

4. APLICAÇÃO

4.1. Base de dados

Conforme dito anteriormente, a fonte de dados para nossa análise é o Censo Escolar do período de 1996 a 2002, produzido pelo Inep/MEC.

O Censo Escolar é um levantamento de informações estatístico-educacionais de âmbito nacional, realizado anualmente. Esse levantamento abrange a Educação Básica em seus diferentes níveis – Educação Infantil, Ensino Fundamental e Ensino Médio – e modalidades – Ensino Regular, Educação Especial e Educação de Jovens e Adultos.

A pesquisa tem como unidade de informação a escola e como informante o diretor ou responsável de cada unidade escolar. Seu principal objetivo é fornecer informações e estatísticas para a realização de diagnósticos e análises sobre a realidade do sistema educacional, subsidiando a definição e a implementação de políticas orientadas para a promoção da equidade, efetividade e qualidade do ensino.

O Questionário do Censo Escolar 2001 tem 50 páginas e está dividido em 11 blocos, que comportam um total de 126 itens.

Os blocos abordam os seguintes temas: Cadastro da Escola; Caracterização Física da Escola; Dados Gerais da Escola; Educação Infantil, Classe de Alfabetização; Ensino Fundamental; Ensino Médio – Curso Normal em Nível Médio; Educação Profissional – Nível Técnico; Educação Especial; Educação de Jovens e Adultos.

Nossa análise baseou-se apenas nas informações de escolas com Ensino Fundamental. E para chegar a um bom modelo de municipalização iniciamos a estimativa com todas as variáveis possíveis, incluindo todas as variáveis dos temas

citados acima. Retiramos as variáveis observadas até que se chegasse a uma melhor combinação entre observações utilizadas e um pseudo-r² bom. No caso de Santa Catarina, por exemplo, o modelo apresentou um r² de 53%, utilizando apenas duas variáveis, o que permitiu a atribuição de um escore de propensão à quase todas as escolas do estado.

A revisão do questionário do Censo Escolar é feita a cada ano, visando aperfeiçoá-lo e atualizá-lo em relação às mudanças na organização e no funcionamento dos sistemas de ensino.

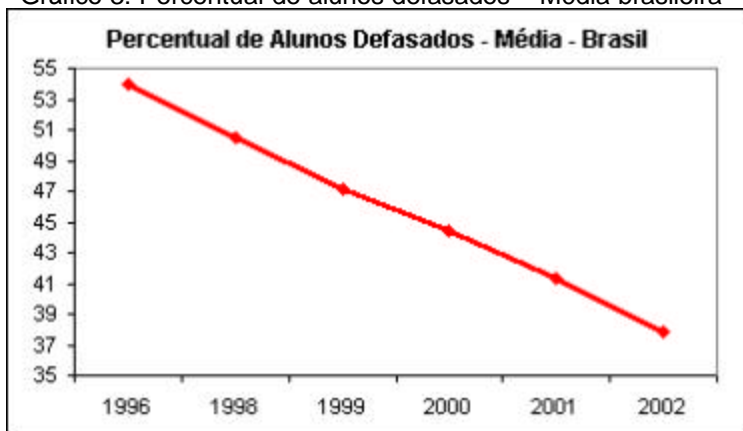
O Censo Escolar atualiza anualmente o Cadastro das Escolas da Educação Básica e as informações referentes à matrícula por dependência administrativa, nível e modalidade de ensino, turno, turmas, séries e sexo, bem como sobre o movimento e o rendimento dos alunos (abandono, aprovação e reprovação). O fato do questionário de 1997 não ter sido tão completo quanto foi para os outros anos nos fez tomar como base o ano de 1996, tanto para análise da municipalização ocorrida de 1996 para 1997, quanto da ocorrida de 1997 para 1998.

4.2. Descrição da análise

O grupo analisado aqui é o das escolas estaduais. Comparamos escolas estaduais que permaneceram estaduais e as escolas estaduais que passaram para uma administração municipal nos anos de 1990, e o objetivo em questão é comparar o desempenho desses dois grupos utilizando variáveis de resultado bastante comuns na literatura do ramo educacional:

- Porcentagem de alunos com defasagem idade-série maior do que dois anos, ou seja, que já repetiram de ano duas ou mais vezes:

Gráfico 3: Percentual de alunos defasados – Média brasileira



O gráfico 3 apresenta claramente a melhora na qualidade do ensino fundamental brasileiro nos últimos anos. Em 1996 o percentual de alunos com repetência maior ou igual a dois anos era de 55%. Este valor foi sofrendo uma queda gradual chegando a 37% em 2002.

A distribuição desta variável no Brasil é dada da seguinte forma:

Eixo Y = Densidade e Eixo X = Percentual de Alunos Defasados.

Gráfico 4: Percentual de alunos defasados – Brasil 1996

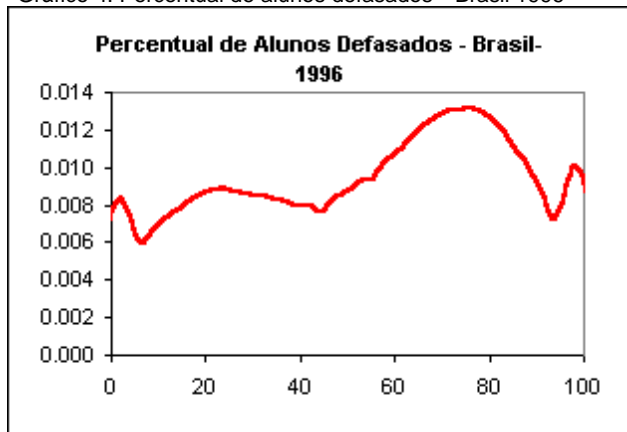


Gráfico 5: Percentual de alunos defasados – Brasil 1998

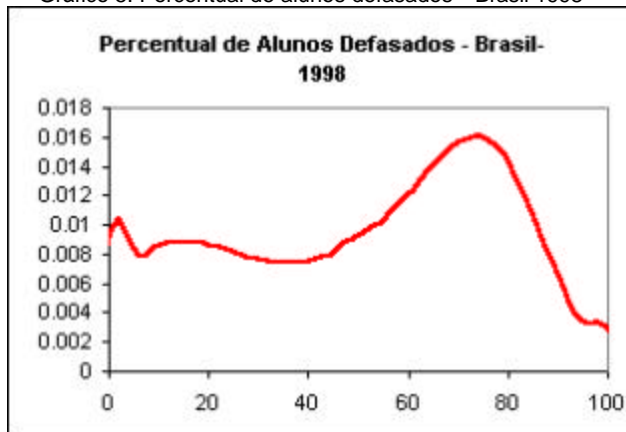


Gráfico 6: Percentual de alunos defasados – Brasil 1999

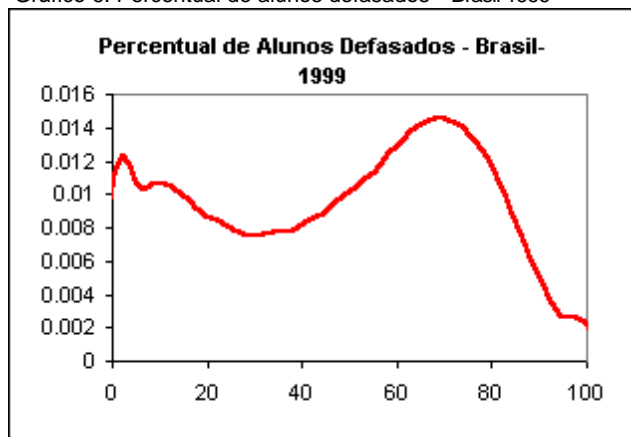


Gráfico 7: Percentual de alunos defasados – Brasil 2000

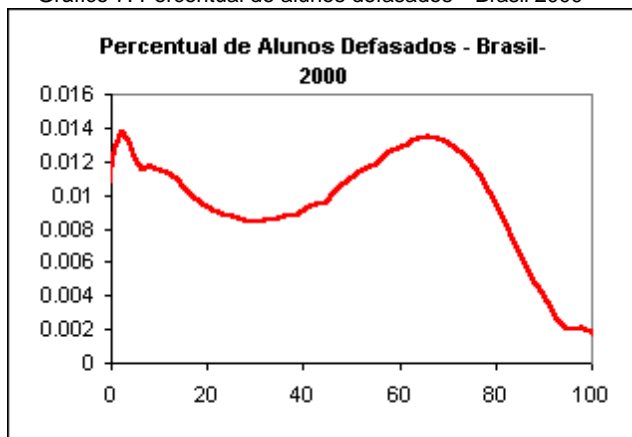


Gráfico 8: Percentual de alunos defasados – Brasil 2001

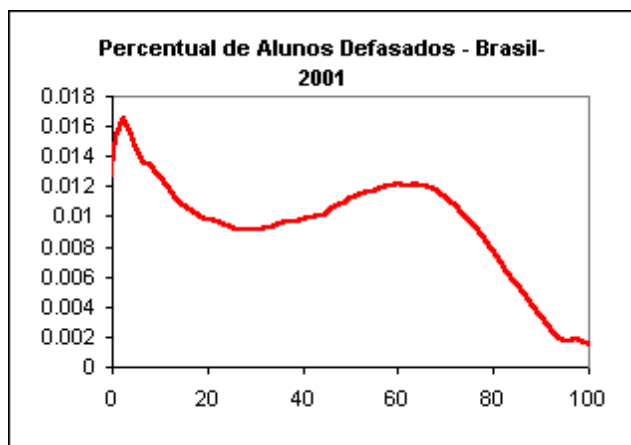
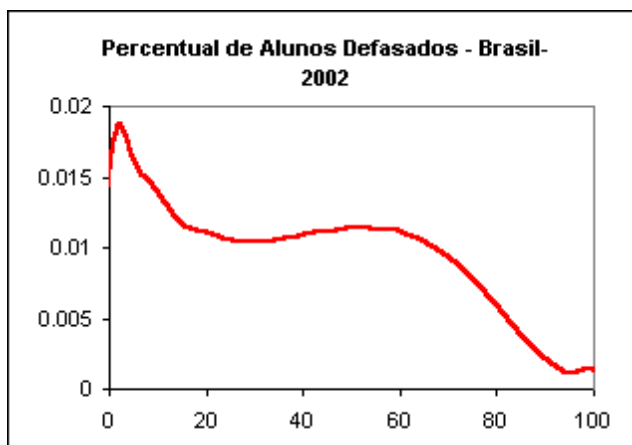


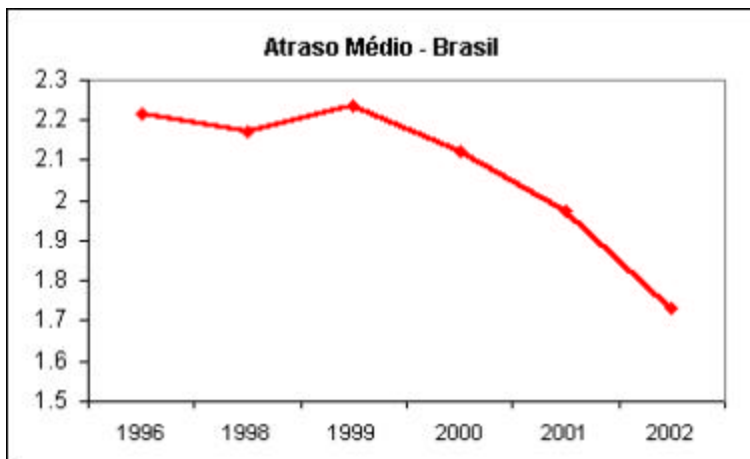
Gráfico 9: Percentual de alunos defasados – Brasil 2002



Observando a distribuição desta primeira variável, escolhida como de resultado, podemos notar que a densidade de escolas com o percentual de alunos com defasagem maior que dois anos vem caindo bastante nos valores acima de 80%. Este fato continua evidenciando a melhora da qualidade no ensino, apresentada no primeiro gráfico.

- Defasagem média – em anos, ou seja, qual o atraso médio nas escolas:

Gráfico 10: Atraso médio – Brasil



Esta é segunda variável de resultado que será avaliada em nossa análise. De maneira mais sutil, esta variável também evidencia a melhora na qualidade do ensino fundamental brasileiro. Em 1996, o atraso médio nas escolas era maior do que dois anos, ficando em torno de 2,2 anos, e em 2002 esta variável cai para 1,7 ano. Analisando as distribuições a seguir podemos constatar que a cada ano a concentração de escolas fica mais próxima do um ano de atraso.

A distribuição desta variável no Brasil é a seguinte:

Eixo Y = Densidade e Eixo X = Atraso Médio

Gráfico 11: Atraso médio – Brasil 1996

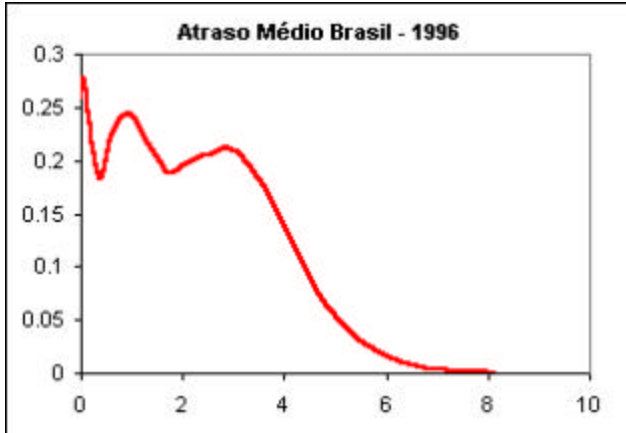


Gráfico 12: Atraso médio – Brasil 1998

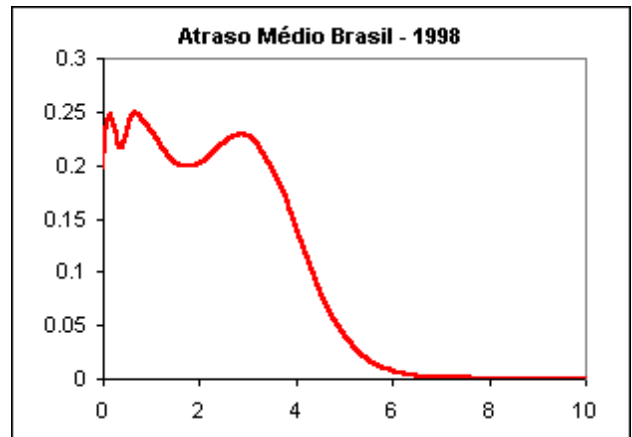


Gráfico 13: Atraso médio – Brasil 1999

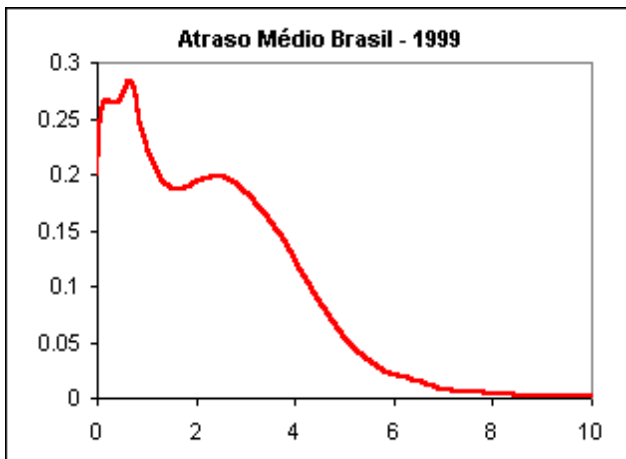


Gráfico 14: Atraso médio – Brasil 2000



Gráfico 15: Atraso médio – Brasil 2001

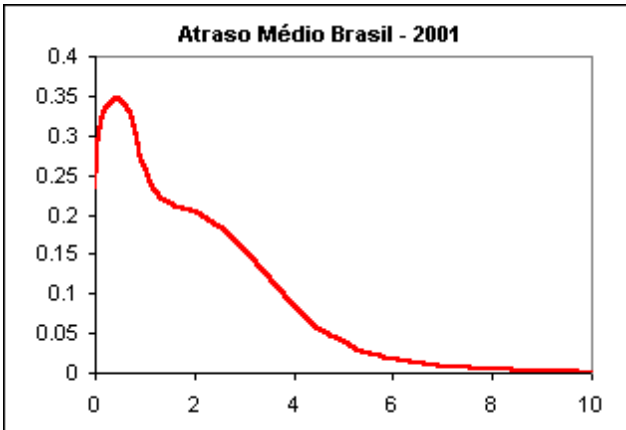


Gráfico 16: Atraso médio – Brasil 2002



Para saber o impacto da municipalização não basta comparar um grupo com o outro; esse seria um resultado viesado, como foi visto acima. Deve-se saber o que aconteceria com as escolas municipalizadas caso tivessem permanecido estaduais. Assim, não podemos apenas utilizar as escolas que permaneceram estaduais como grupo de controle para inferir sobre o que aconteceria com as que foram municipalizadas caso a mudança de rede não tivesse acontecido. Esta técnica assume que as escolas que permaneceram estaduais representam corretamente as que foram municipalizadas caso não tivessem trocado de rede, o que pode ser ou não verdade.

Queremos investigar se a mudança de rede de estadual para municipal influenciou as variáveis de resultado, e, indo além, qual a dimensão desse efeito.

O efeito real da municipalização será a diferença entre o resultado médio das escolas municipalizadas (observado) e a média dos resultados na situação contrafactual de “não participação no programa” (não observado).

$$G = E(\Delta|X, P=1) = E((Y_1 - Y_0)| X, P=1) = E(Y_1| X, P=1) - E(Y_0| X, P=1)$$

Onde,

$E(Y_1|X, P=1)$ = dado que a escola foi municipalizada, qual foi o seu resultado?

$E(Y_0|X, P=1)$ = componente contrafactual = dado que a escola foi municipalizada, qual seria o seu resultado se ela tivesse permanecido estadual?

Para chegar a essa estimativa utilizaremos a metodologia de *matching* apresentada anteriormente, baseada no escore de propensão.

4.3. Procedimentos

Em primeiro lugar é necessário preparar a base de dados de maneira que apenas as escolas estaduais permaneçam no arquivo.

Em seguida, devemos criar uma variável que indicará a troca de rede da escola. Esta variável deve ser criada a cada ano, comparando com a estrutura administrativa da mesma escola no ano anterior. Se a escola trocou de rede, a variável tem valor = 1, e se não trocou, é igual a zero. Essa variável irá se chamar EstMunic97, para os estados onde a municipalização ocorreu de 1996 para 1997, sendo o único estado analisado nesse período o de Santa Catarina. E será chamada EstMunic98 para os estados em que a municipalização ocorreu de 1997 para 1998, no caso a análise será restrita aos estados de Minas Gerais, Ceará e Bahia.

Como explicitado anteriormente, o passo inicial na determinação dos grupos de controle e de tratamento é a formulação do modelo de municipalização que irá gerar um escore de propensão para cada escola estadual, refletindo a probabilidade que essa escola tem de passar para a rede municipal.

A seguir, apresentamos os passos necessários para a implementação do algoritmo utilizado no cálculo do escore de propensão – cálculo programado por Becker S. e Ichino A. (2002) utilizando o programa Stata (pscore.ado):

1) Estimar o modelo *probit* para as escolas que foram municipalizadas:

$$\Pr\{P = 1 | X\} = @ (h(X)),$$

Onde @ denota a c.d.f. e $h(X)$ e a especificação inicial que inclui todas as covariâncias como termos lineares sem as iterações.

2) Ordenar as observações de maneira crescente, por escore de propensão, e dividir a amostra em k blocos.

- 3) Em cada bloco testar se a média do escore de propensão das variáveis de controle ($P=0$) e de tratamento ($P=1$) são iguais.
- 4) Caso o teste falhe para algum bloco/intervalo, dividi-lo em dois; o teste é feito novamente.
- 5) Continuar esse procedimento até que em todos os blocos as médias do escore de propensão sejam iguais.
- 6) Em cada bloco testar se as médias de cada variável observada são iguais para as variáveis de controle e de tratamento. Essa é a condição necessária para a hipótese de balanço.
- 7) Se a média de alguma variável observada diferir, será necessária uma especificação menos parcimoniosa de $h(X)$.

Os passos de 2 a 7 podem ser feitos dentro de uma região de suporte comum. Essa restrição implica que o teste da hipótese de balanço seja feito apenas entre observações cujo escore de propensão pertençam a um mesmo intervalo. A condição de suporte comum melhora a qualidade do *matching* utilizado na estimativa de ATT.

O algoritmo utilizado para definir os grupos de controle e de tratamento e assim calcular o ATT, é definido a seguir.

A função do Stata que estaremos utilizando para a análise será “*psmatch2*”.

Essa função é capaz de implementar todos os métodos de *matching* explicados acima, ajustando, no estado de pré-tratamento, as diferenças das características observadas entre os grupos de controle e de tratamento. O estado de tratamento é identificado pela variável “*depvar = =1*”, e o estado de não tratamento, “*depvar = = 0*”. Neste caso, estaremos fornecendo o escore de propensão tendo em vista que este foi calculado previamente conforme o procedimento I.

O `psmatch2` armazena o efeito do tratamento nos tratados `r(att)` de maneira que podemos facilmente calcular os desvios dessa estimativa através do método de *bootstrap*.

Quando temos mais de uma variável de resultado – que é o caso da nossa análise, já que estamos analisando o resultado ao longo de cinco anos – a função irá armazenar cada ATT separadamente e criar as seguintes variáveis:

- Impacto do programa (ATT) na variável de resultado “Percentual de alunos defasados” nos anos 1996, 1998, 1999, 2000, 2001 e 2002 respectivamente: `r(att_pdefas96)`, `r(att_pdefas98)`, `r(att_pdefas99)`, `r(att_pdefas00)`, `r(att_pdefas01)` e `r(att_pdefas02)`.
- Impacto do programa (ATT) na variável de resultado “atraso médio” nos anos 1996, 1998, 1999, 2000, 2001 e 2002 respectivamente: `r(att_Matraso96)`, `r(att_Matraso98)`, `r(att_Matraso99)`, `r(att_Matraso00)`, `r(att_Matraso01)` e `r(att_Matraso02)`.

Além dessas, a função irá criar algumas outras variáveis para facilitar a análise, sendo elas:

`_treated`: essa é a variável indicadora dos grupos de controle e de tratamento. Será igual a zero quando a observação/escola pertencer ao grupo de controle, e igual a um quando a observação/escola pertencer ao grupo de tratamento.

`_support`: indica se a observação/escola pertence à região de suporte comum (igual a 1) ou se esta for a dessa região (igual a zero). Com esta variável podemos ter uma noção de quanto a amostra foi reduzida ao impormos a restrição de suporte comum.

`_weight`: armazena o peso que foi dado a cada observação/escola do grupo de controle. Com isto, podemos calcular toda a distribuição do resultado da situação contrafactual de interesse, ou seja, a distribuição de qual seria o resultado das escolas municipalizadas caso elas tivessem permanecido estaduais com os devidos pesos para cada escola.

Considerações especiais

Nos estados de Minas Gerais e da Bahia retiramos da base de dados escolas que incluíam o ensino supletivo em sua grade. Isto porque com a implementação do Fundef e a vinculação dos recursos à matrícula no ensino fundamental, em alguns municípios foram observados casos em que alunos do ensino supletivo tiveram sua matrícula transferida para o ensino fundamental regular para que a arrecadação fosse maior. Diante disso, preferimos retirar da base as escolas com ensino supletivo e conseguimos evitar assim que as variáveis de resultado, atraso médio e percentual de alunos com defasagem maior que dois anos, ficassem adulteradas.