

avaliação econômica e projetos sociais

Relatório de Avaliação Econômica

6. Programa Melhoria da educação no Município

2007



O PROGRAMA AVALIAÇÃO ECONÔMICA DE PROJETOS SOCIAIS
conta com a coordenação técnica da Gerência de Avaliação de Projetos do
Banco Itaú.

Equipe responsável pela avaliação:

Banco Itaú-Unibanco:

Lígia Vasconcellos

Consultor externo:

Naercio Menezes-Filho (Insper e USP)

Apresentação

Criado em 2004, o PROGRAMA AVALIAÇÃO ECONÔMICA DE PROJETOS SOCIAIS ocorre em parceria entre a Fundação Itaú Social e o Banco Itaú. Desta forma, o Banco Itaú leva suas competências na área econômica para o campo social.

O PROGRAMA possui duas vertentes principais de atuação, a realização de avaliações de projetos e a disseminação da cultura de avaliação para gestores de projetos sociais e de políticas públicas. No campo da avaliação, tem-se por premissa sua realização para todos os programas próprios. A disseminação da cultura de avaliação é feita tanto através da avaliação de projetos de terceiros, como também de ações de disseminação de conhecimento, por meio de cursos, seminários e informações disponibilizadas na página eletrônica da Fundação Itaú Social.

A avaliação econômica engloba a avaliação de impacto, que verifica se os impactos esperados foram alcançados, e se foram efetivamente causados pelo programa; e o cálculo do retorno econômico, que é fruto de uma análise de custo-benefício do programa.

Acreditando que a participação de todos os interessados na avaliação é o melhor meio de validar e perpetuar a cultura de avaliação, o PROGRAMA procura incluir os gestores do projeto a ser avaliado nas discussões sobre o desenho da avaliação. Este trabalho conjunto possibilita, de um lado, um maior conhecimento do programa em questão, importante para um bom desenho de avaliação, e, por outro, leva à apropriação pelos gestores da cultura de avaliação.

Avaliação econômica

Programa Melhoria da Educação

Sumário Executivo

O programa Melhoria da Educação busca auxiliar os órgãos de governos municipais a planejar, implementar e avaliar seus projetos educativos, visando garantir ensino de qualidade e permanência na escola. O programa existe desde 1999 e já esteve presente em 17 estados brasileiros. Foram oito edições até esta avaliação.

A avaliação de impacto analisa se o programa gerou melhora na média municipal de fluxo escolar antes e depois da participação. Especificamente, os indicadores utilizados são as variações das taxas de aprovação, reprovação e abandono um ano e dois anos após da participação do município no programa. A análise após um ano inclui dados das edições 2 a 5; e após dois anos, apenas dados das edições 2 e 3.

Foram considerados diferentes recortes dos municípios participantes, que definem a dedicação ao programa e também a forma de implementação. A dedicação é medida pela entrega de dois relatórios que constam da formação, a análise de diagnóstico da educação no município, e um plano de ação, planejado para execução durante o período de formação. A implementação considera se a formação foi realizada diretamente pela equipe do Cenpec - Centro de Estudos e Pesquisas em Educação, Cultura e Ação Comunitária, ou se foi ministrada por multiplicadores treinados pelo Cenpec. As estimativas, portanto, foram calculadas para todos os municípios; somente para os que fizeram avaliação diagnóstica; somente para os que desenvolveram plano de ação; e somente para os municípios que receberam formação direta.

A avaliação utiliza dados secundários do Censo Demográfico de 2000 e dos Censos Escolares dos anos de 2000 a 2005. A metodologia é de mínimos quadrados ponderados, sendo o ponderador baseado na probabilidade estimada de participação do município. Para determinação do grupo de municípios de comparação, utilizamos variáveis que definem diferentes perfis de municípios e qualidade da educação.

Em 2000, as taxas médias de fluxo escolar do grupo de tratamento eram 72% de aprovação, 14% de reprovação, e 14% de abandono. A taxa de aprovação cai após um ano de programa, recuperando-se no ano seguinte (aumento de 0,67 pontos percentuais após

dois anos); a taxa de reprovação aumenta nos dois períodos que se seguem ao programa (0,86 pp após dois anos), e a taxa de abandono diminui nos dois períodos (1,53 pp após dois anos). Todas as variações de fluxo são melhores que as do grupo de controle, apesar das variações contrárias ao esperado para aprovação e reprovação.

A avaliação de impacto verifica se essa melhora pode ser atribuída ao programa Melhoria da Educação, e se ela é estatisticamente significativa. Considerando o grupo de municípios tratados que fez avaliação diagnóstica, os resultados são positivos e estatisticamente significativos: o programa aumentou as taxas de aprovação e diminuiu as taxas de reprovação e de abandono em relação às taxas do grupo de controle, um e dois anos após o programa. Nos municípios que elaboraram plano de ação e nos que tiveram formação direta, o programa teve impacto estatisticamente significativo sobre as taxas de aprovação e de abandono, um e dois anos após o programa.

Para a amostra com todos os municípios, contudo, os resultados não são estatisticamente significativos para o impacto um ano após a participação no programa. O impacto é significativo sobre as taxas de aprovação e reprovação dois anos após o programa.

Os resultados, portanto, são mais robustos quando há avaliação diagnóstica ou plano de ação (caso de cerca de 50% dos municípios) ou houve formação direta (70% dos municípios).

Resultados por edição e especificamente para a Paraíba, estado que recebeu atenção especial nas primeiras edições, não apresentaram impactos estatisticamente significativos. O menor tamanho das amostras pode ajudar a explicar a diferença de resultado por edições em relação ao resultado para o total de municípios.

Índice

1	Descrição do programa	7
2	Metodologia.....	8
2.1	Avaliação de impacto	9
2.1.1	Mínimos quadrados ponderados.....	9
2.1.2	Dados	10
2.1.3	Definição dos grupos de tratamento e controle	11
3	Análise descritiva dos dados da pesquisa	12
4	Avaliação de impacto	15
4.1	Estimativas por edição.....	17
4.2	Estimativas para a Paraíba.....	19
5	Anexos.....	21
6	Fontes bibliográficas	23

1 *Descrição do programa*

O programa Melhoria da Educação busca auxiliar os órgãos de governos municipais a planejar, implementar e avaliar seus projetos educativos, visando garantir um ensino de qualidade e a permanência na escola. É realizado pela Fundação Itaú Social, em parceria com a Unicef, a Undime e o Cenpec.

Os participantes do programa são incentivados a desenvolver uma avaliação diagnóstica das principais questões a serem trabalhadas em seus municípios e, posteriormente, um plano de ação para resolver os principais problemas. É oferecida formação direta para alguns municípios. Outros participam de formação indireta, através da formação de técnicos em Núcleos Regionais. Estes técnicos são multiplicadores nos Núcleos Municipais, que englobam municípios de uma mesma região, disseminando as orientações do programa.

O programa é focado nos municípios de pequeno porte (até 20 mil habitantes), com exceção das sedes dos polos regionais (em geral municípios maiores, que também recebem formação). O programa existe desde 1999 e já esteve presente em 17 estados brasileiros. Foram oito edições até esta avaliação, como mostram as tabelas 1 e 2:

Tabela 1

Edição	Estados	Período
1	AL, AM, PA, CE, MA, BA, PB, PE, PI, RR, SE, SP	2º semestre 1999 a 1º semestre 2000
2	AC, MS, SP	2º semestre 2000 a 1º semestre
3	AL, MA, MG, PB, PI, SP	2º semestre 2001 a 1º semestre
4	GO, MG, PB, RN, SE	2º semestre 2002 a 1º semestre
5	BA, PB, PE, PI	2003
6	BA, PE, PI, SP	2004
7	PI, SP	2005-2006
8	MG	2006-2007

Tabela 2 - Formação direta

Edições	Estado	Nº de municípios com formação direta	% dos participantes
2	MS	22	100%
	AC	22	100%
3	AL	20	100%
	MA	21	100%
	MG	11	79%
	PB	15	79%
	PI	18	100%
	SP	3	23%
4	GO	21	100%
	MG	17	85%
	RN	27	100%
	SE	20	100%
5	BA	13	81%
	PE	3	100%
	PI	1	50%

2 Metodologia

Foi realizada a avaliação do impacto do programa através da comparação de alguns indicadores entre os municípios que participaram (grupo de tratamento) e os que não participaram (grupo de controle). Os indicadores municipais de desempenho utilizados foram as variações das taxas de aprovação, reprovação e abandono antes e depois da participação no programa.¹

Na avaliação foram consideradas as edições 2, 3, 4 e 5. Foram excluídos da análise os municípios que participaram da primeira edição, porque esta apresentou características muito diversas das demais. Os municípios que participaram das edições 6, 7 e 8 foram considerados não-participantes (grupo de controle), pois os dados utilizados são anteriores a essas edições.

Além do impacto médio do programa, estimou-se o impacto separadamente para cada edição e para o Estado da Paraíba.

As estimativas foram feitas para diferentes recortes da amostra dos participantes, que definem diferentes níveis de envolvimento com o programa:

- todos os municípios;
- somente os que fizeram a avaliação diagnóstica;

¹ Também foi feita a análise com as variáveis em nível (em anexo).

- somente os que desenvolveram o plano de ação;

Foi feito também um outro recorte para tipos diferentes de formação:

- somente os que participaram de formação direta.

Foram excluídos da amostra os municípios que participaram mais de uma vez do programa. Para a análise específica sobre a Paraíba, foram usados todos os municípios, pois houve muitos municípios reincidentes e sua exclusão poderia afetar o resultado da análise.

2.1 Avaliação de impacto

O objetivo da avaliação de impacto é mensurar o efeito que o programa teve sobre os indicadores escolhidos. Como é impossível saber o que teria ocorrido se o município não tivesse participado do programa, busca-se comparar os municípios que participaram com os não-participantes que sejam mais “parecidos”.

Para determinar o grupo de controle “parecido”, utilizamos variáveis que definem diferentes perfis de municípios e qualidade da educação, com dados do Censo Demográfico de 2000 e Censo Escolar dos anos de 2000 a 2005. As variáveis relativas aos municípios utilizadas foram: porcentagem de professores com nível superior, média de horas-aula das escolas do município, IDH-educação, renda média per capita, tamanho da população e variáveis *dummy* para os diferentes estados do País.

A metodologia econométrica utilizada está explicada no *box* “Mínimos quadrados ponderados”.

2.1.1 Mínimos quadrados ponderados

Imbens (2004) propõe uma combinação do método de *propensity score* com a regressão, que tem o benefício de ser “duplamente robusto”.

A regressão estimada é $Y_i = \alpha + \tau \cdot W_i + \varepsilon_i$, com pesos iguais a:

$$\lambda_i = \sqrt{\frac{W_i}{e(X_i)} + \frac{1 - W_i}{1 - e(X_i)}}, \text{ onde:}$$

Y_i = resultado

W_i = variável de participação no programa (0 ou 1)

$e(X_i)$ = *propensity score*

X_i = características observadas que explicam a participação no programa

λ_i = peso

O coeficiente τ estimado é um estimador consistente para o efeito médio do tratamento (ATE). A regressão tem a vantagem de diminuir o viés e aumentar a precisão da estimativa. Sem os pesos dados pelo *propensity score*, porém, o estimador de mínimos quadrados não seria consistente.

Sendo \hat{e} um estimador consistente do *propensity score* e garantindo que cada termo dentro da raiz quadrada some 1, o estimador por mínimos quadrados ponderados é consistente para o ATE:

$$\hat{\lambda}_i = \sqrt{\frac{W_i / \hat{e}(X_i)}{\sum_{i=1}^N W_i / \hat{e}(X_i)} + \frac{(1-W_i) / (1-\hat{e}(X_i))}{\sum_{i=1}^N (1-W_i) / (1-\hat{e}(X_i))}}$$

em que $\hat{e}(X_i)$ é o *propensity score* estimado.

Para se estimar o efeito-tratamento sobre os tratados (ATT), o ponderador

$$\hat{\lambda}_i = \sqrt{\frac{W_i}{\sum_{i=1}^N W_i} + \frac{(1-W_i) \cdot \hat{e}(X_i) / (1-\hat{e}(X_i))}{\sum_{i=1}^N (1-W_i) \cdot \hat{e}(X_i) / (1-\hat{e}(X_i))}}$$

produz o estimador do coeficiente τ que é consistente para o ATT.

2.1.2 Dados

Foram utilizadas as seguintes fontes de dados:

- base de dados do programa, elaborada pelo Cenpec – municípios participantes, com informação se fizeram avaliação diagnóstica, plano de ação e formação direta.
- Censo Escolar dos anos de 2000 a 2005 – porcentagem de professores com nível superior, horas-aula, taxas de aprovação, reprovação e abandono;
- Censo Demográfico de 2000 – população, renda per capita, IDH educação.

Os dados utilizados são as médias por município. No caso dos dados escolares, são consideradas apenas as escolas municipais.

Para cada edição, foram consideradas as variações nos indicadores de fluxo escolar referentes a um ano antes e um e dois anos após o programa. Assim:

Edição 2, realizada em 2000/2001 → variação 2000-2002 e 2000-2003

Edição 3, realizada em 2001/2002 → variação 2001/2003 e 2001-2004

Edição 4, realizada em 2002/2003 → variação 2002-2004

Edição 5, realizada em 2003 → variação 2002-2004

Nas duas últimas edições, não foi utilizada a informação sobre dois anos após o programa por falta de dados disponíveis.

2.1.3 Definição dos grupos de tratamento e controle

O grupo de tratamento é formado pelos municípios participantes do programa, que participaram só uma vez. Supõe-se que os municípios que participaram mais de uma vez possam apresentar resultados melhores, e portanto afetariam os resultados relativos a uma participação.

Os números de municípios participantes por estado, com e sem os municípios que participaram mais de uma vez, são apresentados nas tabelas 3 e 4.

Tabela 3: Municípios participantes por estado

Edição:	AC	AL	BA	GO	MA	MG	MS	PB	PE	PI	RN	SP	SE
2	22						22					55	
3		20			21	14		19		18		13	
4				21		20		60			27		20
5			16					68	3	2			
% do total de municípios	100%	19%	4%	9%	10%	4%	28%	51%	2%	9%	16%	9%	27%

Tabela 4: Municípios participantes por estado, excluindo os que participaram mais de uma vez

Edição:	AC	AL	BA	GO	MA	MG	MS	PB	PE	PI	RN	SP	SE
2	22						22					48	
3		20			21	11		19		18		6	
4				21		17		28			27		20
5			16					36	3	2			
% do total de municípios	100%	19%	4%	9%	10%	3%	28%	37%	2%	9%	16%	8%	27%

Nos poucos casos em que o estado participou em diferentes edições, considerou-se apenas a edição com maior número de municípios participantes. Foram desprezados, portanto (ver tabela 4, números que não estão em negrito):

- Municípios de São Paulo que participaram na edição 3 (6 municípios);
- Municípios de Minas Gerais que participaram na edição 3 (11 municípios);
- Municípios da Paraíba que participaram nas edições 3 e 4 (47 municípios);
- Municípios do Piauí que participaram na edição 5 (2 municípios).

Portanto, a amostra utilizada considera apenas a participação em uma edição por estado. Esta exclusão de alguns estados ocorreu em função de um problema de identificação. Como não seria possível usar *dummies* de estado e de edição em um painel, optou-se por uma regressão de seção transversal com a identificação dos estados.

O grupo de controle é formado pelos municípios que não participaram do programa nas edições 1 a 5 e os que participaram apenas nas edições posteriores (pois os indicadores relativos a eles são anteriores à sua participação no programa). São considerados somente os estados onde houve alguma participação a partir da segunda edição, logo não foram considerados 14 estados: AM, AP, CE, DF, ES, MT, PA, PR, RJ, RO, RR, RS, SC e TO.

Para os municípios de controle, são consideradas as taxas de fluxo escolar referentes à edição em que o seu estado participou. Assim, para os municípios do Maranhão que não participaram, por exemplo, foi considerada a variação dos indicadores de fluxo escolar referentes à edição 3 (ou seja, 2001/2003 e 2001-2004). Para os municípios da Paraíba que não participaram, foi considerada a variação referente à edição 5.

3 *Análise descritiva dos dados da pesquisa*

A tabela 5 mostra o tamanho das amostras do grupo de tratamento. Alguns municípios foram excluídos da estimativa por não terem informações disponíveis de todas as variáveis explicativas (por isso, os números dos grupos de tratamento não coincidem exatamente com os da tabela 4). Nas estimativas específicas a cada edição, o número de observações nas subamostras de municípios – que fizeram avaliação diagnóstica, plano de ação e que tiveram formação direta – é muito pequeno, impedindo que uma estimativa confiável fosse realizada.

Na estimativa do impacto do programa sobre os indicadores de fluxo escolar dois anos à frente, muitas observações foram desprezadas por falta de dados disponíveis para as edições 4 e 5, como mostra a tabela 6.

Tabela 5: Número de observações

		Todas	Edição 2	Edição 3	Edição 4	Edição 5
Total	tratados	261	63	91	113	57
	controle	2495	264	1576	1299	794
Avaliação diagnóstica	tratados	129	17	26	82	33
	controle	2441	210	711	1299	794
Plano de Ação	tratados	124	16	18	77	33
	controle	2441	210	711	1299	794
Formação Direta	tratados	180	18	87	85	17
	controle	2194	54	1523	1208	703

Tabela 6: Número de observações - com dados para 2 anos após o programa

		Todas
Total	tratados	121
	controle	674
Avaliação diagnóstica	tratados	39
	controle	620
Plano de Ação	tratados	31
	controle	620
Formação Direta	tratados	74
	controle	464

A tabela 7 mostra que os estados que participaram do programa têm menor proporção de professores com nível superior, menor IDH e menor renda per capita, indicando que a focalização do programa nos estados que mais necessitam é boa.

Tabela 7: Médias dos estados

Fonte:	Participaram	sim	não
Censo escolar 2001	Prof. com nível superior	17%	23%
	Horas-aula	4,2	4,1
Censo demográfico 2000	IDH - educação	0,7503	0,8278
	Renda per capita	144	205
	População	21.731	29.268

A tabela 8 mostra as médias das variáveis explicativas para os municípios e a comparação entre os municípios que participaram e os que não participaram. Nota-se que a

porcentagem de professores com nível superior é maior nos municípios que participaram do programa, mas o IDH da educação nestes é menor. Renda per capita e horas-aula apresentam valores similares. Em média, os municípios que participaram são maiores (embora o foco do programa sejam os municípios com menos de 20 mil habitantes, as cidades-polo aumentam a média populacional).

Tabela 8

Participou	% professores com nível superior		Renda per capita 2000 (R\$)		Média de horas-aula		População 2000		IDH educação 2000	
	Média	Desvio-padrão	Média	Desvio-padrão	Média	Desvio-padrão	Média	Desvio-padrão	Média	Desvio-padrão
não	16,2	22,2	141	85	4,2	0,3	20.602	58.145	0,749	0,086
sim	20,1	23,7	140	85	4,2	0,4	35.721	108.413	0,742	0,097
Total	16,5	22,3	141	85	4,2	0,3	21.951	64.370	0,748	0,087

A tabela 9 mostra a evolução dos indicadores de fluxo escolar ao longo do tempo. O indicador que apresentou maior variação ao longo do tempo foi a taxa de abandono e ela caiu um pouco mais entre os participantes do programa. As taxas de aprovação aumentaram, mas em menor magnitude, e o grupo de municípios que participaram de formação direta aparentemente teve melhor desempenho. A taxa de reprovação aumentou em todos os grupos.

A avaliação de impacto, levando em consideração os municípios “parecidos”, irá confirmar ou não o resultado da comparação de médias simples.

Tabela 9

Participou	Amostra	2000	2001	2002	2003	2004	Varição 2000-2004
Médias das taxas de abandono							
não		14,0	10,9	9,7	9,7	10,4	-25,73%
	aval. diagnóstica	16,1	12,0	11,0	10,6	11,5	-28,32%
sim	formação direta	15,6	12,0	11,3	10,9	11,2	-28,03%
	total participantes	14,2	10,8	10,2	9,8	10,3	-27,22%
Total		14,0	10,9	9,7	9,7	10,4	-25,87%
Médias das taxas de aprovação							
não		71,8	75,1	75,7	75,2	73,9	2,87%
	aval. diagnóstica	67,5	71,3	71,9	72,8	70,6	4,71%
sim	formação direta	65,9	69,8	71,9	71,2	69,2	4,89%
	total participantes	69,4	73,0	73,5	73,9	72,0	3,74%
Total		71,6	74,9	75,5	75,1	73,7	2,95%
Médias das taxas de reprovação							
não		14,2	14,0	14,6	15,1	15,8	10,75%
	aval. diagnóstica	16,5	16,7	17,0	16,5	17,9	8,30%
sim	formação direta	18,5	18,1	17,9	17,9	19,6	6,14%
	total participantes	16,4	16,2	16,3	16,3	17,7	7,74%
Total		14,4	14,2	14,8	15,2	15,9	10,44%

4 Avaliação de impacto

4.1. Estimativas para todas as edições

As tabelas abaixo apresentam os resultados das estimativas do impacto do programa sobre os indicadores de fluxo escolar. O símbolo * indica que o resultado é estatisticamente significativo a um nível de 90%. As estimativas para a variação de fluxo escolar de um ano consideram todas as edições (2 a 5). As estimativas para dois anos depois consideram apenas as edições 2 e 3.

A metodologia de mínimos quadrados ponderados (MQP) gera resultados consistentes e mais eficientes que as estimativas por pareamento, apresentadas no anexo.

Para a amostra de todos os municípios os resultados não são estatisticamente significantes para o impacto um ano após a participação no programa. O impacto é significativo sobre as taxas de aprovação e reprovação quando se considera as taxas dois anos após o programa.

Abaixo seguem os resultados da estimativa para a amostra total, um ano após a participação no programa:

O primeiro estágio é uma estimativa *probit* em que a variável dependente é a participação no programa.

Tabela 10:

Regressão <i>probit</i>	Número de obs	=	2756
	LR chi2(16)	=	235.38
	Prob > chi2	=	0.0000
Log verossimilhança = -745.72461	Pseudo R2	=	0.1363

Participação	Coef.	Erro-padrão	z	P> z	[Interv. de conf. 95%]
% prof super	.0003389	.002344	0.14	0.885	-.0042553 .004933
média h-aula	-.0101428	.1401422	-0.07	0.942	-.2848165 .2645309
idh educação	2.388147	.8221426	2.90	0.004	.776777 3.999517
renda p.cap.	-.0018974	.0009151	-2.07	0.038	-.0036909 -.0001039
população	1.33e-06	4.11e-07	3.23	0.001	5.21e-07 2.13e-06
Bahia	-1.100913	.2036408	-5.41	0.000	-1.500041 -.7017838
Goiás	-.7544334	.2460412	-3.07	0.002	-1.236665 -.2722016
Maranhão	-.5133572	.2014715	-2.55	0.011	-.9082341 -.1184802
Minas Gerais	-1.440774	.2173008	-6.63	0.000	-1.866676 -1.014873
Mato G. Sul	.0094801	.2599563	0.04	0.971	-.5000249 .5189852
Paraíba	.1836923	.1931445	0.95	0.342	-.1948639 .5622485
Pernambuco	-1.410087	.2861398	-4.93	0.000	-1.970911 -.8492637
Piauí	-.6106377	.1991088	-3.07	0.002	-1.000884 -.2203917
Rio G. Norte	-.3206802	.2043118	-1.57	0.117	-.7211239 .0797636
Sergipe	.029817	.229969	0.13	0.897	-.4209139 .4805479
São Paulo	-.3486329	.2792754	-1.25	0.212	-.8960027 .1987369
constante	-2.165446	.7385234	-2.93	0.003	-3.612925 -.7179669

O segundo estágio é uma estimativa de mínimos quadrados ponderados, que usa as probabilidades estimadas no primeiro estágio como ponderador para a estimativa da relação entre participação no programa e fluxo escolar. No exemplo abaixo, a variável dependente é a variação da taxa de aprovação. Apesar de a melhora na taxa de aprovação do grupo de tratamento ser maior (0,49pp), este aumento não é estatisticamente significativo:

Tabela 11:

	SS	df	MS		Número de obs =	2778
Modelo	146.954807	1	146.954807		F(1, 2776) =	2.15
Resíduo	189377.257	2776	68.2194729		Prob > F =	0.1423
					R-2 =	0.0008
					R-2 ajustado =	0.0004
Total	189524.212	2777	68.2478256		Root MSE =	8.2595

Taxa aprov.	Coef.	Erro-padrão	t	P> t	[Interv. de conf. 95%]	
Participação	.4946908	.3370514	1.47	0.142	-.1662059	1.155587
constante	-.8356802	.1894862	-4.41	0.000	-1.207228	-.4641321

Na tabela 12 são apresentadas as estimativas considerando as diferentes subamostras: apenas os municípios que implementaram bem o programa (fizeram análise diagnóstica e/ou plano de ação) e apenas os municípios que tiveram formação direta.

No recorte da amostra que considera, entre municípios tratados, só os que fizeram avaliação diagnóstica, os resultados são estatisticamente significantes – o programa aumentou as taxas de aprovação e diminuiu as taxas de reprovação e abandono, um e dois anos após o programa. Nos municípios que elaboraram um plano de ação e nos que tiveram formação direta, o programa teve impacto positivo e significativo sobre as taxas de aprovação e abandono, um e dois anos após o programa.

Tabela 12: Impacto do Programa - todas as edições (pontos percentuais)

amostras	todos	avaliação diagnóstica	plano de ação	formação direta
Varição da aprovação 1 ano depois	0,495	1,292 *	0,937 *	1,190 *
Varição da aprovação 2 anos depois	1,068 *	2,427 *	1,994 *	2,590 *
Varição da reprovação 1 ano depois	-0,316	-0,557 *	-0,228	-0,411
Varição da reprovação 2 anos depois	-0,775 *	-1,056 *	-0,552	-0,947
Varição do abandono 1 ano depois	-0,179	-0,735 *	-0,709 *	-0,779 *
Varição do abandono 2 anos depois	-0,292	-1,371 *	-1,442 *	-1,643 *

* estatisticamente significativa a 90%

Os resultados, portanto, são muito mais robustos quando o programa é bem implementado ou oferece formação direta. Cerca de 50% dos municípios implementaram bem o programa, e 70% receberam formação direta.

4.1 Estimativas por edição

Para a edição 2, não foi possível considerar as amostras de municípios que fizeram avaliação diagnóstica, plano de ação e formação direta porque o número de observações tratadas era muito pequeno.

Os resultados não são estatisticamente significantes para as variações das taxas de aprovação e reprovação. No caso das taxas de abandono, o resultado é significativo mas indesejável: indica que a participação no programa aumentou a taxa de abandono.

Aprovação 1 ano depois	-0,660
Aprovação 2 anos depois	0,430
Reprovação 1 ano depois	-0,844
Reprovação 2 anos depois	-0,622
Abandono 1 ano depois	1,504 *pos
Abandono 2 anos depois	0,192

UFs participantes: AC, MS, SP

* significativos a 90%

Na estimativa referente à edição 3, somente a subamostra dos municípios que tiveram formação direta tinha observações suficientes para fazer as estimativas. Os municípios de estados que participaram em outras edições (6 municípios de SP, 11 de MG e 19 da PB), que foram retirados na amostra total, foram mantidos nesta amostra específica. Em média, a taxa de abandono destes municípios aumentou mais (ou diminuiu menos) que a do restante, logo a inclusão desses municípios contribuiu para que o resultado fosse pior que o resultado da amostra total.

Tabela 14: Impacto do Programa - Edição 3 (pontos percentuais)

amostras	todos	formação direta
Varição da aprovação 1 ano depois	-0,155	-0,278
Varição da aprovação 2 anos depois	-0,700	-0,792 *neg
Varição da reprovação 1 ano depois	-0,353	-0,231
Varição da reprovação 2 anos depois	-0,198	-0,120
Varição do abandono 1 ano depois	0,508 *pos	0,509 *pos
Varição do abandono 2 anos depois	0,898 *pos	0,912 *pos

UFs participantes: AL, MA, MG, PB, PI, SP

* estatisticamente significativa a 90%

Para a edição 4, só havia informação disponível para avaliar o impacto sobre os indicadores referentes a um ano após a participação no programa. Nos municípios que tiveram formação direta, o impacto sobre a taxa de abandono foi significativa.

Tabela 15: Impacto do Programa - Edição 4 (pontos percentuais)

amostras	todos	avaliação diagnóstica	plano de ação	formação direta
Varição da aprovação 1 ano depois	0,508	0,182	0,005	0,748
Varição da reprovação 1 ano depois	-0,362	-0,261	-0,004	-0,196
Varição do abandono 1 ano depois	-0,146	0,079	-0,001	-0,552 *

UFs participantes: GO, MG, PB, RN, SE

* estatisticamente significativa a 90%

Para a edição 5, assim como nas edições 2 e 3, só foram estimados os impactos para a amostra total, em função do tamanho das subamostras. Os resultados não são significativos.

Tabela 16: Impacto do Programa - Edição 5 (pontos percentuais)	
Varição da aprovação 1 ano depois	-0,023
Varição da reprovação 1 ano depois	0,308
Varição do abandono 1 ano depois	-0,285

UFs participantes: BA, PB, PE, PI

* significativos a 90%

4.2 Estimativas para a Paraíba

As tabelas a seguir mostram como evoluíram os indicadores do total de municípios participantes da Paraíba, participantes que tiveram formação direta, e dos não-participantes. Somente a edição 3 ofereceu formação direta (para 15 municípios).

Tabela 17 - Paraíba

Participou	Amostra	2000	2001	2002	2003	2004	Varição 2000-2004
Médias das taxas de aprovação							
não		65,61	68,19	69,02	69,80	66,49	1,34%
sim	formação direta	66,20	69,57	68,95	70,35	66,81	0,92%
	total participantes	65,35	68,88	68,27	68,82	65,57	0,33%
Total		65,47	68,57	68,60	69,25	65,98	0,78%
Médias das taxas de reprovação							
não		16,91	17,98	18,15	17,66	18,48	9,28%
sim	formação direta	15,90	15,54	15,92	15,68	16,35	2,85%
	total participantes	15,56	17,23	17,77	17,66	18,52	19,02%
Total		16,16	17,56	17,94	17,66	18,50	14,50%
Médias das taxas de abandono							
não		17,48	13,82	12,83	12,54	15,04	-14,00%
sim	formação direta	17,90	14,89	15,13	13,97	16,84	-5,94%
	total participantes	19,08	13,90	13,95	13,52	15,91	-16,63%
Total		18,37	13,86	13,46	13,09	15,52	-15,52%

Para estimar o impacto do programa no estado da Paraíba, fizemos duas estimativas. A primeira considera os municípios que participaram das edições 4 e 5, já que a variação dos indicadores de fluxo utilizados para essas edições é referente aos mesmos anos (2002-2004). Não foram desprezados os municípios que participaram mais de uma vez. A outra estimativa utiliza o método de painel, considerando também os municípios que participaram na edição 3. Nesta estimativa, as informações dos municípios que participaram uma vez e não participaram nas edições seguintes foi considerada apenas até o ano de sua participação. Desta forma eles não são utilizados como grupo de controle nos anos seguintes à participação no programa.

Nenhuma das estimativas teve resultado estatisticamente significante.

Tabela 18: Impacto do Programa - Paraíba - painel

amostras	Varição da aprovação 1 ano depois	Varição da reprovação 1 ano depois	Varição do abandono 1 ano depois
total	-0,761	0,741	0,020
avaliação diagnóstica	-0,207	-0,343	0,550
plano de ação	-0,134	-0,252	0,386
formação direta	-0,174	0,198	-0,023

Tabela 19: Impacto do Programa - Paraíba - edições 4 e 5

amostras	Varição da aprovação 1 ano depois	Varição da reprovação 1 ano depois	Varição do abandono 1 ano depois
total	-0,061	0,288	-0,227
avaliação diagnóstica	0,647	-0,602	-0,045
plano de ação	0,726	-0,490	-0,237
formação direta	-	-	-

5 Anexo

O impacto também foi calculado com o método de *propensity score matching*, com *local linear regression*. Seguem abaixo explicação da metodologia e resultados obtidos. As estimativas obtidas por este método não são estatisticamente significativas.

Método: Pareamento pelo método *local linear regression*

Sejam dois resultados potenciais (Y_0, Y_1), onde 1 indica resultado com tratamento. Seja $D = 1$ se ocorre tratamento e $D = 0$ caso contrário. Sejam X variáveis observadas que determinam a participação no tratamento e seu resultado.

Podemos escrever o efeito tratamento sobre tratados (ATT) como:

$$E(Y_1 - Y_0 | X, D=1) \quad (1)$$

Como não temos o contrafactual $E(Y_0 | X, D=1)$, o problema de seleção surge quando queremos utilizar a diferença das médias amostrais observadas para estimar o efeito do tratamento:

$$E(Y_1 | X, D=1) - E(Y_0 | X, D=0) \quad (2)$$

No caso do ATT, o viés gerado equivale à diferença entre (1) e (2):

$$\begin{aligned} & [E(Y_1 | X, D=1) - E(Y_0 | X, D=0)] - [E(Y_1 - Y_0 | X, D=1)] = \\ & = E(Y_0 | X, D=1) - E(Y_0 | X, D=0) \quad (3) \end{aligned}$$

As hipóteses de identificação, ou seja, as hipóteses que garantem que a diferença em (3) é nula, são:

(a) $Y_0 \perp D | X$, isto é, independência de Y_0 em relação a $D | X$

(b) $0 < \Pr(D=1 | X) = P(X) < 1$, isto é, existem observações tratadas e não-tratadas.

Rosenbaum e Rubin (1983) mostram que, dados (a) e (b), também vale:

(c) $Y_0 \perp D | P(X)$, o que reduz a dimensão necessária para resolver o pareamento.

A hipótese (a) poderia ser simplificada para a hipótese de independência de médias, $E(Y_0 | X, D=0) = E(Y_0 | X)$. Neste caso, porém, a condição (c) não segue imediatamente, precisando ser considerada uma hipótese adicional.

Se, por um lado, a hipótese de identificação pode ser forte em casos em que nem todas as variáveis relevantes são observáveis, o pareamento tem a vantagem de não exigir exogeneidade das variáveis explicativas (se o pareamento é bom, as variáveis não

observadas relevantes estão igualmente distribuídas nos grupos de tratados e de controle, e seu efeito, portanto, se cancela) e não exigir restrição de exclusão (que entre as variáveis que explicam a participação no programa haja variáveis que não são correlacionadas com o resultado). A hipótese de identificação depende, portanto, de não haver variáveis não-observadas que afetem o resultado de forma diferenciada nos grupos de tratamento e controle.

O uso de $P(X)$ em vez de X é conhecido como método de *propensity score*. Escolhemos as observações *matched* de acordo com as probabilidades estimadas (*propensity scores*) mais parecidas com as estimadas para o grupo de tratados. Para garantir que as observações sejam realmente parecidas, utilizamos um suporte comum de probabilidades, tal que apenas as observações não-tratadas com *propensity score* dentro da faixa de *propensity score* das observações tratadas são consideradas.

Resultados:

Impacto do Programa - todas as edições (pontos percentuais)

	todos	avaliação diagnóstica	plano de ação	formação direta
Varição da aprovação 1 ano depois	0,122	0,907	0,520	0,561
Varição da aprovação 2 anos depois	-0,600	-0,554	-3,337	0,275
Varição da reprovação 1 ano depois	-0,415	-0,451	-0,180	-0,446
Varição da reprovação 2 anos depois	-0,715	-0,230	0,307	-0,427
Varição do abandono 1 ano depois	0,293	-0,456	-0,340	-0,115
Varição do abandono 2 anos depois	1,315	0,784	3,030	0,152

6 *Fontes bibliográficas*

CENPEC (2005) Os municípios em busca da melhoria da educação. São Paulo.

CENPEC (2006) Avaliação de impacto do programa melhoria da educação no município.

FIRPO, S. (2006) Comentário sobre IMBENS (2004). (manuscrito).

IMBENS, G. (2004) Nonparametric estimation of average treatment effects under exogeneity: a review. *The Review of Economics and Statistics*, Fevereiro, 86(1).