

Tabela A2 - Resultados do Probit para participação no Mais Educação, 9º ano, usando dados pré-tratamento de 2009 (amostra 2)

Variáveis	Coefficientes
População	0,000*** (0,000)
IDHM2000	4,587*** (0,235)
IDEB	-1,106 (0,043)
Saneamento	-0,027 (0,108)
Nu_Salas	-0,005 (0,004)
Nu_Computadores	-0,005*** (0,002)
Ens_Médio	-0,499*** (0,035)
Duração_Média	-0,002*** (0,000)
Nu_Turmas	0,019*** (0,002)
Media_9º_LP	0,032*** (0,002)
Media_9º_MT	-0,027*** (0,002)
Tx_Aprovação	0,009*** (0,002)
Tx_Abandono	0,008*** (0,003)
Número de Observações	11579
LR chi2	2896,57
Prob>chi2	0,00
Pseudo R2	0,2546

Fonte: Resultados da Pesquisa

Nota: Erros padrão em parênteses. *p<0,10, **p<0,05, ***p<0,01

Tabela A3 - Resultados do Probit para participação no Mais Educação, 5º ano, usando dados pré-tratamento de 2011 (amostra 3)

Variáveis	Coefficientes
População	0,000** (0,000)
IDEB	-0,210*** (0,039)
IDHM2010	-5,152*** (0,191)
Saneamento	-0,084 (0,051)
Lab_Ciências	0,101** (0,042)
Quadra	-0,111*** (0,025)
Nu_Salas	-0,006*** (0,010)
Nu_Computadores	0,000 (0,000)
Nu_Funcionários	0,007*** (0,000)
Duração_Média	0,000** (0,000)
Ens_Médio	-0,373*** (0,041)
Nu_Turmas	0,021*** (0,002)
Media_5º_LP	0,001 (0,001)
Media_5º_MT	-0,010*** (0,001)
Tx_Aprovação	-0,000 (0,002)
Tx_Abandono	-0,006 (0,004)
Número de Observações	18704
LR chi2	6434,94
Prob>chi2	0,00
Pseudo R2	0,2646

Fonte: Resultados da Pesquisa

Nota: Erros padrão em parênteses. *p<0,10, **p<0,05, ***p<0,01

Tabela A4 - Resultados do Probit para participação no Mais Educação, 9º ano, usando dados pré-tratamento de 2011 (amostra 3)

Variáveis	Coefficientes
IDEB	-0,604*** (0,035)
IDHM2000	-4,941*** (0,210)
Saneamento	0,053 (0,063)
Quadra	-0,151*** (0,030)
Nu_Computadores	0,000 (0,000)
Nu_Salas	0,005*** (0,002)
Ens_Médio	-0,368*** (0,027)
Duração_Média	-0,002*** (0,000)
Media_9º_LP	0,007*** (0,002)
Media_9º_MT	-0,005*** (0,001)
Tx_Aprovação	0,019*** (0,002)
Tx_Abandono	0,003 (0,003)
Número de Observações	14630
LR chi2	5277,02
Prob>chi2	0,00
Pseudo R2	0,2806

Fonte: Resultados da Pesquisa

Nota: Erros padrão em parênteses. *p<0,10, **p<0,05, ***p<0,01

Tabela A5 - Testes de Média entre escolas do Mais Educação e escolas não participantes, antes e depois do pareamento, 5º ano (amostra 2)

Variável	Amostra	Média		% viés	% viés reduzido	Teste t	
		Trat.	Cont.			t	p> t
População	Não pareada	8,1e ⁰⁸	7,5e ⁰⁵	3,1		1,29	0,197
	Pareada	8,1e ⁰⁵	8,5e ⁰⁵	-2,0	35,6	-0,67	0,501
Saneamento	Não pareada	0,973	0,978	-3,6		-1,74	0,082
	Pareada	0,973	0,965	4,9	-34,6	1,56	0,119
Lab_Ciências	Não pareada	0,129	0,111	5,5		2,57	0,010
	Pareada	0,129	0,133	-1,3	76,4	-0,45	0,656
Biblioteca	Não pareada	0,604	0,554	10,1		4,61	0,000
	Pareada	0,603	0,591	2,5	75,1	0,90	0,366
Quadra	Não pareada	0,539	0,635	-19,6		-9,13	0,000
	Pareada	0,538	0,528	2,1	89,5	0,73	0,468
Nu_Salas	Não pareada	11,458	11,087	6,4		3,02	0,003
	Pareada	11,449	11,279	2,9	54,0	0,97	0,332
Nu_Computadores	Não pareada	9,991	12,438	-22,3		-9,68	0,000
	Pareada	9,991	10,017	-0,2	98,9	-0,10	0,922
Duração_Média	Não pareada	255,73	264,51	-25,2		-11,34	0,000
	Pareada	255,75	254,46	3,7	85,3	1,49	0,137
Nu_Turmas	Não pareada	23,708	20,96	23,6		11,30	0,000
	Pareada	23,643	23,254	3,3	85,8	1,09	0,277
Media_5º_LP	Não pareada	178,07	192,4	-80,3		-34,80	0,000
	Pareada	178,09	177,19	5,0	93,8	1,82	0,069
Media_5º_MT	Não pareada	196,16	215,27	-90,8		-39,16	0,000
	Pareada	196,19	195,18	4,8	94,7	1,81	0,070
Tx_Aprovação	Não pareada	86,906	92,536	-57,2		-28,35	0,000
	Pareada	86,944	87,416	-4,8	91,6	-1,50	0,133
Tx_Abandono	Não pareada	2,467	0,907	44,6		24,14	0,000
	Pareada	2,437	2,20	6,8	84,8	2,08	0,038
R de Rubin	B	12,7					
	R	0,91					
Tratados	Amostra pareada	2574					
Controle	Amostra pareada	5642					

Fonte: Resultados da Pesquisa

Tabela A6 - Testes de Média entre escolas do Mais Educação e escolas não participantes, antes e depois do pareamento, 9º ano (amostra 2)

Variável	Amostra	Média		% viés	% viés reduzido	Teste t	
		Trat.	Cont.			t	p> t
População	Não pareada	8,7e ⁰⁵	9,5e ⁰⁵	-3,5		-1,31	0,192
	Pareada	8,7e ⁰⁵	7,6e ⁰⁵	4,8	-38,8	1,90	0,058
IDHM2000	Não pareada	0,626	0,631	-5,2		-2,07	0,038
	Pareada	0,626	0,619	8,7	-67,2	2,75	0,006
IDEB	Não pareada	3,462	4,260	-108,1		-45,30	0,000
	Pareada	3,462	3,438	3,3	97,0	1,06	0,291
Saneamento	Não pareada	0,979	0,979	-0,1		-0,05	0,957
	Pareada	0,979	0,980	-0,7	-446,9	-0,23	0,815
Nu_Salas	Não pareada	13,899	13,002	14,5		6,37	0,000
	Pareada	13,9	13,84	0,8	94,2	0,25	0,809
Nu_Computadores	Não pareada	16,059	18,555	-19,2		-7,63	0,000
	Pareada	16,061	15,681	2,9	84,8	1,10	0,271
Ens_Médio	Não pareada	0,418	0,553	-27,2		11,53	0,000
	Pareada	0,419	0,399	4,1	85,1	1,37	0,172
Duração_Média	Não pareada	259,25	268,48	-32,1		-12,92	0,000
	Pareada	259,26	258,18	3,8	88,3	1,42	0,157
Nu_Turmas	Não pareada	30,066	26,675	26,9		11,66	0,000
	Pareada	30,068	29,449	4,9	81,7	1,47	0,142
Media_9º_LP	Não pareada	237,34	246,12	-50,8		-20,93	0,000
	Pareada	237,33	237,28	0,3	99,5	0,09	0,931
Media_9º_MT	Não pareada	238,62	251,43	-69,5		-27,86	0,000
	Pareada	238,62	238,77	-0,8	98,8	-0,28	0,777
Tx_Aprovação	Não pareada	79,718	86,705	-57,5		-26,46	0,000
	Pareada	79,721	80,032	-2,6	95,6	-0,78	0,438
Tx_Abandono	Não pareada	5,977	3,066	48,7		23,52	0,000
	Pareada	5,977	5,780	3,3	93,2	0,91	0,362
R de Rubin	B	14,1					
	R	0,68					
Tratados	Amostra pareada	2240					
Controle	Amostra pareada	3930					

Fonte: Resultados da Pesquisa

Tabela A7 - Testes de Média entre escolas do Mais Educação e escolas não participantes, antes e depois do pareamento, 5º ano (amostra 3)

Variável	Amostra	Média		% viés	% viés reduzido	Teste t	
		Trat.	Cont.			t	p> t
População	Não pareada	2,4e ⁰⁵	7,8e ⁰⁵	-31,6	96,4	-18,70	0,000
	Pareada	2,4e ⁰⁵	2,2e ⁰⁵	1,1		1,18	0,240
IDEB	Não pareada	4,269	5,422	-119,5	99,9	-78,60	0,000
	Pareada	4,270	4,268	0,2		0,10	0,922
IDHM2010	Não pareada	0,646	0,732	-118,2	98,7	-79,28	0,000
	Pareada	0,646	0,645	1,6		0,83	0,407
Saneamento	Não pareada	0,916	0,978	-28,1	89,1	-20,14	0,000
	Pareada	0,916	0,923	-3,1		-1,44	0,150
Lab_Ciências	Não pareada	0,057	0,109	-18,7	96,1	-11,77	0,000
	Pareada	0,057	0,059	-0,7		-0,49	0,625
Quadra	Não pareada	0,337	0,637	-62,9	95,7	-41,07	0,000
	Pareada	0,337	0,324	2,7		1,58	0,113
Nu_Salas	Não pareada	9,579	11,572	-29,6	98,5	-19,23	0,000
	Pareada	9,576	9,546	0,4		0,26	0,792
Nu_Computadores	Não pareada	13,529	19,36	-20,0	98,0	-13,00	0,000
	Pareada	13,53	13,414	0,4		0,24	0,807
Nu_Funcionários	Não pareada	42,341	45,465	-12,6	78,3	-8,15	0,000
	Pareada	42,297	41,618	2,7		1,36	0,175
Ens_Médio	Não pareada	0,062	0,120	-20,3	98,4	-12,73	0,000
	Pareada	0,062	0,062	0,3		0,22	0,829
Duração_Média	Não pareada	251,98	267,24	-40,1	98,1	-25,50	0,000
	Pareada	251,99	251,69	0,8		0,53	0,595
Media_5º_LP	Não pareada	175,36	198,55	-113,9	99,0	-74,81	0,000
	Pareada	175,38	175,62	-1,2		-0,66	0,506
Media_5º_MT	Não pareada	192,65	219,93	-115,7	98,6	-75,82	0,000
	Pareada	192,68	193,07	-1,7		-0,96	0,336
Tx_Aprovação	Não pareada	86,681	93,724	-68,3	92,8	-47,27	0,000
	Pareada	86,694	87204	-5,0		-2,46	0,014
Tx_Abandono	Não pareada	2,459	0,811	49,2	96,9	34,55	0,000
	Pareada	2,456	2,404	1,5		0,74	0,458
R de Rubin	B	11,0					
	R	0,94					
Tratados	Amostra pareada	6624					
Controle	Amostra pareada	7190					

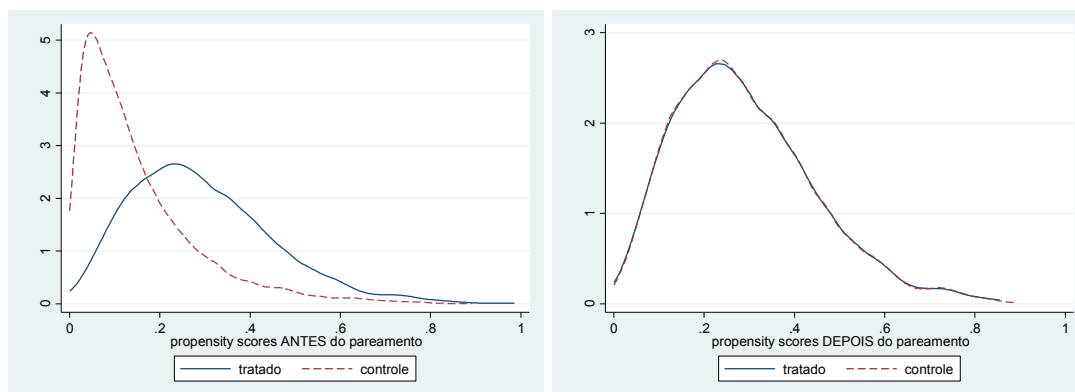
Fonte: Resultados da Pesquisa

Tabela A8 - Testes de Média entre escolas do Mais Educação e escolas não participantes, antes e depois do pareamento, 9º ano (amostra 3)

Variável	Amostra	Média		% viés	% viés reduzido	Teste t	
		Trat.	Cont.			t	p> t
IDEB	Não pareada	3,571	4,396	-105,7	97,3	-60,46	0,000
	Pareada	3,571	3,549	2,8		1,39	0,164
IDHM2000	Não pareada	0,647	0,735	-121,1	99,4	0,34	0,000
	Pareada	0,647	0,647	0,07		0,34	0,734
Saneamento	Não pareada	0,930	0,980	-24,2	90,0	15,23	0,000
	Pareada	0,930	0,925	2,4		0,96	0,338
Quadra	Não pareada	0,462	0,807	-76,7	98,5	-45,74	0,000
	Pareada	0,462	0,467	-1,1		-0,52	0,605
Nu_Salas	Não pareada	11,183	13,451	-29,2	95,9	-16,87	0,000
	Pareada	11,183	11,275	-1,2		-0,58	0,565
Nu_Computadores	Não pareada	16,512	24,223	-23,8	98,2	-13,66	0,000
	Pareada	16,512	16,373	0,4		0,26	0,794
Ens_Médio	Não pareada	0,285	0,549	-55,7	100,0	-31,51	0,000
	Pareada	0,285	0,285	-0,0		-0,01	0,989
Duração_Média	Não pareada	254,35	271,67	-53,8	97,9	-29,73	0,000
	Pareada	254,35	254,72	-1,1		-0,72	0,471
Media_9º_LP	Não pareada	228,4	246,77	-97,5	97,8	-56,32	0,000
	Pareada	228,4	228,8	-2,1		-1,02	0,306
Media_9º_MT	Não pareada	235,49	255,26	-93,3	95,1	-53,70	0,000
	Pareada	235,49	236,46	-4,6		-2,27	0,023
Tx_Aprovação	Não pareada	86,795	87,43	-5,9	-7,6	-3,40	0,001
	Pareada	86,795	86,111	6,3		3,05	0,002
Tx_Abandono	Não pareada	5,373	2,950	43,3	87,2	26,17	0,000
	Pareada	5,373	5,683	-5,5		-2,32	0,020
R de Rubin	B	11,3					
	R	0,94					
Tratados	Amostra pareada	5011					
Controle	Amostra pareada	5346					

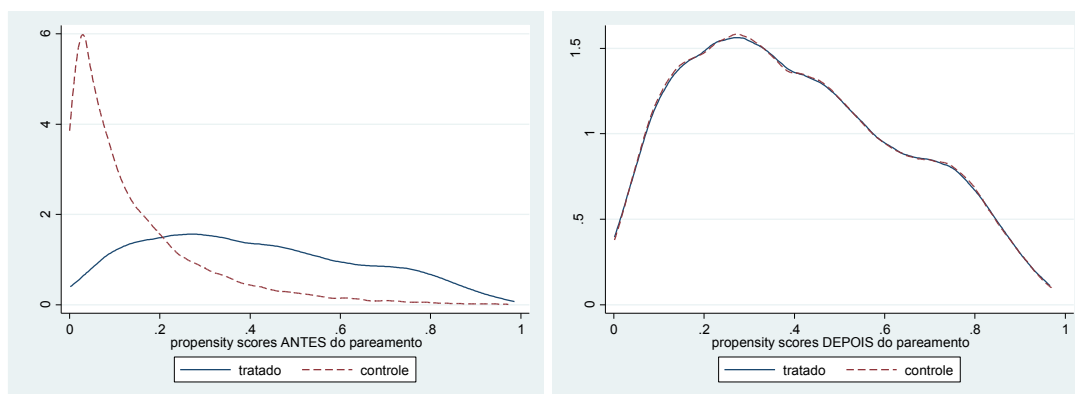
Fonte: Resultados da Pesquisa

Figura A1 - Distribuição dos escores de propensão estimados para as escolas tratadas e controle de 5º ano, antes e depois do pareamento, amostra 2



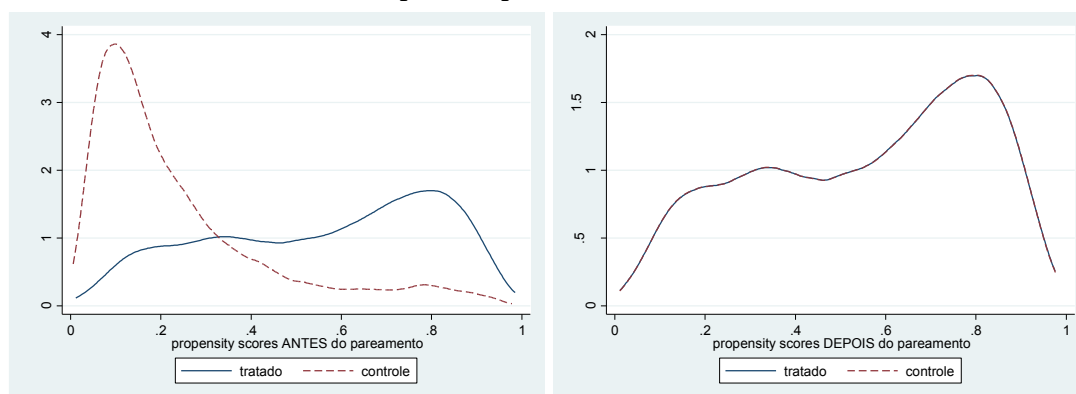
Fonte: Resultados da Pesquisa

Figura A2 - Distribuição dos escores de propensão estimados para as escolas tratadas e controle de 9º ano, antes e depois do pareamento, amostra 2



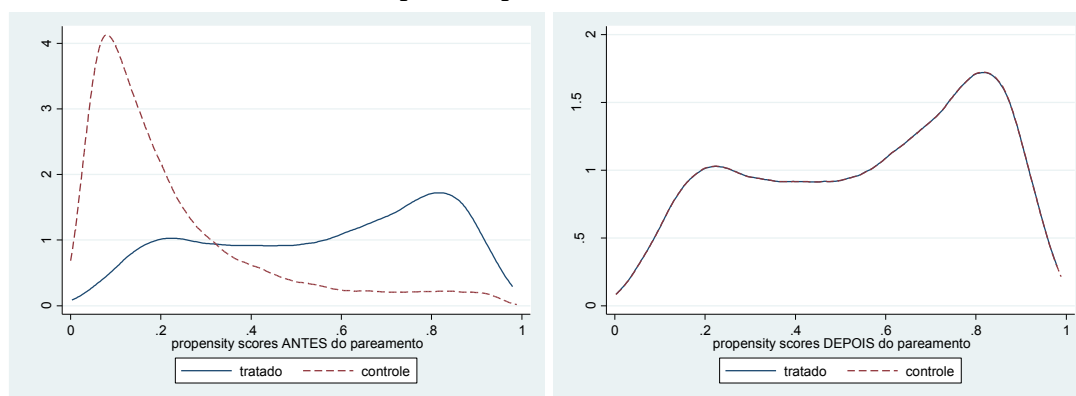
Fonte: Resultados da Pesquisa

Figura A3 - Distribuição dos escores de propensão estimados para as escolas tratadas e controle de 5º ano, antes e depois do pareamento, amostra 3



Fonte: Resultados da Pesquisa

Figura A4 - Distribuição dos escores de propensão estimados para as escolas tratadas e controle de 9º ano, antes e depois do pareamento, amostra 3



Fonte: Resultados da Pesquisa