

RELATÓRIO METODOLÓGICO

Processo de aleatorização da avaliação de impacto do Jovem de Futuro

■ **autores:**
Ricardo Paes Barros // Samuel Franco

INSTITUTO UNIBANCO

Relatório metodológico - Processo de aleatorização da avaliação de impacto do Jovem de Futuro

Autores: Ricardo Paes Barros // Samuel Franco

ISBN 978-85-60154-04-3

RESUMO

O presente relatório tem como objetivo explicitar as escolhas metodológicas e processuais adotadas para aleatorização das escolas do Jovem de Futuro e o método de avaliação empregado para estimação do impacto do programa.

Após o término da segunda geração (executada entre 2011 e 2014), o Jovem de Futuro passou por uma reestruturação em suas frentes de trabalho e isso levou ao desenvolvimento de uma terceira geração, implementada nos estados do Espírito Santo, Piauí e Pará entre 2015 e 2017. As mudanças propostas acarretaram na necessidade de pactuação de um método que viabilizasse a realização de avaliações de impacto nos três estados abrangidos pela terceira geração.

A opção por um método experimental de seleção via aleatorização das escolas e diferenciação do tempo de entrada no programa criou bases sólidas para um processo de avaliação metodologicamente estruturado e passível de validação. Deste modo, as escolas selecionadas para receber o benefício no início do Programa (1º ano) foram denominadas escolas de tratamento e aquelas selecionadas para o último ano de implementação (3º ano), escolas de controle.

PALAVRAS-CHAVE: Aleatorização; Avaliação Experimental; Grupos de tratamento e controle; Jovem de Futuro; Gestão Escolar; Ensino Médio.

ABSTRACT

This paper has the objective to explicit the methodological and processual choices applied for the randomization of the Jovem de Futuro schools and the evaluation method adopted to estimate the program impact.

After the end of the second generation (developed between 2011 e 2014), the Jovem de Futuro was restructured and their working processes changed, which led to the developed of a third generation implemented in the states of Espírito Santo, Piauí and Pará between 2015 and 2017.

The changes proposed led to the need to agree on a method that proportionated the realization of impact evaluation in the three states of the third generation.

The option for an experimental method of selection by the randomization of the schools and differentiation of the time of entry made possible an evaluation methodologically structured and subjected of validation. Therefore, the schools selected to receive the benefit of the beginning of the program (first year) was denominated schools of treatment, and those selected for the last year of implementation (third year), control schools.

KEYWORDS: Randomization; Experimental Evaluation; Control and treatment groups; Jovem de Futuro; School management; High School.

SUMÁRIO

1. Desenho original da seleção de escolas	4	3. Aleatorização por praça: Pará	17
A escolha do método	4	Testes	17
Grupos de tratamento e controle	4	Delimitação de universos e estratos	18
Homogeneidade entre os grupos	4	Validação	20
Testes para definição do método	5	4. Aleatorização por praça: Espírito Santo	22
Aleatorização e pareamento	6	Testes	22
Validação da aleatorização	9	Delimitação de universos e estratos	23
2. Aleatorização por praça: Piauí	10	Validação	26
Testes	10	Referências	29
Delimitação de universos e estratos	11		
Validação	15		

1. Desenho original da seleção de escolas

A escolha do método

O Programa Jovem de Futuro foi desenvolvido considerando que “uma gestão eficiente, participativa e com foco nos resultados positivos de aprendizagem pode influenciar de maneira decisiva a qualidade da educação oferecida na escola” (UNIBANCO, 2012).

O Instituto Unibanco implantou em 2007 o Programa em um conjunto de escolas públicas de Ensino Médio de Belo Horizonte e Porto Alegre. Em 2010, o Programa foi expandido para Rio de Janeiro, São Paulo e Vale do Paraíba. A implantação do Programa nessas cinco praças, com recursos financeiros e tecnologia do Instituto Unibanco, ficou conhecida como piloto. Nos anos seguintes ela serviu de base para a realização do Programa em parceria com o Ministério da Educação (MEC), integrando o Programa Jovem de Futuro ao Programa Ensino Médio Inovador (PRoEMI) – conhecida como fase de disseminação ou 2ª Geração – que possibilitou a implantação em larga escala do Programa nos estados do Ceará, Goiás, Mato Grosso do Sul, Pará e Piauí com recursos do Governo Federal, tecnologia do Instituto Unibanco e parceria com as Secretarias Estaduais de Educação de cada estado.

O Programa Jovem de Futuro (PJF), a avaliação e o processo de aleatorização das escolas participantes passaram na 2ª Geração (2011-2014) por uma reestruturação após o acompanhamento dos resultados de impacto das praças que receberam o Programa. Essa reestruturação, denominada 3ª Geração, foi desenvolvida para o Espírito Santo, cuja intervenção começou no ano de 2015, e foi replicada para as novas fases das praças de Piauí (2015-2017) e Pará (2015-2017). Abaixo, destacamos as principais reestruturações realizadas:

- O *Programa*, agora desvinculado do ProEMI, possui metas para as escolas, atenção especial para as escolas vulneráveis e uma sistemática de acompanhamento;
- A *Avaliação de Impacto* sob a 3ª série do Ensino Médio das escolas de intervenção passará a ser realizada durante dois anos ao invés de três anos;

Para a implantação do Programa era preciso escolher um método que garantisse que as escolas selecionadas pudessem ser avaliadas e o impacto do Programa ser mensurado (método experimental).

Para isso, foi adotado um método experimental de seleção por meio da aleatorização.¹ Dessa forma, o Programa pôde ser implantado sobre bases sólidas de um processo de seleção metodologicamente estruturado e passível de validação para escolha do grupo de escolas participantes e também com um mecanismo fundamental para a avaliação de impacto dos resultados do Programa, uma vez que a seleção permite a classificação das escolas de tratamento e controle.

Grupos de tratamento e controle

As escolas selecionadas para receberer o benefício no início do Programa (1º ano) foram denominadas escolas de *tratamento* e aquelas selecionadas para o último ano de tratamento, (3º ano), escolas de *controle*, pois receberão o benefício apenas após o término da intervenção nas escolas de tratamento. Tendo em vista esta estrutura, a intervenção fica organizada da seguinte forma:



Homogeneidade entre os grupos

É necessário considerar a possibilidade de sortear aleatoriamente, sem reposição, as escolas inscritas e incorrer em um evento raro no qual determinada característica marcante esteja mais presente em um dos dois grupos; por exemplo, o grupo de escolas selecionadas para o tratamento apresentar melhor desempenho em avaliações anteriores do que o grupo de controle, ou o inverso.

Para evitar este problema, considera-se utilizar um procedimento em que se formam grupos de escolas semelhantes entre si e, posteriormente, efetua-se o sorteio dentro de cada grupo. Desta forma, garante-se que entre as escolas semelhantes, uma formará o grupo que, de tratamento e outra o grupo de contro-

¹ “Randomization, in which the selection into the treatment and control groups is random within some well-defined set of people. In this case there should be no difference (in expectation) between the two groups besides the fact that the treatment group had access to the program (...).” (BAKER, 2000)

le, mantendo assim o “equilíbrio” de características entre as escolas que serão comparadas.

A aleatorização deve, portanto, gerar grupos de controle e tratamento ideais para a captação do impacto. Partimos da hipótese de que os grupos considerados ideais são aqueles cuja evolução de seu desempenho é igual na ausência do Programa.

Essa evolução é controlada por duas características: (i) a nota média das escolas nas provas de Língua Portuguesa e Matemática na linha de base do Programa (para garantir que escolas com a mesma evolução mas notas iniciais muito distantes não façam parte do mesmo grupo); e (ii) pelo Nível Socioeconômico das Escolas (NSE), um índice que mede a vulnerabilidade das escolas, em respeito ao redesenho do Programa.

Testes para definição do método

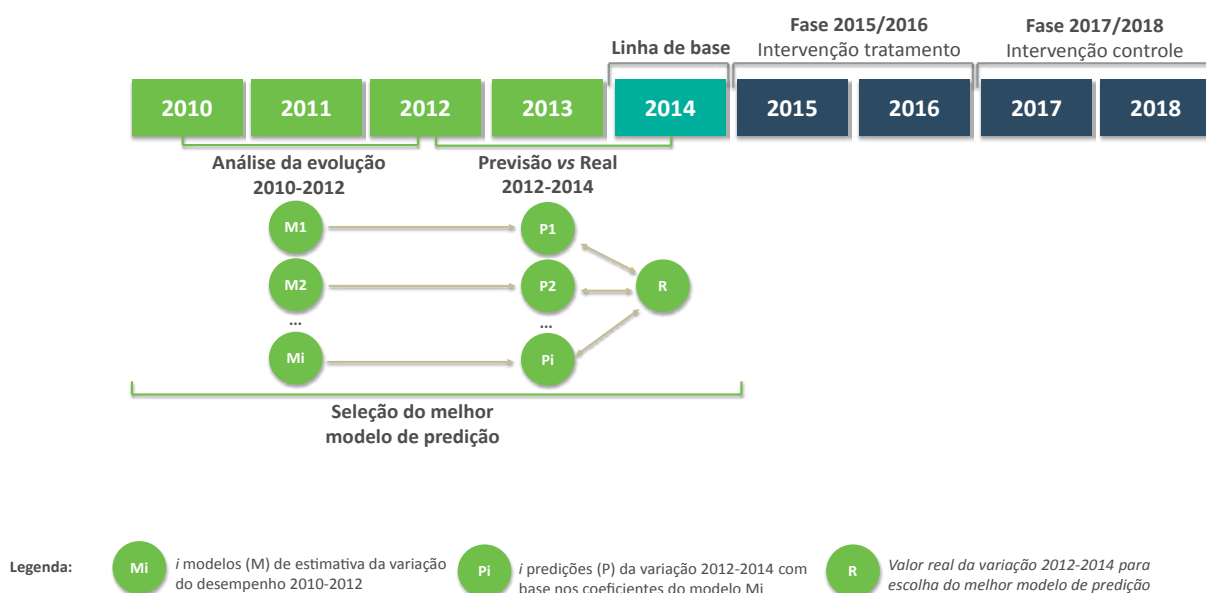
Estabelecemos a hipótese de que os grupos considerados ideais são aqueles cuja evolução de seu desempenho é igual na ausência do Programa.

- **Problema:** a igualdade da evolução que estabelecemos para a hipótese refere-se à Fase 2015/2016 do Programa, ou seja, é um momento futuro;
- **Solução:** precisamos estimar a evolução dos grupos para a Fase 2015/2016 caso não houvesse a intervenção do Programa.

Para estimar a evolução futura desses grupos temos que nos basear nas informações passadas sobre o desempenho das escolas. Sendo 2014 a linha de base do Programa, e 2015 e 2016 os anos da primeira fase, utilizaremos a evolução 2010-2012² para prever a evolução 2012-2014, dois períodos com informações disponíveis. Ao final, comparamos a evolução real observada entre 2012-2014 com a evolução prevista pelo modelo 2010-2012. O método escolhido para o pareamento de escolas será aquele que garantir a melhor previsão.

Com o melhor modelo de predição escolhido para cada praça inicia-se, efetivamente, o processo de aleatorização, pareamento e seleção da amostra.

QUADRO TEÓRICO DO MODELO DE PREVISÃO: ESTRUTURA DO TESTE



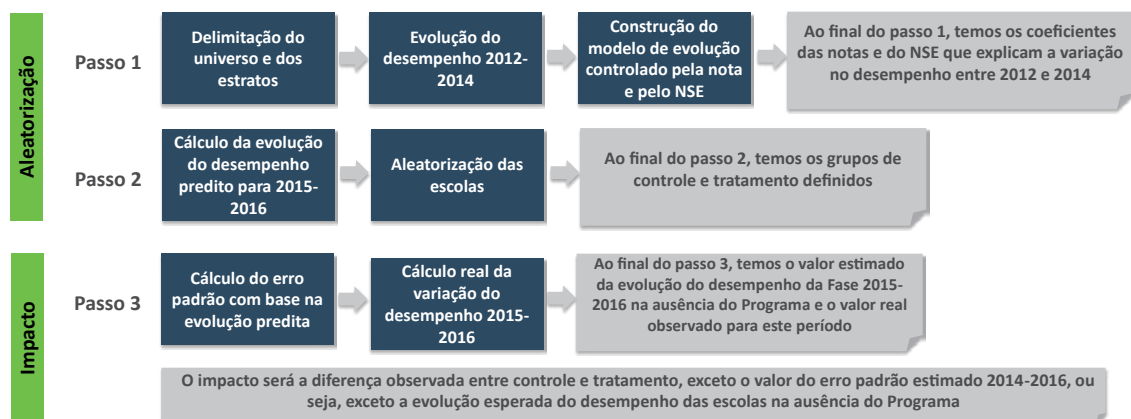
² O desenho original do teste considera a evolução 2010-2012 para a predição 2012-2014. Entretanto, o intervalo utilizado varia conforme a disponibilidade de dados da praça.

Aleatorização e pareamento

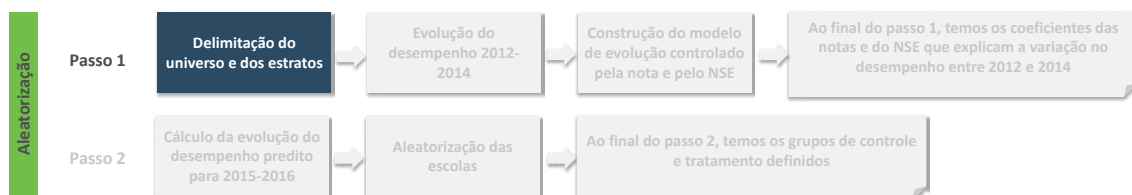
Respeitando o desenho do Programa, a característica experimental da avaliação e a necessidade de homogeneidade dos grupos de controle e trata-

mento, a aleatorização foi realizada em dois passos, e o terceiro passo é o cálculo do impacto, conforme esquema abaixo:

PANORAMA GERAL DO PROCESSO



Passo 1: Delimitação do universo e dos estratos



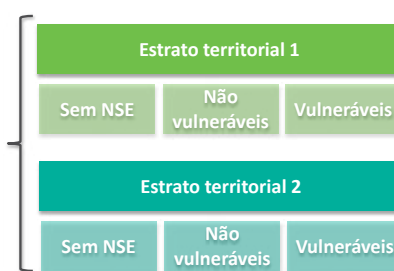
O passo 1 consistiu em delimitar o universo e os estratos de escolas que participarão do Programa na Fase 2015-2016 e na avaliação de impacto ao final de 2016. O universo e a estratificação de escolas foram definidos de acordo com a capacidade de gestão e prioridade de atendimento de cada Secretaria. Sendo assim, cada praça de intervenção do Programa possui uma particularidade distinta. A única contrapartida exigida pelo Instituto Unibanco foi que a Secretaria mantivesse um número mínimo de escolas não participantes da Fase 2015-2016 para viabilizar a avaliação de impacto ao final de 2016. Este grupo não participante da Fase 2015-2016 receberá o Programa nos dois anos seguintes (Fase 2017-2018).

Os estratos foram organizados em dois níveis. O primeiro respeitando a característica territorial do estado (conjunto de municípios, conjunto de regionais de ensino etc.) de acordo com a realidade de cada Secretaria. O segundo nível, após a delimitação territorial, foi a estratificação segundo vulnerabilidade das escolas, realizado da seguinte forma:

Estratificação Nível 2 – segundo vulnerabilidade da escola:

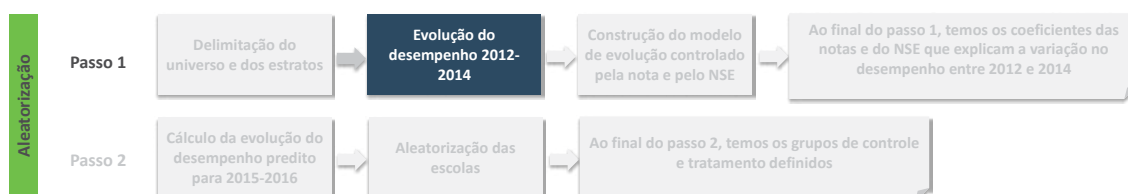
- i. A Secretaria e o Instituto Unibanco determinaram a quantidade de escolas vulneráveis que tem condições de atender durante a intervenção do Programa;
- ii. Para cada escola, calculou-se o percentil ao qual ela pertence para três diferentes indicadores: Nível Socioeconômico da Escola (NSE), nota média entre Língua Portuguesa e Matemática em 2014 (linha de base do Programa) e variação da nota média entre essas matérias no período de 2012 e 2014 (evolução recente observada do desempenho);
- iii. Calculou-se o valor máximo entre os percentis dos três indicadores;
- iv. Dentro de cada estrato territorial, ordenou-se as escolas segundo o menor valor observado nos percentis dos três indicadores; e
- v. Selecionou-se as escolas até atingir a quantidade de escolas vulneráveis determinada por cada Secretaria.

- i. Escolas que não possuem NSE ou alguma das notas são retiradas dessa seleção de vulnerabilidade e formam o estrato Sem NSE.
- ii. Escolas com NSE e não selecionadas como vulneráveis formam o estrato Não Vulneráveis.
- iii. Escolas com NSE e selecionadas como vulneráveis formam o estrato Vulneráveis.



A aleatorização é feita dentro de cada estrato do nível 2

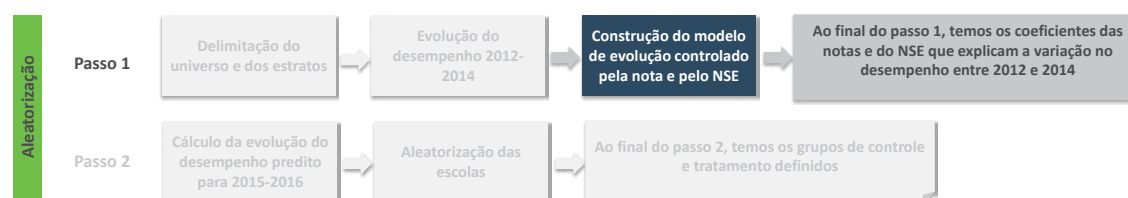
Passo 1: Evolução do desempenho 2012-2014



Com o universo escolhido, calculamos o desempenho das escolas nas avaliações estaduais no período de, no mínimo, dois anos antes da implantação da primeira fase do Programa. Sabendo que a primeira fase teria início em 2015, a evolução do desempenho considerada deveria ser, ao menos, 2012-2014.

Com base na variação do desempenho entre 2012 e 2014, é possível determinar a semelhança entre as escolas de controle e tratamento, garantindo a homogeneidade destes grupos. Isto é necessário porque partimos da hipótese de que escolas semelhantes são aquelas que possuem evolução semelhante no desempenho escolar.

Passo 1: Construção do modelo de evolução controlado pela nota e pelo NSE



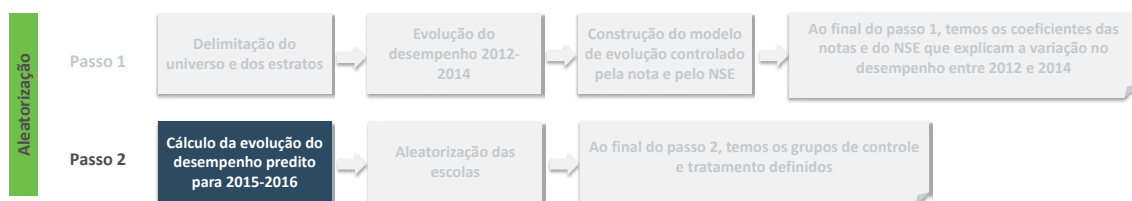
Para garantir que escolas com evolução do desempenho semelhante fossem efetivamente parecidas, a análise desta evolução foi controlada por três variáveis: (i) as notas das provas de Língua Portuguesa e Matemática em 2012, garantindo assim que as escolas partiram de um ponto semelhante do desempenho, e (ii) o Nível Socioeconômico da Escola (NSE), medido pelo índice NSE, garantindo assim que as escolas tinham realidades socioeconômicas semelhantes.³

Dessa forma, as escolas com evolução semelhantes entre 2012 e 2014 também tinham desempenho

semelhante em 2012 e características socioeconômicas semelhantes. Com isso, construímos um modelo de evolução controlado pela nota nas disciplinas de Língua Portuguesa e Matemática e pelo NSE: a variação do desempenho médio nas provas entre 2012 e 2014 (variável dependente) foi determinado pela nota em Língua Portuguesa em 2012 (ano inicial), pela nota em Matemática em 2012, e pelo NSE (variáveis independentes). Os coeficientes desta regressão correspondem, portanto, ao quanto da variação do desempenho entre 2012 e 2014 pode ser explicado pelas variáveis independentes.

³ Escolas que não possuem notas para cálculo dos coeficientes do modelo permanecem na aleatorização, assumindo o valor do coeficiente estimado considerando as demais escolas.

Passo 2: Cálculo da evolução do desempenho predito para 2015-2016

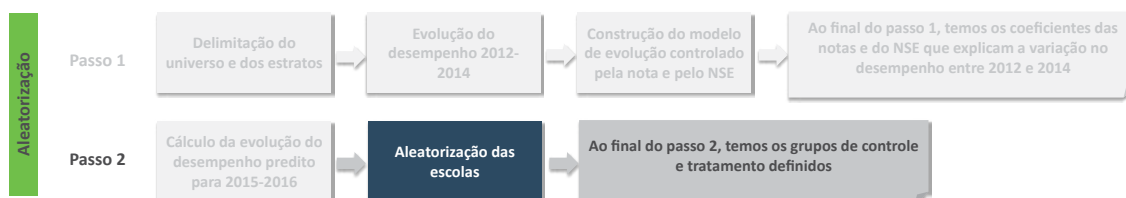


Com os coeficientes do modelo 2012-2014, assumimos que a evolução futura (2015-2016) terá comportamento semelhante à observada em 2012-2014. Assim, com base nestes coeficientes, estimamos a *evolução do desempenho predito para 2015 e 2016* de cada escola caso não houvesse a intervenção do Programa. Esta estimativa é controlada pelas notas de Língua Portuguesa e Matemática da linha de base do Programa (2014) e pelo NSE das escolas. Dessa forma, cada escola do universo passa a ter uma evolução de desempenho esperada (predita) para a Fase 2015-2016 caso não houvesse a intervenção do Programa. Nos testes realizados o predito foi eleito o melhor modelo de previsão da

variação futura dos resultados da escola na ausência do programa.

Com o valor predito da variação do desempenho em 2016, chegamos à evolução que as escolas teriam na ausência do Programa. Estimando um erro padrão com base neste valor podemos, ao final da intervenção, isolar o impacto do Programa que será dado pela diferença entre o desempenho 2015-2016 observada entre os grupos de controle e tratamento. O impacto torna-se identificável quando a evolução é superior ao erro padrão. Escolas sem a informação das notas de 2014 (linha de base) não farão parte da avaliação.

Passo 2: Aleatorização das escolas

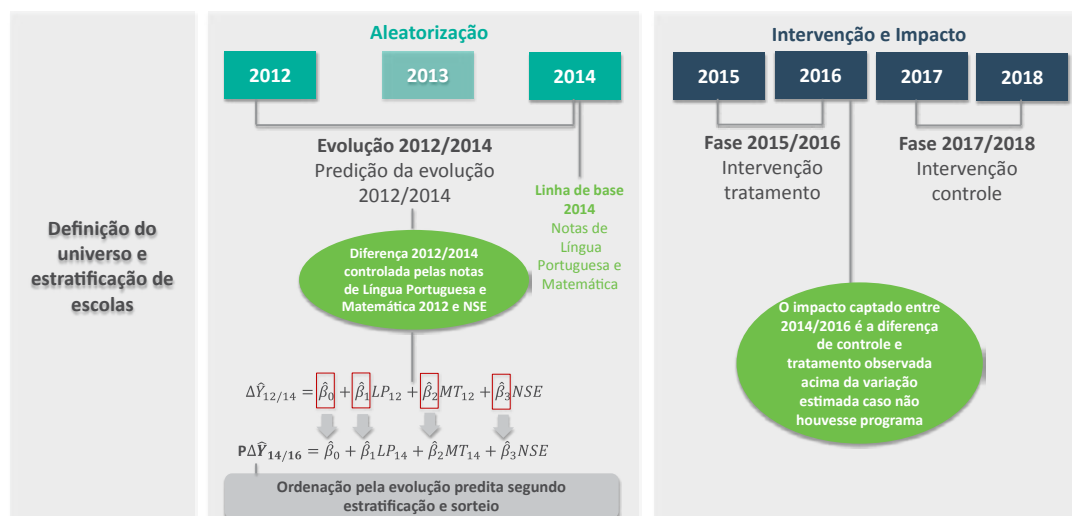


A evolução predita do desempenho é utilizada para ordenar as escolas dentro de cada estrato. A quantidade de grupos formados dentro de cada estrato foi determinada pela quantidade de escolas cujas Secretarias estavam dispostas a não participar da intervenção da Fase 2015-2016 do Programa.

A aleatorização foi realizada mantendo um controle para cada grupo de escolas e os tratamentos foram distribuídos de forma equilibrada entre os grupos.

Exemplo: se o universo era composto por 100 escolas e a Secretaria determinou que apenas 20 seriam controle, os grupos de escolas seriam compostos por quatro escolas de tratamento e uma de controle (totalizando 20 grupos). Se o universo era composto por 90 escolas e apenas 25 seriam controle, os grupos seriam formados de duas formas: por duas escolas de tratamento e um de controle (10 grupos) e por três escolas de tratamento e uma de controle (15 grupos).

QUADRO RESUMO



Validação da aleatorização

A validação da aleatorização foi feita com base em Testes T de igualdade entre médias e Kolmogorov-Smirnov. A diferença observada para os indicadores de desempenho das escolas de

tratamento e controle devem ser estatisticamente não significantes de forma que os grupos de tratamento e controle possam ser considerados iguais, a um nível de significância de 5%. Resultados diferentes entre os testes resultam em análise inconclusiva.

No Teste T de igualdade entre médias

Consideramos d_{ij} a diferença observada nos rendimentos dos alunos entre as escolas i e j . Admitindo como hipótese nula que a diferença entre as médias é zero ($H_0: m_d=0$) e, por conseguinte, como hipótese alternativa a existência da diferença entre as médias ($H_0: m_d \neq 0$), o Teste T é dado por:

$$T = \frac{\bar{d} - m_d}{\frac{s_d}{\sqrt{n}}}$$

Onde \bar{d} é a média das diferenças entre os pares, s_d é o desvio padrão estimado com base na variância amostral das diferenças e acertar no total de observações pareadas, com base em uma significância de 10% (confiança de 90%). Aceitando a hipótese nula, concluímos que os pares são estatisticamente iguais, ou seja, a diferença entre as médias pode ser considerada nula.

Teste de Kolmogorov-Smirnov (KS)

Baseado na distância máxima vertical entre as frequências acumuladas da função de distribuição de cada amostra aleatória independente de tamanhos n_1 e n_2 e funções de distribuição $Y(X_1)$ e $Z(X_2)$. O teste verifica se existe ou não diferença estatística entre as variáveis estudadas com base em duas hipóteses: H_0 – não há diferença entre as amostras [$Y(X_1) = Z(X_2)$]; e H_1 – há diferença entre as amostras [$Y(X_1) \neq Z(X_2)$].

Como o teste é sobre a máxima distância vertical [$KS = \max |Y(X_1) - Z(X_2)|$], ele demonstra se a maior distância observável entre as escolas é estatisticamente significativa para que elas sejam consideradas diferentes, comparado ao valor ks tabulado ou p -valor da estatística. Se aceitarmos a hipótese nula concluímos que a maior diferença observada pode ser considerada nula e, portanto, as amostras podem ser consideradas estatisticamente iguais.

2. Aleatorização por praça: Piauí

Testes

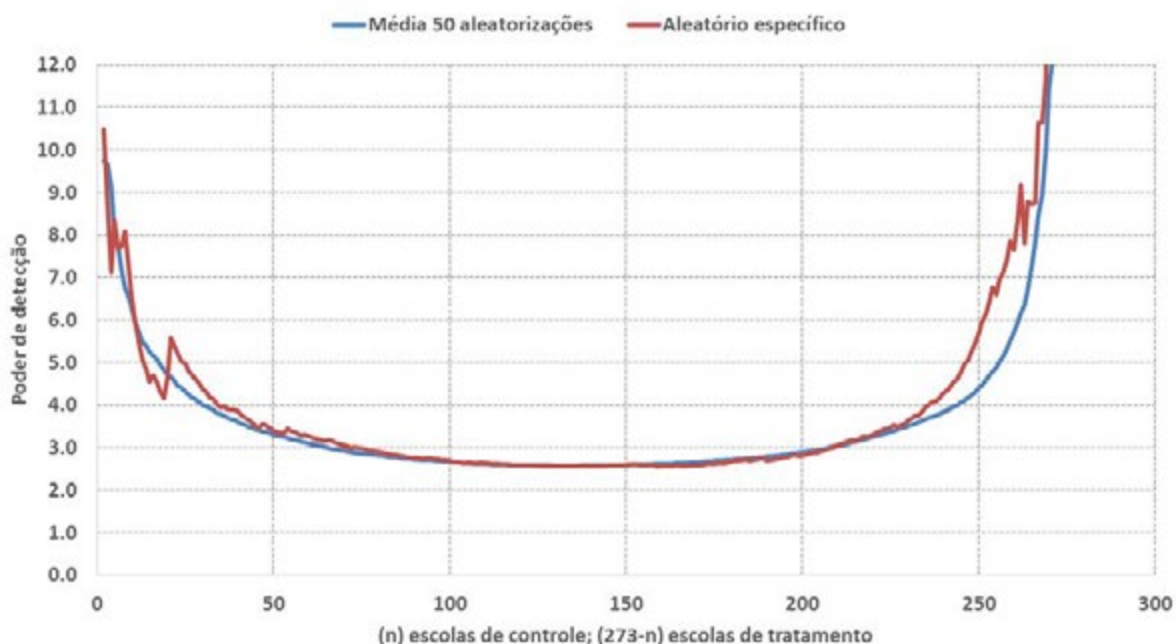
O método de pareamento via predição da evolução do desempenho 2014 a 2016 foi o método escolhido para o processo de sorteio das escolas do Piauí. Para o teste foi considerado um conjunto de 273 escolas e os modelos testados foram: (i) o de predição, controlado pela nota e pelo NSE das escolas para o pareamento e seleção de grupos de tratamento e controle; e (ii) o aleatório de seleção de tratamento e controle.

O critério de escolha do melhor modelo se deu com base no menor poder de detecção do impacto, ou

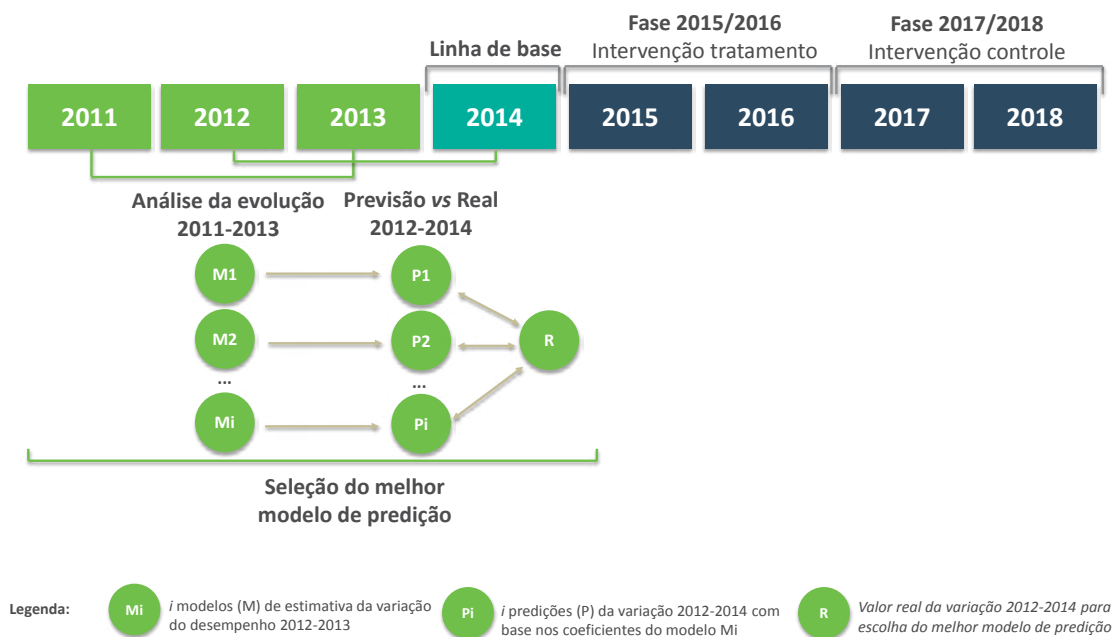
seja, aquele modelo que captasse de forma mais sensível o impacto do Programa. Sendo assim, a diferença futura observada entre controle e tratamento acima do valor do poder de detecção poderá ser atribuída ao impacto do Programa.

Considerando 123 escolas de controle e 150 de tratamento, o poder de detecção do modelo de predição (i) foi de 2,8. Já o modelo aleatório, para a mesma quantidade de escolas de controle e tratamento, obteve um poder de detecção de 2,6. Devido à proximidade entre os resultados, optou-se por manter o método de predição para o pareamento a fim de utilizar a metodologia que melhor se ajustou às demais praças. Assim, o modelo escolhido para dar continuidade ao processo de aleatorização foi o (i) modelo de predição controlado pela nota e pelo NSE das escolas.

PODER DE DETECÇÃO DE ACORDO COM BALANCEAMENTO ENTRE O NÚMERO DE CONTROLES E DE TRATAMENTOS: PIAUÍ, 273 ESCOLAS – PAREAMENTO ALEATÓRIO



QUADRO TEÓRICO DO MODELO DE PREVISÃO: ESTRUTURA DO TESTE, PIAUÍ



Note que para o Piauí, o desenho original do modelo de teste de previsão precisou ser construído para o intervalo de dois anos 2011-2013 para o cálculo teste do predito de 2012-2014.

Esta sobreposição dos intervalos de testes foi necessária pela ausência de dados de 2010 e para manter o desenho de 2 anos de evolução do modelo.

Delimitação de universos e estratos

Nesta etapa de redesenho do Programa Jovem de Futuro (Fases 2015-2016 e 2017-2018), o acordo entre

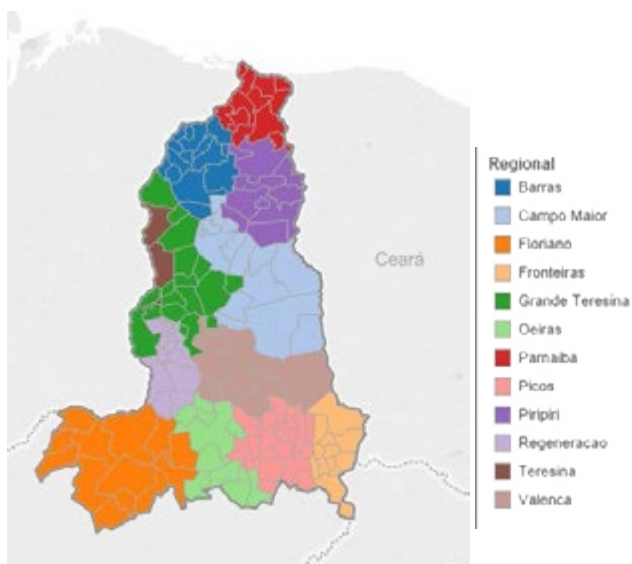
o Instituto Unibanco e a Secretaria de Estado de Educação culminou em um grupo de 275 escolas aptas a participarem do Programa.

Os critérios utilizados pela Secretaria para a delimitação do universo foram, nesta ordem:

- (i) Fazer parte de alguma das 15 regionais de ensino escolhidas (segundo a capacidade de atuação da Secretaria);
- (ii) Escolas com mais de 100 alunos matriculados no Ensino Médio regular em 2015.

Tabela – Universo de escolas e Atendimento: Piauí, avaliação JF

Seleção	Número de escolas	Número de grupos			Atendimento		Vulneráveis	
		Total	Pares	Trios	T	C	Pares	Escolas
Total de escolas no arquivo	468	-	-	-	-	-	-	-
Entre as 15 regionais selecionadas	371	-	-	-	-	-	-	-
Com pelo menos 100 alunos no EM em 2015	275	134	131	3	137	134	10	20
Teresina	73	35	33	2	37	35	0	0
Norte	60	29	29	0	29	29	2	4
Centro	65	32	32	0	32	32	4	8
Sul	77	38	37	1	39	38	4	8



Regional	Região	
1	PARNAÍBA	NORTE
2	BARRAS	NORTE
3	PIRIPIRI	NORTE
5	CAMPO MAIOR	CENTRO
7	VALENÇA	CENTRO
18	GRANDE TERESINA	CENTRO
6	REGENERAÇÃO	SUL
8	OEIRAS	SUL
9	PICOS	SUL
10	FLORIANO	SUL
16	FRONTEIRAS	SUL
4	TERESINA CENTRO NORTE	TERESINA
19	TERESINA SUL	TERESINA
20	TERESINA NORDESTE	TERESINA
21	TERESINA SUDESTE	TERESINA

Inicialmente, estas 275 escolas foram estratificadas em quatro grupos, cada um formado por um conjunto de regionais de ensino.

Dentre as 275 escolas selecionadas para o universo, quatro não possuíam evolução do desempenho 2012-2014. Com isso, a linha de base não pôde ser determinada e, portanto, elas deixam de fazer parte da avaliação, reduzindo o grupo de escolas avaliadas para 271 escolas. As escolas retiradas da avaliação estão listadas a seguir. Elas permanecem na

aleatorização apenas para definir em qual fase receberão o Programa.

- 22017011 – U.E. Presidente Castelo Branco (NORTE);
- 22001557 – U.E. Honorina Tito (NORTE);
- 22017992 – U.E. Afonso Mafrense (CENTRO);
- 22028021 – U.E. Cristino Castelo Branco (TERESINA).

Tabela – Estratificação nível 1 – critério territorial: Piauí

Região	Escolas			Grupos	
	Total	Controle	Tratamento	Pares	Trios
Teresina	72	35	37	33	2
Norte	58	29	29	29	0
Centro	64	32	32	32	0
Sul	77	38	39	37	1
Total	271	134	137	131	3

No segundo nível de estratificação, determinação de escolas vulneráveis, a Secretaria determinou que o atendimento deveria ser limitado a 20 escolas.

- Teresina – não possui estrato vulnerável;
- Norte – 2 pares, 4 escolas;
- Central – 4 pares, 8 escolas;
- Sul – 4 pares, 8 escolas.

Tabela – Estratificação nível 2 – critério de vulnerabilidade: Piauí

Código da Escola	Escola	Região	Regional (GRE)
22135383	UNID ESC PEDRO MENDES PESSOA	CENTRO	GRANDE TERESINA
22029869	UNID ESC PROFA ELISA SOUSA	CENTRO	GRANDE TERESINA
22132740	UNID ESC MENINO JOÃO PEDRO	CENTRO	GRANDE TERESINA
22030158	UNID ESC FILINTO REGO	CENTRO	GRANDE TERESINA
22019294	UNIDADE ESCOLAR RAIMUNDO MARTINS	CENTRO	GRANDE TERESINA
22048456	UNID ESC JOÃO DE DEUS CARVALHO	CENTRO	VALENÇA
22135901	UNIDADE ESCOLAR CEZAR LEAL	CENTRO	GRANDE TERESINA
22034072	UNID ESC BRIOLANJA OLIVEIRA	CENTRO	CAMPO MAIOR
22011170	UNID ESC SÃO JOSÉ	NORTE	BARRAS
22126880	UNID ESC SANTA TERESINHA	NORTE	BARRAS
22007431	UNID ESC JOSÉ AMAVEL	NORTE	BARRAS
22137556	UNIDADE ESCOLAR JOÃO ODORICO	NORTE	BARRAS
22083480	UNID ESC DOM EDILBERTO DINKELBORG	SUL	OEIRAS
22135030	CENTRO EDUCACIONAL SEBASTIÃO DE SOUSA	SUL	FRONTEIRAS
22136509	UNID ESC JOÃO ANTONIO DA VERA	SUL	PICOS
22134484	UNIDADE ESCOLAR HELVIDIO NUNES	SUL	PICOS
22086528	UNIDADE ESCOLAR NOSSA SENHORA DO PATROCÍNIO	SUL	FRONTEIRAS
22060596	UNIDADE ESCOLAR JOSÉ SALUSTIANO DA SILVA	SUL	FLORIANO
22083065	UNID ESC CLEMENTINO MARTINS	SUL	PICOS
22083189	UNID ESC JOAQUIM BORGES DE OLIVEIRA	SUL	PICOS

A delimitação final do universo de escolas para aleatorização de grupos de controle e trata-

mento contabilizou 271 escolas divididas em 134 grupos:

Tabela – Estratificação nível 1 – critério territorial: Piauí

Região	Escolas			Grupos		
	Total	Controle	Tratamento	Pares não vulneráveis	Tríos não vulneráveis	Pares vulneráveis
Teresina	72	35	37	33	2	0
Norte	58	29	29	27	0	2
Centro	64	32	32	28	0	4
Sul	77	38	39	33	1	4
Total	271	134	137	121	3	10

Dentre as 271 escolas participantes do processo de aleatorização para seleção dos grupos de controle e tratamento; 232 foram pareadas com base na predição da variação do desempenho para o período 2014-2016; 20 com base na vulnerabilidade; 19 por não terem informações suficientes para a predição 2014-2016 e não serem vulneráveis, foram pareadas de forma aleatória.

O modelo 2012-2014 para a estimativas dos coeficientes para o preditor da evolução do desempenho entre 2014 e 2016 gerou as seguintes estatísticas. Com o valor predito calculado, as escolas são ordenadas dentro de seus estratos e aleatorizadas.

Resumo de resultados

Estatística de regressão	
R múltiplo	0.332
R-Quadrado	0.110
R-Quadrado ajustado	0.099
Erro padrão	12.127
Observações	246

ANOVA

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	3	4417.952	1472.651	10.014	0.000
Resíduo	242	35589.096	147.062		
Total	245	40007.048			

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95.0%</i>	<i>Superior 95.0%</i>
Interseção	57.10134	12.76004	4.47501	0.00001	31.96641	82.23627	31.96641	82.23627
NSE	-1.57311	3.57610	-0.43990	0.66040	-8.61736	5.47114	-8.61736	5.47114
LP 2012	-0.32077	0.09797	-3.27413	0.00121	-0.51375	-0.12778	-0.51375	-0.12778
MT 2012	0.08149	0.10106	0.80627	0.42088	-0.11759	0.28056	-0.11759	0.28056

Validação

Seguem abaixo tabelas com os valores do Teste T de igualdade de médias e o teste Kolmogorov-Smirnov.

Tabela – Testes de Média e de Kolmogorov para os resultados da linha de base de Matemática entre as escolas de Tratamento e Controle do Projeto Jovem de Futuro: Piauí, 2014

Indicadores	Média (%)		Diferença	P-valor do teste T de Igualdade médias	Estatísticas do teste de Kolmogorov	P-valor do teste T Kolmogorov
	Tratamento	Controle				
Indicadores de desempenho						
Média	242.57	242.89	-0.32	88	0.10	53
Mediana	238.82	238.95	-0.13	95	0.11	35
1ª quartil	213.32	214.22	-0.90	64	0.10	45
3ª quartil	268.25	267.17	1.07	69	0.10	56
Desvio padrão	40.46	39.96	0.50	56	0.12	28
Melhor desempenho						
Média entre os 25% melhores	296.95	297.15	-0.20	95	0.13	24
Média entre os 50% melhores	274.98	274.45	0.53	84	0.11	39
Média entre os 75% melhores	259.00	258.96	0.05	98	0.12	31
Porcentagem de alunos com nota acima de 317 (meta do Todos pela educação)	7.05	5.95	1.10	27	0.07	86
Porcentagem de alunos com nota acima de 300	10.93	10.15	0.78	52	0.08	81
Porcentagem de alunos com nota acima de 350	2.55	1.61	0.94	14	0.06	96
Porcentagem de alunos com nota acima de 275	20.94	21.13	-0.19	90	0.10	47
Pior desempenho						
Média entre os 25% piores	196.09	197.30	-1.20	40	0.13	22
Média entre os 50% piores	211.59	212.69	-1.10	52	0.14	13
Porcentagem de alunos com nota abaixo de 200	17.76	16.06	1.70	9	0.10	52
Porcentagem de alunos com nota abaixo de 175	2.75	2.69	0.06	86	0.05	99
Porcentagem de alunos com nota abaixo de 150	0.00	0.00	0.00	0	0.00	100
Número médio de alunos que fizeram a prova	61	64	-4	50	0.09	60

DIFERENÇA DE DESEMPENHO NA 3ª SÉRIE EM MATEMÁTICA ENTRE OS ESTRATOS DE TRATAMENTO E CONTROLE: PIAUÍ, LINHA DE BASE – 2014

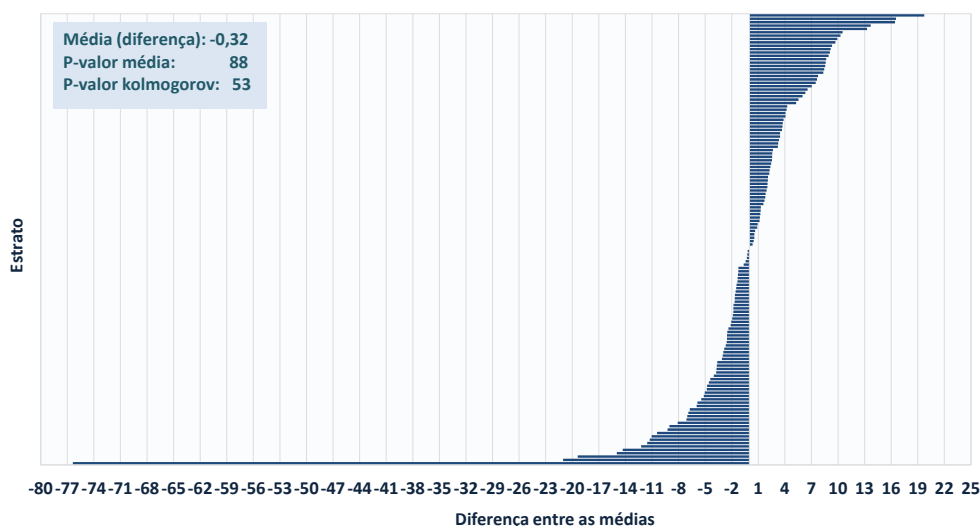
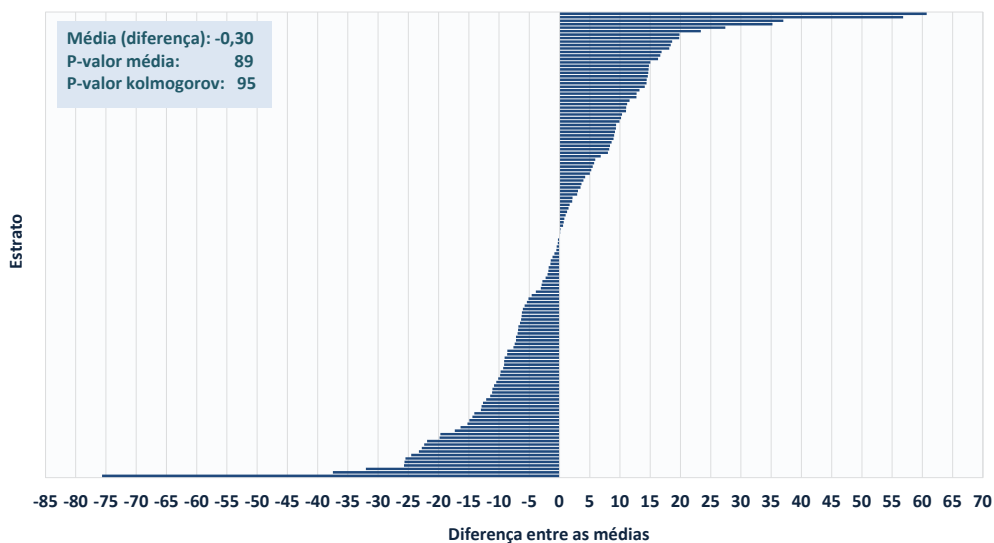


Tabela – Testes de Média e de Kolmogorov para os resultados da linha de base de Língua Portuguesa entre as escolas de Tratamento e Controle do Projeto Jovem de Futuro: Piauí, 2014

Indicadores	Média (%)		Diferença	P-valor do teste T de Igualdade médias	Estatísticas do teste de Kolmogorov	P-valor do teste T Kolmogorov
	Tratamento	Controle				
Indicadores de desempenho						
Média	238,68	238,99	-0,30	89	0,06	95
Mediana	237,67	236,11	1,57	55	0,09	68
1ª quartil	206,63	206,14	0,48	84	0,08	77
3ª quartil	268,20	269,86	-1,67	53	0,13	19
Desvio padrão	43,45	44,49	-1,04	19	0,16	5
Melhor desempenho						
Média entre os 25% melhores	295,33	298,15	-2,83	26	0,14	14
Média entre os 50% melhores	273,96	275,27	-1,31	60	0,12	26
Média entre os 75% melhores	257,04	257,74	-0,70	77	0,09	67
Porcentagem de alunos com nota acima de 283 (meta do Todos pela educação)	17,81	19,88	-2,06	18	0,13	20
Porcentagem de alunos com nota acima de 275	22,39	24,31	-1,92	26	0,11	39
Porcentagem de alunos com nota acima de 300	10,28	11,88	-1,61	17	0,13	18
Porcentagem de alunos com nota acima de 350	40,29	39,97	0,32	87	0,06	95
Pior desempenho						
Média entre os 25% piores	187,07	186,46	0,61	76	0,06	98
Média entre os 50% piores	205,43	204,58	0,85	69	0,06	99
Porcentagem de alunos com nota abaixo de 200	22,92	23,70	-0,78	63	0,06	94
Porcentagem de alunos com nota abaixo de 175	8,51	8,60	-0,09	92	0,10	48
Porcentagem de alunos com nota abaixo de 150	1,59	1,11	0,48	14	0,14	17
Número médio de alunos que fizeram a prova	53	56	-3	53	0,09	69

DIFERENÇA DE DESEMPENHO NA 3ª SÉRIE EM LÍNGUA PORTUGUESA ENTRE OS ESTRATOS DE TRATAMENTO E CONTROLE: PIAUÍ, LINHA DE BASE – 2014



3. Aleatorização por praça: Pará

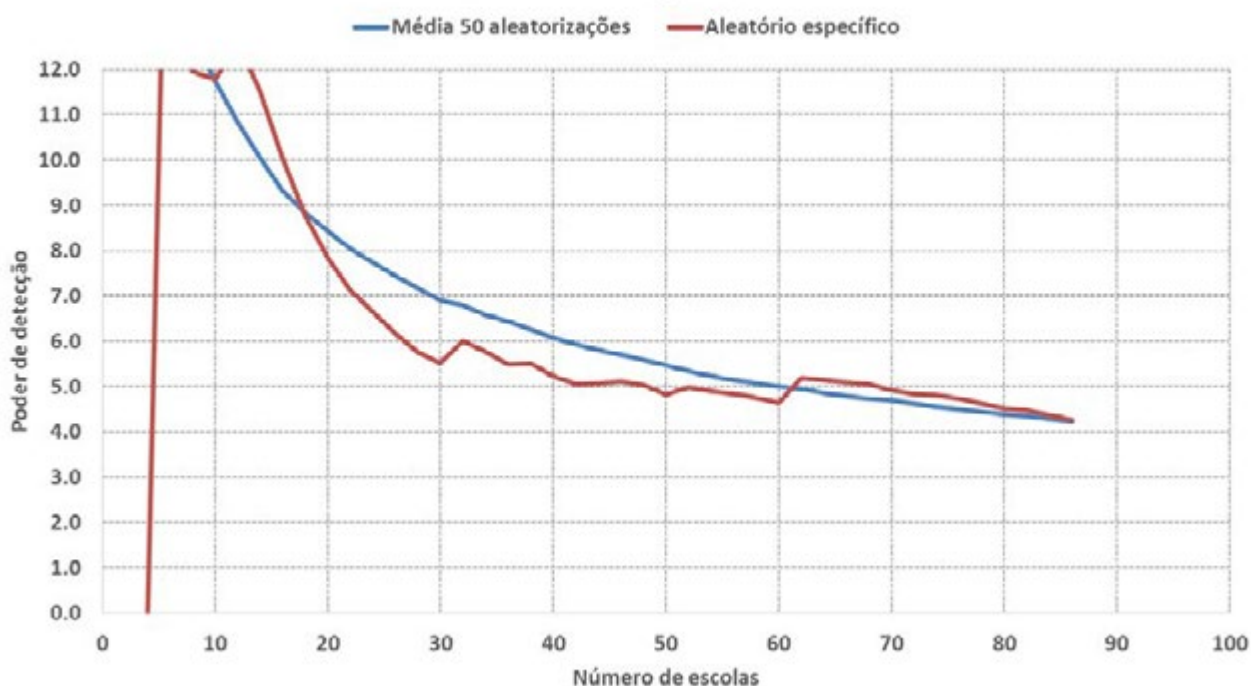
Testes

O método de pareamento via predição da evolução do desempenho 2014 a 2016 foi considerado o melhor método de pareamento das escolas do Pará. Para o teste foi considerado um conjunto de 86 escolas e os modelos testados foram: (i) o de predição, controlado pela nota e pelo NSE das escolas, para o pareamento e seleção de grupos de tratamento e controle; e (ii) o aleatório de seleção de tratamento e controle.

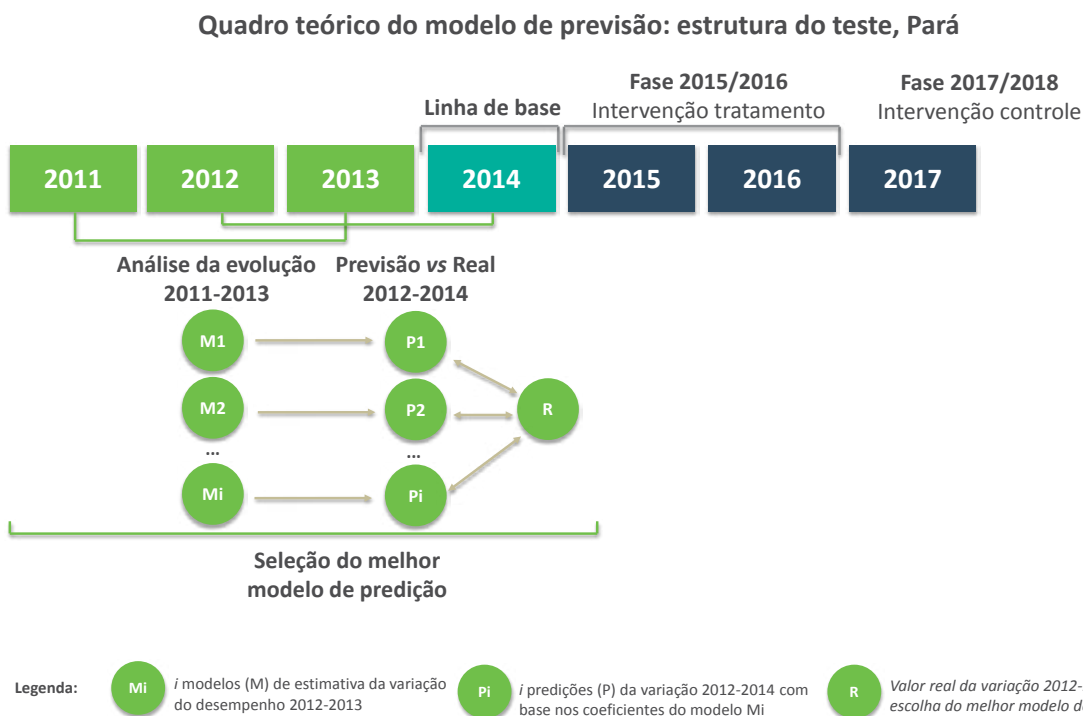
O critério de escolha do melhor modelo deu-se com base no menor poder de detecção do impacto, ou seja, aquele modelo que captasse de forma mais sensível o impacto do Programa. Sendo assim, a diferença futura observada entre controle e tratamento acima do valor do poder de detecção poderá ser atribuída ao impacto do Programa.

Considerando 42 escolas de controle e 44 de tratamento, o poder de detecção do modelo de predição (i) foi 4,2. Já o modelo aleatório, para a mesma quantidade de escolas de controle, obteve um poder de detecção de 5,9. Assim, o modelo escolhido para dar continuidade ao processo de aleatorização foi o (i) o de predição controlado pela nota e pelo NSE das escolas.

PODER DE DETECÇÃO SEGUNDO O NÚMERO DE ESCOLAS CONSIDERANDO PAREAMENTO ENTRE CONTROLE E TRATAMENTO: PARÁ - PAREAMENTO ALEATÓRIO



QUADRO TEÓRICO DO MODELO DE PREVISÃO: ESTRUTURA DO TESTE, PARÁ



Delimitação de universos e estratos

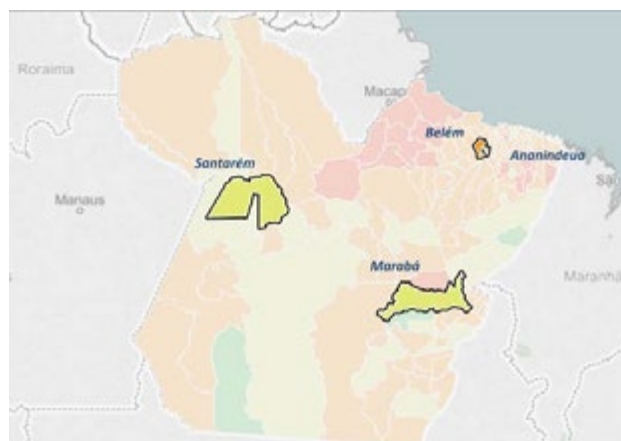
Nesta etapa de redesenho do Programa Jovem de Futuro (Fases 2015-2016 e 2017-2018), o acordo entre o Instituto Unibanco e a Secretaria de Estado de Educação culminou em um grupo de 87 escolas aptas a participarem do Programa.

às escolas vulneráveis nos quatro municípios escolhidos para participar do Programa, optou-se, nesta praça, por não estratificar as escolas por vulnerabilidade.

Dessa forma, no Pará, o único nível de estratificação é o territorial (nível 1).

Tabela – Universo de escolas e Atendimento: Pará, avaliação JF

Seleção	Número de escolas	Número de grupos		
		Total	Pares	Trios
Total de escolas elegíveis	87	42	39	3
Belém	47	23	22	1
Ananindeua	16	8	8	0
Marabá	13	6	5	1
Santarém	11	5	4	1



Inicialmente, estas 87 escolas foram estratificadas em quatro grupos, cada um formado por um município.

Devido à quantidade limitada de supervisores disponíveis no estado do Pará para dar suporte

A delimitação final do universo de escolas para aleatorização de grupos de controle e tratamento contabilizou 87 escolas divididas em 42 grupos. As 87 escolas foram aleatorizadas utilizando o critério de previsão.

Tabela – Universo de escolas e Atendimento: Pará, avaliação JF

Seleção	Número de escolas			Número de grupos		
	Total	Controle	Tratamento	Total	Pares	Trios
Belém	47	23	24	23	22	1
Ananindeua	16	8	8	8	8	0
Marabá	13	6	7	6	5	1
Santarém	11	5	6	5	4	1
Total	87	42	45	42	39	3

O modelo 2012-2014 para as estimativas dos coeficientes para o preditor da evolução do desempenho entre 2014 e 2016 gerou as seguintes

estatísticas. Com o valor predito calculado, as escolas são ordenadas dentro de seus estratos e aleatorizadas.

Resumo de resultados

Estatística de regressão	
R múltiplo	0.538
R-Quadrado	0.290
R-quadrado ajustado	0.263
Erro padrão	10.024
Observações	84

ANOVA

	gl	SQ	MQ	F	F de significação
Regressão	3	3282.854	1.094.285	10.890	0.000
Resíduo	80	8038.974	100.487		
Total	83	11321.827			

	Coefficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores	Inferior 95.0%	Superior 95.0%
Interseção	96.1622	21.3991	4.4938	0.0000	53.5766	138.7477	53.5766	138.7477
NSE	-0.5705	0.1825	-3.1256	0.0025	-0.9337	-0.2073	-0.9337	-0.2073
LP 2012	0.1425	0.1926	0.7400	0.4615	-0.2407	0.5257	-0.2407	0.5257
MT 2012	20.3065	9.2912	2.1856	0.0318	1.8163	38.7966	1.8163	38.7966

Validação

Seguem abaixo tabelas com os valores do Teste T de igualdade de médias e o teste Kolmogorov-Smirnov.

Tabela – Testes de Média e de Kolmogorov para os resultados da linha de base de Matemática entre as escolas de Tratamento e Controle do Projeto Jovem de Futuro: Pará, 2014

Indicadores	Média (%)		Diferença	P-valor do teste T de Igualdade médias	Estatísticas do teste de Kolmogorov	P-valor do teste T Kolmogorov
	Tratamento	Controle				
Indicadores de desempenho						
Média	237	239	-1,87	48	0,12	90
Mediana	233	237	-3,23	26	0,21	32
1ª quartil	208	210	-1,59	51	0,12	93
3ª quartil	262	264	-2,41	49	0,16	64
Desvio padrão	39	39	0,25	84	0,14	81
Melhor desempenho						
Média entre os 25% melhores	290	292	-1,59	69	0,17	54
Média entre os 50% melhores	269	271	-2,05	55	0,16	69
Média entre os 75% melhores	253	255	-2,13	49	0,14	82
Porcentagem de alunos com nota acima de 317 (meta do Todos pela educação)	0,04	0,04	0,00	91	0,13	89
Porcentagem de alunos com nota acima de 300	0,08	0,08	0,00	90	0,11	95
Porcentagem de alunos com nota acima de 350	0,01	0,01	0,00	39	0,10	99
Porcentagem de alunos com nota acima de 275	0,17	0,18	-0,01	66	0,16	69
Pior desempenho						
Média entre os 25% piores	193	195	-1,24	46	0,24	19
Média entre os 50% piores	207	209	-1,93	35	0,15	69
Porcentagem de alunos com nota abaixo de 200	0,19	0,19	0,00	79	0,16	67
Porcentagem de alunos com nota abaixo de 175	0,03	0,02	0,01	25	0,20	39
Porcentagem de alunos com nota abaixo de 150	0,00	0,00	0,00	0	0,00	100
Número médio de alunos que fizeram a prova	71	100	-29	3	0,23	20

DIFERENÇA DE DESEMPENHO NA 3ª SÉRIE EM MATEMÁTICA ENTRE OS ESTRATOS DE TRATAMENTO E CONTROLE: PARÁ, LINHA DE BASE – 2014

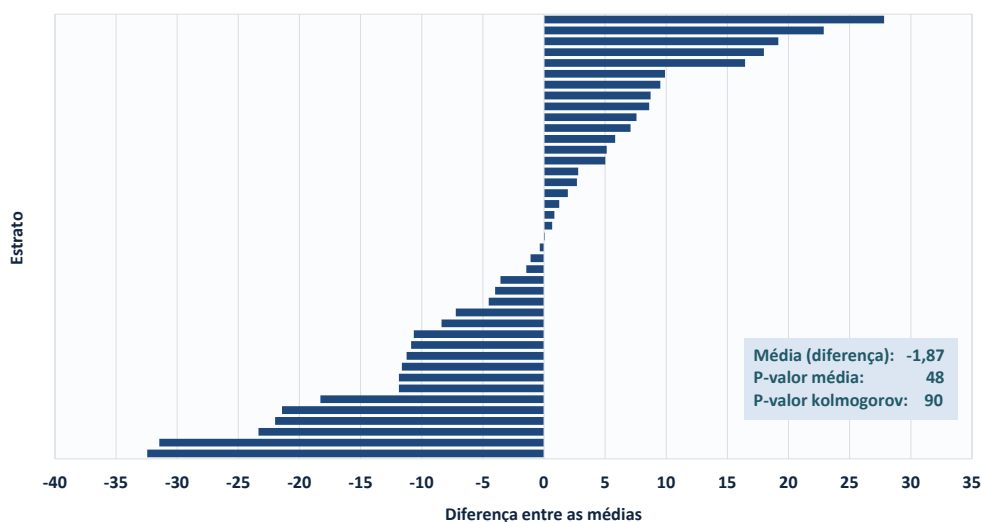
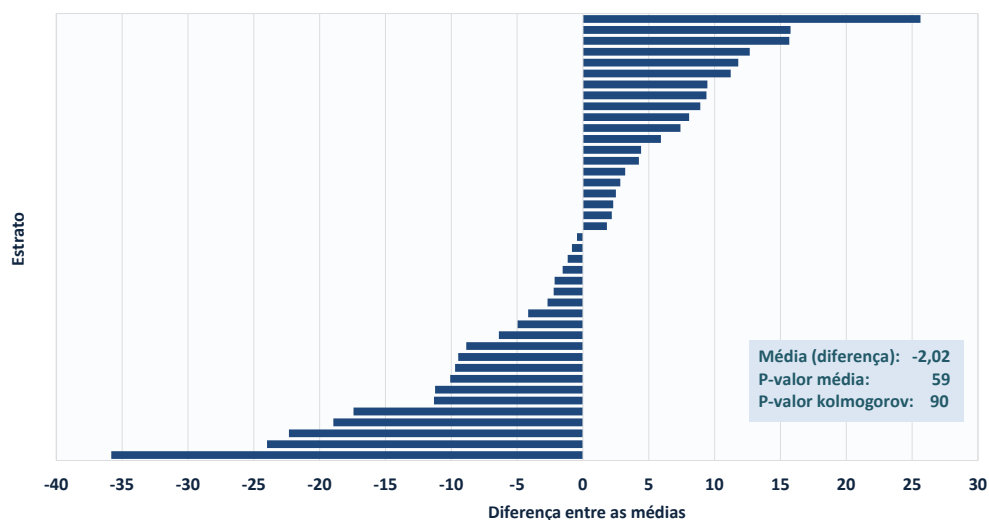


Tabela – Testes de Média e de Kolmogorov para os resultados da linha de base de Língua Portuguesa entre as escolas de Tratamento e Controle do Projeto Jovem de Futuro: Pará, 2014

Indicadores	Média (%)		Diferença	P-valor do teste T de Igualdade médias	Estatísticas do teste de Kolmogorov	P-valor do teste T Kolmogorov
	Tratamento	Controle				
Indicadores de desempenho						
Média	226	228	-2,02	59	0,12	90
Mediana	222	226	-4,02	38	0,13	87
1ª quartil	190	191	-1,40	70	0,15	75
3ª quartil	259	262	-2,48	60	0,14	77
Desvio padrão	48	48	-0,18	89	0,14	79
Melhor desempenho						
Média entre os 25% melhores	290	292	-1,48	76	0,10	99
Média entre os 50% melhores	265	267	-2,28	62	0,14	78
Média entre os 75% melhores	245	248	-2,79	51	0,15	75
Porcentagem de alunos com nota acima de 283 (meta do Todos pela educação)	0,16	0,16	0,00	98	0,12	91
Porcentagem de alunos com nota acima de 275	0,19	0,19	0,00	94	0,12	92
Porcentagem de alunos com nota acima de 300	0,09	0,09	0,00	80	0,15	74
Porcentagem de alunos com nota acima de 250	0,31	0,34	-0,03	35	0,14	78
Pior desempenho						
Média entre os 25% piores	170	170	-0,30	91	0,15	72
Média entre os 50% piores	188	191	-2,09	51	0,14	83
Porcentagem de alunos com nota abaixo de 200	0,35	0,33	0,01	60	0,11	97
Porcentagem de alunos com nota abaixo de 175	0,18	0,16	0,02	40	0,19	43
Porcentagem de alunos com nota abaixo de 150	0,04	0,04	0,00	80	0,09	99
Número médio de alunos que fizeram a prova	71	100	-29	3	0,23	20

DIFERENÇA DE DESEMPENHO NA 3ª SÉRIE EM LÍNGUA PORTUGUESA ENTRE OS ESTRATOS DE TRATAMENTO E CONTROLE: PARÁ, LINHA DE BASE – 2014



3. Aleatorização por praça: Espírito Santo

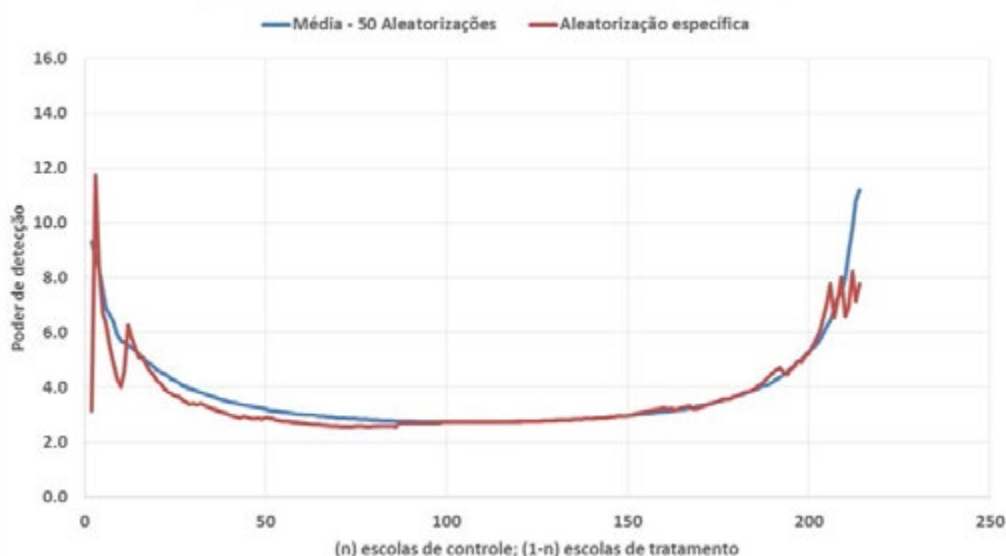
Testes

O método de pareamento via predição da evolução do desempenho 2014 a 2016 foi considerado o melhor para as escolas do Espírito Santo. Para o teste foi considerado um conjunto de 216 escolas e os modelos testados foram: (i) o modelo de predição, controlado pela nota e pelo NSE das escolas, para o pareamento e seleção de grupos de tratamento e controle; e (ii) modelo aleatório de seleção de tratamento e controle.

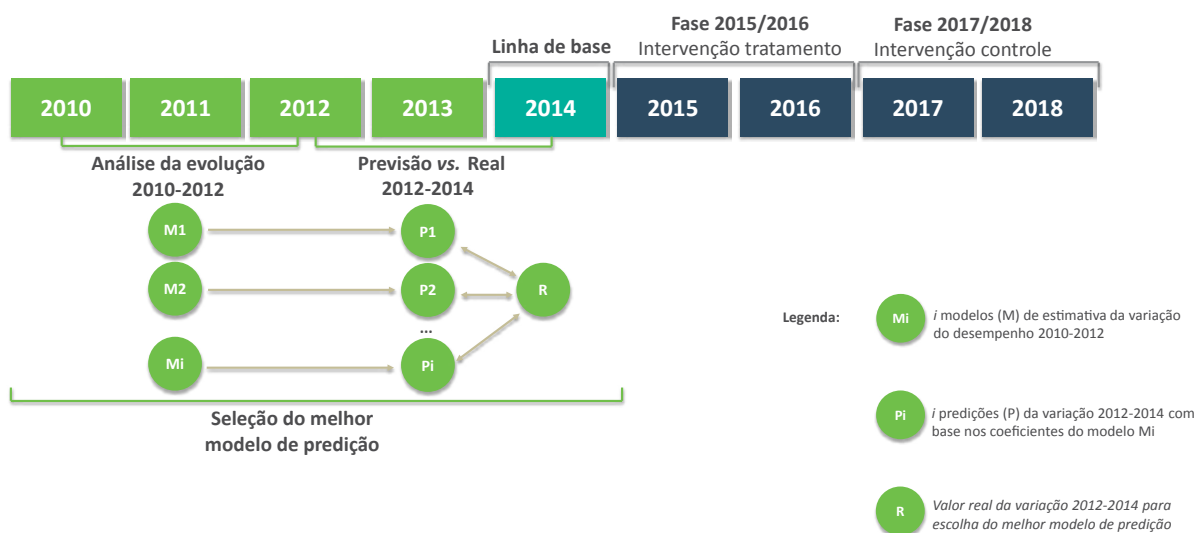
O critério de escolha do melhor modelo deu-se com base no menor poder de detecção do impacto, ou seja, aquele modelo que captasse de forma mais sensível o impacto do Programa. Sendo assim, a diferença futura observada entre controle e tratamento acima do valor do poder de detecção poderá ser atribuída ao impacto do Programa.

Considerando 60 escolas de controle, o poder de detecção do modelo de predição (i) foi 2,7. Já o modelo aleatório, para a mesma quantidade de escolas de controle, obteve um poder de detecção de 3,1. Assim, o modelo escolhido para dar continuidade ao processo de aleatorização foi o (i) de predição controlado pela nota e pelo NSE das escolas.

PODER DE DETECÇÃO NA VARIAÇÃO 2012-2014 DA NOTA MÉDIA (LP, MT) DE ACORDO COM A QUANTIDADE DE ESCOLAS DE CONTROLE: ESPÍRITO SANTO, 3º ANO DO EM



QUADRO TEÓRICO DO MODELO DE PREVISÃO: ESTRUTURA DO TESTE, ESPÍRITO SANTO



Para o Espírito Santo, pela disponibilidade de dados, os modelos de teste foram construídos para a variação 2010-2012 controlados pelo NSE e pela nota de 2011 e 2010.

Assim, o modelo de predição final (2014-2016) será controlado pelo NSE e pelas notas de 2012, 2013 e 2014.

Delimitação de universos e estratos

Nesta etapa de implantação do Programa Jovem de Futuro (Fases 2015-2016 e 2017-2018), o acordo entre

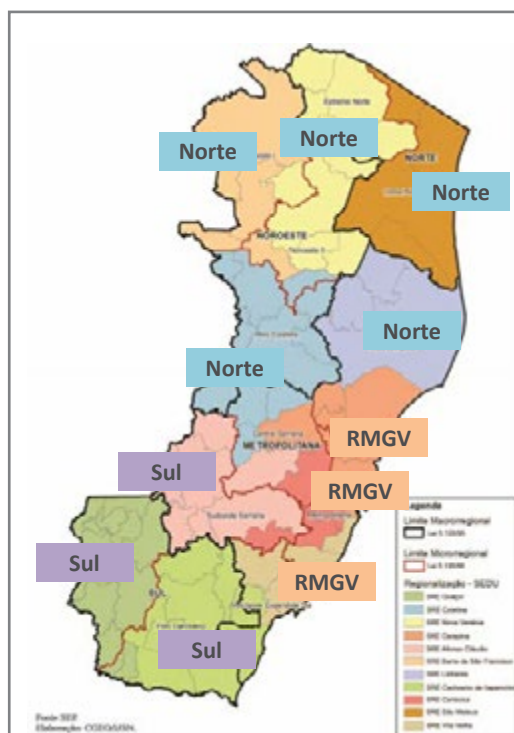
o Instituto Unibanco e a Secretaria de Estado de Educação culminou em um grupo de 222 escolas aptas a participar do Programa.

O critério utilizado pela Secretaria para a delimitação do universo era a escola ter, ao menos, 120 alunos matriculados no Ensino Médio regular em 2015 (ano de início da primeira fase do Programa).

Inicialmente, estas 222 escolas foram estratificadas em 4 grupos, cada um formado por um conjunto de regionais de ensino.

Tabela – Universo de escolas e Atendimento: Pará, avaliação JF

Seleção	Número de escolas	Número de grupos		
		Total	Grupos de 3	Grupos de 4
Total de escolas no arquivo	288	–	–	–
Com pelo menos 120 alunos no EM em 2015	222	70	59	11
Norte	69	22	19	3
RMGV	92	29	24	5
Sul	60	19	16	3
Sem nota em 2014	1	–	–	–



Regional	Região
SRE BARRA DE SÃO FRANCISCO	Norte
SRE COLATINA	Norte
SRE LINHARES	Norte
SRE NOVA VENÉCIA	Norte
SRE SÃO MATEUS	Norte
SRE CARAPINA	RMGV
SRE CARIACICA	RMGV
SRE VILA VELHA	RMGV
SRE AFONSO CLÁUDIO	Sul
SRE CACHOEIRO DE ITAPEMIRIM	Sul
SRE COMENDADORA JUREMA MORETZ SOHN	Sul

Dentre as 222 escolas selecionadas para o universo, uma não possuía nota. Com isso, a linha de base não pôde ser determinada e, portanto, elas deixam de fazer parte da avaliação, reduzindo o grupo de escolas avaliadas para 221 escolas. As escolas retiradas da avaliação estão listadas abaixo. Ainda está em processo de definição se esta escola sem nota entrará na primeira ou na segunda fase de intervenção do Programa:

- 32037163 – EEEFM Francisco Nascimento (RMGV);

A quantidade de controles foi, também, determinada pela Secretaria de Educação no acordo com o Instituto Unibanco. Como o Programa será implantado em toda a rede de acordo com os critérios de estratificação, a quantidade de escolas de controle será 70, a fim de garantir a possibilidade de avaliação de impacto do Programa.

Tabela – Estratificação – nível 1: Espírito Santo, avaliação JF

Seleção	Número de escolas	Número de grupos			Atendimento	
		Total	Grupos de 3	Grupos de 4	T	C
Total de escolas no arquivo	288	–	–	–	–	–
Com pelo menos 120 alunos no EM em 2015	222	70	59	11	151	70
Norte	69	22	19	3	47	22
RMGV	92	29	24	5	63	29
Sul	60	19	16	3	41	19
Sem nota em 2014	1	–	–	–	–	–

O segundo nível de estratificação, determinação de escolas vulneráveis, a Secretaria determinou que o atendimento deveria ser limitado a 30 escolas.

- Norte – 5 trios, 15 escolas;

- RMGV – 2 trios, 6 escolas; e
- Sul – 3 trios, 9 escolas.

Tabela – Estratificação nível 2 – critério de vulnerabilidade: Espírito Santo

Código da Escola	Escola	Região	Regional (GRE)
32015631	EEEM NOSSA SENHORA DE LOURDES	Norte	SRE NOVA VENÉCIA
32075936	EEEFM PROFª ANTONIETA BANHOS FERNANDES	Norte	SRE LINHARES
32013728	EEEFM SEBASTIANA GRILO	Norte	SRE COLATINA
32007876	EEEFM SÃO GABRIEL DA PALHA	Norte	SRE NOVA VENÉCIA
32019459	EEEFM NESTOR GOMES	Norte	SRE SÃO MATEUS
32010699	EEEFM PROFª NÉA MONTEIRO COSTA	Norte	SRE COLATINA
32000499	EEEFM DERMEVAL LEITE RIBEIRO	Norte	SRE BARRA DE SÃO FRANCISCO
32020333	EEEFM ERMENTINA LEAL	Norte	SRE LINHARES
32001916	EEEFM PROF ASCENDINA FEITOSA	Norte	SRE BARRA DE SÃO FRANCISCO
32078528	EEEM SANTINA MOROSINI CUPERTINO	Norte	SRE LINHARES
32007175	EEEFM ALARICO JOSÉ DE LIMA	Norte	SRE NOVA VENÉCIA
32009402	EEEFM DR. JONES DOS SANTOS NEVES	Norte	SRE COLATINA
32005652	EEEFM ANTONIO DOS SANTOS NEVES	Norte	SRE NOVA VENÉCIA

Continua ▼

Código da Escola	Escola	Região	Regional (GRE)
32013906	EEEFM JANUARIO RIBEIRO	Norte	SRE COLATINA
32030584	EEEFM EURICO SALLES	Norte	SRE COLATINA
32063199	EEEFM ELZA LEMOS ANDREATA	RMGV	SRE CARAPINA
32065043	EEEFM EWERTON MONTENEGRO GUIMARÃES	RMGV	SRE CARIACICA
32076410	EEEFM DR. JOSÉ MOYSES	RMGV	SRE CARIACICA
32079230	EEEM MARIO GURGEL	RMGV	SRE VILA VELHA
32043686	EEEM DR SILVA MELLO	RMGV	SRE VILA VELHA
32034016	EEEFM ANA LOPES BALESTRERO	RMGV	SRE CARIACICA
32052405	EEEFM LIONS SEBASTIÃO PAIVA VIDAURRE	Sul	SRE CACHOEIRO DE ITAPEMIRIM
32045360	EEEFM PROFESSOR PEDRO SIMÃO	Sul	SRE COMENDADORA JUREMA MORETZ SOHN
32049242	EEEFM P AFONSO BRAZ	Sul	SRE COMENDADORA JUREMA MORETZ SOHN
32025920	EEEFM AFONSO CLÁUDIO	Sul	SRE AFONSO CLÁUDIO
32048459	EEEFM BERNARDO HORTA	Sul	SRE COMENDADORA JUREMA MORETZ SOHN
32046103	EEEFM SIRENA REZENDE FONSECA	Sul	SRE COMENDADORA JUREMA MORETZ SOHN
32046022	EEEFM JOSÉ CORRENTE	Sul	SRE COMENDADORA JUREMA MORETZ SOHN
32053622	EEEFM PROFESSOR DOMINGOS UBALDO	Sul	SRE CACHOEIRO DE ITAPEMIRIM
32046197	EEEFM PROF CELIA TEIXEIRA DO CARMO	Sul	SRE COMENDADORA JUREMA MORETZ SOHN

A delimitação final do universo de escolas para aleatorização de grupos de controle e

tratamento contabilizou 221 escolas divididas em 70 grupos:

Tabela – Estratificação – nível 1: Espírito Santo, avaliação JF

Região	Escolas			Grupos		
	Total	Controle	Tratamento	Trios não vulneráveis	Quartetos não vulneráveis	Trios vulneráveis
Norte	69	22	47	14	3	5
RMGV	92	29	63	22	5	2
Sul	60	19	41	13	3	3
Sem nota em 2014	1	-	-	-	-	-
Total	222	70	151	49	11	10

Dentre as 221 escolas participantes do processo de aleatorização para seleção dos grupos de controle e tratamento, 188 foram pareadas com base na predição da variação do desempenho para o período 2014-2016, 30 com base na vulnerabilidade, 3, de forma aleatória. A escola sem nota em 2014 permanece aguardando definição.

O modelo 2012-2014 para as estimativas dos coeficientes para o preditor da evolução do desempenho entre 2014 e 2016 gerou as seguintes estatísticas. Com o valor predito calculado, as escolas são ordenadas dentro de seus estratos e aleatorizadas.

Resumo dos resultados

Estatística de regressão	
R múltiplo	0.547
R-Quadrado	0.299
R-quadrado ajustado	0.275
Erro padrão	10.350
Observações	215

ANOVA

	gl	SQ	MQ	F	F de significação
Regressão	7	94446.940	1349.563	12.599	0.000
Resíduo	207	22172.740	107.115		
Total	214	31619.680			

	Coefficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores	Inferior 95.0%	Superior 95.0%
Interseção	61.473	15.708	3.913	0.000	30.505	92.441	30.505	92.441
NSE	-0.257	1.359	-0.189	0.850	-2.936	2.422	-2.936	2.422
LP 2010	-0.107	0.116	-0.923	0.357	-0.336	0.122	-0.336	0.122
LP 2011	0.037	0.123	0.297	0.767	-0.207	0.280	-0.207	0.280
LP 2012	-0.286	0.110	-2.601	0.010	-0.503	-0.069	-0.503	-0.069
MT 2010	0.297	0.104	2.852	0.005	0.092	0.503	0.092	0.503
MT 2011	0.212	0.119	1.778	0.077	-0.023	0.447	-0.023	0.447
MT 2012	-0.385	0.099	-3.879	0.000	-0.581	-0.189	-0.581	-0.189

Validação

Seguem abaixo tabelas com os valores do Teste T de igualdade de médias e o teste Kolmogorov-Smirnov.

Tabela – Testes de Média e de Kolmogorov para os resultados da linha de base de Matemática entre as escolas de Tratamento e Controle do Projeto Jovem de Futuro: Espírito Santo, 2014

Indicadores	Média (%)		Diferença	P-valor do teste T de Igualdade médias	Estatísticas do teste de Kolmogorov	P-valor do teste T Kolmogorov
	Tratamento	Controle				
Indicadores de desempenho						
Média	275	278	-2,55	35	0,10	74
Mediana	275	278	-2,38	42	0,10	69
1ª quartil	242	245	-2,84	36	0,12	52
3ª quartil	308	311	-3,37	25	0,15	21
Desvio padrão	47	48	-0,36	62	0,10	73
Melhor desempenho						
Média entre os 25% melhores	335	338	-2,69	32	0,11	66
Média entre os 50% melhores	313	316	-2,76	31	0,13	37
Média entre os 75% melhores	296	298	-2,67	33	0,13	44
Porcentagem de alunos com nota acima de 317 (meta do Todos pela educação)	0,21	0,24	-0,03	9	0,14	28
Porcentagem de alunos com nota acima de 300	0,32	0,34	-0,02	31	0,12	52
Porcentagem de alunos com nota acima de 350	0,07	0,08	-0,01	45	0,09	83
Porcentagem de alunos com nota acima de 275	0,50	0,53	-0,02	35	0,10	68
Pior desempenho						
Média entre os 25% piores	216	218	-1,96	48	0,10	73
Média entre os 50% piores	238	240	-2,32	41	0,11	59
Porcentagem de alunos com nota abaixo de 200	0,08	0,08	0,00	70	0,07	98
Porcentagem de alunos com nota abaixo de 175	0,02	0,02	0,00	82	0,07	97
Porcentagem de alunos com nota abaixo de 150	0,00	0,00	0,00	62	0,02	100
Número médio de alunos que fizeram a prova	96	85	11	21	0,18	10

DIFERENÇA DE DESEMPENHO NA 3ª SÉRIE EM MATEMÁTICA ENTRE OS ESTRATOS DE TRATAMENTO E CONTROLE: ESPÍRITO SANTO, LINHA DE BASE – 2014

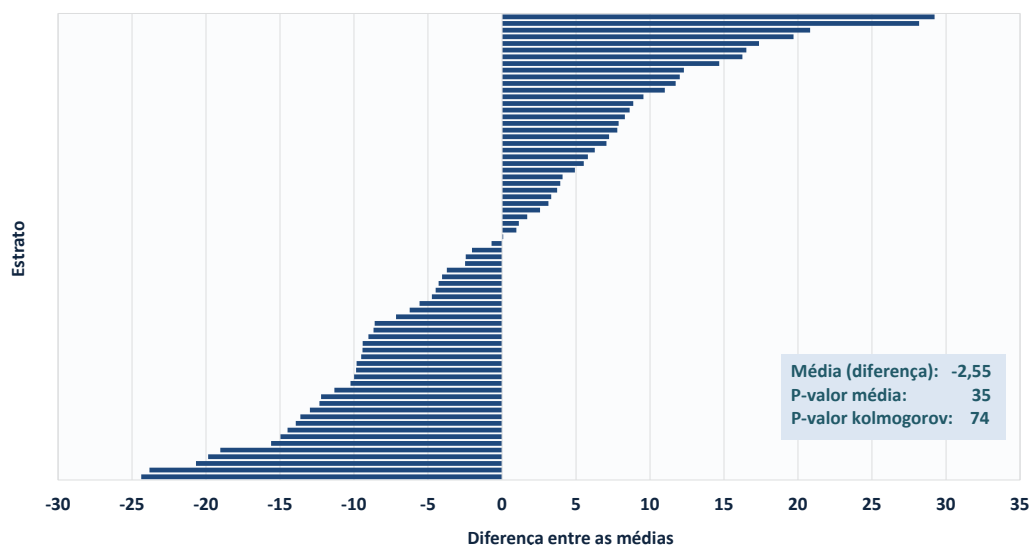
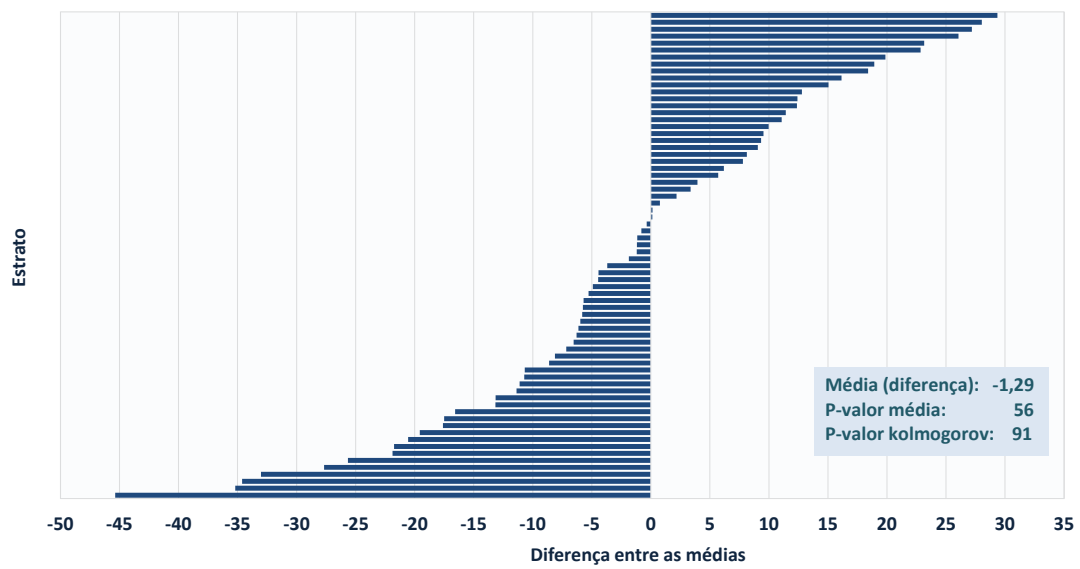


Tabela – Testes de Média e de Kolmogorov para os resultados da linha de base de Língua Portuguesa entre as escolas de Tratamento e Controle do Projeto Jovem de Futuro: Espírito Santo, 2014

Indicadores	Média (%)		Diferença	P-valor do teste T de Igualdade médias	Estatísticas do teste de Kolmogorov	P-valor do teste T Kolmogorov
	Tratamento	Controle				
Indicadores de desempenho						
Média	266	267	-1,29	56	0,08	91
Mediana	268	270	-1,24	61	0,09	82
1ª quartil	235	236	-1,65	55	0,09	85
3ª quartil	298	301	-2,69	22	0,15	24
Desvio padrão	46	46	-0,49	51	0,14	35
Melhor desempenho						
Média entre os 25% melhores	322	324	-1,72	38	0,10	68
Média entre os 50% melhores	303	304	-1,72	40	0,12	53
Média entre os 75% melhores	286	288	-1,56	47	0,09	83
Porcentagem de alunos com nota acima de 283 (meta do Todos pela educação)	0,39	0,41	-0,02	29	0,12	51
Porcentagem de alunos com nota acima de 275	0,45	0,47	-0,02	28	0,13	37
Porcentagem de alunos com nota acima de 300	0,25	0,27	-0,02	19	0,17	14
Porcentagem de alunos com nota acima de 250	0,65	0,65	0,00	93	0,09	85
Pior desempenho						
Média entre os 25% piores	208	208	-0,26	92	0,07	99
Média entre os 50% piores	230	231	-0,86	72	0,08	95
Porcentagem de alunos com nota abaixo de 200	0,10	0,10	0,00	95	0,06	100
Porcentagem de alunos com nota abaixo de 175	0,04	0,04	0,00	97	0,07	97
Porcentagem de alunos com nota abaixo de 150	0,01	0,01	0,00	82	0,07	98
Número médio de alunos que fizeram a prova	96	85	11	21	0,18	10

DIFERENÇA DE DESEMPENHO NA 3ª SÉRIE EM LÍNGUA PORTUGUESA ENTRE OS ESTRATOS DE TRATAMENTO E CONTROLE: ESPÍRITO SANTO, LINHA DE BASE – 2014



Referências

BAKER, Judy (2000) Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners. The World Bank, Washington. Disponível em: <https://openknowledge.worldbank.org/bitstream/handle/10986/13949/207450PUB0Repl0xtop0146970same0info.pdf?sequence=1>



Agência Brasileira do ISBN
ISBN 978-85-60154-04-3



9 788560 154043