

## IMPACTO DA VIOLÊNCIA NAS ESCOLAS PAULISTAS SOBRE O DESEMPENHO ACADÊMICO DOS ALUNOS

EVANDRO CAMARGOS TEIXEIRA \*  
ANA LÚCIA KASSOUF †

### Resumo

O objetivo desse artigo é mensurar o impacto da violência sobre o desempenho escolar dos alunos do estado de São Paulo no ano de 2007. Para tal, é utilizado um modelo logit multinível, que possibilita a análise hierárquica dos dados. Os resultados demonstram que a violência nas escolas diminui em 0,54% a probabilidade de os alunos da terceira série do ensino médio apresentarem um desempenho satisfatório em matemática. A partir desse resultado, espera-se que políticas públicas sejam implementadas no intuito de diminuir a violência e elevar a probabilidade de desenvolvimento econômico do estado e consequentemente do país.

**Palavras-chave:** Violência; Desempenho Escolar; Desenvolvimento Econômico.

### Abstract

The aim of this paper is to measure the impact of violence on school performance of students in the state of São Paulo in 2007. To this end, we use a multilevel logit model which enables the analysis of hierarchical data. The results show that violence in schools decreased by 0.54% the likelihood of third graders from the school present a satisfactory performance in Mathematics. From this result, it is expected that public policies are implemented in order to reduce violence and increase the probability of state economic development and consequently the country.

**Keywords:** Violence; School Performance; Economic Development.

**JEL classification:** I21, K42

**DOI:** <http://dx.doi.org/10.1590/1413-8050/ea124436>

---

\* Professor Adjunto II do Departamento de Economia Rural da UFV. Viçosa, Minas Gerais, Brasil. E-mail: [evandro.teixeira@ufv.br](mailto:evandro.teixeira@ufv.br)

† Professora Titular do Departamento de Economia, Administração e Sociologia da Esalq/USP. Piracicaba, São Paulo, Brasil. E-mail: [anakassouf@usp.br](mailto:anakassouf@usp.br)

## 1 Introdução

A educação é uma variável chave na determinação do progresso econômico de qualquer país. Nesse sentido, indivíduos mais instruídos elevam o nível de capital humano e conseqüentemente as taxas de crescimento econômico. No entanto, um relevante fenômeno social vem diminuindo o nível de capital humano e afetando, em conseqüência, o desenvolvimento socioeconômico dos países: a violência nas escolas.

De acordo com Abramovay & Rua (2002), em todo o mundo ocidental moderno, a ocorrência de violência nas escolas não é um fenômeno recente. Desde os primeiros estudos realizados sobre o assunto, nos Estados Unidos, na década de 1950, diversas dimensões desse fenômeno passaram por grandes mudanças e os problemas decorrentes assumiram maior gravidade. Algumas dessas notáveis transformações foram: o surgimento de armas nas escolas; a disseminação do uso de drogas; e a expansão do fenômeno das gangues, influenciando na rotina das escolas, eventualmente associadas ao narcotráfico.

Guimarães (1998) afirma que, no Brasil, a violência nas escolas tem aumentado bruscamente nos últimos anos, sendo caracterizada principalmente pela presença do narcotráfico, de gangues e de pobreza acentuada. Nesse sentido, em um estudo realizado em diversas capitais do país, Abramovay & Rua (2002) concluíram que a violência física e a violência contra a propriedade são predominantes nas escolas dentre todos os tipos de violência, tendo diversos desdobramentos negativos sobre a qualidade do ensino e a aprendizagem. Tal efeito tem proporções elevadas para o país ao diminuir o nível de capital humano, tão importante no processo de desenvolvimento econômico.

Diante da importância do tema, este trabalho tem como objetivo analisar o impacto da violência nas escolas sobre o desempenho acadêmico dos alunos paulistas da terceira série do ensino médio em matemática no ano de 2007. O trabalho está dividido em mais cinco seções, além desta introdução. Na próxima seção serão apresentadas as evidências empíricas que concernem o tema e em seguida as estatísticas descritivas do banco de dados utilizado. Posteriormente, serão descritos o método e as variáveis a serem utilizadas. Logo em seguida, são apresentados os resultados da estimação do modelo proposto. Por fim, serão apresentadas as considerações finais.

## 2 Evidências Empíricas

A violência escolar tem despertado muita atenção por parte das autoridades públicas. Mas, apesar da importância do tema, poucas pesquisas têm sido realizadas. A maioria dos trabalhos foi realizada nos últimos anos e apresenta diversas facetas acerca da violência nas escolas: alguns analisam o impacto da violência dentro das escolas, outros a violência no entorno das escolas e os demais analisam ambos os impactos.

O primeiro trabalho relevante que versa o tema foi o de Grogger (1997). O autor estimou as conseqüências da violência nas escolas sobre o desempenho acadêmico dos alunos do ensino médio de escolas públicas nos Estados Unidos. O autor classificou cada tipo de violência ocorrido nas escolas como sério, moderado, menor ou não existente. Além da violência escolar, Grogger (1997) levou em consideração características dos estudantes que aumentam a propensão à violência, tais como cor, *background* familiar, além de características das escolas e da região onde a mesma está inserida. O autor concluiu

que níveis moderados de violência diminuam a probabilidade de conclusão do ensino médio em 5,1%, assim como diminuem a frequência escolar em 6,9%.

Seguindo a mesma linha do trabalho de Grogger (1997), Hurt et al. (2001), Henrich et al. (2004) e Ratner et al. (2006) também examinaram a associação entre exposição à violência nas escolas e desempenho escolar. Os primeiros autores citados levaram em consideração em sua amostra alunos afro-americanos de seis a sete anos. Os resultados mostraram que a exposição à violência está positivamente relacionada à baixa *performance* escolar, sintomas de ansiedade e baixa autoestima.

Por sua vez, Henrich et al. (2004) utilizaram como amostra 759 alunos norte-americanos numa análise *cross-section*. Todavia, diferentemente dos trabalhos até então realizados, esse estudo concluiu que a vitimização não está relacionada a baixos níveis de proficiência. Entretanto, a vitimização apresentou uma relação inversa com a sensação de segurança nas escolas, no caso específico de meninos que relataram não ter suporte da família.

Já Ratner et al. (2006) utilizaram como amostra filhos de mulheres que receberam assistência pré-natal e foram continuamente examinadas pelos hospitais da cidade de Detroit (EFUA). Os resultados corroboraram com a maior parte da literatura, que concerne o tema, ou seja, a violência está negativamente relacionada às medidas de proficiência.

No mesmo ano, McGarvey et al. (2006), diferentemente dos trabalhos até aqui apresentados, levaram em consideração o crime tanto dentro como no entorno das escolas e seu impacto sobre o desempenho acadêmico de alunos do ensino fundamental e médio em Atlanta (EUA) entre 1999 e 2002. Os autores utilizaram dois métodos: Mínimos Quadrados Ordinários — MQO e estimação via variáveis instrumentais (número total de adultos na escola, distância da escola à habitação pública mais próxima, número de estações ferroviárias no bairro e distância da escola às estações de trânsito ferroviário). Os resultados obtidos pelos dois métodos foram similares. Um incidente de violência em uma escola faz com que o desempenho acadêmico caia em 2,5% (pelo método de MQO) e em 4,0% (pela estimação via variáveis instrumentais)

Ainda nesse mesmo ano, em um estudo menos abrangente, Carroll (2006) calculou o efeito de atos violentos e criminosos em escolas públicas da Carolina do Norte (EUA), entre 2003 e 2004, sobre os níveis de desempenho acadêmico de alunos da oitava série, controlando por uma série de fatores pertinentes. A fim de desmembrar o efeito da violência sobre o desempenho dos alunos em leitura e matemática e corrigir possíveis correlações estatísticas causadas pela possibilidade de omissão de variáveis, o autor usou o método de Regressões Aparentemente Não Correlacionadas — SUR. Como resultado, verificou-se que um crime ou violência diminuam marginalmente o desempenho do aluno em 0,138% e 0,143% em matemática e leitura, respectivamente.

Ammermueller (2007) analisou os efeitos da violência escolar sobre a *performance* dos estudantes em um âmbito diferente dos trabalhos até então realizados: em onze países da Europa. Assim como verificado na literatura, o autor também concluiu que a violência afeta negativamente o desempenho acadêmico dos alunos. No entanto, o autor chegou a outras conclusões interessantes. Em primeiro lugar, Ammermueller (2007) verificou que alunos da oitava série, do sexo feminino e oriundos de municípios de pequeno porte apresentaram menor probabilidade de serem vítimas da violência escolar. Além disso, ao analisar a amostra da quarta série o autor ainda percebeu que o fato de o

aluno ser imigrante aumenta a possibilidade de vitimização, enquanto que o *background* familiar diminui tal probabilidade.

Nesse sentido, enfatizando a importância das características do aluno e de seu *background* familiar, Aizer (2008) destacou que as vítimas da violência possuem várias características em comum: em sua maioria são negras, pobres e de famílias com baixa escolaridade. Assim, a autora procurou resposta para o seguinte questionamento: a baixa *performance* escolar é melhor explicada pela exposição à violência ou pelas características e o *background* familiar dos alunos? Aizer (2008) utilizou efeitos fixos a fim de considerar características não observáveis das famílias dos alunos, assim como dos bairros onde os mesmos residem, concluindo que as características e o *background* familiar são os maiores responsáveis pelo decréscimo no desempenho acadêmico dos alunos.

No mesmo ano, Hjalmarsson (2008) examinou empiricamente um aspecto diferente dentro do tema: os efeitos de apreensões e encarceramento juvenis aos 16 anos sobre a conclusão do ensino médio aos 19 anos nos Estados Unidos. O autor concluiu que jovens que são presos têm 25% menor probabilidade de completar o ensino médio em relação a jovens com a mesma idade que não foram presos.

No Brasil, a literatura é ainda mais incipiente, sendo raros os trabalhos realizados que tratam desse tema. Severnini (2007) utilizou os indicadores de violência escolar estabelecidos no Sistema de Avaliação da Educação Básica — SAEB de 2003, concluindo que alunos oriundos de escolas mais violentas tiveram, em média, pior desempenho em matemática e português. O autor constatou também que existe uma relação indireta entre violência e proficiência acadêmica, que se evidencia por meio da elevação da rotatividade dos professores.

Gama (2009) analisou a relação entre violência e proficiência escolar de alunos da quarta série do ensino fundamental no município de São Paulo em 2005, encontrando resultados muito diferentes de Severnini (2007). O autor concluiu de forma surpreendente que alunos de escolas públicas de regiões violentas têm melhor desempenho em matemática e português para crimes violentos, crimes contra o patrimônio, e roubo consumado. Além disso, utilizando equações quantílicas, Gama (2009) ainda constatou que a violência afeta mais os alunos da cauda superior da distribuição de notas.

Becker & Kassouf (2012) analisaram um aspecto diferente com relação aos trabalhos realizados até então no país: o impacto da violência no entorno das escolas sobre a violência dentro das mesmas. Para isso, as autoras utilizaram dados da Prova Brasil para estimar um modelo *logit* de efeitos fixos, concluindo que a possibilidade de observar um ato agressivo de um aluno é maior em ambientes escolares com traços da violência, onde ocorreram crimes contra patrimônio, contra a pessoa, tráfico de drogas ou atuação de gangues. Becker & Kassouf (2012) também observaram que a possibilidade de um aluno cometer uma agressão física é 3,54 vezes maior em escolas onde foi observado o mesmo comportamento por parte de um professor.

Diante da literatura que concerne o tema apresentada acima, este trabalho inova ao analisar a relação entre violência e desempenho escolar no estado de São Paulo em 2007 por meio de um modelo multinível. Tal método considera a hierarquia existente nos dados (nível individual e nível agregado) e permite ao pesquisador verificar o impacto individual da inclusão de cada variável contextual (dentre elas, a violência nas escolas) no modelo sobre a *performance* acadêmica dos alunos.

### 3 Estatísticas Descritivas

Nesta seção, será apresentada uma breve descrição da base de dados utilizada neste trabalho, tendo como fonte o Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (Saresp) 2007.<sup>1</sup> No que tange ao desempenho dos alunos da terceira série do ensino médio na prova de matemática, verifica-se que somente 3,7% apresentaram um resultado considerado “avançado”. O desempenho dos demais alunos foi considerado “básico” (24,79%) ou ainda “abaixo do básico” (71,51%).

Com relação ao gênero dos alunos, percebe-se que há uma predominância das mulheres em relação aos homens: 54,73% são do sexo feminino e 45,27% do sexo masculino. No que tange à cor, percebe-se a maioria dos alunos é branca (54,37%) ou parda (31,79%) em relação a negros (10,39%), amarelos (2,30%) e indígenas (1,15%).

A Tabela 1 apresenta o nível socioeconômico dos estudantes por meio de alguns indicadores relevantes, o que compõe o denominado *background* familiar, ou seja, características socioeconômicas individuais dos alunos da terceira série do ensino médio, que podem afetar seus respectivos níveis de desempenho escolar na prova de matemática. Dada a indisponibilidade do nível de renda dos alunos diretamente, tais indicadores têm o objetivo de verificar a presença ou não de alguns bens e condições mínimas de moradia.

Percebe-se que a grande maioria dos alunos possui em seus domicílios, principalmente, os seguintes bens: tv em cores, rádio e geladeira. Em contrapartida, a maior parte dos alunos não possui aspirador de pó. Com relação às condições de moradia, a maioria dos alunos as possui minimamente. Cabe salientar que rua sem calçamento é o problema estrutural mais recorrente entre os alunos.

Uma variável relevante a ser analisada é a inserção do aluno no mercado de trabalho. Aproximadamente 54,46% dos alunos trabalham e 45,64% não trabalham. Mais importante que isso, do total de alunos que trabalham, 44,76% relatam que tais atividades atrapalham seus estudos.

A Tabela 2 apresenta, respectivamente, o grau de escolaridade dos pais dos alunos (pais e mães), que também compõe o *background* familiar. Verifica-se que o nível de escolaridade dos pais é relativamente baixo, pois a maior parte terminou apenas o ensino fundamental (57,54% e 56,27%, respectivamente) e um contingente relativamente pequeno possui curso superior (8,61% e 9,26%, respectivamente).

Nas Tabelas 3, 4 e 5, são apresentadas as condições estruturais, assim como os principais problemas enfrentados pelos diretores das escolas onde os alunos estudam. A Tabela 3 descreve como são as condições dos principais equipamentos da escola. As quadras e os espaços para recreação e refeições são os que menos atendem aos anseios dos alunos, segundo os diretores. Já em relação aos laboratórios, a situação é mais dramática, pois simplesmente 62,57% dos diretores afirmam que os mesmos não existem em suas escolas.

Com relação às condições e conservação do ambiente escolar, vide Tabela 4, conclui-se que a maioria dos diretores as considera boas ou regulares. A maior demanda por parte dos alunos, de acordo com os diretores, é que seja melhorada a limpeza e a conservação da área externa.

---

<sup>1</sup>Ver São Paulo (2007).

**Tabela 1:** Nível socioeconômico dos alunos — proporção (%)

| Indicadores Socioeconômicos | Não Possui | Possui |
|-----------------------------|------------|--------|
| Tv em cores                 | 2,92       | 97,08  |
| Videocassete ou DVD         | 13,87      | 86,13  |
| Computador                  | 44,57      | 55,43  |
| Rádio                       | 7,84       | 92,16  |
| Máquina de lavar roupas     | 20,18      | 79,82  |
| Aspirador de pó             | 67,93      | 32,07  |
| Telefone fixo               | 27,38      | 72,62  |
| Telefone celular            | 11,03      | 88,98  |
| Geladeira                   | 5,28       | 94,72  |
| Carro                       | 45,01      | 54,99  |
| Banheiro em casa            | 2,03       | 97,97  |
| Rua com calçamento          | 14,87      | 85,13  |
| Água encanada em casa       | 3,25       | 96,75  |
| Eletricidade em casa        | 1,44       | 98,56  |

Fonte: Elaborada pelos autores a partir de informações de São Paulo (2007).

**Tabela 2:** Grau de escolaridade dos pais dos alunos — proporção (%)

| Grau de parentesco | Níveis de escolaridade dos pais |              |                 |                  |
|--------------------|---------------------------------|--------------|-----------------|------------------|
|                    | Ensino fundamental              | Ensino médio | Ensino superior | Nunca frequentou |
| Mães               | 57,54                           | 31,08        | 8,61            | 2,77             |
| Pais               | 56,27                           | 31,18        | 9,26            | 3,29             |

Fonte: Elaborada pelos autores a partir de informações de São Paulo (2007).

A Tabela 5 apresenta as principais condições adversas relatadas pelos diretores das escolas. Os diretores, em sua maioria, concordam que existe precariedade em termos de infraestrutura física nas escolas (65,27%), falta de motivação por parte dos professores (77,20%) e apoio insuficiente da comunidade (59,26). Com relação à variável foco deste trabalho — insegurança na escola (vandalismo, roubos, etc.) — 41,71% dos diretores concordam que existem tais transtornos nas escolas em que trabalham. Assim, um percentual substancial das escolas que fazem parte da base de dados que compõe este trabalho é acometido por violência, o que denota a importância do presente estudo.

#### 4 Metodologia

Segundo Riani & Rios-Neto (2004), existem três conjuntos de fatores que afetam de forma distinta o desempenho escolar de um indivíduo: familiar, escolar e comunitário. O primeiro altera a demanda intradomiciliar por educação, enquanto que os outros dois alteram o lado da oferta educacional. Muitos trabalhos utilizam variáveis relacionadas aos três conjuntos de fatores citados no mesmo nível, estimando os determinantes educacionais por meio do método de Mínimos Quadrados Ordinários (MQO). Esse tipo de análise apresenta di-

**Tabela 3:** Condições dos principais equipamentos das escolas — proporção (%)

| Indicadores              | Condições |           |       |            |                |
|--------------------------|-----------|-----------|-------|------------|----------------|
|                          | Precárias | Regulares | Boas  | Excelentes | Não existe (m) |
| Prédio escolar           | 8,00      | 36,47     | 49,74 | 5,79       | -              |
| Salas de aula            | 5,58      | 37,41     | 52,07 | 4,94       | -              |
| Laboratórios             | 9,03      | 13,96     | 12,87 | 1,57       | 62,57          |
| Sala de informática      | 10,94     | 24,95     | 46,38 | 12,16      | 5,57           |
| Biblioteca               | 10,87     | 24,38     | 36,00 | 10,19      | 18,56          |
| Quadras                  | 26,40     | 32,25     | 27,77 | 5,61       | 7,97           |
| Sala para os professores | 9,14      | 32,55     | 47,07 | 8,29       | 2,95           |
| Espaço para recreação    | 13,26     | 32,41     | 33,94 | 7,38       | 13,01          |
| Espaço para refeições    | 14,47     | 32,68     | 35,04 | 7,40       | 10,41          |

Fonte: Elaborada pelos autores a partir de informações de São Paulo (2007).

**Tabela 4:** Condições e conservação do ambiente escolar — proporção (%)

| Indicadores                                  | Condições |           |       |            |
|--|-----------|-----------|-------|------------|
|  | Precárias | Regulares | Boas  | Excelentes |
| Condições dos vidros, lâmpadas, encanamentos | 12,12     | 39,71     | 42,54 | 5,63       |
| Limpeza e conservação dos banheiros          | 5,77      | 29,17     | 55,22 | 9,84       |
| Conforto e conservação do mobiliário         | 6,72      | 40,72     | 48,22 | 4,34       |
| Limpeza e conservação da área externa        | 4,82      | 27,99     | 56,90 | 10,29      |
| Limpeza e conservação das paredes e muros    | 6,86      | 34,87     | 51,66 | 6,61       |
| Aparência geral da escola                    | 6,76      | 30,57     | 53,35 | 9,32       |

Fonte: Elaborada pelos autores a partir de informações de São Paulo (2007).

**Tabela 5:** Condições adversas observadas nas escolas — proporção (%)

| Indicadores                                     | Concordam | Discordam |
|---|-----------|-----------|
| Precariedade da infra-estrutura física          | 65,27     | 34,73     |
| Insegurança na escola (vandalismo, roubos, etc) | 41,71     | 58,29     |
| Poucas oportunidades de desenvolvimento         | 32,56     | 67,44     |
| Falta de motivação dos professores              | 77,20     | 22,80     |
| Apoio insuficiente da equipe escolar            | 36,96     | 63,04     |
| Apoio insuficiente da comunidade                | 59,26     | 40,74     |

Fonte: Elaborada pelos autores a partir de informações de São Paulo (2007).

ficuldades tanto conceituais como estatísticas.

De acordo com Shirasu & Arraes (2014), a estrutura do sistema educacional é organizada de modo hierárquico ou em níveis. Assim, um grupo de alunos constitui uma turma, o agrupamento de turmas compõe as escolas, o conjunto destas está inserido em um município, os municípios estão agrupados em estados e assim por diante. Nesse sentido, diversos fatores afetam a distribuição dos alunos entre as escolas, como questões socioeconômicas, disponibilidade de vagas ou mesmo a distância da escola ao local de residência. Tais fatores implicam na não aleatoriedade dessa distribuição ao fazerem com que os alunos de uma mesma escola sejam mais parecidos do que os de escolas diferentes. A semelhança advém desse contexto, onde as escolas estão inseridas, além do fato de compartilharem os mesmos professores e se beneficiarem da mesma infraestrutura.

Dessa forma, segundo os autores, não se pode tratar cada indivíduo de uma escola como independente dos demais alunos da mesma. Nesses casos, há violação de uma importante hipótese assumida pelo modelo de regressão linear clássico: de que os termos de erro são independentes e identicamente distribuídos, o que torna a estimação por mínimos quadrados ordinários inadequada. Tal inadequação, segundo Lee (2001), ocorre em função dos seguintes problemas: vies de agregação, estimação errônea do erro padrão, heterogeneidade dos coeficientes da regressão e diferenças entre agrupamentos.

Lee (2010) afirma que os trabalhos que não levam em consideração a natureza multinível da escola quase sempre apresentam resultados enviesados e, geralmente, a direção desse vies tende a subestimar o efeito-escola. Por levar em consideração a natureza hierárquica dos dados educacionais e a violação da hipótese de independência do termo de erro, alguns estudos na literatura têm utilizado os modelos multiníveis.

Assim, conforme salientado por Ferrão et al. (2002), na modelagem de dados educacionais, a aplicação dos modelos de regressão multinível possui duas vantagens em relação aos modelos de regressão clássica. Em primeiro lugar, na presença de correlação intraclasse, a estimação dos parâmetros do modelo via regressão clássica produz estimativas do erro padrão muito pequenas. Pode-se dizer assim que as estimativas produzidas pelos modelos de regressão multinível são geralmente mais conservadoras.

Além disso, ao decompor a variância do erro segundo os níveis hierárquicos, o modelo de regressão multinível permite ao analista a melhor compreensão e/ou explicação do processo que está sendo modelado. Nesse sentido, torna-se mais simples, por exemplo, estudar a capacidade explicativa de variáveis intraescolares face às extraescolares.

Conclui-se, portanto, que ao se trabalhar com dados em diferentes níveis é imprescindível o uso dos modelos hierárquicos. Nesse sentido, no intuito de analisar o impacto da violência nas escolas sobre o desempenho dos alunos será utilizado um modelo logístico hierárquico. Como já salientado, quando se incorpora a modelagem hierárquica no modelo logístico, pressupõe-se a existência de um conjunto de dados hierárquicos, onde a variável dependente é medida no nível mais baixo e as variáveis explicativas em todos os níveis existentes. Neste trabalho, o nível mais baixo contempla o indivíduo — nível individual — e o nível mais alto, as características da escola — nível agregado.<sup>2</sup>

---

<sup>2</sup>Ver Hox (1995).

Para a especificação do modelo de nível individual,  $y_{ij}^*$  é considerada uma variável contínua latente, não observável, que representa o desempenho acadêmico do aluno  $i$  na escola  $j$  (nível agregado). Essa variável pode ser definida a partir da seguinte relação linear:

$$y_{ij}^* = \beta_{0j} + X_{kij}\beta_{kj} + \epsilon_{ij} \quad (1)$$

onde:

$X_{kij}$  = vetor das  $k$  variáveis independentes medidas no nível individual;  
 $\beta_{0j}$  = intercepto;  $\beta_{kj}$  = vetor dos  $k$  parâmetros a serem estimados pelo modelo;  
 $\epsilon_{ij}$  = termo aleatório, com distribuição normal, com média zero e variância constante.

Considere uma variável binária observada  $y_{ij}$ , que mede o desempenho acadêmico do aluno, definido por:

$$y_{ij} = 1 \text{ se } y_{ij}^* > 0, \\ y_{ij} = 0 \text{ caso contrário}$$

tal que:

$$E(y_{ij}^* | X'_{ijk}) = X'_{ijk}\beta_{kj} \quad (2)$$

Ao se especificar uma função de probabilidade logística obtém-se o “modelo de chance proporcional” com coeficientes randômicos da seguinte forma:

$$P(y_{ij} = 1 | X_{ij}) = \frac{\exp(X'_{ijk}\beta_{kj})}{1 + \exp(X'_{ijk}\beta_{kj})} \quad (3)$$

Esse modelo é semelhante ao modelo de regressão logistic padrão. A diferença é que a análise multinível considera a natureza hierárquica dos dados e permite analisar a relação entre variáveis medidas no nível agregado e outras medidas no nível individual. Alguns ou todos os parâmetros estimados podem conter um termo aleatório que varia segundo a unidade de análise medida no nível 2, ou seja, no nível agregado. Neste trabalho, apenas o intercepto é suposto aleatório, no qual o modelo de nível 2 é especificado pelas equações a seguir:

$$\beta_{0j} = \gamma_{00} + \mu_{0j} \quad (4)$$

$$\beta_{kj} = \gamma_{k0}, \text{ com } k \neq 0 \quad (5)$$

onde:  $\mu_{0j}$  = termo residual no nível agregado com distribuição normal  $\mu_j \approx (0, \sigma_\mu^2)$ .

Definido e justificado o método a ser utilizado, é importante analisar a base de dados utilizada, que neste estudo tem como fonte o Saresp 2007 — (São Paulo 2007). Essa base de dados possui características que corroboram com a estrutura hierárquica das relações. O Saresp é um teste aplicado desde 2001 no estado de São Paulo que avalia as habilidades cognitivas dos alunos nas disciplinas de matemática e português. O desempenho do estudante é mensurado de acordo com a Teoria de Resposta ao Item, comumente chamada

de TRI. Essa variável é o indicador das habilidades e competências dos estudantes, ou seja, é o indicador de qualidade na sua formação. São avaliados alunos do ensino fundamental e do ensino médio de escolas públicas e privadas. Juntamente aos resultados dos testes são incorporadas as características socioeconômicas dos alunos, assim como informações relevantes a respeito das escolas.

A variável dependente é uma medida categórica do desempenho acadêmico dos alunos da terceira série do ensino médio em matemática, tendo como fonte primária os dados do Saesp (2007). Para classificar o desempenho dos alunos é necessário estabelecer um ponto de corte entre essas categorias. Neste trabalho, para o aluno que teve seu resultado na prova de matemática classificado como “avançado” ou “básico” atribui-se (1) na variável profmat. Em contrapartida, no caso do aluno que obteve resultado na prova classificado como “abaixo do básico” atribui-se (0) para tal variável.

A opção em se utilizar uma variável dependente categórica é bastante plausível. A nota do aluno, calculada a partir da TRI, é uma estimativa dos acertos, levando-se em consideração que o aluno tenha respondido a todas questões da prova, o que na prática não ocorre. Assim, no Saesp nunca é utilizada a média de acerto pura e simples, mas sim a proficiência média do aluno gerada a partir da TRI. Por isso, a nota individual do aluno não é divulgada, pois ela não corresponde exatamente aos seus acertos nas provas específicas.

Após o cálculo da média de proficiência de todos os alunos da escola, é realizada a distribuição dessas médias nos níveis de desempenho: abaixo do básico, básico e avançado para cada disciplina e ano/série avaliados. Assim, é determinado em quais escolas predominam alunos que estão nos níveis mais avançados ou básicos. Acredita-se que a distribuição dos alunos nos níveis expressa de forma mais adequada a realidade em termos de aprendizagem nas escolas, pois demonstra mais claramente a quantidade de alunos que precisam de maior atenção por parte dos professores e/ou direção das escolas.

A média simples da proficiência dos alunos é um resultado muito limitado, como qualquer medida de tendência central, e oferece menos subsídios no intuito de verificar quais alunos necessitam de mais atenção. Assim, no cálculo do Índice de Desenvolvimento da Educação do estado de São Paulo (IDESP) utiliza-se a distribuição dos alunos nos níveis como forma de caracterizar melhor a população da escola, obtendo-se, assim, informações se a escola possui mais alunos nos níveis inferiores ou superiores, o que determinará projetos pedagógicos dirigidos a serem encaminhados à mesma (São Paulo 2014).

Com relação às variáveis de controle, estas foram escolhidas de acordo com a literatura sobre os determinantes do desempenho educacional. Dessa forma, no nível individual foram utilizadas medidas que caracterizam a demanda escolar, ou seja, variáveis que representam características específicas dos alunos e de sua família.

No nível agregado são incluídas variáveis que caracterizam a oferta educacional, que possam influir no desempenho dos alunos. Com esse objetivo, serão utilizadas as seguintes variáveis: violência na escola, percentual de professores efetivos, participação da comunidade na escola, conservação das instalações físicas e dos equipamentos da escola, e existência e quantidade de equipamentos em condições de uso na escola.

#### 4.1 Variáveis a nível individual

A maior parte da literatura brasileira, que analisa os determinantes do desempenho acadêmico dos alunos, demonstra que os fatores que melhor explicam o desempenho escolar são as características familiares e ambientais do aluno.

Neste trabalho, as duas primeiras variáveis consideram características físicas dos alunos: “sexo” e “cor”. Assim, inicialmente, foi incluída uma variável binária que identifica o sexo do aluno: (0) para feminino e (1) para masculino. Posteriormente, foram incluídas cinco variáveis binárias que denotam a cor do estudante: branco, negro, pardo, amarelo e indígena.

Uma variável de extrema relevância, mas que apresenta uma dificuldade inerente ao questionário aplicado junto aos alunos é o nível socioeconômico, pois não existe nenhuma variável que denote o valor da renda obtida pela família do estudante. Assim, será construída por Análise de Componentes Principais — ACP,<sup>3</sup> a variável “nível socioeconômico” dos alunos, que inclui os seguintes elementos, tendo como base os critérios de classificação socioeconômica da Associação Brasileira de Empresas de Pesquisas — ABEP (2007):<sup>4</sup>

- número de automóveis;
- número de aparelhos de TV em cores;
- número de rádios;
- número de empregadas domésticas;<sup>5</sup>
- número de banheiros;
- posse de máquina de lavar roupa;
- posse de geladeira e freezer;
- posse de vídeo cassete ou dvd;
- nível de instrução do chefe de família.

Nesse sentido, com o intuito de construir o indicador de nível socioeconômico dos alunos foram criadas:

- três variáveis binárias que denotam a posse de geladeira, freezer, vídeo cassete ou dvd e máquina de lavar roupa.

---

<sup>3</sup>A Análise de Componentes Principais — ACP é uma técnica de análise multivariada que consiste em transformar um conjunto original de variáveis em outro conjunto — os Componentes Principais — CP. Os CP's são combinações lineares das variáveis originais e são estimados de forma a captar o máximo de variação total dos dados. O processo de estimação é tal que o primeiro CP capta o máximo de variância possível, o segundo capta o máximo possível do restante de variância, o terceiro o máximo possível do restante de variância, e assim sucessivamente. Segundo Mingoti (2005), o principal objetivo é o de explicar a estrutura de variância-covariância de um vetor aleatório, composto de p-variáveis aleatórias, por meio da construção de combinações lineares das variáveis originais. Tais combinações lineares são chamadas de componentes principais e não são correlacionadas entre si. Se existirem p-variáveis originais é possível obter-se p componentes principais. No entanto, em geral deseja-se obter 'redução do número de variáveis a serem avaliadas e interpretação das combinações lineares construídas.

<sup>4</sup>A única variável não incluída foi o número de empregadas domésticas, que não está disponível em São Paulo (2007).

<sup>5</sup>Com as informações disponibilizadas em São Paulo (2007) não foi possível a construção dessa variável.

- variáveis que identificam o número de automóveis, de rádios, de aparelhos de TV em cores e de banheiros no domicílio do aluno com a seguinte ponderação: (0) se o aluno não possui, (1) se o aluno possui uma unidade, e (2) se o aluno possui mais de uma unidade.
- variáveis que correspondem ao nível de instrução do pai e da mãe<sup>6</sup> do aluno com a seguinte ponderação de acordo com o nível de escolaridade: (0) nunca frequentou a escola, (1) ensino fundamental — até a quarta série, (2) ensino fundamental — até a oitava série, (3) ensino médio e (4) ensino superior.

Percebe-se pelo processo de construção das variáveis que compõem a Análise de Componentes principais — ACP que quanto maiores seus valores, melhores tendem a ser as condições socioeconômicas dos alunos. Nesse sentido, espera-se uma relação diretamente proporcional entre a variável nível socioeconômico e o desempenho escolar.

As últimas variáveis a serem incluídas no nível individual são aquelas que identificam se o aluno exerce algum tipo de trabalho. Assim, foram incluídas seis variáveis binárias:

- Trabalho1: aluno nunca trabalhou ou não está trabalhando atualmente;
- Trabalho2: aluno trabalha sem jornada fixa, até 10 horas semanais;
- Trabalho3: aluno trabalha de 11 a 20 horas semanais;
- Trabalho4: aluno trabalha de 21 a 30 horas semanais;
- Trabalho5: aluno trabalha de 31 a 40 horas semanais;
- Trabalho6: aluno trabalha mais de 40 horas semanais.

Espera-se que quanto maior o número de horas de trabalho exercidas pelo aluno, menor tenda a ser o seu desempenho escolar. Segundo Kassouf (2007), boa parte dos prejuízos na educação ocorre em função do trabalho desde a infância. De acordo com a autora, a baixa escolaridade e o desempenho escolar mais baixo, consequências do trabalho infantil, restringem as oportunidades de emprego a postos de baixa qualificação e, portanto, de baixa remuneração.

#### 4.2 Variáveis a nível agregado

As variáveis incluídas no nível agregado possuem a finalidade de captar as condições de oferta escolar, que impactam sobre o desempenho escolar dos alunos e terão como base o questionário respondido pelo diretor da escola.

A variável “violência” avalia se na escola ocorrem roubos, vandalismo, agressões, e etc. Caso o diretor reporte que há violência na escola esta variável assume o valor 1. Caso contrário, seu valor é 0. A existência de violência na escola diminuiu principalmente a frequência escolar e a concentração dos alunos. Nesse sentido, espera-se que quanto maior a violência nas escolas menor tenda a ser a performance escolar.

---

<sup>6</sup>Essas duas variáveis caracterizam a escolaridade apresentada pelo chefe de família, já que não é possível identificá-lo diretamente no questionário.

De acordo com Carroll (2006), a qualidade dos professores pode ter um efeito importante sobre o desempenho dos alunos na prova. Nesse caso, professores mais experientes, com maior titulação, e efetivos corroboram para um desempenho escolar mais elevado. Assim, foram incluídas cinco variáveis binárias que denotam qual o percentual de professores efetivos na escola:

- Profefet1: se até 25% dos professores são efetivos;
- Profefet2: de 26% a 50% dos professores são efetivos;
- Profefet3: de 51% a 75% dos professores são efetivos;
- Profefet4: de 76% a 90% dos professores são efetivos;
- Profefet5: de 91% a 100% dos professores são efetivos;

Essas cinco variáveis binárias incluídas medem também a rotatividade dos professores, pois um professor efetivo tende a mudar menos de trabalho. Como especificado por Abramovay & Rua (2002), a excessiva rotatividade dos professores nas escolas é um dos fatores que explicam o mau desempenho escolar. Nesse sentido, espera-se que quanto maior o percentual de professores efetivos mais elevado tenda a ser o desempenho escolar.

No caso das três demais variáveis de controle do nível agregado (comunidade, conservação e equipamento) também será utilizada a técnica de ACP a fim de se extrair um componente principal. A variável “comunidade” assinala a frequência em que se dá a participação da comunidade nos seguintes aspectos: colaboração dos pais na solução dos problemas da escola; presença e apoio dos pais nos eventos e festas que a escola promove; participação dos pais no Conselho de Escola; e prestação de serviços à escola (pintura, reparos, outros cuidados, doação de bens). Para cada um desses itens foi criada uma variável com a seguinte ponderação: (0) raramente ou nunca, (1) regularmente, (2) muitas vezes. Assim, espera-se que quanto maior o valor da variável “comunidade” mais elevado tenda a ser o desempenho educacional.

Por sua vez, a variável “conservação” avalia a conservação e o funcionamento das seguintes instalações e equipamentos da escola: prédio escolar; salas de aula; laboratórios; sala de informática; biblioteca; salas para exibição de filmes e DVD's; quadras, salas para os professores; sala específica para o professor coordenador; salas destinadas ao grêmio estudantil; espaço para recreação; espaço para as refeições; condições dos vidros, lâmpadas e encanamentos; limpeza e conservação dos banheiros; condições de conforto e conservação do mobiliário escolar; limpeza e condição da área externa (pátio, jardins, e etc.); condições de limpeza e conservação das paredes e muros; e aparência geral da escola (cortinas, filtros, objetos de decoração). Para cada um desses itens também foi criada uma variável com a seguinte ponderação: (0) precárias, (1) regulares, (2) boas e (3) excelentes. Nesse caso, quanto maior a conservação das instalações e dos equipamentos da escola e, portanto, o valor da variável “conservação”, mais elevada tende a ser a *performance* escolar.

Por fim, inclui-se a variável “equipamento”, que identifica se existem e quantos são os equipamentos em condições de funcionamento na escola, tais como: televisores; videocassetes e/ou dvd's; retroprojetores; computadores para uso dos professores; computadores para uso dos alunos; computadores para uso administrativo; impressoras; máquinas copadoras (xerox); antena

parabólica; e aparelhos de som. A partir de cada uma dessas informações uma variável será criada com a seguinte ponderação: (0) não existem, (1) nenhum, (2) alguns, (3) metade, e (4) todos. Assim, quanto maior o valor da variável “equipamento”, a tendência é que o desempenho escolar seja mais elevado.

## 5 Resultados

Na Tabela 6, são apresentados os resultados<sup>7</sup> da estimação dos efeitos fixos e aleatórios estimados pelo modelo *logit* multinível de intercepto randômico. Por meio dos resultados, percebe-se que todos os coeficientes foram estatisticamente significativos a um nível de significância de 1% e apresentaram os sinais esperados.

Foram testadas quatro especificações, concomitante à inclusão gradativa das variáveis contextuais. Nesse sentido, a especificação mais simples é o modelo não condicional (modelo 1), na qual nenhuma variável do nível agregado é considerada na análise, e o modelo 4 é o mais complexo, no qual todas as variáveis contextuais fazem parte da análise.

Uma questão que surge nesse momento é a seguinte: se cada modelo representa a inserção gradativa de variáveis contextuais, qual a importância da inclusão das mesmas na análise? Para responder a tal questionamento será utilizado o método proposto por Raudenbush & Bryk (2002). Esses autores especificaram uma forma de verificar em quanto a inserção das variáveis do nível agregado contribuem para reduzir o componente da variância estimada relativo ao intercepto. Quanto menor tal componente de variância, mais elevado tende a ser o poder de explicação (percentual de variância explicada) das variáveis contextuais em relação ao desempenho escolar individual dos alunos. Assim, Raudenbush & Bryk (2002) calculam o percentual da variância explicada pelas medidas contextuais, de acordo com a seguinte expressão:

$$\%VE = \left( \frac{\tau_{00}(\text{não condicional}) - \tau_{00}(\text{condicional})}{\tau_{00}(\text{não condicional})} \right) \times 100 \quad (6)$$

sendo:

$\%VE$  = percentual da variância explicada;  $T_{00}$  = componente da variância do intercepto estimado pelo modelo não condicional e pelo modelo condicional (com pelo menos uma variável de nível agregado).

Com o intuito de auxiliar na escolha do modelo mais adequado, os percentuais de variância explicada foram calculados e reportados na última linha da Tabela 1. Comparando-se o modelo 2 (condicional) com o modelo 1 (não-condicional), chega-se à conclusão de que a inclusão da violência escolar explica 1,66% da variância do intercepto, ou seja, 1,66% da variação no desempenho escolar dos alunos da terceira série do ensino médio em matemática são explicados pelos diferentes níveis de violência entre as escolas, dado o efeito fixo (variáveis do nível individual).

Em seguida, no modelo 3, acrescentam-se as variáveis que especificam o percentual de professores efetivos nas escolas. Consequentemente, o percen-

<sup>7</sup>O software estatístico utilizado na estimação do modelo foi o *Hierarchical Linear and Non-linear Modeling* — HLM 6.08. O modelo utilizado não permite o controle da possível presença de endogeneidade entre algumas variáveis explicativas e a variável dependente, configurando-se assim uma limitação do trabalho. No entanto, apesar de tal limitação, acredita-se que o modelo apresenta resultados razoáveis.

**Tabela 6:** Resultados da estimação do modelo *logit* para o desempenho em matemática da terceira série do ensino médio

| Variáveis                          | Modelo 1   | Modelo 2   | Modelo 3   | Modelo 4    | Efeitos Marginais <sup>8</sup> |
|------------------------------------|------------|------------|------------|-------------|--------------------------------|
| Nível individual                   |            |            |            |             |                                |
| Intercepto                         | -5,140162* | -5,23954*  | -5,649791* | -5,693699*  |                                |
| Sexo (ref. fe-<br>minino)          | 0,665155*  | 0,66598*   | 0,664569*  | 0,665355*   | 1,43                           |
| Cor (ref. branco)                  |            |            |            |             |                                |
| Negro                              | -0,393777* | -0,394746* | -0,395059* | -0,396588*  | -1,23                          |
| Amarelo                            | 0,173192*  | 0,173596*  | 0,173084*  | 0,174005*   | 0,95                           |
| Pardo                              | -0,237264* | -0,237522* | -0,237262* | -0,237695*  | -0,50                          |
| Indígena                           | -0,693242* | -0,694945* | -0,695509* | -0,69989*   | -6,93                          |
| Trabalho (ref. Trabalho1)          |            |            |            |             |                                |
| Trabalho2                          | -0,711313* | -0,71114*  | -0,71239*  | -0,713012*  | -2,30                          |
| Trabalho3                          | -0,262729* | -0,263154* | -0,261916* | -0,262328*  | -0,89                          |
| Trabalho4                          | -0,303945* | -0,304479* | -0,303328* | -0,304124*  | -1,13                          |
| Trabalho5                          | -0,302434* | -0,303063* | -0,301528* | -0,301662*  | -0,94                          |
| Trabalho6                          | -0,369175* | -0,369692* | -0,367484* | -0,368164*  | -1,00                          |
| Nível Socioe-<br>conômico          | 0,103155*  | 0,103086*  | 0,102863*  | 0,102801*   | 0,06                           |
| Nível agregado                     |            |            |            |             |                                |
| Violência                          | -          | -0,261226* | -0,257718* | -0,141005*  | -0,54                          |
| Professor Efetivo (ref. Profefet1) |            |            |            |             |                                |
| Profefet2                          | -          | -          | 0,472538*  | 0,450266*   | 3,61                           |
| Profefet3                          | -          | -          | 0,737299*  | 0,699328*   | 5,33                           |
| Profefet4                          | -          | -          | 0,739263*  | 0,714058*   | 5,39                           |
| Profefet5                          | -          | -          | 0,83447*   | 0,775984*   | 7,84                           |
| Comunidade                         | -          | -          | -          | 0,032077*   | 0,04                           |
| Conservação                        | -          | -          | -          | 0,061518*   | 0,05                           |
| Equipamento                        | -          | -          | -          | 0,024145*** | 0,03                           |
| Efeito Aleatório                   |            |            |            |             |                                |
| Componente<br>da Variância         | 0,759      | 0,747      | 0,730      | 0,712       |                                |
| % Variação<br>Explicada            |            | 1,66       | 3,88       | 6,25        |                                |

Legenda: \*, \*\*, \*\*\* denotam níveis de significância de, respectivamente, 1%, 5% e 10%.

<sup>8</sup> Dado que o modelo *logit* é log-linear, para que sejam analisados os efeitos marginais de cada variável sobre a probabilidade de o aluno apresentar um desempenho escolar avançado, foi calculada a razão de chance de o evento ocorrer, que é igual à exponencial do coeficiente estimado. Segundo Noronha & Andrade (2007), esse valor é igual a:  $[\exp(\text{coeficiente estimado} \times \text{desvio padrão do coeficiente estimado}) - 1] \times 100$ .

tual de variância explicada alcança um patamar mais elevado: 3,88. Por fim, no modelo 4, são incluídas as variáveis que denotam a qualidade da escola: participação da comunidade, conservação de instalações e equipamentos, e existência de equipamentos importantes na escola. A adição destas variáveis foi responsável, em conjunto com as demais variáveis contextuais do nível agregado, pela explicação de 6,25% da variância do intercepto. Portanto, percebe-se que o modelo 4 deve servir de base para analisar o impacto da violência sobre o desempenho escolar, pois o mesmo é o que melhor explica a variância do intercepto.

A partir dessa definição, inicialmente, será analisado o impacto da variável foco deste trabalho — violência nas escolas — sobre o desempenho dos alunos da terceira série do ensino médio em matemática. Conclui-se que uma elevação na violência reduz em 0,54% a probabilidade de que o desempenho dos alunos da terceira série do ensino médio no estado de São Paulo seja satisfatório em matemática. Esse resultado vai ao encontro das análises realizadas por McGarvey et al. (2006), e Carroll (2006). No primeiro trabalho, a violência reduz o desempenho dos alunos em 2,5% (pelo método de MQO) e em 4,0% (pela estimação via variáveis instrumentais); e já no segundo em 0,138% e 0,143% em matemática e leitura, respectivamente.

O que explica tal resultado e existem outras consequências? Como salientado anteriormente por Abramovay & Rua (2002), pode-se concluir que a violência escolar tem consequências perversas sobre a qualidade do ensino e o nível de aprendizagem dos alunos, materializando-se na diminuição no nível de concentração, menor frequência às aulas, além da elevação na desmotivação e rotatividade por parte dos professores. Além disso, a violência escolar afeta indiretamente o nível de capital humano, pois jovens com menor nível de desempenho e aprendizagem tendem a se tornar profissionais menos qualificados, que consequentemente terão menos condições de contribuir no processo de desenvolvimento econômico do país.

Cabe salientar que a literatura mais recente tem levado em consideração um aspecto adicional: a violência pode causar impactos de acordo com o horizonte temporal considerado. Mesmo não analisando especificamente a violência escolar, Chamarbagwala & Mórán (2011) e Justino et al. (2014) utilizaram dados em *coorte* para avaliar efeitos de curto e longo prazo das guerras civis sobre o desempenho dos alunos. Os primeiros autores concluíram que o impacto dos 36 anos de guerra civil na Guatemala reduziu de forma mais severa a escolaridade de homens e mulheres residentes nas zonas rurais e no longo prazo, considerando três *coortes*: 1960-1978, 1979-1984 e 1985-1996. Assim, a escolaridade de homens e mulheres nas zonas rurais da Guatemala decresceu nos três *coortes* citados em 6, 15 e 23%; e 3, 12 e 30%, respectivamente.

Por sua vez, Justino et al. (2014) avaliaram o impacto da violência no Timor Leste em 1999, durante a ocupação da Indonésia, sobre o desempenho dos alunos no ensino primário em 2001 (curto prazo) e 2007 (longo prazo). Os autores concluíram que o impacto é maior sobre os meninos no longo prazo, onde boa parte destes abandonou a escola para se dedicar às atividades militares.

Além da violência escolar, outro aspecto contextual que impacta na proficiência dos alunos em matemática é o percentual de professores efetivos na escola, representado pelas variáveis binárias *profefet2*, *profefet3*, *profefet4* e *profefet5*, sendo omitida a variável de referência — *profefet1* —, que identifica as escolas onde apenas até 25% dos professores são efetivos. Como esperado, todos os coeficientes estimados das variáveis binárias apresentaram

um sinal positivo, o que significa que escolas com uma maior proporção de professores efetivos elevam a probabilidade de os alunos apresentarem um desempenho escolar avançado.

Para se ter uma ideia da importância dessas variáveis sobre a proficiência escolar, o efeito marginal calculado para a variável *profefet5* foi de 7,84, ou seja, o fato de uma escola possuir de 91 a 100% de professores efetivos aumenta em 7,84% a probabilidade de o aluno ter um bom desempenho em matemática.

Cabe ainda sublinhar que se levarmos em consideração que o percentual de professores efetivos é uma boa *proxy* da rotatividade dos professores, confirma-se o resultado encontrado por Severnini (2007). Nesse sentido, uma elevação da rotatividade dos professores diminuiu o nível de performance escolar.

Ainda no nível agregado, as variáveis comunidade, conservação e equipamento, ambas construídas por meio de Análise de Componentes Principais — ACP, possuem a função de qualificar a escola onde o aluno estuda. A variável comunidade representa a participação e contribuição da comunidade nas atividades implementadas pelas escolas. Por sua vez, conservação descreve qual é o estado de conservação das instalações e equipamentos que a escola possui. Já a variável equipamento denota a existência e quantidade de equipamentos disponíveis na escola para utilização.

Os sinais dos coeficientes estimados para essas três variáveis contextuais apresentaram o sinal esperado, ou seja, positivo. Porém, os valores estimados dos efeitos marginais foram relativamente pequenos. Assim, uma elevação das variáveis comunidade, conservação e equipamento é responsável por um aumento de 0,04 0,05%, e 0,03% na probabilidade de o aluno apresentar um desempenho satisfatório em matemática, respectivamente.

Analisados os impactos das variáveis contextuais, é importante verificar os efeitos fixos, ou seja, a relação entre as variáveis do nível individual e a probabilidade de o aluno apresentar um desempenho satisfatório em matemática. A variável sexo do aluno assume 1 se o aluno for do sexo masculino e 0 se for do sexo feminino. Percebe-se por meio da Tabela 1 que o fato de o aluno ser do sexo masculino eleva em 1,43% a probabilidade do mesmo apresentar um desempenho avançado em matemática.

Considerando agora a cor do aluno, foram criadas cinco variáveis binárias: branco (omitida), negro, amarelo, pardo e indígena. Pode-se concluir que o fato de o aluno ser branco eleva a probabilidade de um melhor nível de proficiência em matemática em relação a negros, pardos e indígenas da ordem de 1,23%, 0,5% e 6,93%, respectivamente. Já em relação ao aluno da cor amarela a conclusão é inversa. Se o aluno for amarelo, sua probabilidade de apresentar um melhor desempenho em matemática aumenta em 0,95%.

As variáveis que representam a inserção do aluno no mercado de trabalho são muito importantes, pois como especificado anteriormente por Kasouf (2007), trabalhar desde a infância tende a diminuir a proficiência escolar. Foram incluídas seis variáveis binárias: trabalho1 (omitida), trabalho2, trabalho3, trabalho4, trabalho5 e trabalho6. Como esperado, as variáveis apresentaram um sinal negativo, podendo-se concluir que o fato de o aluno trabalhar diminui a sua probabilidade de apresentar um desempenho satisfatório em matemática. Tal efeito é potencializado no caso da variável trabalho2. Nesse caso, o fato de o aluno trabalhar até 10 horas semanais diminui a probabilidade do mesmo alcançar um resultado adequado em matemática em 2,3%, quando comparado ao aluno que não trabalha.

Por fim, assim como nos trabalhos de Grogger (1997) e Aizer (2008), leva-se em consideração neste estudo o *background* familiar, representado pela variável nível socioeconômico. Tal variável foi construída também via Análise de Componentes Principais — ACP, levando-se em consideração atributos socioeconômicos dos alunos, como escolaridade dos pais, além da posse e quantidade de eletrodomésticos, como rádio, tv, máquina de lavar, freezer, geladeira, etc. O coeficiente estimado dessa variável apresentou um sinal esperado positivo. Nesse sentido, segundo o modelo, um maior nível socioeconômico do aluno eleva sua probabilidade de apresentar desempenho satisfatório em matemática em 0,06%.

## 6 Considerações finais

O principal objetivo deste trabalho foi o de mensurar o impacto da violência nas escolas sobre o desempenho dos alunos paulistas da terceira série do ensino médio em matemática no ano de 2007. Quando são analisados os fatores que determinam a proficiência dos alunos verifica-se que existe uma vasta gama de fatores explicativos: características individuais dos alunos (sexo e cor), *background* familiar (renda domiciliar e escolaridade dos pais), e características da escola (infraestrutura, qualidade dos professores, e violência).

Percebe-se, assim, que o desempenho escolar dos alunos pode ser influenciado por variáveis de níveis hierárquicos distintos, o que foi negligenciado por diversos trabalhos já realizados. Tal negligência teve como consequência estimações imprecisas, onde ignorou-se a possível existência de falácia ecológica e de falácia atomística.

Diante disso, este estudo contribui de forma relevante na literatura econômica do crime, pois utiliza um modelo *logit* multinível, controlando a relação existente entre violência nas escolas e o desempenho escolar. Nesse caso, considerou-se dois níveis: individual e agregado.

Os resultados do trabalho demonstraram que a violência escolar diminui a probabilidade de o aluno da terceira série do ensino médio apresentar um desempenho satisfatório na prova de matemática. Nesse sentido, a modelagem multinível permite que se conclua o seguinte: um aluno do mesmo sexo, mesma cor e nível socioeconômico similar pode apresentar um rendimento escolar menor caso estude em uma escola com mais violência.

Porém, não é possível afirmar com exatidão de que forma o impacto da violência se materializa sobre o desempenho escolar. No entanto, pode-se inferir que a violência nas escolas diminui o nível de concentração dos alunos, reduz a frequência às aulas e também aumenta a desmotivação e rotatividade por parte dos professores.

Dessa forma, a violência escolar afeta ainda o nível de capital humano, pois crianças e adolescentes que apresentam um menor nível de desempenho escolar e aprendizagem tendem a se tornar profissionais menos qualificados. Consequentemente, eles terão menos condições de contribuir no processo de desenvolvimento econômico do país.

Diante dessa grave consequência surge a necessidade urgente de que sejam implementadas políticas públicas, que reduzam a violência nas escolas, aperfeiçoem o nível de ensino e, consequentemente, elevem o nível de capital humano, tão importante no processo de desenvolvimento econômico de qualquer país.

## Referências Bibliográficas

- ABEP (2007), *O novo critério padrão de classificação econômica Brasil/2008*, number 41 p., Campinas.
- Abramovay, M. & Rua, M. (2002), *Violências nas escolas*, Vol. 1 of 400 p., 2. ed. edn, UNESCO, Brasília.
- Aizer, A. (2008), Neighborhood violence and urban youth, Working Paper NBER 13773, National Bureau of Economic Research - NBER, New York. 47p.
- Ammermueller, A. (2007), Violence in european schools: victimization and consequences, Discussion Paper 07-004, Centre for European Economic Research - ZEW, Mannheim. p. 1-40.
- Becker, K. L. & Kassouf, A. L. (2012), Violência nas escolas: uma análise da relação entre o comportamento agressivo dos alunos e o ambiente escolar, in 'Encontro Nacional de Economia. Porto de Galinhas, XL, 2012. Anais... Porto de Galinhas', p. 15 p.
- Carroll, B. (2006), *The effects of school violence and crime on academic achievement*, 33 p., Davidson College.
- Chamarbagwala, R. & Mórán, H. E. (2011), 'The human capital consequences of civil war: Evidence from guatemala', *Journal of Development Economics* 94(1), 41-61.
- Ferrão, M., K., B. & Santos, D. (2002), 'O impacto de políticas de não repetência sobre o aprendizado dos alunos da quarta série', *Pesquisa e Planejamento Econômico* 32(3), 495-514. Rio de Janeiro.
- Gama, V. (2009), Uma análise de relação entre violência escolar e proficiência no município de são paulo, Mestrado em economia aplicada, Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto.
- Grogger, J. (1997), 'Local violence and educational attainment', *Journal of Human Resources* 32(4), 659-682. Wisconsin.
- Guimarães, A. (1998), *Escola, galeras e narcotráfico*, 248 p., Rio de Janeiro: UFRJ.
- Henrich, C. C., Schwab-Stone, M., Fanti, K., Jones, S. M. & Ruchkin, V. (2004), 'The association of community violence exposure with middle-school achievement: a prospective study', *Journal of Applied Developmental Psychology* 25(3), 327-348. Atlanta.
- Hjalmarsson, R. (2008), 'Criminal justice involvement and high school completion', *Journal of Urban Economics* 63(2), 613-630. Atlanta.
- Hox, J. (1995), *Applied multilevel analysis*, 126 p., T.T. Publikatie, Amsterdam.
- Hurt, H., Malmud, E., Brodsky, N. L. & Giannetta, J. (2001), 'Exposure to violence. psychological and academic correlates in child witnesses', *Archives of Pediatrics Adolescent & Medicine* 155(12), 1351-1356.

Justino, P., Leone, M. & Salardi, P. (2014), 'Short and long-term impact of violence on education: The case of timor leste', *World Bank Economic Review* 28(2), 320–353.

Kassouf, A. L. (2007), 'O que conhecemos sobre o trabalho infantil?', *Nova Economia* 17(2), 323–350. Belo Horizonte.

Lee, V. E. (2001), What are multilevel questions, and how might we explore them with quantitative methods?, in 'Estudos em Avaliação Educacional, São Paulo', Vol. 24, pp. 31–45.

Lee, V. E. (2010), 'A necessidade dos dados longitudinais na identificação do efeito-escola', *Revista Brasileira de Estudos Pedagógicos* 91(229), 471–480. Brasília.

McGarvey, M., Smith, W. J. & Walker, M. B. (2006), The interdependence of school outcomes and school and neighborhood crime, Working Paper 07-19, Georgia State University, Department of Economics, Nebraska. 22 p.

Mingoti, S. (2005), *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*, 292 p., Belo Horizonte: UFMG.

Noronha, K. S. & Andrade, M. (2007), 'O efeito da distribuição de renda sobre o estado de saúde individual no brasil', *Pesquisa e Planejamento Econômico* 37(3), 521–556. Rio de Janeiro.

Ratner, H. H., Chiodo, L., Covington, C., Sokol, R. J., Ager, J. & Delaney-Black, V. (2006), 'Violence exposure, iq, academic performance, and children's perception of safety: evidence of protective effects', *Merrill-Palmer Quarterly* 52(2), 264–287.

Raudenbush, S. W. & Bryk, A. (2002), *Hierarchical linear models: applications and data analysis methods*, 485 p., 2 nd ed. edn, London: Sage.

Riani, J. L. R. & Rios-Neto, E. (2004), Impacto dos fatores familiares, escolares e comunitários na quantidade e qualidade do ensino no estado de minas gerais, in 'SEMINÁRIO SOBRE ECONOMIA MINEIRA, 11., 2004, Diamantina. Anais... Diamantina: Cedeplar', 22 p.

Severnini, E. (2007), A relação entre violência nas escolas e proficiência dos alunos, Mestrado em economia, Pontifícia Universidade Católica do Rio de Janeiro.

**URL:** Disponível em: <[www.maxwell.lambda.ele.puc-rio.br](http://www.maxwell.lambda.ele.puc-rio.br)>. Acesso em: 20 ago. 2010.

Shirasu, M. R. & Arraes, R. A. E. (2014), Determinantes da evasão e repetência escolar, in 'X Encontro Economia do Ceará em Debate. Fortaleza'.

São Paulo (2007), 'Secretaria de educação do estado de são paulo. sistema de avaliação de rendimento escolar do estado de são paulo (saresp). são paulo'. 1 cd-rom.

São Paulo (2014), 'Secretaria de educação do estado de são paulo. sistema de avaliação de rendimento escolar do estado de são paulo (saresp). são paulo'.  
**URL:** Disponível em:< [www.educacao.sp.gov.br/saresp](http://www.educacao.sp.gov.br/saresp)>.