



Universidade Federal da Paraíba
Centro de Ciências Sociais Aplicadas
Programa de Pós-Graduação em Economia

Avaliação do impacto da greve universitária federal de 2012 sobre a nota média das IES no ENADE

Trabalho apresentado ao Programa de Pós-Graduação
em Economia Universidade Federal da Paraíba, em
cumprimento às exigências da disciplina de Avaliação de
Políticas Públicas, ministrada pelo professor Dr. Aléssio
Tony C. Almeida.

Pedro Milreu Cunha,
Rafael de Sousa Araújo

João Pessoa - PB
2020

Sumário

1	INTRODUÇÃO	4
2	REVISÃO DE LITERATURA	7
3	METODOLOGIA	9
3.1	Modelo empírico	9
3.1.1	Modelo de dados em painel	9
3.1.2	<i>Propensity Score Matching (PSM)</i>	10
3.1.3	Estimação pelo método de Diferenças-em-Diferenças	11
3.1.3.1	Modelo estimado	13
3.2	Dados	14
4	RESULTADOS	15
4.1	Análise de robustez	15
4.1.1	Propensity score matching	15
4.1.1.1	Hipótese do suporte comum	15
4.2	Balanceamento das covariadas	19
4.2.1	Diferenças em diferenças	21
4.2.1.1	Tendências paralelas - inspeção gráfica	21
4.2.1.2	Teste de placebo	21
4.3	Análise de impacto	24
5	CONCLUSÕES	29
	REFERÊNCIAS	30
6	APÊNDICE	32
6.1	Lista de instituições de ensino superior utilizadas na estimação . . .	32
6.2	Balanceamento das covariadas - Testes de placebo	34

Resumo

O presente trabalho teve como propósito verificar o impacto, se é que ele existiu, da greve universitária federal de 2012 sobre a nota média das instituições afetadas na prova do ENADE em 2012 e 2013. Para tanto foram utilizados dois painéis longitudinais com informações sobre as IES para os ciclos 2009-2012 e 2010-2013 e foi estimado um modelo de Diferenças-em-Diferenças com efeitos fixos individuais e aplicação da técnica de *Propensity Score Matching* (PSM) para pareamento prévio dos grupos de tratado e controle, garantindo maior robustez aos resultados. Encontrou-se que a greve não teve efeito contemporâneo, tendo, por outro lado, impactado negativamente de maneira significativa o desempenho médio dos discentes no ENADE 2013, com uma redução marginal média de 0,3798 na nota média para as unidades afetadas no período pós-intervenção (ciclo 2010-2013).

Palavras-chave: Greve universitária federal. PSM. Entropia. Diff-in-Diff. Impacto.

1 Introdução

Ao longo da história da humanidade, diversos movimentos sociais têm acontecido em prol da conquista de direitos e da busca pela equidade e pelo respeito. No caso das classes trabalhadoras, essas repercussões se fortaleceram e se intensificaram com a chegada dos sindicatos.

Os sindicatos têm papel primordial no que diz respeito à luta constante entre classes. É a partir dessas organizações e de seus movimentos que muitos direitos têm sido garantidos. Nesse contexto, quando não há um acordo entre a classe trabalhadora e classe empregadora, ou ainda quando determinado acordo firmado entre ambas as partes não é posto em prática, pode vir a ocorrer um movimento grevista.

A greve, de modo geral, traz consigo efeitos ambíguos. Esses movimentos, por um lado, podem trazer retornos positivos aos membros da associação envolvida e àqueles que dependem direta ou indiretamente do produto/serviço, uma vez que um melhor reconhecimento ou melhor qualidade de trabalho pode estar associada a um bem final de qualidade superior; por outro lado - e analogamente - podem provocar efeitos negativos enquanto acontecem.

Com base nisso, pode-se destacar os casos das greves de professores no âmbito das Instituições de Ensino Superior (IES) públicas do Brasil. A paralisação de professores em busca do atendimento de inúmeras demandas como salário e benefícios inerentes à carreira são responsáveis por uma drástica mudança no calendário das IES, comprometendo não só os projetos individuais dos alunos, mas repercutindo na própria configuração posterior das aulas, na execução de projetos acadêmicos e de extensão e no desempenho em exames importantes, seja de caráter acadêmico ou profissional. Além disso, também pode ser percebido um efeito evasão, onde parte dos alunos (por terem uma condição favorável e por não quererem perder tempo) acabam mudando de IES (BELOT; WEBBINK, 2010).

Nesse sentido, toda paralisação nas atividades desempenhadas pelos docentes (espinha dorsal das atividades universitárias) tem o condão de afetar o planejamento pedagógico e estender o calendário acadêmico, não sendo raras as situações em que surge a necessidade de se promover uma condução mais rápida das atividades, de modo a se manter com a alteração no planejamento das ações da IES.

De acordo com Greenglass et al. (2002) e Wickens et al. (2006), as greves de professores do ensino superior, enquanto acontecem, provocam nos alunos uma quebra de rotina, tornando-os mais improdutivos. Além disso, esses prejuízos entram em descompasso com relação a prazos de formatura, ofertas de emprego, viagens e outras seleções posteriores.

Como ressaltado anteriormente, é importante consignar que as alterações no calendário e as eventuais mudanças no plano pedagógico (após o período de greve) também

afetam diretamente um importante momento da vida universitária, qual seja, a avaliação dos cursos superiores e do egresso por meio de exames como o Exame Nacional de Desempenho dos Estudantes (ENADE). O referido exame representa uma oportunidade de se conhecer a qualidade do ensino ofertado e de que modo o mesmo preparou adequadamente o profissional para o exercício de suas atividades. Cabe destacar que ele é aplicado em ciclos que têm três anos de duração, de modo a dividir a avaliação em três grandes áreas de conhecimento, que serão avaliadas individualmente em cada um dos anos de um ciclo.

Não é muito trivial inferir que algum serviço teve uma melhora ou uma piora no decorrer dos anos sem que se faça alguma avaliação. Para Januzzi (2016, p. 46), o termo avaliação refere-se ao esforço analítico de produzir informação e conhecimento sobre determinada atividade ou ocorrência [...] com a finalidade de aprimorar os objetivos e a gestão das intervenções, bem como analisar os efeitos e seus custos tanto no curto quanto a longo prazo.

Nesse caso, o ENADE serve justamente para isso, ou seja, para avaliar a qualidade dos cursos de ensino superior e, com base nessa avaliação, a instituição de ensino deve realizar um planejamento adequado de modo que atue na promoção de uma melhor qualidade de ensino dentro dos seus limites.

Assim, contextualizada a importância do desempenho dos estudantes no ENADE e o possível efeito da greve dos docentes sobre esse exame, esta pesquisa pretende avaliar qual o impacto da greve universitária federal de 2012 sobre o desempenho médio das IES que aderiram ao movimento no ENADE nos anos de 2012 e de 2013.

A escolha do período se deu de forma a tentar captar tanto os possíveis efeitos contemporâneos de uma greve, o quê pode ser analisado no ciclo 2009-2012, quanto os efeitos de longo-prazo, que foram estimado utilizando o ciclo 2009-2013.

Um outro fator importante foi a escolha de anos pré-intervenção nos quais fossem avaliadas as mesmas áreas do que seus respectivos pares pós-intervenção. Isso é importante para garantir que a heterogeneidades dos diversos corpos discentes existentes no cursos de educação superior não torne as análises enviesadas.

Além de buscar esclarecer sobre os principais efeitos da greve dos professores, este artigo também tem fundamental importância no que se refere a trabalhar com uma metodologia alternativa, que considera não somente o efeito da greve no ano corrente, mas também no ano posterior à mesma. Esse tipo de análise é defendida por Baker (2013). Para o autor, estudos que consideram apenas um corte transversal são muito suscetíveis a críticas.

Para verificar a possível existência de um efeito da greve, fez-se uso de técnicas de pareamento, visando tornar os grupos de controle e tratamento o mais semelhantes possível; além disso, o formato dos dados em painel possibilitou o uso de efeitos fixos individuais e, por fim, a disponibilidade de dados temporais compatíveis, permitiu a utilização do

estimador de Diferenças-em-Diferenças, de modo a ter realizado uma análise de impacto robusta, que apontou para a existência de um efeito causal negativo de longo-prazo.

A presente pesquisa considera apenas as IES que de fato aderiram à greve de 2012 como unidades de tratamento, de modo a fazer com que a variável de tratamento reflita a ocorrência de tratamento e não apenas a possibilidade de tratamento. Segue, portanto, que quaisquer instituições de ensino que não aderiram à greve são consideradas como unidades de controle.

Entre os principais resultados, pode-se mencionar que, embora não tenha sido percebido nenhum efeito significativo sobre a nota do ENADE em 2012, estimou-se que a greve causou uma redução de 0,3798 na nota média das universidades afetadas no exame em 2013, considerando apenas o período pós-intervenção.

O restante do trabalho encontra-se dividido em Revisão de Literatura, seção na qual é feita uma breve revisão da literatura sobre os efeitos de greves de docente; Metodologia, onde são apresentados o modelo empírico utilizado, bem como os dados e suas estatísticas descritivas; Resultados, seção na qual se realizou a apresentação e discussão dos resultados, traçando paralelos com a literatura e indicando possíveis razões para os efeitos observados; Considerações finais, onde é sumarizado o que foi feito no trabalho, quais os principais resultados, suas aplicações e possíveis caminhos futuros. Além disso, o trabalho possui ainda uma seção de Referências e um Apêndice.

2 Revisão de Literatura

As greves, de modo geral, provocam mudanças no relacionamento entre os agentes envolvidos. É o momento em que a classe desfavorecida dá voz às suas ideias de justiça. Por outro lado, a depender da essencialidade do bem ou do serviço prestado por esses, envolve outros agentes.

De acordo com Belot e Webbink (2010), as greves de professores geralmente são motivadas por questões relacionadas à remuneração, e a importância dessa garantia está associada à qualidade do ambiente escolar. Todavia, não somente as questões contratuais importam, além disso – e por exemplo –, pode-se mencionar as próprias características da população estudantil.

No tocante a isso, Hoxby (1996) faz um estudo relacionado às greves de professores e seus efeitos no ambiente escolar e nas condições contratuais. Segundo o autor, os resultados indicam que os efeitos positivos de uma greve estão mais relacionados ao corpo docente do que ao discente, uma vez que as negociações visam mais um salário coerente e a proporção professor-aluno. Já no que diz respeito ao desempenho dos alunos, a percepção é de que, pelo menos no curto prazo pós-greve, o desempenho dos alunos sofra um efeito negativo.

Na contramão, Zwerling (2008) também faz um estudo para avaliar os efeitos das greves de professores da Pensilvânia no desempenho acadêmico dos alunos, mas não encontra relação significativa, inclusive para a taxa de frequência. De acordo com o autor, o baixo desempenho está mais relacionado às questões demográficas e socioeconômicas. Esse resultado também pode estar relacionado ao fato de que, por lei, o período de paralisação deve ser compensado posteriormente. Sendo assim, os efeitos negativos sobre o ensino são minimizados.

Lovenheim (2009) também fez um estudo semelhante para três distritos dos Estados Unidos, sendo eles: Iowa, Indiana e Minnesota. Seu objetivo foi analisar o efeito dos sindicatos de professores na alocação de recursos do distrito escolar e no desempenho dos alunos daquelas regiões. Para isso, utilizou-se de dados do Censo escolar realizado pelo Governo e referentes aos anos de 1970 e 1991. Foram utilizados dados de certificação eleitoral sindical, salário real e emprego dos professores em tempo integral, proporção de alunos por professor, taxa de abandono e despesas operacionais reais por aluno. Os principais resultados encontrados foram que os sindicatos não possuem efeito algum sobre o salário dos professores. Por outro lado, afeta negativamente a contratação de mais professores em tempo integral. Quanto à proporção professor-aluno, o efeito sobre as despesas por aluno é positivo apenas no curto prazo. Também não foi percebido nenhuma evidência clara sobre o impacto da sindicalização sobre a taxa de abandono.

Belot e Webbink (2010) ressaltam sobre a ausência de estudos que avaliem os

efeitos de longo prazo das greves de professores no ambiente escolar. Então, no seu estudo, eles objetivaram avaliar isso, além dos efeitos de curto e médio prazo. Para isso, utilizaram o método de diferenças em diferenças e seus principais resultados foram que as greves diminuíram o desempenho educacional dos estudantes, principalmente no tocante à repetência devido às reprovações, com o atraso médio de um semestre. Além disso, que também foi percebido um efeito migração entre cursos de modalidades semelhantes de ensino superior.

Wills (2014) fez um estudo semelhante ao analisar a greve de professores ocorrida em 2007 na África do Sul, onde avaliou os efeitos dessa na aprendizagem dos alunos que cursavam a sexta série do ensino fundamental. Para isso, ele utilizou um modelo cross-section com efeitos fixos semelhante ao de dados em painel, onde são considerados que os alunos e as escolas são invariantes, sendo que o que muda nesse caso não é o tempo, e sim a notas dos alunos em cada disciplina em um ano específico. Segundo o autor, um problema desse método é que ele elimina possíveis viés relacionados aos alunos e às escolas, mas não garante o mesmo quanto aos professores. É importante ressaltar que nem todos os professores aderiram a greve. Com base nisso, o autor evidencia que os alunos tiveram desempenho inferior naquelas disciplinas em que os professores aderiram à greve, e ainda que esse movimento teve efeito principalmente sobre os alunos que já tinham um baixo desempenho e que estudavam em escolas rurais (quando comparadas às escolas urbanas).

Um estudo mais recente feito por Anegues, Porto Júnior e Figueiredo (2017) avaliou o impacto da greve de professores da Universidade Federal da Paraíba sobre o rendimento dos alunos dessa instituição. Os autores utilizaram um modelo de dados em painel balanceado por efeitos fixos, bem como do método de regressões quantílicas para dados em painel. Os principais resultados foram que, na média, durante o período afetado pela greve, os alunos tiveram um rendimento inferior, principalmente sobre aqueles que já não tinham um bom desempenho anteriormente.

Assim, portanto, pode-se perceber que não há um consenso sobre os efeitos da greve na produção acadêmica, principalmente quanto ao desempenho de alunos em suas atividades. No entanto, pelo diagnóstico apresentado pela literatura vigente, um dos resultados que aparenta mais consiste é a ideia de que as greves possuem impacto negativo maior para com aqueles alunos que já possuíam um menor desempenho acadêmico. Por fim, também é importante ressaltar sobre as divergências legais por região. A depender do país, os professores não possuem o direito de aderirem a greves.

Além disso, existem países onde o período em atraso devido a esses movimentos de paralisação devem ser compensados em igual importância à posteriori, ou seja, para esse caso, espera-se que os efeitos de uma greve sejam minimizados. Em resumo, resultados divergentes podem ter motivos que não somente dizem respeito à paralisação da classe de professores, mas sim de variáveis endógenas não controladas (ou não controláveis).

Na próxima seção são apresentados a metodologia e os dados utilizados na análise.

3 Metodologia

3.1 Modelo empírico

3.1.1 Modelo de dados em painel

O termo “dados em painel” diz respeito à forma como os dados são organizados, ou seja, observações em cross-section de indivíduos, países, firmas ou qualquer outro tipo de unidade organizacional ao longo de dois ou mais períodos de tempo. Diferentemente do modelo cross-section ou das séries temporais, os dados em painel permitem que fatores fixos no tempo ou no espaço possam ser incorporados aos resultados, diminuindo possíveis vieses de variável omitida. A grande vantagem desse método é que os primeiros permitem modelar diferenças entre as unidades observadas. Além disso, outras vantagens seriam, por exemplo, que nos dados em painel há mais graus de liberdade, menor colinearidade entre as variáveis, maior eficiência dos estimadores e maior variabilidade, devido a maior quantidade de informações (BALTAGIA, 2005).

De especial importância para o presente artigo é a utilização dos **efeitos fixos**. A intenção por trás do seu uso é captar as características invariantes no tempo de cada unidade observacional, permitindo, assim, um controle da heterogeneidade amostral.

Com o uso de efeitos fixos, a equação a ser estimada utilizando dados em painel teria o seguinte formato

$$y_{it} = \gamma_i + \mathbf{X}\beta + \mu \quad (3.1.1)$$

onde i representa a unidade individual de análise, t o tempo, γ_i os efeitos fixos individuais de cada unidade, \mathbf{X} é uma matriz de covariadas e μ é um vetor de distúrbios.

Apesar das vantagens citadas acima, o modelo também tem algumas restrições. Uma das dificuldades em se utilizar dados em painel é a coleta das informações, ou seja, a falta de cobertura completa acerca das variáveis de interesse. Há a possibilidade de erros de medição, que podem surgir devido a respostas falhas para perguntas mal formuladas, distorção de respostas e viés do entrevistador. Além disso, se existir algum tipo de dependência entre as unidades organizacionais e essa dependência não for incorporada ao modelo, isso poderá gerar inferências enganosas (BALTAGIA, 2005).

Uma característica importante de um painel que deve ser abordada diz respeito ao seu balanceamento. Diz-se que um painel é balanceado quando há observações para todas as unidades organizacionais ao longo de todos os períodos de tempo, de modo que para um painel com n indivíduos e T unidades de tempo, haverá $n \cdot T$ observações. Um painel não-balanceado é aquele para o qual nem todos os n indivíduos apresentam observações para todos os T período de tempo, de tal forma que o número de observações será $k < n \cdot T$.

Um importante questionamento que deve ser feito é qual a razão da ausência das observações. Há um padrão nos dados faltantes ou essa ausência ocorre de forma aleatória? Apenas o primeiro caso é problemático, pois ele pode causar viés na análise, caso esse padrão esteja relacionado ao erro idiossincrático, que é o erro que varia ao longo do tempo e das unidades organizacionais (HSIAO, 2014). Como os painéis utilizados no presente trabalho são todos balanceados, esse problema não gera questionamentos acerca da validade do artigo.

3.1.2 *Propensity Score Matching (PSM)*

A de impactos tem como propósito verificar se algum determinado experimento ou programa está atingindo os resultados esperados, além de permitir quantificar esse efeito, dando-lhe sinal e magnitude (CAMERON; TRIVEDI, 2005). Como é preciso fazer inferências acerca do que teria acontecido com um agente que foi tratado caso não tivesse sido, ou seja, qual seria o seu resultado em um "mundo paralelo" no qual a única diferença é a ausência de tratamento, é necessário construir o contrafactual. Tendo por base essa dita "especulação" (Caliendo e Kopeinig (2008)), torna-se clara a importância de se ter uma grande gama de informações acerca das variáveis de interesse para os dois grupos.

Para realizar um estudo de efeito causal, é necessário dispor de informações sobre os indivíduos, o tratamento e os resultados potenciais. A grande dificuldade em se estimar π , o efeito do tratamento, consiste na inexistência de dados sobre o resultado potencial na ausência de tratamento, visto que esse cenário não é observável. Assim, na equação (3.1.2), apenas $Y_i(1)$ é observável, ou seja, somente conseguimos obter o resultado da unidade tratada no cenário de tratamento:

$$\pi_i = Y_i(1) - Y_i(0) \quad (3.1.2)$$

O grande problema é que a variável de tratamento pode muito bem ser correlacionada com o termo de erro da equação a ser estimada por um modelo de regressão qualquer, o que tornaria os resultados enviesados e não-consistentes, podendo existir, inclusive, um viés relacionado ao problema de autoseleção (HECKMAN, 1979). Um dos métodos que se presta a resolver isso, utilizando o conjunto de dados observáveis, é justamente o *Propensity Score Matching* (PSM), proposto por Rosenbaum e Rubin (1983) e utilizado no presente artigo. O PSM tem como intenção resumir as características observáveis de cada indivíduo de cada grupo em um único *propensity score*, o que possibilita que o pareamento dos grupos seja feito, por algum algoritmo de minimização de distância ou "erro"¹, que cria um grupo de controle próximo ao grupo de tratamento em questão de seus atributos observáveis.

¹ O artigo fez uso da distância baseada nos PS calculados utilizando o modelo logit multinomial e dos pesos calculados pelo modelo de entropia.

De acordo com o exposto por Becker e Ichino (2002), ao fazer uso do PSM, há a possibilidade da correção do viés incorrido pela ausência de aleatoriedade na formação dos grupos ao torná-los o mais semelhante possível. Há um relativo consenso na literatura de avaliação de impacto no que tange ao fato do variável de tratamento satisfazer a hipótese de exogeneidade, o que implica que as diferenças entre os dois grupos são devidas apenas aos tratamentos (CALIENDO; KOPEINIG, 2008). Desse modo, podemos definir o *propensity score* como a chance condicional de um indivíduo receber o tratamento baseado em suas características **pré-tratamento** observáveis, como na equação 3.1.3

$$p(X) \equiv \text{Prob}(P|X) = \Phi(g(X_i)) \quad (3.1.3)$$

Em que T é a variável dummy de tratamento, assumindo valor 1 para as unidades tratadas e 0 para as demais; X é o vetor de características observáveis; $g(X_i)$ é a função que inclui as covariadas e Φ é uma função de distribuição acumulada. Se for feita a suposição de que Φ segue uma distribuição de probabilidade logística, $p(X)$ pode ser estimado por meio de um modelo com variável dependente binária ou de resposta qualitativa logit.

De acordo com Rosenbaum e Rubin (1983), se o *propensity score* for conhecido, o **efeito médio de tratamento sobre os tratados**, π_{ATT} pode ser estimado por meio da equação 3.1.4

$$\pi_{ATT} = E\{E[Y_{1i}|T_i = 1, P(X_i)] - E[Y_{0i}|T_i = 0, P(X_i)]\} \quad (3.1.4)$$

onde Y_{1i} e Y_{0i} são os resultados observados na situação de tratamento para os tratados e o contrafactual.

Entretanto, para que o cálculo de π_{ATT} seja válido é preciso que duas suposições sejam válidas: i) independência condicional, isto é, os fatores não-observados não devem afetar a participação no programa; ii) existência de suporte comum ou grande interseção entre a distribuição dos *propensity scores* nos dois grupos (KHANDKER; KOOLWAL; SAMAD, 2010, p. 55).

Por fim, como forma de garantir mais robustez na estimação dos *ps*, optou-se por realizar o cálculo de pesos fazendo uso, para além do modelo mais simples de pareamento logit, de um modelo de entropia. Mais informações podem ser encontradas em Hainmueller (2012), artigo que deu origem ao modelo .

Na próxima seção é abordado o estimador de Diferenças-em-Diferenças, seus pressupostos e suas vantagens.

3.1.3 Estimação pelo método de Diferenças-em-Diferenças

Conforme aponta Lechner (2010), o design de Diferenças-em-Diferenças tem uma longa história na análise empírica de causalidade, tanto dentro quanto fora da econometria. Intuitivamente, o modelo argumenta que deve ser feita a comparação entre os *outcomes*

do grupo de tratamento e controle antes e depois da intervenção, de modo que seja possível comparar a trajetória temporal deles. Sob a hipótese de identificação do modelo, ou seja, de que as trajetórias temporais de ambos os grupos são iguais (hipótese das trajetórias paralelas), o modelo de Diferenças-em-Diferenças (DD) permite controlar por fatores fixos no tempo, sendo esses observáveis ou não, e, assim, estimar o efeito causal do programa ou intervenção.

Para verificar a validade dessa hipótese, empregam-se algumas técnicas: uma primeira forma, mais ingênua e visual, é caso se tenha os dados, simplesmente plotar a trajetória temporal do outcome de interesse para ambos os grupos e verificar se são semelhantes; outra possibilidade é a realização de testes placebo, estimando modelos com diferentes outcomes, grupos de comparação ou localização temporal, com o intuito de verificar se a intervenção teve algum efeito sobre variáveis não relacionadas com o propósito do programa, ou variáveis mensuradas antes do período de início de intervenção. Por fim, uma última técnica utilizada é a realização de um processo de pré-tratamento dos dados, por meio do uso de *matching*, buscando deixar os dois grupos o mais semelhantes possível antes da estimação.

A ideia básica da estimação por DD se baseia nas quatro médias a seguir:

$$\begin{aligned} \delta_{DiD} = & [E(y_i|T_i = 1, t = 1) - E(y_i|T_i = 1, t = 0)] \\ & - [E(y_i|T_i = 0, t = 1) - E(y_i|T_i = 0, t = 0)] \end{aligned} \quad (3.1.5)$$

onde $T_i, t \in \{0, 1\}$ são variáveis binárias representando a variável de tratamento e a variável de tempo (pós-intervenção).

Veja que partindo de um modelo de regressão linear dado por

$$y_{it} = \alpha_0 + \alpha_1 \cdot T_i + \alpha_2 \cdot t + \alpha_3 \cdot T_i \cdot t + \mu_{it} \quad (3.1.6)$$

Note que o coeficiente α_3 representa o efeito do estimador de DD. Para observar isso, faça uso da tabela 1:

Tabela 1 – Estimador de Diferenças-em-Diferenças em modelos de regressão

	Tratamento ($T = 1$)	Controle ($T = 0$)	Diferença
Antes ($t = 0$)	$\alpha_0 + \alpha_1$	α_0	α_1
Depois ($t = 1$)	$\alpha_0 + \alpha_1 + \alpha_2 + \alpha_3$	$\alpha_0 + \alpha_2$	$\alpha_1 + \alpha_3$
Diferença	$\alpha_2 + \alpha_3$	α_2	α_3

onde as células da tabela são dadas pelos valores esperados de y condicional a T e t .

Notadamente, o modelo pode ser expandido de modo a incluir covariadas importantes e, em um cenário de dados em painel, efeitos fixos invariantes no tempo. Além disso, conforme já mencionado, pode-se fazer uso do estimador de Diferenças-em-Diferenças

Condicional, que é a combinação das técnicas de *propensity score matching* (PSM) com o DD, resultando na hipótese dada pela equação 3.1.7.

$$E[Y_{t=0}(0) - Y_{t=1}(0)|P(x), T = 1] = E[Y_{t=0}(0) - Y_{t=1}(0)|P(x), T = 0] \quad (3.1.7)$$

onde $t = 0$ e $t = 1$ referem-se, respectivamente, aos períodos pré e pós-intervenção; T é a dummy de tratamento; e $P(x)$ é algum *propensity score* arbitrário.

O modelo estimado no presente trabalho fez o uso do estimador de DD Condicional, bem como se aproveitou dos dados em painel para controlar efeitos fixos individuais, realizando também as estimações de placebo com vistas de tornar os resultados obtidos mais robustos. Ele é apresentado na subseção seguinte.

3.1.3.1 Modelo estimado

Combinando os dados em painel, com a técnica de PSM e a estimação pelo método de Diferenças-em-Diferenças, estimou-se a equação (3.1.8)

$$\begin{aligned} \ln \text{Nota Geral}_{it} = & \beta_0 \cdot \text{Tempo}_{it} + \beta_1 \cdot \text{Idade Média}_{it} \\ & + \beta_2 \cdot \text{Idade Média}^2_{it} + \beta_3 \cdot \text{Prop. brancos}_{it} \\ & + \beta_4 \cdot \text{Prop. casados}_{it} + \beta_5 \cdot \text{Prop. homens}_{it} \\ & + \beta_6 \cdot \text{Prop. Ensino Médio+}_{it} + \beta_7 \cdot \text{Prop. Renda até 3SM}_{it} \\ & + \gamma_i + \delta \cdot \text{Impacto}_{it} + \mu_{it}, \quad \mu \sim \mathcal{N}(0, \sigma^2) \end{aligned} \quad (3.1.8)$$

onde os subscritos i e t referem-se, respectivamente, a instituição de ensino superior i no tempo t , onde $t \in \{2009, 2010, 2012, 2013\}$. **Nota Geral** é a nota média geral da IES na prova do ENADE; **Tempo** é uma variável binária que assume valor 1 se $t \geq 2012$ e 0 caso contrário; **Idade Média** é a idade média dos discentes da amostra para cada IES; **Prop. brancos**, **Prop. casados**, **Prop. homens**, **Prop. Ensino Médio+**, **Prop. Renda até 3SM** são as proporções nas amostras de cada IES de, respectivamente: discentes brancos; discentes casados; discentes do sexo masculino; discentes cujas mães têm ensino médio ou mais e discentes cuja renda familiar é de até três (3) salários mínimos. γ_i representa o efeito fixo da IES i ; Por fim, a variável **Impacto** é o produto da variável **Tempo** e da variável **Tratamento** (assume valor 1 para as unidades tratadas e 0 para o controle) e, no modelo de Diferenças-em-Diferenças, **é o estimador do impacto do tratamento** e μ é o termo de erro distribuído normalmente com média 0 e variância constante.

É muito importante lembrar também que a regressão foi estimada utilizando os pesos amostrais obtidos à partir da utilização do *matching*, com o propósito de deixar os grupos de tratamento e controle mais similares.

Esperou-se que β_1, β_6 tivessem sinais positivos, enquanto que β_2, β_7 e δ fossem negativos. Ou seja, esperava-se que discentes mais velhos tivessem, em média, notas mais

altas, muito embora esse retorno fosse decrescente no tempo. Além disso, esperava-se que a escolaridade da mãe afetasse positivamente o desempenho acadêmico do filho. Por outro lado, a proporção de estudantes com baixa renda e a greve deveriam afetar negativamente o resultado esperado.

Para as demais variáveis não havia nenhuma expectativa *à priori*.

3.2 Dados

A base de dados utilizada no presente trabalho foram as bases de microdados do ENADE disponibilizadas pelo INEP (2020). Foram utilizados os dados referentes aos anos de 2009, 2010, 2012 e 2013. Inicialmente, buscou-se deixar apenas as IES que estavam presentes em 2009 e 2012 e 2010 e 2013, de forma a possibilitar a construção do painel para a estimação por Diferenças-em-Diferenças. Após o processo de limpeza, recodificação e criação do painel, obteve-se um total de 510 observações, com os painéis de cada ciclo, 2009-2012 e 2010-2013, tendo 170 observações cada um.

Destaca-se que ao final do processo de tratamento dos dados, restavam 85 IES únicas na base. Seus nomes encontram-se no apêndice, bem como se foram consideradas como tratadas ou não, sendo essa designação baseada no fato da instituição ter aderido à greve federal universitária de 2012. Há aproximadamente 55% de tratados e 45% de não tratados na base de dados utilizada.

Por fim, as estatísticas descritivas são apresentadas na tabela 2. Pode-se observar que as variáveis apresentam grande amplitude e variabilidade, o que faz sentido, uma vez que as instituições de ensino superior são compostas por diversos quadros discentes e docentes distintos, o que se traduz nas estatísticas vistas na tabela.

Tabela 2 – Estatísticas descritivas

Estatística	N	Média	Desv. pad	Mín	Mediana	Máx
Idade Média	510	25,630	2,760	18,143	25,270	40,271
Idade Média ²	510	664,499	150,668	329,163	638,550	1621,748
Nota média	510	45,971	6,282	2,621	45,887	65,930
Prop. brancos	510	0,512	0,212	0,000	0,496	0,915
Prop. casados	510	0,132	0,094	0,000	0,109	0,542
Prop. homens	510	0,418	0,111	0,077	0,424	0,795
Prop. mães ensino médio+	510	0,611	0,153	0,113	0,639	0,947
Prop. renda até 3SM	510	0,311	0,210	0,000	0,247	1,000
Tratamento	510	0,541	0,499	0	1	1

Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

4 Resultados

4.1 Análise de robustez

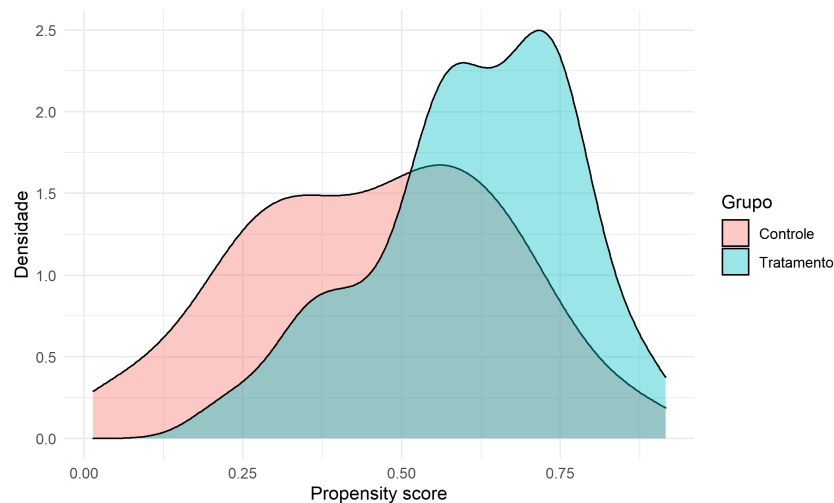
Nessa seção são apresentados os passos tomados com intuito de garantir a robustez dos resultados obtidos. Inicialmente, trataram-se das hipóteses referentes ao Propensity Score Matching (PSM); após, foram analisados os pressupostos referente à estimação por Diferenças-em-Diferenças (DD).

4.1.1 Propensity score matching

4.1.1.1 Hipótese do suporte comum

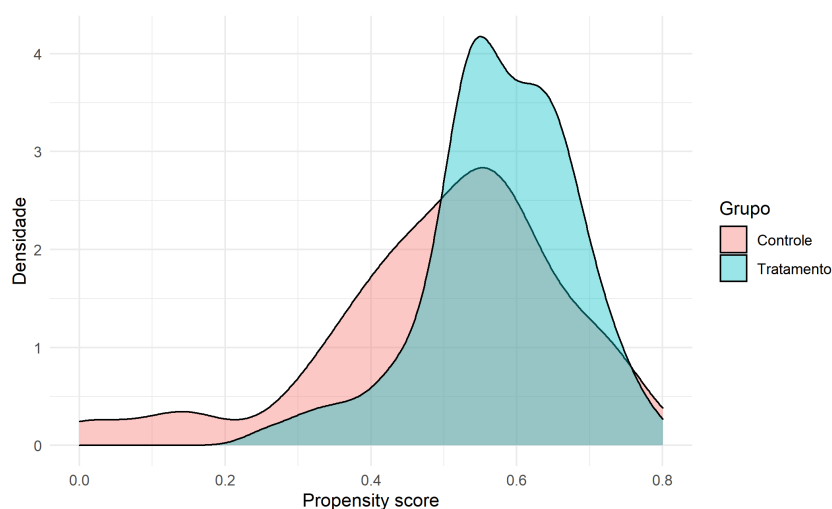
Para que os resultados do PSM sejam válidos, é preciso que o modelo atenda ao critério de existência de suporte comum entre os grupos de tratamento e de controle. Pode-se verificar isso analisando a função densidade de probabilidade do propensity score (PS) para cada grupo. Isso é feito para os Ciclos 2009-2012 e 2010-2013, respectivamente, nas Figuras 1 e 2:

Figura 1 – Suporte comum - Ciclo 2009-2012 - PSM (logit)



Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Figura 2 – Suporte comum - Ciclo 2010-2013 - PSM (logit)



Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Em seguida, conforme recomendado em Garrido et al. (2014), fez-se uma análise dessa mesma hipótese porém analisando agora os dados separados em quantis. Optou-se por utilizar cinco (5) quantis, sendo o último desses descartado devido ao pequeno número de observações (duas observações por grupo), ficando, portanto com quatro (4) quantis.

O resultado dessa divisão pode ser visualizado na Tabela 3, na qual tem-se o valor médio do PS em cada um dos quatro quantis para cada um dos grupos:

Tabela 3 – Propensity score médio por quantil - Tratados x Controle

Modelo	Quantil	Grupo	Propensity score
Ciclo 2009-2012	(0,014;0,4]	Controle	0,251
	(0,014;0,4]	Tratamento	0,324
	(0,4;0,562]	Controle	0,485
	(0,4;0,562]	Tratamento	0,507
	(0,562;0,69]	Controle	0,625
	(0,562;0,69]	Tratamento	0,622
	(0,69;0,889]	Controle	0,753
	(0,69;0,889]	Tratamento	0,756
	(0,000005;0,476]	Controle	0,350
	(0,000005;0,476]	Tratamento	0,389
Ciclo 2010-2013	(0,476;0,554]	Controle	0,524
	(0,476;0,554]	Tratamento	0,526
	(0,554;0,632]	Controle	0,586
	(0,554;0,632]	Tratamento	0,593
	(0,632;0,757]	Controle	0,694
	(0,632;0,757]	Tratamento	0,675

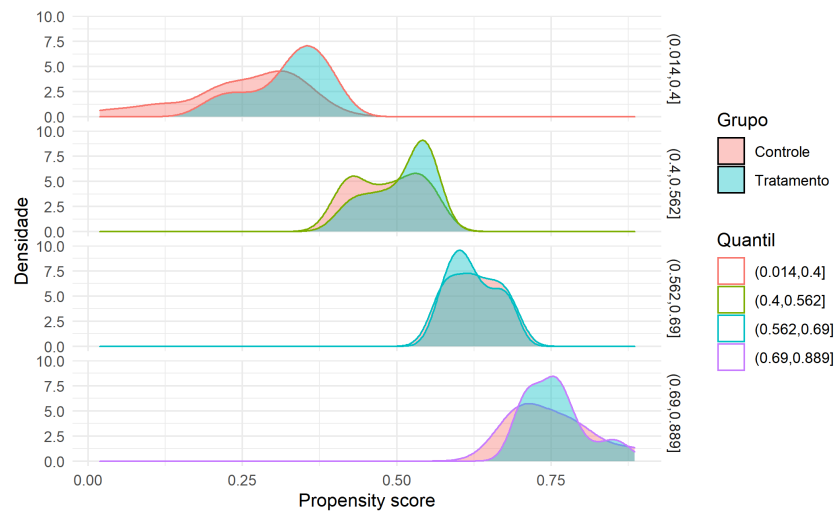
Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Note que para ambos os ciclos as médias são extremamente próximas, com as diferenças sendo em torno de 5%, com exceção do primeiro quantil, que em ambas amostras se mostrou com a maior diferença relativa entre grupos. Apesar disso, considerou-se que a diferença não é grande o suficiente para ser um problema grave¹.

Para corroborar essa hipótese, procedeu-se à construção de distribuições de densidade do PS para cada quantil, separados em grupo de tratamento e grupo de controle. Os resultados estão nas figuras 3 e 4

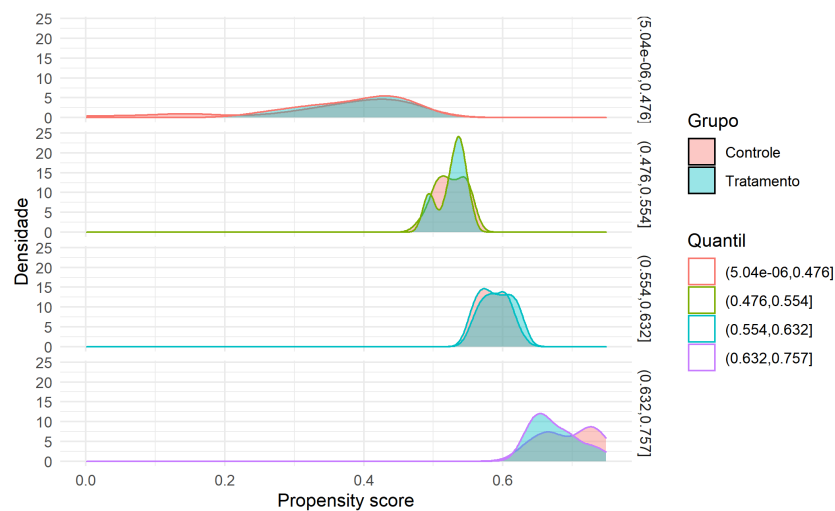
¹ Uma possibilidade de garantir ainda mais robustez aos resultados seria realizar uma estimação quantílica, separando, assim, o quantil "problemático" dos demais. Os autores do presente artigo crêem que a abordagem utilizada aqui é confiável o suficiente para ser útil.

Figura 3 – Suporte comum por quantil - Ciclo 2009-2012 - PSM (logit)



Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Figura 4 – Suporte comum por quantil - Ciclo 2010-2013 - PSM (logit)



Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

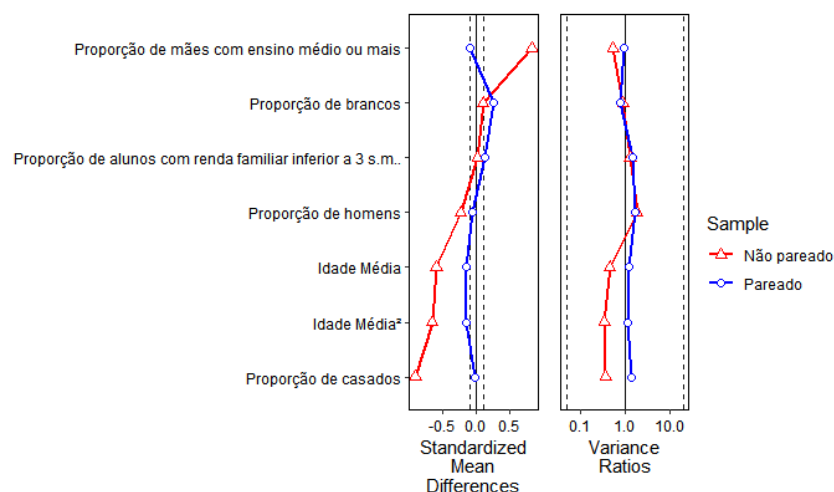
De fato, vê-se que a análise visual apoia a hipótese da existência de suporte comum entre os grupos de tratamento e controle, ocorrendo isso até mesmo em nível de quantil.

Na próxima seção, a análise do balanceamento das covariadas dos modelos, ou seja, do resultado da utilização do PSM, é realizada. Para tanto, foram utilizados gráficos que indicam, visualmente, como eram e como ficaram as diferenças entre as médias e a razão das variâncias de cada covariada entre os dois grupos.

4.2 Balanceamento das covariadas

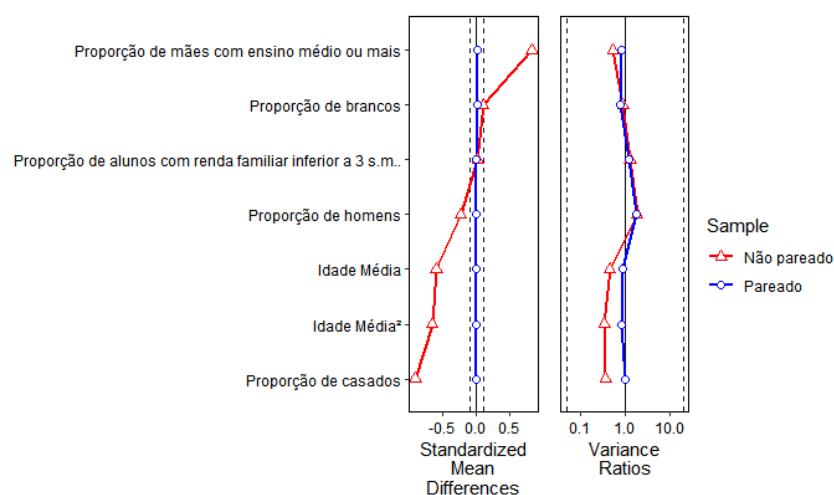
Utilizando o comando *love.plot* do pacote *Cobalt* (Greifer (2020)) da linguagem *R* (R Core Team (2020)), criaram-se os diversos gráficos citados no parágrafo anterior. Eles encontram-se dispostos nas figuras 5, 6, 7 e 8.

Figura 5 – Balanceamento das covariadas - Ciclo 2009-2012 - PSM (logit)



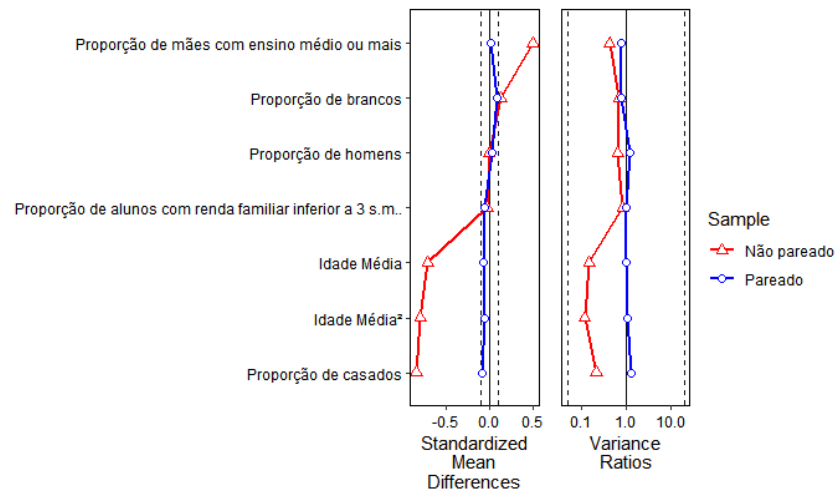
Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Figura 6 – Balanceamento das covariadas - Ciclo 2009-2012 - PSM (entropia)



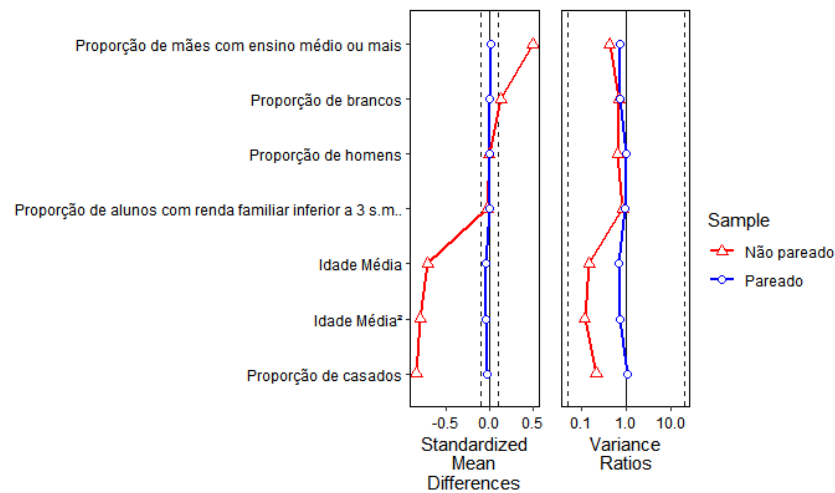
Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Figura 7 – Balanceamento das covariadas - Ciclo 2010-2013 - PSM (logit)



Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Figura 8 – Balanceamento das covariadas - Ciclo 2010-2013 - PSM (entropia)



Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Notadamente, como se pode perceber em todas as figuras, os grupos eram muito diferentes entre si, com quase nenhuma das variáveis tendo médias semelhantes, embora suas variâncias fossem parecidas. Apesar disso, os métodos empregados tiveram sucesso em tornar os dois grupos comparáveis.

Veja que até o ajuste relativamente menos complexo realizado pelo PSM com *logit* já conseguiu praticamente zerar as diferenças entre médias e variâncias em ambos os ciclos. Não obstante, a utilização do cálculo de **entropia** serviu para melhorar ainda

mais os resultados, conseguindo deixar praticamente todas as variáveis dentro dos limiares aceitáveis tradicionais.

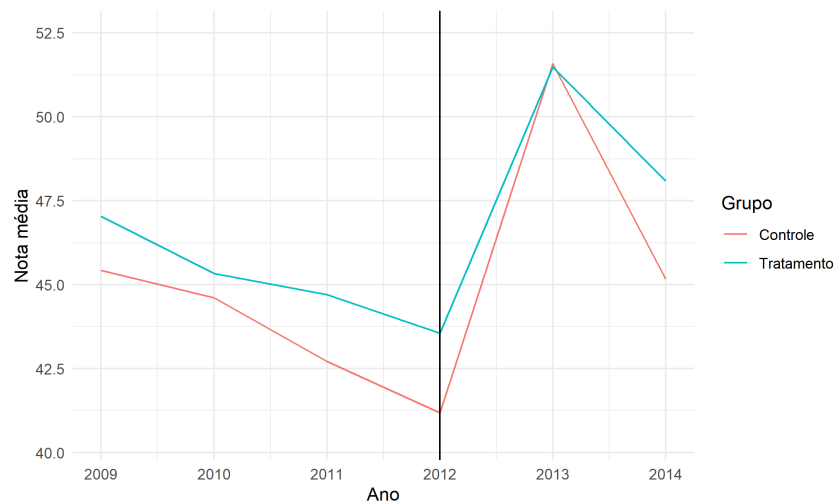
Na próxima seção é tratada a hipótese crucial do modelo de Diferenças-em-Diferenças (DD), a saber, a de que a variável de interesse tem trajetória igual (ou próxima) para ambos os grupos antes do tratamento. Para isso, foi utilizada uma inspeção gráfica, bem como a estimação de dois modelos de placebo.

4.2.1 Diferenças em diferenças

4.2.1.1 Tendências paralelas - inspeção gráfica

Iniciando a análise com uma inspeção ingênua, puramente gráfica, tem-se a figura 9

Figura 9 – Tendências temporais - Nota média por grupo de tratamento



Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Note que ao menos visualmente a trajetória para ambos os grupos é **muito próxima**, indicando que possivelmente o *matching* teve sucesso e temos resultados confiáveis. A seguir tratará-se da estimação de dois modelos placebos para cada ciclo: o primeiro, considerando a proporção de casados como variável dependente, e o segundo, tratando a educação da mãe como variável dependente. Como já foi abordado, espera-se que o tratamento não tenha efeito sobre essas variáveis, o que indicaria que não houve nenhum choque no período que possa ter feito com que as trajetórias se desviassem.

4.2.1.2 Teste de placebo

Nas tabelas 4 e 5, têm-se o resultado das estimações dos modelos de placebo.²

² Por uma questão de espaço, optou-se por colocar os gráficos de análise de balanceamento dos testes de placebo no apêndice. Esteja seguro, entretanto, de que todos os *matchings* obtiveram sucesso.

Tabela 4 – Placebo - Ciclo 2009-2012

	<i>Variável dependente:</i>	
	Proporção de casados	Proporção de mães com ensino médio+
Pós-tratamento	−0,037*** (0,009)	0,079*** (0,016)
Idade média	−0,008 (0,018)	−0,018 (0,034)
Idade média ²	0,001* (0,0004)	−0,0003 (0,001)
Prop. brancos	0,096 (0,080)	0,238 (0,153)
Prop. casados		−0,175 (0,219)
Nota média	−0,001* (0,001)	0,001 (0,001)
Prop. renda até 3SM	0,006 (0,030)	−0,082 (0,058)
Prop. homens	0,019 (0,044)	0,023 (0,082)
Prop. Ensino Médio+	−0,049 (0,060)	
Diff-and-diff/Impacto	−0,001 (0,006)	0,011 (0,011)
Observações	170	170
Estatística F (df = 9; 76)	28,765***	10,020***
<i>Nota</i> ¹ : *p<0,1; **p<0,05; ***p<0,01 <i>Nota</i> ² : Os dois modelos foram estimados utilizando DD+FE+PSM (entropia)		

Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Tabela 5 – Placebo - Ciclo 2010-2013

	<i>Variável dependente:</i>	
	Proporção de casados	Proporção de mães com ensino médio+
Pós-tratamento	−0,035** (0,015)	0,041 (0,046)
Idade média	−0,004 (0,030)	−0,126 (0,078)
Idade média ²	0,0005 (0,001)	0,002 (0,002)
Prop. brancos	0,008 (0,046)	0,751*** (0,103)
Prop. casados		0,191 (0,335)
Nota média	−0,001 (0,001)	0,006*** (0,002)
Prop. renda até 3SM	0,078*** (0,026)	0,143* (0,075)
Prop. homens	−0,129** (0,052)	−0,037 (0,170)
Prop. Ensino Médio+	0,026 (0,039)	
Diff-and-diff/Impacto	0,003 (0,006)	0,034* (0,018)
Observações	170	170
Estatística F (df = 9; 76)	36,792***	38,078***
<i>Nota</i> ¹ :		*p<0,1; **p<0,05; ***p<0,01
<i>Nota</i> ² :		Os dois modelos foram estimados utilizando DD+FE+PSM (entropia)

Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Veja que em três (3) dos quatro (4) modelos estimados, a variável de impacto não teve significância estatística. Uma possível explicação para a variável ter sido significativa no último modelo é que, em alguma medida, os alunos com mães que têm maior nível de educação participaram menos, relativamente, das provas do ENADE 2013, do que os demais discentes em razão da greve. Apesar disso, o problema não parece causar nada

muito grave e podemos concluir pela validade da nossa estimação de placebo.

Findada a apresentação das medidas tomadas para garantir a robustez dos resultados obtidos, na próxima subseção eles são apresentados e discutidos.

4.3 Análise de impacto

Inicialmente foi feita a análise dos resultados obtidos para o Ciclo 2009-2012 e depois para o Ciclo 2010-2013. A tabela 6 apresenta as estimações realizadas para o primeiro Ciclo.

Nota-se que as variáveis já consagradas na literatura têm os sinais esperados e são significativas: a idade é positivamente correlacionada com a nota, apresentando, entretanto, concavidade, indicando que seu efeito é decrescente no tempo. A proporção de estudantes com renda familiar de até três (3) salários tem relação negativa com a nota, possivelmente pela melhor educação que o dinheiro pode proporcionar.

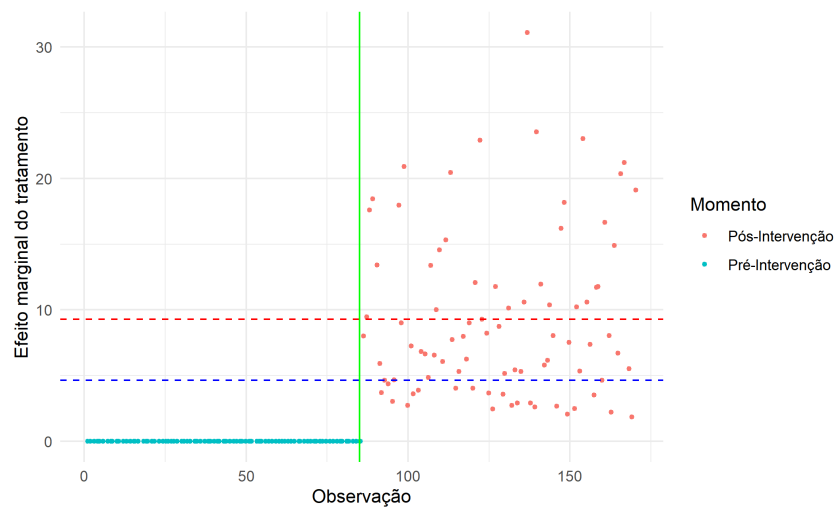
A variável que captura a proporção de casados nas instituições é estatisticamente significativa e apresenta relação negativa com a nota. Uma possível razão para isso é que discentes casados têm menos tempo para dedicarem aos seus estudos, o que se traduziria em uma nota menor para esse grupo.

É importante frisar como os resultados mudam, se comparados o modelo ingênuo de DD com o modelo final, mais robusto, de DD + FE + PSM (entropia). A variável que captura a proporção de brancos em cada IES muda de sinal e aumenta expressivamente em magnitude, indicando que uma maior proporção de brancos é negativamente correlacionada com a nota média da instituição.

A variável de impacto, entretanto, não ganha significância. Seu coeficiente aponta uma relação positiva com a nota média. Possivelmente a variável não foi significativa pois os maiores efeitos da greve levam um certo intervalo de tempo para serem sentidos, hipótese que é retomada na discussão a seguir. Supondo que ela fosse significativa nesse ciclo, a distribuição de seu efeito marginal ao longo das observações, considerando apenas o modelo mais robusto, seria como está representado na figura 10³, onde a linha vertical representa o limiar de ocorrência da intervenção.

³ No caso do modelo log-lin estimado, o efeito marginal não é linear e depende do valor da variável dependente.

Figura 10 – Efeitos marginais da variável de impacto - Ciclo 2009-2012

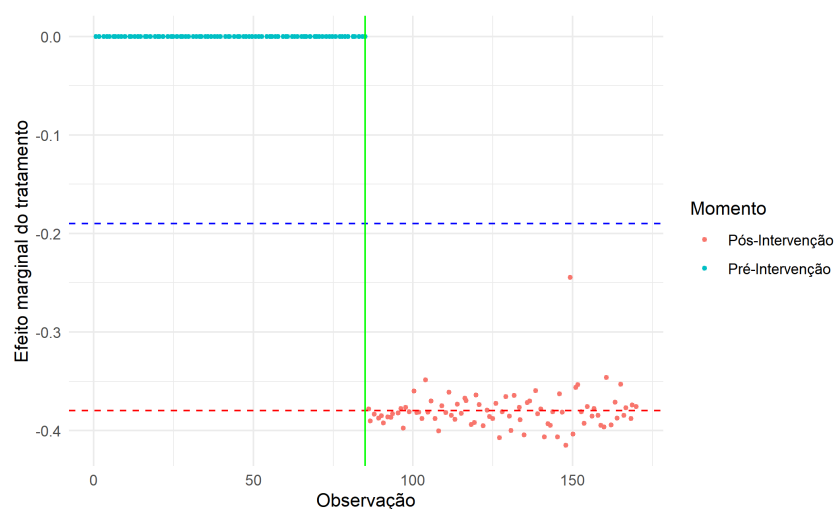


Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Fazendo agora a análise dos resultados para o segundo ciclo, cujas estimações encontram-se na tabela 7, tem-se que muita coisa mudou. Todas as variáveis perderam significância, com exceção da variável de impacto. Entretanto, seu sinal se inverteu, indicando que a greve teve um efeito de longo-prazo sobre os rendimentos dos alunos. Isso faz sentido, quando se considera o impacto que as greves tiveram sobre os calendários acadêmicos subsequentes. Esse resultado é semelhante ao obtido por Annegues, Porto Júnior e Figueiredo (2017), que encontraram que as greves tiveram um impacto negativo estatisticamente significativo sobre o desempenho dos discentes da Universidade Federal da Paraíba (UFPB) em seus respectivos cursos, medidos pelas suas notas.

Em magnitude, o efeito marginal médio da variável de impacto foi de uma redução de 0,3798 na nota média das instituições, considerando o período pós-greve ($t = 1$). A distribuição de seu efeito marginal ao longo das observações é como está representada na figura 11

Figura 11 – Efeitos marginais da variável de impacto - Ciclo 2010-2013



Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Por outro lado, a educação da mãe ganhou significância estatística e apresentou um valor positivo relativamente elevado. Isso aliado à não significância estatística da variável de renda, talvez possa indicar que os cursos analisados no segundo ciclo possam ter uma composição estudantil diferentes dos cursos do primeiro ciclo.

Na seção seguinte, fez-se uma breve recapitulação do problema de pesquisa, dos passos tomados para a estimação e dos resultados e tecem-se as considerações finais sobre o trabalho, indicando possíveis caminhos futuros e aplicações.

Tabela 6 – Efeito médio de tratamento sobre os tratados - Ciclo 2009-2012

	<i>Variável dependente:</i>			
	log(nota média)			
	Diff-diff	Diff-diff + FE	Diff-diff + FE + PSM (logit)	Diff-diff + FE + PSM (entropia)
Constante	−0,193 (1,375)			
Tratamento	−0,028 (0,054)			
Pós-tratamento	−0,020 (0,079)	−0,023 (0,090)	−0,008 (0,083)	−0,045 (0,085)
Idade média	0,315*** (0,099)	0,606*** (0,148)	0,737*** (0,171)	0,681*** (0,164)
Idade média ²	−0,005*** (0,002)	−0,010*** (0,003)	−0,013*** (0,004)	−0,012*** (0,003)
Prop. brancos	0,060 (0,103)	−3,106*** (0,663)	−3,370*** (0,709)	−3,338*** (0,709)
Prop. casados	−1,427** (0,586)	−2,836*** (0,965)	−2,074** (1,018)	−2,436** (1,035)
Prop. homens	−0,004 (0,275)	0,475 (0,393)	0,201 (0,418)	0,301 (0,404)
Prop. Ensino Médio+	−0,270 (0,262)	−0,057 (0,566)	−0,273 (0,551)	−0,211 (0,557)
Prop. Renda até 3SM	−0,599*** (0,216)	−1,113*** (0,301)	−1,310*** (0,285)	−1,184*** (0,274)
Diff-and-diff/Impacto	0,010 (0,075)	−0,044 (0,057)	0,002 (0,052)	0,006 (0,053)
Observações	170	170	170	170
R ²	0,183			
R ² ajustado	0,132			
Erro-padrão residual	0,239 (df = 159)			
Estatística F	3,563*** (df = 10; 159)	8,605*** (df = 9; 76)	9,071*** (df = 9; 76)	8,719*** (df = 9; 76)
<i>Nota:</i>	*p<0,1; **p<0,05; ***p<0,01			

Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Tabela 7 – Efeito médio de tratamento sobre os tratados - Ciclo 2010-2013

	<i>Variável dependente:</i>			
	log(nota média)			
	Diff-diff	Diff-diff + FE	Diff-diff + FE + PSM (logit)	Diff-diff + FE + PSM (entropia)
Constante	3,017*** (0,379)			
Tratamento	−0,007 (0,019)			
Pós-tratamento	0,131*** (0,043)	0,091* (0,048)	0,045 (0,051)	0,072 (0,050)
Idade média	0,057** (0,025)	0,142*** (0,045)	0,181* (0,098)	0,130 (0,089)
Idade média ²	−0,001*** (0,0004)	−0,003*** (0,001)	−0,003 (0,002)	−0,002 (0,002)
Prop. brancos	0,038 (0,041)	−0,137 (0,140)	−0,050 (0,147)	−0,083 (0,152)
Prop. casados	−0,080 (0,244)	0,100 (0,326)	−0,129 (0,369)	−0,133 (0,377)
Prop. homens	0,048 (0,063)	0,225 (0,173)	0,144 (0,179)	0,236 (0,188)
Prop. Ensino Médio+	0,248*** (0,081)	0,380*** (0,117)	0,325*** (0,121)	0,340*** (0,126)
Prop. Renda até 3SM	−0,021 (0,064)	−0,043 (0,078)	0,051 (0,088)	0,012 (0,088)
Diff-and-diff/Impacto	−0,019 (0,027)	−0,040* (0,020)	−0,046** (0,019)	−0,042** (0,020)
Observações	170	170	170	170
R ²	0,577			
R ² ajustado	0,551			
Erro-padrão residual	0,084 (df = 159)			
Estatística F	21,706*** (df = 10; 159)	26,009*** (df = 9; 76)	26,288*** (df = 9; 76)	24,104*** (df = 9; 76)
<i>Nota:</i>				*p<0,1; **p<0,05; ***p<0,01

Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

5 Conclusões

O intuito do presente trabalho era verificar se a greve federal de 2012 das instituições de ensino superior teve algum efeito sobre o desempenho dos discentes na prova do ENADE, tanto daquele ano quanto de 2013. Para tanto, empregaram-se, inicialmente, métodos de pareamento, visando deixar os grupos de tratamento e controle o mais parecidos possível. Tendo sido feito isso, procedeu-se a estimação do modelo utilizado Diferenças-em-Diferenças com controle de efeitos fixos individuais de cada IES.

Os resultados mostram que no ano da greve, 2012, o seu efeito sobre as notas médias das instituições não foi significativo, muito provavelmente pelo pequeno intervalo de tempo entre o final da greve e a prova. Por outro lado, ao se analisar a prova do ano seguinte, 2013, obteve-se um impacto negativo estatisticamente significativo, apontando para a existência de um efeito deletério não-contemporâneo da greve sobre o desempenho dos discentes.

Novos trabalhos devem buscar dados mais específicos sobre as greves, como a duração dela em cada instituição, possibilitando assim a criação de uma variável que capture a intensidade do tratamento. Além disso, uma vez que as IES que escolheram aderir ao movimento grevista têm características distintas das demais, seria interessante utilizar a correção de Heckman (1979), visando modelar a equação de autosseleção das instituições.

Referências

- ANNEGUES, A. C.; PORTO JÚNIOR, S.; FIGUEIREDO, E. O impacto das greves dos professores universitários sobre o desempenho dos alunos da ufpb. *XXII Encontro Regional de Economia - ANPEC*, 2017. Disponível em: <http://www.anpec.org.br/nordeste/2017/submissao/arquivos_identificados/173-4f11b45b9ea3cb10d55a7d79224ad576.docx>.
- BALTAGIA, B. H. *Econometric analysis of panel data*. 3. ed. [S.l.]: John Wiley and Sons, 2005.
- BECKER, S. O.; INCHINO, A. Estimation of average treatment effects based on propensity scores. *The Stata Journal*, v. 2, n. 4, p. 358–377, 2002.
- BELOT, M.; WEBBINK, D. Do Teacher Strikes Harm Educational Attainment of Students? *LABOUR*, v. 24, n. 4, p. 391–406, 2010.
- 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.
- CAMERON, A. C.; TRIVEDI, P. K. *Microeconometrics: methods and applications*. [S.l.]: Cambridge University Press, 2005.
- GARRIDO, M. M. et al. Methods for Constructing and Assessing Propensity Scores. *Health Services Research*, v. 49, p. 1701–1720, 2014.
- GREENGLASS, E. et al. Stressful effects of a university faculty strike on students: Implications for coping. *Interchange*, v. 33, p. 261–279, 2002.
- GREIFER, N. *cobalt: Covariate Balance Tables and Plots*. [S.l.], 2020. R package version 4.2.4. Disponível em: <<https://CRAN.R-project.org/package=cobalt>>.
- HAINMUELLER, J. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, v. 20, p. 25–46, 2012.
- HECKMAN, J. J. Sample selection bias as a specification error. *Econometrica*, v. 47, n. 1, p. 153–161, 1979.
- HOXBY, C. M. How teachers' unions affect education production. *Quarterly Journal of Economics*, p. 671–718, 1996.
- HSIAO, C. *Analysis of panel data*. 3. ed. [S.l.]: Cambridge University Press, 2014.
- INEP. Microdados - enade. 2020. Disponível em: <<http://www.portal.inep.gov.br/microdados>>.
- JANUZZI, P. *Monitoramento e Avaliação de Programas Sociais: uma introdução aos conceitos e técnicas*. [S.l.]: Editora Alínea, 2016.
- KHANDKER, R.; KOOLWAL, G. B.; SAMAD, H. A. *Handbook on Impact Evaluation: Quantitative Methods and Practices*. [S.l.]: The World Bank, 2010.

- LECHNER, M. The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics*, v. 4, n. 3, p. 165–224, 2010.
- LOVENHEIM, M. F. The effect of teachers' unions on education production: Evidence from union election certifications in three midwestern states. *Journal of Labor Economics*, v. 27, n. 4, p. 525–587, 2009.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2020. Disponível em: <<https://www.R-project.org/>>.
- ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. *Biometrika*, v. 70, n. 1, p. 41–55, 1983.
- WICKENS, C. et al. Student stress and coping following a university strike in canada. *Journal of Collective Negotiations*, v. 31, p. 1–19, 2006.
- WILLS, G. The effects of teacher strike activity on student learning in south african primary schools. *Economic Research Southern Africa Working Paper*, 2014.
- ZWERLING, H. L. Pennsylvania teachers' strikes and academic performance. *Journal of Collective Negotiations*, v. 32, n. 2, p. 151–172, 2008.

6 Apêndice

6.1 Lista de instituições de ensino superior utilizadas na estimação

Tabela 8 – Lista de instituições de ensino superior com dummy de aderência à greve

Nome	Aderiram
UNIVERSIDADE FEDERAL DE MATO GROSSO	Não
UNIVERSIDADE DE BRASÍLIA	Não
UNIVERSIDADE FEDERAL DE SERGIPE	Sim
UNIVERSIDADE FEDERAL DO AMAPÁ	Sim
UNIVERSIDADE FEDERAL DO PIAUÍ	Sim
UNIVERSIDADE FEDERAL DE OURO PRETO	Sim
UNIVERSIDADE FEDERAL DE SÃO CARLOS	Sim
UNIVERSIDADE FEDERAL DE VICOSA	Sim
UNIVERSIDADE ESTADUAL DE LONDRINA	Não
UNIVERSIDADE FEDERAL DO RIO GRANDE	Sim
UNIVERSIDADE FEDERAL DE UBERLÂNDIA	Sim
UNIVERSIDADE ESTADUAL DE SANTA CRUZ	Não
UNIVERSIDADE ESTADUAL DO CEARÁ	Não
UNIVERSIDADE DO ESTADO DA BAHIA	Não
FUNDAÇÃO UNIVERSIDADE DO ESTADO DE SANTA CATARINA	Não
UNIVERSIDADE ESTADUAL DE GOIÁS	Não
UNIVERSIDADE ESTADUAL PAULISTA JÚLIO DE MESQUITA FILHO	Não
UNIVERSIDADE ESTADUAL DE MARINGÁ	Não
UNIVERSIDADE DO ESTADO DO RIO GRANDE DO NORTE	Não
UNIVERSIDADE ESTADUAL DO VALE DO ACARAÚ	Não
UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL REI	Sim
UNIVERSIDADE ESTADUAL DE MONTES CLAROS	Não
UNIVERSIDADE DE PERNAMBUCO	Não
UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO	Não
UNIVERSIDADE FEDERAL DO MARANHÃO	Sim
UNIVERSIDADE ESTADUAL DA PARAIBA	Não
UNIVERSIDADE ESTADUAL DO MARANHÃO	Não
UNIVERSIDADE FEDERAL DO PARÁ	Sim
UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE	Não
UNIVERSIDADE FEDERAL DO PARANÁ	Sim
UNIVERSIDADE FEDERAL FLUMINENSE	Sim
UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO	Sim
UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO	Sim
UNIVERSIDADE FEDERAL DE MINAS GERAIS	Sim
UNIVERSIDADE FEDERAL DE JUIZ DE FORA	Sim
UNIVERSIDADE FEDERAL DE ALAGOAS	Sim
UNIVERSIDADE FEDERAL DA BAHIA	Sim
UNIVERSIDADE FEDERAL DA PARAIBA	Sim
UNIVERSIDADE FEDERAL DE PERNAMBUCO	Sim
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL	Sim
UNIVERSIDADE FEDERAL DE SANTA MARIA	Sim
UNIVERSIDADE FEDERAL DO CEARÁ	Sim

Continua na próxima página

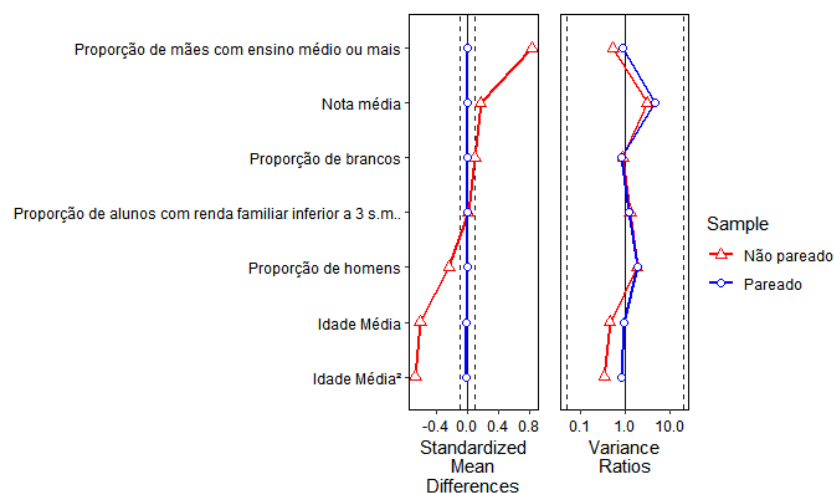
Tabela 8 – continuação da página anterior

Nome	Aderiram
UNIVERSIDADE FEDERAL DE GOIAS	Sim
UNIVERSIDADE FEDERAL DE SANTA CATARINA	Não
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO	Sim
UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO	Sim
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ	Não
UNIVERSIDADE FEDERAL RURAL DO SEMIÁRIDO	Não
UNIVERSIDADE FEDERAL DE LAVRAS	Sim
CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA CELSO SUCKOW DA FONSECA	Sim
UNIVERSIDADE FEDERAL DOS VALES DO JEQUITINHONHA E MUCURI	Sim
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DA BAHIA	Não
UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ	Não
UNIVERSIDADE FEDERAL DE PELOTAS	Sim
UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA	Não
UNIVERSIDADE ESTADUAL DO SUDOESTE DA BAHIA	Não
UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO	Sim
UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL	Não
FUNDAÇÃO UNIVERSIDADE FEDERAL DE RONDONIA	Sim
UNIVERSIDADE DO ESTADO DE MATO GROSSO	Não
UNIVERSIDADE ESTADUAL DE PONTA GROSSA	Não
UNIVERSIDADE ESTADUAL DO PIAUÍ	Não
UNIVERSIDADE FEDERAL DE RORAIMA	Sim
UNIVERSIDADE DO TOCANTINS	Não
UNIVERSIDADE ESTADUAL DE MATO GROSSO DO SUL	Não
UNIVERSIDADE ESTADUAL DO CENTRO OESTE	Não
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DA PARAIBA	Sim
UNIVERSIDADE FEDERAL DE CAMPINA GRANDE	Sim
UNIVERSIDADE DO ESTADO DO AMAPÁ	Não
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO NORTE DE MINAS GERAIS	Não
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE MINAS GERAIS	Sim
UNIVERSIDADE ESTADUAL DO RIO GRANDE DO SUL	Não
FUNDAÇÃO UNIVERSIDADE FEDERAL DO TOCANTINS	Sim
FUNDAÇÃO UNIVERSIDADE FEDERAL DO VALE DO SÃO FRANCISCO	Sim
FUNDAÇÃO UNIVERSIDADE FEDERAL DA GRANDE DOURADOS	Sim
UNIVERSIDADE ESTADUAL DE RORAIMA	Não
UNIVERSIDADE ESTADUAL DE ALAGOAS UNEAL	Não
FUNDAÇÃO UNIVERSIDADE FEDERAL DO PAMPA UNIPAMPA	Não
UNIVERSIDADE FEDERAL DO ACRE	Sim
UNIVERSIDADE REGIONAL DO CARIRI	Não
UNIVERSIDADE FEDERAL DO AMAPÁ	Sim
UNIVERSIDADE FEDERAL DO RECONCAVO DA BAHIA	Sim
UNIVERSIDADE FEDERAL DE SÃO PAULO	Sim
UNIVERSIDADE FEDERAL DO TRIÂNGULO MINEIRO	Sim
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ	Sim

Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

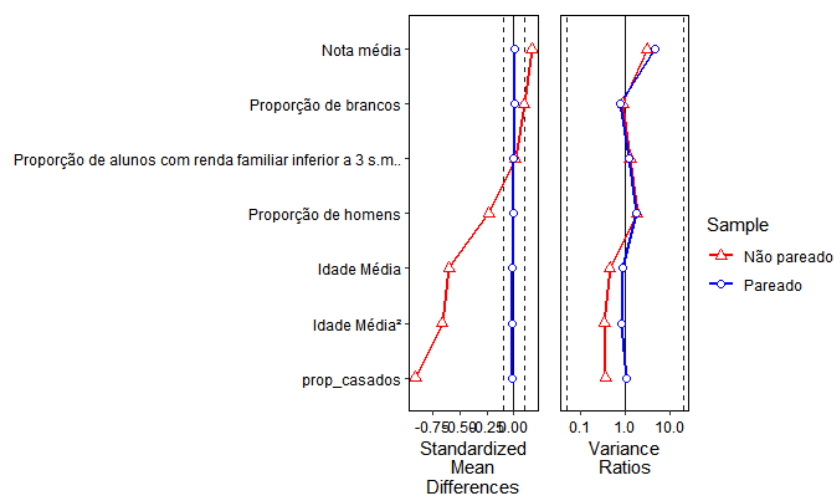
6.2 Balanceamento das covariadas - Testes de placebo

Figura 12 – Balanceamento das covariadas - Ciclo 2009-2012 - Teste de placebo 1



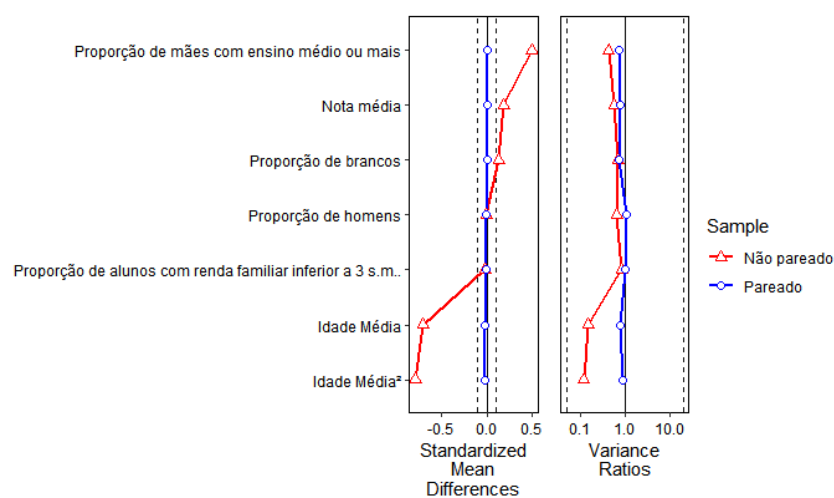
Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Figura 13 – Balanceamento das covariadas - Ciclo 2009-2012 - Teste de placebo 2



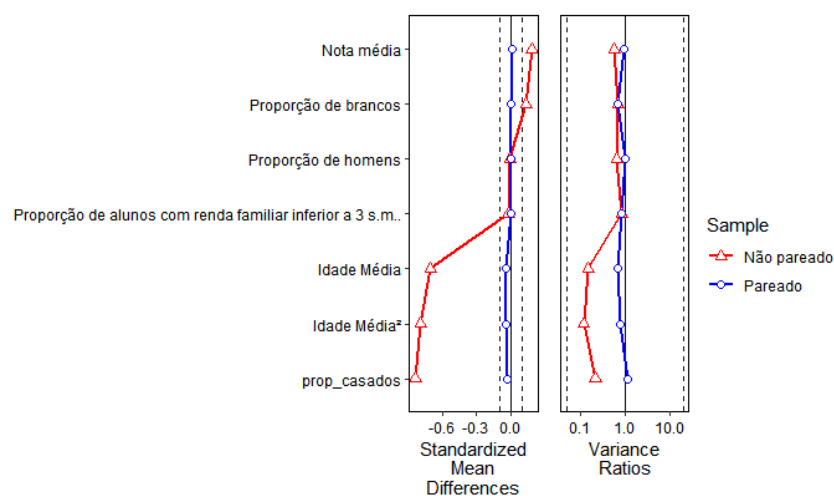
Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Figura 14 – Balanceamento das covariadas - Ciclo 2010-2013 - Teste de placebo 1



Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).

Figura 15 – Balanceamento das covariadas - Ciclo 2010-2013 - Teste de placebo 2



Fonte: Elaboração própria à partir dos Microdados do ENADE disponíveis em (INEP, 2020).