

## Área 10 - Economia Regional e Urbana

### A Qualidade do Ambiente Urbano afeta o Desempenho Escolar? Uma Análise do Caso das Favelas da Cidade do Recife

Julia Rocha Araujo – UFPE/PIMES

Raul da Mota Silveira Neto – UFPE/PIMES

**Resumo:** O presente estudo tem como objetivo analisar empiricamente a influência das favelas sobre o desempenho escolar dos alunos da rede pública do Recife. Para tanto, utilizamos a base de dados oriunda da pesquisa realizada pela Fundaj em 2013, em que entrevistou diretores, professores, responsáveis e alunos do 6º ano das escolas públicas dessa cidade. Essa é a primeira base a oferecer informações georreferenciadas dos alunos e das escolas do Recife, o que permitiu distinguir os alunos entre moradores ou não das favelas. A partir da identificação dos moradores da favela foi possível empregar as técnicas de *Propensity Score Matching*, os métodos de reponderação e o estimador de Oaxaca-Blinder para tentar contornar os problemas ocasionados pela distribuição não aleatória dos alunos ao longo do tecido urbano. Os resultados encontrados sugerem a influência negativa das favelas sobre o desempenho escolar.

**Palavras chaves:** Efeito vizinhança, desempenho escolar, favelas.

**Abstract:** The aim of this paper is to empirically analyse the influence of living in slums on scholar performance of public system students of Recife, Brazil. For that, we use Fundaj's data stem from a research that interviewed principals, teachers, parents and 6th grade students. These data are the first one to offer geocoded addresses of the students and schools of Recife, and we identify, for the first time, the students of Recife that live in slums. Then, we use Propensity Score Matching technique, reweighting methods and Oaxaca-Blinder's estimator to deal with the problems due to the non random distribution of the students over the city. The results suggest the negative influence of the slums on the scholar performance.

**Keywords:** Neighborhood effect, school performance, slums.

**Código JEL:** I24, R23.

## 1 Introdução

Existe um consenso na literatura de que as diferenças nas características socioeconômicas entre os indivíduos contribuem para a reprodução da desigualdade educacional (Paes de Barros *et al.*, 2006; Gonçalves e França, 2008). Mais especificamente, o perfil social de uma família tende a ser reproduzido entre as gerações que se seguem: filhos de pais pobres e pouco escolarizados tendem a apresentar essas mesmas características. Constitui-se, assim, um ciclo que dificulta a mobilidade social dos mais pobres.

Se o espaço social exercer um importante papel sobre a trajetória e o desempenho individual, então o ciclo da pobreza pode ser agravado dentro de um contexto de economia urbana na qual as classes sociais são territorialmente segregadas. O problema surge quando o isolamento territorial se transforma no isolamento sociocultural da população mais carente, o que conduziria “à formação de conjunto de valores e visões de mundo muitas vezes desconectado da cultura dominante” (Sant’Anna, 2009, p. 2). Dentro desse enfoque, a segregação residencial pode potencializar diversos problemas sociais, tais como a atividades criminosas (Case e Katz, 1991; Bursik Jr e Grasmic, 1999; Kline *et al.*, 2005), gravidez na adolescência (Anderson, 1991; Crane, 1991; Evans *et al.*, 1992), status de emprego e diferencial de rendimentos (Elliot, 1999; Vartanian, 1999; Bolster *et al.* 2007; Garcia e Nicodemo, 2013), e baixo desempenho escolar (Case e Katz, 1991, Rosenbaum, 1995; Goux e Maurin, 2007, Gibbons *et al.*, 2013a). Essa influência do contexto ao redor do domicílio de um indivíduo (sua vizinhança) sobre diferentes aspectos da sua vida tem sido denominada na literatura por efeito vizinhança (Wilson, 1987).

Devido à relevância do tema, uma vasta literatura internacional sobre efeito vizinhança se desenvolveu nas últimas décadas. Tais estudos procuraram investigar, por um lado, se, de fato, esse efeito é significativo, e, por outro lado, quais os potenciais mecanismos pelos quais a vizinhança influenciaria os resultados individuais. Mayer e Jencks (1989), Ellen e Turner (1997), Durlauf (2004) e, mais recentemente, Sharkey e Faber (2014) realizaram uma revisão detalhada desses estudos.

As pesquisas têm sugerido três canais principais pelos quais a vizinhança afetaria o desempenho individual, quais sejam: a influência dos pares (“modelo epidêmico”), a influência dos adultos (*role model* ou papel social) e a influência das instituições (modelo institucional ou de socialização institucional). O “modelo epidêmico” prevê que uma criança é fortemente influenciada pelo comportamento dos seus pares da vizinhança. O modelo *role model*, por sua vez, parte do pressuposto que os adultos de uma determinada vizinhança serviriam de exemplos para os mais novos. Por fim, o modelo institucional sugere que resultados individuais podem ser afetados pela qualidade do serviço de infraestrutura de uma vizinhança (Jencks e Mayer, 1990; Ainsworth, 2002; Ribeiro e Koslinski, 2009). Em comum, os três mecanismos supõem que as chances de um morador de uma determinada vizinhança ascender socialmente são tão menores quanto pior for o nível socioeconômico da população ali residente (Jencks e Mayer, 1990; Ainsworth, 2002).

Apesar do aumento do número de pesquisas, algumas questões empíricas sobre o tema em tela ainda são controversas. A primeira questão se refere à definição geográfica da vizinhança (Jencks e Mayer, 1990; Gibbons *et al.*, 2013; Sharkey e Faber, 2014). Devido à inexistência de uma barreira natural que demarque territorialmente os vizinhos, não há um consenso sobre qual recorte geográfico que se deve considerar para definir a vizinhança (Gibbons *et al.*, 2013). Note-se que isso é especialmente crítico para os estudos que adotam metodologias que controlam pelo efeito fixo da vizinhança como, por exemplo, em Becker *et al.* (2008) e Helmers e Patnam (2014).

Outra dificuldade que os estudiosos quase sempre se deparam é com a indisponibilidade de bancos de dados que forneça informações suficientes tanto dos indivíduos, quanto do local de residência que permitam a análise da relação causal entre o contexto espacial e os resultados individuais (para mais detalhes, veja-se Durlauf, 2004).

A relação causal dificilmente é garantida em um trabalho empírico nessa temática devido ao viés de seleção espacial. Conforme Gibbons *et al.* (2013), os resultados similares obtidos por crianças que vivem em uma determinada vizinhança podem ser decorrência da semelhança em termos de *background* familiar. As características das crianças são estritamente relacionadas com as dos seus pais, e essas, por sua vez, estão relacionadas com as características dos seus vizinhos através de fatores comuns na escolha residencial. Goux e Maurin (2007), no mesmo sentido, ainda afirmam que as crianças e famílias que vivem na mesma vizinhança tendem a ter resultados semelhantes, no entanto, não sendo claro se isso é

porque eles influenciam uns aos outros ou porque eles compartilham das mesmas características não observáveis.

Ainda que esforços tenham sido feitos para separar a contribuição do contexto da vizinhança da contribuição das características socioeconômicas das famílias para explicar o desempenho individual a partir de diferentes estratégias empíricas, tais como efeito fixo da família com base na comparação de resultados obtidos por irmãs (Plotnick e Hoffman, 1996), *propensity score* (Harding, 2003), variável instrumental (Goux e Maurin, 2007), experimentos (Ludwig *et al.* 2010; Chetty *et al.*, 2015) e econometria espacial (Helmert e Patnam, 2014), essa questão dificilmente é tratada de forma incontestável. Isso porque as características das famílias também estão sujeitas às influências da vizinhança (Sharkey e Faber, 2014). Pode-se argumentar, por exemplo, que a estrutura familiar e o status do emprego dos pais das crianças que influenciam o desempenho escolar também são afetados pelo do local de moradia (Jencks e Mayer, 1990).

Mesmo quando os trabalhos conseguem utilizar dados e metodologias adequadas, como já exposto, surge outra dificuldade que é referente à definição das variáveis de vizinhança que realmente importam. Não existe um consenso sobre quais variáveis da vizinhança devem ser consideradas em um estudo empírico. O trabalho recente de Gibbons *et al.* (2013), por exemplo, conseguiu lidar com problema ocasionado pelo *sorting* das famílias ao empregar uma metodologia denominada como “engenharia reversa”. Intuitivamente, os autores analisaram o efeito da alteração da composição da vizinhança (a partir do fluxo migratório dos estudantes em um dado período de tempo) sobre os alunos que não migraram, não encontrando evidências de que exista uma relação causal entre os pares e o desempenho escolar. Embora os autores tivessem acesso a um rico banco de dados e a uma metodologia robusta, as variáveis adotadas para refletir a qualidade da vizinhança (percentual de alunos homens, portadores de necessidade especiais e beneficiários de programas sociais) podem ser questionáveis. Pode-se argumentar que os resultados obtidos pelos autores poderiam ser decorrentes da escolha das *proxys* para a qualidade da vizinhança que não conseguiram representar de forma satisfatória as características da vizinhança que realmente são importantes para explicar o desempenho escolar.

Diferentemente do cenário internacional, no Brasil poucos estudos analisaram a relação entre o contexto espacial e a trajetória individual, o que indica que as discussões são mais embrionárias nesse país. Especificamente no caso da educação, a escassez de dados que identificam a localização exata dos alunos e das escolas serve como um entrave para as pesquisas. Temos conhecimento de apenas duas bases de dados que ofereceram recentemente informações georreferenciadas dos alunos, a Prova Rio para a cidade do Rio de Janeiro e a Fundaj (2013) para o caso do Recife. Acreditamos que explorar as informações desses bancos seja importante para fomentar e impulsionar a discussão nacional, ainda que não seja possível mitigar todos os problemas mencionados anteriormente.

O presente estudo, então, utilizará dos dados da Fundaj (2013) para analisar empiricamente a relação existente entre as favelas, onde a segregação residencial se expressa da forma mais determinada, e o desempenho escolar das escolas públicas do Recife no ano de 2013. Nessa perspectiva, o presente artigo se insere na literatura que investiga a influência da moradia em uma vizinhança desfavorecida sobre o processo cognitivo dos alunos. Como documentado por Cira (2002), as favelas são fenômenos muito presentes nas cidades brasileiras e latino-americanas e são caracterizadas por reunir uma população pobre e pouco qualificada dentro de um território desorganizado que, em geral, é marcado pela violência associada ao tráfico de drogas.

Conforme Pasternak e D'Ottaviano (2015), apesar de o Brasil testemunhar uma diminuição da pobreza na primeira década dos anos 2000, as condições de moradia nas áreas urbanas são ainda bastante preocupantes. Por exemplo, em 2010, 5,61% da população brasileira viviam em favelas, percentual maior que aquele observado em 2000, quando registrou 3,04%<sup>1</sup>. Em Recife, esse percentual é ainda mais representativo, atingindo o patamar de 22,8%, ficando atrás apenas de Belém, Salvador e São Luis no ranking das capitais brasileiras com maior proporção de pessoas residindo nas áreas urbanas irregulares em 2010 (IBGE, 2011).

---

<sup>1</sup> Esse aumento pode também refletir a melhoria da coleta de dados sobre aglomerados subnormais no último Censo Demográfico (2010).

Para atingir o objetivo, adotamos a técnica de *propensity score matching* (PSM) e análise de sensibilidade proposta por Ichino et al (2008). Adicionalmente, utilizamos dois métodos de ponderação e o coeficiente de Oaxaca-Blinder com o intuito de verificar se o efeito tratamento (isto é, morar na favela) é sensível a diferentes ponderações e ao método utilizado na estimação. Os resultados encontrados apontam para a influência negativa das favelas sobre o desempenho escolar, dado pela nota de matemática. Em média, 50% da diferença das médias incondicionais da nota de matemática entre os dois grupos de alunos pode ser atribuída às favelas. Isto é, a média incondicional dos alunos favelados é 2,3 pontos a menos que a dos demais alunos no teste de matemática<sup>2</sup>, após controlarmos pelas características observáveis, essa diferença diminui para aproximadamente 1,2 pontos. Tais evidências sugerem que os alunos que moram na favela têm um desempenho acadêmico inferior ao obtido pelos estudantes com características similares, mas que moram fora da favela.

Além dessa introdução, esse estudo está organizado em mais cinco seções. A segunda seção tem por objetivo relacionar os fatores que explicariam o efeito adverso das favelas sobre o aprendizado de um aluno. Já a terceira seção realiza uma breve análise das favelas do Recife. A quarta seção, por sua vez, elucida a estratégia empírica e a base de dados adotada. Os resultados encontrados das estimações econométricas são apresentados na quinta seção. Na sexta e última seção, serão oferecidas as considerações finais.

## 2 Favela e desempenho escolar

Os estudos internacionais que se preocupam em analisar a relação entre segregação residencial e os resultados individuais têm documentando os efeitos adversos de se crescer nos guetos, cuja formação é orientada por questões de raça e etnias. Por exemplo, Crane (1991) e Cutler e Glaeser (1997) encontraram evidências de que essas áreas mais segregadas podem acarretar em menores níveis escolaridade, maiores riscos de gerar filhos fora do casamento e piores condições de empregos para a população que vive nas áreas urbanas mais isoladas.

No Brasil, o conceito que mais se aproxima dos guetos é o da favela. Todavia, as favelas se diferenciam dos guetos por ter suas origens guiadas por questões de classes sociais, em vez de raças e etnias (Wacquant, 2004; Costa, 2013). Sob a luz da teoria sobre o efeito vizinhança (Wilson, 1987; Jencks e Mayer, 1990), podemos investigar se os efeitos negativos de se crescer nas áreas urbanas mais segregadas também se repetem no Brasil. Para tanto, levantamos a hipótese da existência do “efeito-favela” que comprometeria a ascensão social dos moradores das áreas urbanas mais precárias. Tal efeito explicaria as diferenças dos resultados obtidos entre moradores e não moradores de favelas que possuiriam características produtivas idênticas.

O “efeito-favela” pode operar através dos três mecanismos potenciais citados na introdução (influência da qualidade das infraestruturas, influência dos adultos, influência dos colegas) e dificilmente são identificados isoladamente em uma análise empírica. As consequências dessa segregação urbana sobre o indivíduo podem se manifestar ainda na fase escolar, quando crianças e adolescentes começam a ser capacitados para ingressar no mercado de trabalho na idade adulta (Pero *et al.*, 2005). Como resumido por Torres, Ferreira e Gomes (2004), a segregação espacial pode ter reflexos negativos sobre os resultados escolares das crianças e adolescentes que residem nas favelas, sendo essas penalizadas pelo seu baixo nível socioeconômico, por não conviver com colegas de nível mais elevado e pela interação entre seu baixo nível socioeconômico e do meio que ele vive.

Pela própria definição, as favelas possuem infraestrutura urbana inadequada, com vielas estreitas e irregulares, que geram dificuldades legais, ambientais, de engenharia e de alocação de profissionais, que acabam por influenciar diretamente a oferta de serviços públicos básicos, como educação, saúde e segurança pública para a população que ali reside (Rodrigues, 2005).

No caso específico da educação, isso é particularmente verdade na medida em que a qualidade da escola está estritamente relacionada com a qualidade da vizinhança (Jencks e Mayer, 1990; Ainsworth, 2002; Ribeiro e Koslinski, 2009), de tal forma que os problemas como a carências de profissionais capacitados da educação, superlotação das salas de aula, infraestrutura física precária das escolas se

---

<sup>2</sup> A escala de nota varia entre 0 a 100 pontos.

tornam especialmente mais graves nas favelas do que nas demais áreas urbanas (Paiva, 2009).

Adicionalmente, a pouca cultura escolar dos pais pode potencializar os problemas vivenciados nessas escolas por esses estarem alheios ao cotidiano escolar dos filhos. Trata-se, então, da primeira evidência de que os adultos também poderiam influenciar negativamente o desempenho escolar dos mais novos por não exercerem o papel de monitoramento e supervisão, como sugerido pelo modelo do papel social (Ainsworth, 2002). O modelo do papel social ainda prevê que as crianças aprendem sobre quais comportamentos são considerados adequados por meio da interação com os adultos de sua vizinhança, de tal modo que os resultados obtidos pelos adultos em suas vidas profissionais serviriam de motivação para os mais novos (Jencks e Mayer, 1990; Ainsworth, 2002; Ribeiro e Koslinski, 2009).

Essa questão fica evidente no âmbito das favelas em um estudo realizado por Paiva e Burgos (2009), que subsidiados por entrevistas realizadas com professores e diretores de escolas na favela do Rio de Janeiro, verificaram que existe uma descrença por parte dos alunos favelados em relação à promessa de um futuro promissor através da escola, em que o sistema educacional não seria capaz de assegurar um lugar no mercado de trabalho e, por consequência, a ascensão social, o que acaba tornando a escola desinteressante aos alunos.

Ainda sobre o mercado de trabalho, uma possível discriminação sofrida pelos trabalhadores moradores das favelas pode ainda ser um fator que desmotive as crianças a se dedicarem aos estudos. Nesse contexto, os residentes na favela teriam maior dificuldade de obter um emprego formal e/ou receberiam salários menores quando comparados com outros trabalhadores com habilidades similares que moram nas áreas urbanas regulares (Pero *et al.*, 2005; Rocha *et al.*, 2011; Westphal, 2014).

Outro canal pelo qual a favela influenciaria o desempenho escolar seria através dos pares, quando o estudante tende a reproduzir os comportamentos dos seus colegas. Nesse aspecto, o desempenho do aluno seria influenciado pelas condições socioeconômicas desfavoráveis dos seus vizinhos de idade similar. De fato, os estudos empíricos que analisaram o caso das favelas do Rio de Janeiro encontraram indícios de um significativo *peer effect* sobre a decisão individual de frequentar a escola (Vasconcellos e Rocha, 2006) e sobre a distorção idade-série (Alves *et al.*, 2008).

Esse comportamento pode ser reforçado pelo ambiente escolar devido à homogeneidade socioeconômica das escolas localizadas nas áreas mais carentes que dificultaria a interação dos mais pobres com colegas que possuem características socioeconômicas diferentes das suas (Jencks e Mayer, 1990; Ribeiro e Koslinski, 2009; Soares *et al.* 2008). Dificilmente, os alunos que moram fora da favela frequentariam uma unidade escolar dentro da favela.

A conjuntura das favelas pode também facilitar a presença de grupos armados organizados, sobretudo, derivados do tráfico de drogas. Como documentado por Rocha e Monteiro (2013) e Ribeiro (2013a, 2013b), as atividades ligadas ao tráfico de drogas e o bom desempenho escolar caminham em direções opostas. Isso porque, diante de um cenário de pobreza, presença precária do Estado, ausência de exemplos de sucesso profissional via escolaridade, as crianças podem sentir atraídas pelas atividades do tráfico e passar a ter comportamentos indesejados perante a escola que vão desde eventos que envolvem indisciplina e o não reconhecimento da autoridade e das hierarquias escolares até mesmo a evasão escolar (Rodrigues, 2005; Ribeiro, 2013a).

Além disso, junto com o tráfico vem a banalização da violência urbana. Nas favelas não são raros os casos em que os moradores testemunham conflitos envolvendo facções criminosas e polícia. Esses conflitos podem impactar no aprendizado dos alunos na medida em que se cria um clima de insegurança na população, força o fechamento das escolas em dias letivos e influencia a contratação e a rotatividade dos professores e diretores das escolas (Rocha e Monteiro, 2013; Ribeiro 2013a, 2013b).

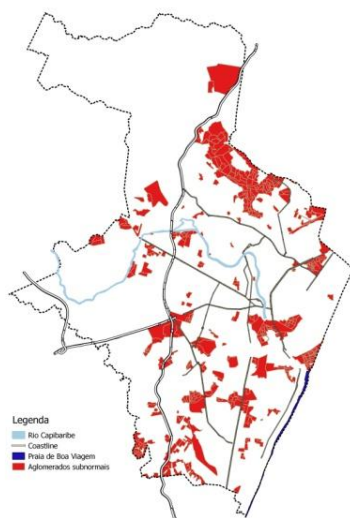
Ressalte-se, contudo, que, além dos poucos trabalhos existentes citados acima, que tem como foco principalmente para cidade do Rio de Janeiro, de acordo com nosso melhor conhecimento, ainda não foram realizados estudos que procuram mensurar o impacto da moradia em favela sobre os resultados sociais de seus residentes no caso da Cidade do Recife. Ou seja, pouco ou nada é conhecido, particularmente, sobre a potencial influência que a moradia localizada na favela tem sobre o desempenho escolar das crianças ou adolescentes na referida cidade. A presente investigação pretende iniciar o preenchimento desta lacuna.

### 3 As favelas do Recife

No presente estudo, são considerados como favelas os aglomerados subnormais definidos pelo IBGE (2010), o que corresponde as áreas constituídas de, no mínimo, 51 unidades habitacionais carentes, em sua maioria, de serviços públicos essenciais, ocupando ou tendo ocupado, até período recente, terreno de propriedade alheia (pública ou particular) e estando dispostas, em geral, de forma desordenada e/ou densa. O mapa a seguir mostra a localização das favelas do Recife em 2010, quando essas abrigavam 22,8% da população recifense.

Como se pode observar na figura 1, as favelas estão espalhadas por todo o Recife, sendo muito difícil encontrar raios de um quilômetro que não contenha nenhuma porção de favela em seu interior (Souza, 2003). Assim, podemos encontrar exemplos de favelas tanto nas áreas localizadas em morros, quanto nas áreas planas. Essa configuração espacial é resultante do processo de urbanização associado às características geográficas tão peculiares da cidade<sup>3</sup>.

Inicialmente, o desenvolvimento urbano foi orientado pela economia açucareira, em que as famílias mais abastadas, donas de engenhos, ocuparam as áreas mais planas e pouco alagáveis às margens do Rio Capibaribe, restando aos mais pobres abrigarem-se em construções de palafitas nas várzeas inundáveis, dando origens aos mocambos, que mais tarde passaria a ser reconhecidos como favelas (Neta, 2005; Cavalcanti *et al.*, 2010). Ao longo do desenvolvimento da cidade, a atividade açucareira entrou em declínio, aumentou o fluxo de imigrantes dos meios rurais e escravos recém-libertos, quase sempre trabalhadores braçais e pouco qualificados. Diante de um mercado imobiliário formal excludente, esse fato culminou no aumento da demanda por habitações mais precárias. Os mocambos, então, passaram a abrigar uma parcela significativa da população recifense (Souza, 2003).



**Figura1: Aglomerados Subnormais de Recife em 2010**  
**Fonte: Elaboração própria com base nos dados do IBGE (2010)**

Entre as décadas de 1930 e 1970, sob o argumento das péssimas condições higiene e insalubridade, várias intervenções políticas foram feitas com o objetivo remover os mocambos e deslocar seus habitantes para as áreas mais periféricas e próximas às encostas dos morros, dando início à ocupação das áreas de relevo mais elevado. Todavia, essas ações não foram suficientes para extinguir as favelas das áreas centrais. Por essa razão, as políticas de demolição das habitações precárias foram gradualmente substituídas pela discussão da necessidade de urbanização das favelas (Cavalcanti *et al.*, 2010).

É nesse contexto que se introduz, em 1983, o conceito das Zonas Especiais de Interesse Social (ZEIS) na Lei de Uso e Ocupação do Solo do Recife com o intuito de incluir as áreas populares no planejamento urbano (Souza, 2003; Neta, 2005; Cavalcanti *et al.*, 2010). Deve-se destacar, entretanto, que as ZEIS não foram capazes em promover mudanças estruturais no que se refere às condições das habitações mais precárias (Cavalcanti *et al.*, 2010). Após 30 anos implementação das ZEIS, Recife ainda conta com áreas densas que não possuem acesso às infraestruturas urbanas.

<sup>3</sup> Apelidado como “Veneza brasileira”, Recife é formado por ilhas, rios e mangues. Para mais detalhes, veja –se Neta (2005).

**Tabela 1: Características dos setores censitários do Recife (2010)**

	Favela	Não favela	Diferença
Características dos domicílios			
% com renda domiciliar per capita até 1/2 salário mínimo	0,569	0,342	0,227***
% com banheiro de uso exclusivo	0,287	0,479	-0,192***
% energia elétrica e medidor de uso exclusivo	0,758	0,895	-0,137***
% mulher responsável	0,479	0,454	0,025***
Características das pessoas			
% de homens	0,477	0,463	0,0139***
% de alfabetizados de 7 a 14 anos	0,89	0,943	-0,053***
% de alfabetizados com 25 anos ou mais	0,842	0,915	-0,073***
% com idade entre 5 e 14 anos	0,185	0,143	0,0416***
% de brancos	0,296	0,425	-0,129***

Fonte: Elaboração própria com base nos dados do Censo/IBGE(2010)

Para ilustrar a precariedade das condições urbanas das áreas mais segregadas, a Tabela 1 trás o perfil dos setores censitários do Recife no ano de 2010, diferenciando-os por favela ou não favela. Note-se que, em média, apenas 28% (76%) dos domicílios localizados na favela possuem acesso a banheiros de uso exclusivo com rede geral de esgoto (energia com medidor de uso exclusivo), percentual muito inferior que aquele observado para domicílios nas áreas regulares. Ademais, os dados confirmam que a população residente na favela é mais pobre e menos escolarizada.

## 4 Metodologia

### 4.1 Estratégia empírica

O presente estudo pretende analisar a diferença das notas entre dois grupos de alunos, os que moram na favela (A), e os que não moram na favela (B), a fim de obter evidências que indiquem que a favela é capaz exercer influência sobre o aprendizado do aluno. Todavia, essa não é uma tarefa simples de ser executada devido à distribuição não aleatória dos alunos ao longo do tecido urbano. Como discutido na seção 2, as pessoas com as condições socioeconômicas mais desfavoráveis tendem a residir nas áreas urbanas mais segregadas. Na presença de um potencial viés de seleção espacial, o “efeito favela” não pode ser calculado simplesmente a partir da comparação das médias (valores esperados) condicionadas ao fato do aluno pertencer ou não à favela, caso contrário, atribuiríamos ao local de moradia as diferenças causadas pela diferença no perfil dos alunos.

Formalmente, seja  $Y_{Ai}$  e  $Y_{Bi}$  os resultados potenciais do estudante  $i$ , caso ele pertencesse ao grupo A e B, respectivamente. O efeito causal de morar na favela sobre o desempenho escolar é definido como a diferença entre esses dois resultados potenciais,  $Y_{Ai} - Y_{Bi}$ . O desafio empírico deriva do fato que o aluno  $i$  é observado somente em um dos dois grupos, o que significa dizer que não somos capazes de observar a nota de um determinado aluno que mora na favela, caso ele morasse nas áreas urbanas regulares. Nesse caso, a nota observada do aluno pode ser representada como  $Y_i = Y_{Bi} + (Y_{Ai} - Y_{Bi})T_i$ , em que  $T_i$  é igual a 1 se o aluno for tratado, isto é, se o aluno morar na favela e 0, caso contrário. A diferença da esperança da nota observada condicional ao fato do aluno ser tratado e controle é dada por:

$$E[Y_i|T_i = 1] - E[Y_i|T_i = 0] = (E[Y_{Ai}|T_i = 1] - E[Y_{Bi}|T_i = 1]) - (E[Y_{Bi}|T_i = 1] - E[Y_{Bi}|T_i = 0]) \quad (1)$$

em que o termo  $(E[Y_{Ai}|T_i = 1] - E[Y_{Bi}|T_i = 1])$  é o efeito médio do tratamento sobre os tratados – ATT, doravante identificado por  $\tau_{ATT}$ , ao passo que o termo  $(E[Y_{Bi}|T_i = 1] - E[Y_{Bi}|T_i = 0])$  representa o viés de seleção espacial que explicamos anteriormente.

Nosso objetivo é obter uma estimativa robusta do ATT. Como  $E[Y_{Bi}|T_i = 1]$  é desconhecido, não podemos calcular  $\tau_{ATT}$  sem que sejam feitas hipóteses adicionais. Por essa razão, recorreremos às técnicas quasi-experimentais de avaliação de impacto em que são imperativas as hipóteses de suporte comum e ignorabilidade, que explicaremos a seguir, e controlaremos o viés de seleção a partir das características observáveis dos alunos.

Uma das estratégias mais conhecidas na literatura para esse fim é o método de *propensity score matching*, proposto por Rosenbaum e Rubin (1983). Intuitivamente, esse método propõe que o viés seria pelo menos reduzido se a comparação for baseada entre os alunos tratados e os de controle que sejam o mais similar possível. Uma forma de resumir as características semelhantes dos dois grupos de alunos (favelados e não favelados) e, ao mesmo tempo, criar critério de comparação é considerar apenas os alunos dos dois grupos que tenham probabilidades semelhantes de morarem na favela (Angrist e Pischke, 2009). Para tanto, precisamos calcular a probabilidade de um estudante morar na favela a partir de suas características observáveis, isto é,  $P(x) = \text{Prob}[T_i = 1|X_i]$ . Após obter  $P(x)$ , o ATT será calculado segundo a equação (2):

$$\tau_{ATT} = E[Y_{Ai} - Y_{Bi}|T_i = 1] = E[E(Y_{Ai}|P(x), T_i = 1) - E(Y_{Bi}|P(x), T_i = 0)|T_i = 1] \quad (2)$$

Fundamentalmente, duas hipóteses são exigidas para que a diferença acima forneça uma estimativa sem viés do efeito da condição analisada. A hipótese do suporte comum assegura que para cada aluno tratado (morar na favela) exista outro estudante não tratado (que não moram nas favelas) que apresente valores similares das covariadas, isto é,  $0 < P(x) = \text{Prob}[T_i = 1|X_i] < 1$ . No nosso estudo, isso implica a necessidade de haver estudantes que não moram nas favelas com características muito semelhantes aos que moram.

Já a hipótese da ignorabilidade assegura que não há viés sistemático quando comparamos indivíduos semelhantes quanto a determinadas características observáveis (Rubin, 1974). Isto é, condicional às características observáveis, a alocação dos grupos de tratamento e controle é feita de forma aleatória, implicando na independência entre o tratamento e os resultados potenciais. Como mostraram Rosenbaum e Rubin (1985), formalmente, essa condição pode ser representada a partir das probabilidades como  $Y_{Ai}, Y_{Bi} \perp T_i | P(x)$ . No nosso caso, isso significa dizer que a distribuição das variáveis não observáveis que afetam o desempenho escolar entre os alunos que moram e não moram na favela é a mesma, se condicionarmos a um vetor de variáveis observáveis.

Sob as hipóteses de ignorabilidade e suporte comum, assumimos que, condicionadas às probabilidades de participação na condição analisada (no nosso caso, morar na favela) obtidas a partir de variáveis observáveis, são também eliminadas as influências de variáveis não observadas sobre os resultados de interesse (no nosso caso, a nota do estudante), o que permite utilizar a nota dos controles pareados como um contrafactual para a estimação do  $\tau_{ATT}$ .

Uma clara dificuldade na operacionalização do método é encontrar indivíduos com a mesma probabilidade de pertencer à favela. Assim, utilizaremos três algoritmos para realizamos o *matching* entre os alunos tratados (moradores de favelas) e não tratados (não moradores de favela) comumente utilizados na literatura empírica: 1 vizinho mais próximo, 5 vizinhos mais próximos, e *kernel matching*.

Conforme Caliendo e Kopeing (2008), a hipótese de ignorabilidade requer que todas as variáveis que afetam o tratamento e/ou o resultado sejam especificadas no modelo. Empiricamente, essa exigência é difícil de ser garantida, mesmo com o rico conjunto de informação como aquele disponibilizado pela Fundaj (2013), devido à possibilidade de omissão de variáveis não observáveis (*confounders*) que afetam tanto a decisão de morar na favela quanto o desempenho de matemática. Na presença dos *confounders*, a hipótese de ignorabilidade não seria satisfeita e, portanto, o  $\tau_{ATT}$  seria viesado.

Para analisar a potencial influência deste problema sobre os resultados, utilizamos a sugestão de análise de sensibilidade sugerida por Ichino *et al.* (2008). A análise de sensibilidade desenvolvida por estes autores busca verificar a influência de potencial variável omitida sobre o ATT estimado caso a hipótese da ignorabilidade não seja satisfeita, o que seria equivalente a dizer que a alocação ao tratamento não é aleatória dado o conjunto de variáveis observáveis, isto é:

$$Pr(T_i = 1|Y_{Ai}, Y_{Bi}, X) \neq Pr(T_i = 1|X).$$

A suposição central do teste proposto por Ichino *et al.* (2008) é que a hipótese de ignorabilidade seria garantida dado um conjunto de variáveis observáveis,  $X$ , e uma variável binária (hipótese simplificadora) não observada,  $U$ . Nesse caso, se  $U$  fosse conhecido, a hipótese de independência poderia ser reescrita como:

$$Pr(T_i = 1|Y_{Ai}, Y_{Bi}, X, U) = Pr(T_i = 1|X, U)$$



Na impossibilidade de se conhecer a distribuição deste fator não observável, Ichino et al (2008) propuseram uma caracterização dessa distribuição baseado em quatro parâmetros definidos a partir das combinações do status de tratamento e dos valores dos resultados, assumindo a variável  $U$  valores de acordo com a distribuição de uma variável observada específica (categórica). Para tanto, é construída uma nova variável binária que possui valor 1, quando o desempenho escolar for maior que a média, e 0, caso contrário e a simulação dos valores de  $U$  é feita então a partir dos quatros parâmetros dados por:

$$p_{ij} \equiv Pr(U = 1 | T = i, I(Y > \bar{y}) = j) \quad (3)$$

em que  $i, j \in \{0,1\}$ ,  $I$  é uma função indicadora e  $\bar{y}$  é o valor médio da variável de resposta, no caso, média obtida da nota de matemática. A partir da equação (3), é obtida a probabilidade de ocorrência de  $U=1$  em cada um dos quatro grupos definidos pelos status de tratamento e pelos valores dos resultados. Então, dados os valores obtidos dos parâmetros  $p_{ij}$ , um valor  $U$  é atribuído para cada aluno da amostra. Desse modo, a variável  $U$  pode ser tratada como se fosse uma variável observada a ser incorporada ao conjunto de variáveis utilizadas para estimar o *score propensity*, e, conseqüentemente, calcular o ATT. Cabe destacar a imputação dos valores para  $U$  e que, dado o conjunto de valores dos parâmetros de sensibilidade, a estimação do ATT são repetidas  $n$  vezes (no caso desse estudo, 200 vezes) para se obter uma estimativa do ATT, que é uma média dos ATTs sobre a distribuição de  $U$ .

Adicionalmente, no sentido de caracterizar a potencial falha da CIA associada ao *cofounder* simulado  $U$ , Ichino et al (2008) ainda propuseram uma simulação da influência da variável  $U$  sobre a probabilidade relativa (*odds ratio*) de se obter um efeito positivo sobre a variável de interesse nos casos em que não for alocado ao tratamento, *outcome-effect*. Tal influência é obtida a partir da estimação de um modelo *logit* para  $Pr(Y|T_i = 0, X, U)$  em cada interação. De modo análogo, o impacto de  $U$  sobre a probabilidade relativa de ser alocada ao tratamento (morar na favela), *treatment-effect*, é obtida a partir da estimação da probabilidade ser escolhido para o tratamento no modelo  $Pr(T_i = 1 | X, U)$ .

Além deste teste de sensibilidade, também se verifica no trabalho se o efeito estimado de se morar na favela sobre o desempenho escolar é sensível a diferentes métodos de estimação e ponderação. Especificamente, duas técnicas de reponderação baseadas no escore de propensão também são empregadas no nosso exercício econométrico, quais sejam: a ponderação pelo inverso do *propensity score* - IPW (Hirano et al, 2003) e da regressão ajustada ponderada pelo inverso do *propensity score* (IPWRA). A motivação para se utilizar o inverso do *propensity score* como peso numa regressão é remover a influência associado ao fato de que o aluno só é observado em apenas uma das situações, morando ou não na favela (tratado ou não tratado), ou seja, a estratégia tenta atenuar o fato de que há informações insuficientes. Em adição, como documentado por Imbens e Wooldridge (2009), a combinação da ponderação com a regressão procura contornar o problema da má especificação, seja ele derivado do modelo de regressão ou da equação de *propensity score*.

O IPW é um estimador eficiente e pondera os estudantes tratados pelo fator  $1/P(x)$ , onde, como definimos anteriormente,  $P(x)$  representa a probabilidade de pertencer à favela dada as características observáveis (Hirano et al, 2003). Então, com o fito de eliminar ou atenuar o fato de as observações (estudantes) não são observadas nos dois estados, pesos maiores são atribuídos aos alunos tratados que possuem as menores probabilidades de pertencer à favela. Em sentido oposto, os alunos do grupo de controle são ponderados pelo fator  $1/(1 - P(x))$  de tal modo que pesos maiores aos estudantes com maior probabilidade de morar na favela. Deve-se ressaltar, contudo, que essa técnica estima a probabilidade do tratamento sem qualquer suposição sobre a forma funcional de determinação da variável de interesse (impacto), ou seja, os estimadores IPW se limitam a modelar o tratamento para explicar a atribuição não aleatória ao tratamento.

Por outro lado, os estimadores IPWRA consideram dois modelos, um para variável de interesse e outro para a probabilidade de tratamento, utilizando os pesos correspondentes ao inverso das probabilidades de tratamento na regressão ajustada da variável de interesse. Uma vantagem importante diz respeito à consistência, garantida mesmo quando apenas um deles deve estar correto (ou seja, uma estratégia duplamente robusta) (Imbens e Wooldridge, 2009)

Por fim, o método de decomposição de Oaxaca-Blinder também é utilizado para separar as contribuições das diferenças nas características das contribuições das diferenças nos retornos dessas características. No caso desse estudo, pretende-se investigar o diferencial de nota entre os alunos que

moram e que não moram na favela. Fortin *et al* (2010) descrevem o método de decomposição em detalhes. De forma sucinta, a decomposição de Oaxaca-Blinder pode ser escrita como segue:

$$\Delta_0^\mu = \underbrace{(E[X_A|T_i = 1] - E[X_B|T_i = 0])\beta_B}_{\text{Explicado}} + \underbrace{E[X_A|T_i = 1](\beta_{Ai} - \beta_B)}_{\text{não explicado}} \quad (4)$$

$$\Delta_0^\mu = \Delta_X^\mu + \Delta_S^\mu$$

em que  $\Delta_0^\mu$  é a diferença bruta das notas entre os dois grupos analisados,  $\Delta_X^\mu = (E[X_A|T_i = 1] - E[X_B|T_i = 0])\beta_B$  é também identificado como “efeito explicado” ou “efeito composição” e representa a parte do diferencial das notas que é decorrente das diferenças nas covariadas entre os dois grupos. O termo  $\Delta_S^\mu = E[X_A|T_i = 1](\beta_{Ai} - \beta_B)$  é identificado como “efeito não explicado” ou “efeito estrutura” e capta os retornos diferenciados pelas mesmas características obtidos pelos dois grupos. No caso desse estudo, esse termo capta o quanto do diferencial de nota é explicado pelo fato do estudante morar na favela e não nas áreas urbanas regulares.

O último termo é frequentemente denominado de “efeito discriminação” nos estudos que têm como objetivo analisar o diferencial de rendimentos no mercado de trabalho. Contudo, as pesquisas mais recentes como o Fortin *et al* (2011), Kline (2011) e Słoczyński (2015) observaram que esse termo pode ser analisado como sendo o efeito médio o tratamento sobre os tratados (ATT). Kline (2011) mostrou que este particular efeito da decomposição de Oaxaca-Blinder padrão é equivalente a um estimador reponderado baseado em modelos lineares para a probabilidade de ser tratado. Como mostra o próprio autor, tal estimador também teria propriedade de ser duplamente robusto (Robins *et al.*, 1994), isto é, para ser consistente, bastaria que apenas um modelo (modelo básico de regressão ou o modelo para a probabilidade do tratamento) seja bem especificado. Conforme Kline (2011) e Słoczyński (2015), a equação (4) pode ser reescrita da seguinte forma:

$$\begin{aligned} \Delta_0^\mu &= (E[Y_A|T_i = 1] - E[Y_B|T_i = 0]) + E[Y_{Ai} - Y_{Bi}|T_i = 1] \\ &= (E[Y_A|T_i = 1] - E[Y_B|T_i = 0]) + \tau_{ATT} \end{aligned} \quad (5)$$

Dessa maneira, o componente não explicado seria equivalente ao ATT.

## 4.2 Bases de dados e descrição das variáveis

Os dados utilizados neste estudo são oriundos da pesquisa realizada pela Fundaj em 2013, em que entrevistou diretores, professores, responsáveis e alunos do 6º ano das escolas públicas de Recife-PE. De maneira complementar, foram utilizados os dados do INEP (2014) e do Censo escolar/INEP (2013) para reunir informações das características das escolas analisadas. Por fim, utilizamos os *shapefiles* disponibilizados pelo IBGE (2010) para identificar os territórios das favelas. De forma semelhante ao que foi feito por Ribeiro (2013b), os alunos que moram nas favelas foram identificados ao sobrepor as informações georreferenciadas dos alunos com o *shapefile* das favelas fornecido pelo IBGE (2010).

A amostra final de alunos, após retirar os estudantes que tinham *missing* em pelo menos uma variável de controle e de interesse, é composta por 2.570 alunos de 117<sup>4</sup> escolas da rede pública de ensino, distribuídos em 142 turmas diferentes do sexto ano. Em relação a variável de interesse, adotamos como variável dependente a nota da segunda prova de matemática como uma *proxy* para o desempenho escolar dos alunos.

A base de informações da Fundaj se destaca por ser a primeira base de dados a oferecer informações georreferenciadas dos alunos do Recife, possibilitando que os estudantes possam ser distinguidos entre morador ou não morador da favela. Outra vantagem dessa base de dados para nosso estudo é que ela é composta majoritariamente por aluno com 11 anos de idade, isto é, formada por alunos muito novos. Tal fato é importante porque o contexto social das crianças mais novas é formado principalmente pelo ambiente familiar e pela vizinhança mais próxima, diferentemente dos alunos mais velhos, que pode ter seu meio social ampliado (Ribeiro e Koslinski, 2009).

Além disso, o banco de dados oferece informações detalhadas sobre as características

<sup>4</sup> Nesse trabalho não consideremos as unidades escolares federais, uma vez que essas escolas possuem processo seletivo para os ingressos de novos alunos. Além disso, não foi inserida uma escola estadual na análise por não ter dados disponíveis do Censo Escolar e do INEP relativos ao ano de 2013, já que essa teve seu funcionamento interrompido no ano de 2014.

socioeconômicas dos alunos e das infraestruturas das escolas que são amplamente citadas na literatura da Economia da Educação como fatores importantes para explicar o desempenho escolar (Menezes-Filho, 2007; Raposo et al, 2015). Na Tabela 2 estão expostas as variáveis utilizadas nesse estudo.

**Tabela 2: Descrição das variáveis**

Variáveis	Descrição das variáveis	Fonte
Variável dependente		
Nota 2	Nota de Matemática no final do ano	Fundaj (2013)
Características dos alunos		
Idade	idade em anos.	Fundaj (2013)
Sexo masculino	<i>dummy</i> igual a 1 se o aluno é do sexo masculino.	Fundaj (2013)
Cor branca	<i>dummy</i> igual a 1 se o aluno se auto declara branco.	Fundaj (2013)
Repetente	<i>dummy</i> igual a 1 se o aluno já foi reprovado pelo menos uma vez.	Fundaj (2013)
Creche	<i>dummy</i> igual a 1 se o aluno frequentou creche.	Fundaj (2013)
Gosta de ir para a escola	<i>dummy</i> igual a 1 se o aluno gosta de ir para a escola	Fundaj (2013)
Características dos responsáveis e do domicílio		
Escolaridade do responsável	anos de estudos do responsável.	Fundaj (2013)
Estado civil do responsável	<i>dummy</i> igual a 1 se o responsável é casado ou possui união estável.	Fundaj (2013)
O responsável possui emprego informal	<i>dummy</i> igual a 1 se o responsável possui emprego informal	Fundaj (2013)
Bolsa família	<i>dummy</i> igual a 1 se é beneficiário do Programa Bolsa Família.	Fundaj (2013)
Computador com internet	<i>dummy</i> igual a 1 se na residência do aluno tem computador com internet.	Fundaj (2013)
Mãe e pai moram no domicílio	<i>dummy</i> igual a 1 se a mãe e o pai moram no domicílio.	Fundaj (2013)
Número de pessoas	Número de pessoas que moram no domicílio.	Fundaj (2013)
Qualidade	<i>dummy</i> igual a 1 se adotou o critério "qualidade".	Fundaj (2013)
Distância da casa do aluno até o centro	distância entre a residência do aluno e o centro	Fundaj (2013)
Condição financeira	<i>dummy</i> igual a 1 se adotou o critério "condição financeira".	Fundaj (2013)
Violência	<i>dummy</i> igual a 1 se o responsável considera a violência como sendo um problema grave no bairro.	Fundaj (2013)
Características das escolas		
Escola estadual	<i>dummy</i> igual a 1 se a escola pertence a rede estadual de ensino.	Censo Escolar (2013)
Biblioteca	<i>dummy</i> igual a 1 se a escola possui biblioteca em quantidade suficiente e condição adequada.	Fundaj (2013)
Quadra de esporte	<i>dummy</i> igual a 1 se a escola possui quadra de esporte em quantidade suficiente e condição adequada.	Fundaj (2013)
Laboratório	<i>dummy</i> igual a 1 se a escola possui laboratório em quantidade suficiente e condição adequada.	Fundaj (2013)
Acesso de internet aos professores	<i>dummy</i> igual a 1 há acesso à internet para o uso dos professores em quantidade suficiente e condição adequada.	Fundaj (2013)
Carência de pessoal de apoio pedagógico	<i>dummy</i> igual a 1 se há carência de pessoal de apoio pedagógico (coordenador, supervisor e orientador educacional).	Fundaj (2013)
Complexidade da gestão escolar	Assume valor igual a 2, se o indicador de complexidade é 2; igual a 3, se o indicador de complexidade é 3; igual a 4, se o indicador de complexidade é 4; igual a 5, se o indicador de complexidade é 5; igual a 6, se o indicador de complexidade é 6.	INEP (2014)
Indicador de esforço do docente*	Proporção de professores de elevado esforço.	INEP (2014)
Turno manhã	<i>dummy</i> igual a 1 se as aulas acontecem no período da manhã.	Fundaj (2013)

Fonte: Elaboração própria

Nota: \* São considerados professores de alto esforço, os docentes que possuem mais de 400 alunos, trabalham nos três turnos e em mais de uma escola.

Para controlar para as influências das características individuais, inserimos no modelo as informações sobre o sexo, cor, idade, informações sobre se aluno gosta de ir pra a escola, se frequentou a pré-escola e um indicador de repetência dos estudantes. As variáveis utilizadas para captar as influências do *background* familiar são o estado civil, escolaridade e ocupação dos responsáveis pelos alunos, uma *dummy* que identifica se o pai e a mãe moram no domicílio, uma *dummy* que identifica se família é beneficiária do Programa Bolsa Família, uma *dummy* que identifica se possui computador com internet e o número de pessoas residentes no domicílio.

Além destas variáveis tradicionais, a riqueza do banco de dados da Fundaj permitiu a inclusão de variáveis associadas a critérios de escolhas da escola e do local de residência, além de informação sobre a importância da presença de violência no bairro: foram incluídas na análise *dummies* para os motivos de escolha da escola (pela qualidade ou não), para os motivos de escolha da moradia (pelo critério de condição financeira ou não) e para o fato da violência ser um problema grave no bairro ou não. Note-se que tais variáveis, raramente disponíveis nos bancos de dados regulares, representam controles potencialmente importantes para possível presença de *sorting* espacial pelas famílias. Pode-se imaginar, por exemplo, que entre os responsáveis com níveis socioeconômicos semelhantes, é provável que aqueles

mais preocupados com a educação de seus filhos desenvolvam estratégias para garantir uma vaga nas escolas públicas mais qualificadas (Retamoso e Kaztman, 2008; Soares et al, 2008).

Também foi incluída a distância da residência até o centro para tentar captar a heterogeneidade existente ao longo da cidade, tendo em vista que as áreas mais próximas ao centro são mais abastadas do que áreas mais distantes (Alves et al, 2008).

As influências das características das escolas foram controladas pelas *dummies* que identificam se a escola é da rede estadual de ensino, se as escolas possuem biblioteca, laboratório, internet para o professor em quantidades suficientes e condições adequadas, se a escola possui carência de profissionais de apoio pedagógico.

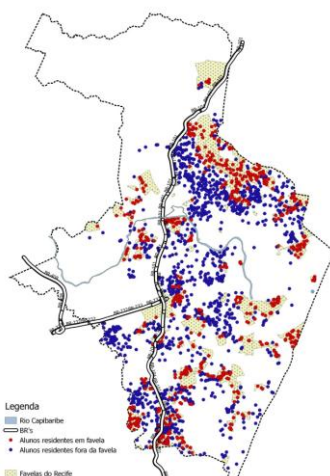
Além destas, incorporamos no modelo o índice de Complexidade de Gestão escolar a partir do indicador calculado pelo INEP (2014) que resume, em uma única medida, as informações de porte, turnos de funcionamento, nível de complexidade das etapas e quantidade de etapas ofertadas. O INEP (2014) definiu seis níveis de complexidade, em que os níveis mais elevados indicam maior complexidade. Nenhuma escola recifense analisada possui nível de complexidade 1.

Por fim, adicionamos ao modelo o indicador do esforço docente que reúne os aspectos do trabalho do professor que contribuem para a sobrecarga no exercício da profissão. É mensurado a partir do percentual de docentes por escola cujo esforço para o exercício da profissão é considerado elevado. Consideram-se docentes com esforço elevado nos anos finais aqueles que atendem mais de 400 alunos, atuam em turmas que funcionam nos três turnos, em duas ou mais etapas e em duas ou mais escolas.

## 5 Resultados

### 5.1 *P propensity score matching*

A pesquisa realizada em 2013, ao captar os endereços dos alunos, ofereceu a oportunidade inédita de comparar o desempenho escolar entre os alunos que moram e que não moram na favela do Recife. De forma similar ao que foi feito por Ribeiro (2013b), os alunos que moram nas favelas foram identificados a partir do georreferenciamento e dos *shapefiles* disponibilizados pelo IBGE (2010), conforme ilustrado na figura 2. Por meio desse procedimento, foi possível constatar que os alunos residentes nas favelas respondem por 35% da amostra.



**Figura2: Distribuição espacial dos alunos.**

**Fonte: Elaboração própria com base nos dados da Fundaj(2013), Censo/IBGE(2010) e Open Street Maps (2016).**

Os alunos que moram na favela, além de apresentarem as características socioeconômicas mais desfavoráveis, também estão expostos às escolas públicas que possuem infraestruturas mais precárias. Por exemplo, os estudantes que moram nas áreas urbanas irregulares são menos propensos a ter frequentado as creches (68% contra 72%), possuem responsáveis menos escolarizados. Em relação à renda, os dados indicam 62% dos estudantes que moram na favela é beneficiário do Programa Bolsa Família, percentual maior que o observado para o grupo de não favelados (56%). Os alunos favelados estudam em uma unidade escolar com alta complexidade da gestão escolar, com maior dificuldade de contratar professores

e profissionais de apoio pedagógico e conta com os piores serviços de bibliotecas, laboratórios e internet para professores. Esse resultado está de acordo com o que foi exposto na seção 2.

**Tabela 3: Características dos alunos**

	Antes do Matching			Depois do Matching		
	Favela	Não Favela	Diferença	Favela	Não Favela	Diferença
Nota de matemática no final do ano	37,436	39,783	-2,347***	37,436	38,667	-1,222*
Características dos indivíduos						
Homem	0,482	0,494	-0,012	0,482	0,484	-0,002
Branco	0,179	0,198	-0,019	0,179	0,178	0,002
Idade do aluno	11,388	11,297	0,091**	11,388	11,377	0,011
Frequentou creche	0,676	0,717	-0,041**	0,676	0,676	0,000
Já foi reprovado pelo menos uma vez	0,274	0,256	0,018	0,274	0,273	0,001
Gosta da escola	0,318	0,369	-0,050***	0,318	0,319	-0,001
Características dos responsáveis e dos domicílios						
Responsável é casado	0,527	0,542	-0,015	0,527	0,527	0,000
Escolaridade dos responsáveis (anos de estudo)	8,435	8,958	-0,524***	8,435	8,407	0,028
Mãe e pai moram domicílio	0,437	0,440	-0,003	0,437	0,433	0,004
Beneficiário do Bolsa Família	0,617	0,559	0,058***	0,617	0,616	0,001
Possui computador com internet em casa	0,488	0,562	-0,074***	0,488	0,488	0,000
Numero de pessoas no domicílio	4,675	4,583	0,092	4,675	4,673	0,002
Responsável é trabalhador informal	0,324	0,292	0,032*	0,324	0,318	0,006
Escolheu a escola pela qualidade	0,260	0,270	-0,010	0,260	0,251	0,009
Escolheu o local de moradia pela condição financeira	0,142	0,104	0,038**	0,142	0,139	0,004
Violência é um problema grave no bairro	0,271	0,220	0,051***	0,271	0,269	0,002
Distância da residência até o centro	7,709	7,991	-0,282**	7,709	7,746	-0,037
Características das escolas						
Estuda no período da manhã	0,722	0,786	-0,064***	0,722	0,719	0,003
A escola possui carência de profissionais de apoio pedagógico	0,541	0,399	0,142***	0,541	0,541	-0,001
Escola estadual	0,774	0,804	-0,029*	0,774	0,780	-0,005
A escola possui internet para o professor	0,359	0,401	-0,042***	0,359	0,367	-0,008
A escola possui biblioteca	0,657	0,705	-0,048**	0,657	0,666	-0,009
A escola possui laboratório	0,347	0,396	-0,050**	0,347	0,365	-0,019
Complexidade da gestão escolar	4,973	4,839	0,134*	4,973	4,973	0,000
Porcentagem de docentes de alto esforço <sup>1</sup>	0,102	0,096	0,007*	0,102	0,102	0,000

Fonte: Elaboração própria com base nos dados da Fundaj(2013).

Notas: \*\*\* p<0,01, \*\* p<0,05, \* p<0,1

1-São considerados professores de alto esforço, os docentes que possuem mais de 400 alunos, trabalham nos três turnos e em mais de uma escola.

Diante dessas evidências não é surpreendente a constatação de que os estudantes residentes da favela apresentam desempenhos escolares inferiores àqueles que não moram na favela. Em média, os alunos residentes nas áreas mais precárias obtiveram 2,3 pontos a menos que os alunos que moram fora da favela no teste de matemática aplicado pela Fundaj (2013), cuja escala de nota pode variar entre 0 e 100 pontos. Como foi mostrado, nas favelas residem indivíduos com características menos favoráveis ao estudo, logo, a simples comparação entre as medias obtidas pelos dois grupos se torna inapropriada para afirmar que exista o “efeito favela”, no sentido que a favela estaria penalizando os estudantes, uma vez que essa diferença poderia ser decorrente simplesmente do diferencial do perfil dos alunos analisados.

Por esta razão é desejável que se encontre um grupo de comparação que possuam características semelhantes. As três últimas colunas mostram as médias das covariadas depois de efetuar o *matching* a

partir do procedimento de *kernel*. Não foi encontrado nenhum aluno fora do suporte comum. Como podemos observar, as características dos dois grupos de alunos tornam-se bastante similares depois do *matching*, sugerindo a boa qualidade do pareamento.

**Tabela4: Resultados do Propensity Score – modelo logit**

	Coefficiente	Erro padrão
Homem	-0,0946	(0,0865)
Branco	-0,0509	(0,109)
Idade	0,0668	(0,0496)
Frequentou creche	-0,0666	(0,0932)
Já reprovou pelo menos uma vez	-0,0579	(0,117)
Gosta de ir para escola	-0,267***	(0,0909)
Responsável é casado	-0,0156	(0,0951)
Escolaridade do Responsável	-0,0252*	(0,0130)
Mãe e pai moram no domicílio	0,0450	(0,0990)
Beneficiário do Bolsa Família	0,149*	(0,0882)
Possui computador com internet	-0,186**	(0,0877)
Número de pessoas no domicílio	0,0127	(0,0259)
Responsável é trabalhador informal	0,118	(0,0923)
Escolheu a escola pela qualidade	0,0788	(0,0988)
Escolheu o bairro pela condição financeira	0,387***	(0,128)
Violência é um problema grave no bairro	0,254**	(0,0989)
Distância da residência ao centro	-0,0382**	(0,0162)
Estuda pela manhã	-0,180*	(0,101)
Carência de profissionais de apoio pedagógico	0,557***	(0,0896)
Escola estadual	-0,207*	(0,120)
Internet para os professores	-0,170*	(0,0919)
Biblioteca	-0,0902	(0,0961)
Laboratório	-0,252***	(0,0935)
Complexidade da gestão escolar	0,134***	(0,0495)
Percentual de professores de alto esforço <sup>1</sup>	-0,593	(0,632)
Constante	-1,493**	(0,648)
Número de observações	2.570	

Fonte: Elaboração própria com base nos dados da Fundaj (2013).

Notas: \*\*\* p<0,01, \*\* p<0,05, \* p<0,1

<sup>1</sup>São considerados professores de alto esforço, os docentes que possuem mais de 400 alunos, trabalham nos três turnos e em mais de uma escola.

A Tabela 4 reporta os valores estimados para os coeficientes do modelo *logit* utilizado para obter o *propensity score*. Observa-se que a maioria dos coeficientes estimados possuem os sinais esperados, principalmente as variáveis associadas à renda. Por exemplo, ser beneficiário do Bolsa Família e não ter computador em casa com internet aumenta a probabilidade de morar na favela. Além disso, escolher o bairro de moradia segundo o critério condição financeira também está relacionado à maior probabilidade de morar na favela. A pesquisa ainda questionou os responsáveis sobre o problema da violência no bairro em que moram. As evidências expostas na Tabela 4 sugerem que a percepção de que a violência é um problema grave está positivamente correlacionada com o fato de morar na favela.

Com o intuito de avaliar a qualidade do pareamento, nós aplicamos os testes propostos por Dehejia e Wahba (2002) cujos resultados são reportados na Tabela 5.

**Tabela 5: Teste da qualidade do Pareamento**

Amostra	Pseudo R <sup>2</sup>	LR chi2	P > chi2	Viés Médio	Viés Mediano
Não pareado	0,082	140,65	0	9,6	9,7
1 vizinho mais próximo					
Pareado	0,006	15,11	0,955	2,7	2,4
5 vizinhos mais próximo					
Pareado	0,003	7,13	1	1,8	0,9
Kernel					
Pareado	0,002	4,51	1	1,1	0,6

Fonte: Elaboração própria com base nos dados da Fundaj (2003).

Nota-se que para qualquer que seja o algoritmo adotado para parear a amostra, os valores obtidos dos pseudo-R<sup>2</sup> são bem próximos de zero, o que sugere que o modelo após o *matching* tem menor poder para explicar a condição de tratamento. O teste da Razão de Verossimilhança (LR) aponta para a insignificância conjunta dos regressores quando utilizada a amostra pareada, sendo uma evidência

adicional para a boa qualidade do *matching*. Por fim, nota-se a expressiva redução do viés, dado pela diferença de média e de mediana entre as características observáveis do grupo controle e tratamento, antes e depois do pareamento. Todas essas evidências indicam que construímos grupos de alunos bastante similares.

## 5.2 Influência da condição de favelado sobre o desempenho escolar

Nessa seção apresentamos o efeito estimado de morar na favela sobre o desempenho escolar via OLS e utilizando o *propensity score matching*. Como exposto na Tabela 6, os resultados dos três algoritmos do *propensity score* (vizinho mais próximo, 5 vizinhos mais próximos e Kernel) foram significativos e variam -1,22 e -1,44, sugerindo que exista um “efeito-favela”. Esse valor é bastante representativo e corresponde entre 52% e 60% da diferença média incondicional apresentada na Tabela 3.

**Tabela 6: Influência da favela sobre o desempenho escolar - *Propensity score matching***

	Coeficiente	Bootstrap Std. Err.	z	P>z	[95% Conf. Interval]	
OLS	-1,067989	0,5234474	-2,04	0,041	-2,09392	-0,042050
Nn (1)	-1,447513	0,7328965	-1,98	0,048	-2,88396	-0,011062
Nn (5)	-1,168433	0,6968318	-1,68	0,094	-2,53419	0,197332
Kernel	-1,222924	0,576496	-2,12	0,034	-2,35283	-0,093012
Controles						
Aluno	sim					
Responsáveis	sim					
Escola	sim					

Fonte: Elaboração própria com base nos dados da Fundaj (2003).

Nota: Erro Bootstrap calculado após 100 replicações.

Como explicado na metodologia, esses resultados serão críveis se a hipótese de independência condicional for respeitada, suposição essa que não é possível de ser testada empiricamente. Por essa razão, empregamos uma metodologia de análise de sensibilidade desenvolvida por Ichino et al (2006), a qual simula a possibilidade da presença de uma variável não observada que viole hipótese de ignorabilidade.

**Tabela 7: Influência da favela sobre o desempenho escolar - Análise de Sensibilidade: Efeito da calibração dos *confounders***

	(p11)	(p10)	(p01)	(p00)	ATT	Out. Eff.	Sel. Eff.
Sem Confounder	0	0	0	0	-1,447	-	-
Gosta de estudar	0,30	0,33	0,36	0,37	-1,551	0,958	0,792
Bolsa Família	0,61	0,62	0,53	0,58	-1,405	0,808	1,267
Computador com internet	0,56	0,42	0,61	0,51	-1,324	1,443	0,74
Escolheu a escola pela qualidade	0,27	0,26	0,28	0,26	-1,47	1,139	0,986
Escolheu o bairro pela condição financeira	0,15	0,14	0,1	0,11	-1,426	0,96	1,516
Violência é um problema no bairro	0,26	0,28	0,22	0,22	-1,457	1,03	1,305
Carência de profissionais de apoio pedagógico	0,52	0,56	0,39	0,41	-1,340	0,912	1,765
Internet para professores	0,38	0,36	0,42	0,39	-1,452	1,123	0,862
Biblioteca	0,65	0,66	0,69	0,72	-1,492	0,859	0,812
Laboratório	0,39	0,31	0,39	0,4	-1,478	0,94	0,806

Fonte: Elaboração própria com base nos dados da Fundaj (2003).

Nota: Resultados obtidos a partir do pacote SENSATT do Stata disponibilizado por Nanicinni(2007).

A Tabela 7 expõe os resultados obtidos a partir da calibragem de fatores não observados. Essa calibragem se deu por meio da simulação das variáveis binárias independentes do modelo, isto é, geramos variáveis que “imitam” o comportamento das variáveis binárias que foram utilizadas no pareamento original.

As quatro primeiras colunas reportam os valores estimados para os parâmetros  $p_{ij}$  utilizados para caracterizar a distribuição do fator  $U$ , em que  $i$  se refere ao status de tratamento, ao passo que  $j$  indica o status do resultado. Por exemplo,  $p_{11}$  indica a fração de alunos que moram na favela e que obteve nota

maior que a média. A quarta coluna informa o ATT estimado depois de considerar a distribuição de  $U$ . Por fim, as duas últimas colunas apresentam o efeito da variável  $U$  sobre a probabilidade relativa de se obter um efeito positivo sobre a variável de interesse, *outcome-effect*, e sobre a probabilidade relativa do aluno morar na favela, *treatment-effect*, respectivamente.

Para facilitar a comparação entre os resultados obtidos sem e com *confounders*, seguimos a sugestão de Ichino et al (2006) e reportamos na primeira linha os resultados o ATT estimado sem a inclusão de nenhuma outra variável além daquelas apresentadas na Tabela 3.2. Nas demais linhas, inserimos os resultados estimados quando o fator  $U$  é calibrado para imitar diferentes covariadas observáveis. Por exemplo, na terceira linha, a distribuição de  $U$  foi desenhada para ser semelhante à distribuição dos beneficiários do programa Bolsa Família.

Como se pode observar, a análise de sensibilidade indica que fatores não observáveis correlacionados com as covariáveis binárias empregadas neste estudo não parecem ser suficientes para direcionar o ATT estimado para zero e nem mesmo geram estimativas significativamente diferentes daquelas da Tabela 6. Na verdade, as novas estimativas para o ATT correspondem a valores entre 1,34 e 1,55, bastante próximos ao valor de -1,4, obtido desconsiderando-se qualquer destas variáveis omitidas simuladas. Nota-se, também, que, em geral, a influência da variável omitida simulada sobre as probabilidades relativas de receber o tratamento (morar na favela), *treatment-effect*, e de apresentar variável de impacto acima da média (nota acima da média), *outcome effect*, estão próximas ao valor unitário.

**Tabela 8: Análise de sensibilidade: caracterizando “killer” confounders**

		s=0,1	s=0,2	s=0,3	s=0,4	s=0,5	s=0,6	s=0,7
Out. Eff.\Sel. Eff.		[0,538; 1,325]	[0,821; 1,987]	[1,243; 3,056]	[1,921; 4,752]	[2,931; 7,413]	[4,938; 13,162]	[9,797; 27,314]
d=-0,1	[0,556; 0,657]	-1,328 (0,071)	-1,035 (0,116)	-0,731 (0,183)	-0,392 (0,274)	-0,13 (0,408)	0,172 (0,434)	0,626 (0,637)
d=-0,2	[0,324; 0,437]	-1,371 (0,083)	-0,872 (0,136)	-0,332 (0,183)	0,27 (0,242)	0,779 (0,332)	1,426 (0,458)	2,304 (0,526)
d=-0,3	[0,193; 0,285]	-1,511 (0,112)	-0,913 (0,152)	-0,165 (0,183)	0,635 (0,240)	1,385 (0,320)	2,428 (0,381)	3,586 (0,426)
d=-0,4	[0,121; 0,177]	-1,748 (0,145)	-1,016 (0,167)	-0,116 (0,205)	0,894 (0,242)	1,895 (0,311)	3,036 (0,379)	4,334 (0,434)
d=-0,5	[0,072; 0,105]	-2,23 (0,168)	-1,209 (0,201)	-0,136 (0,243)	1,041 (0,253)	2,205 (0,307)	3,593 (0,322)	5,047 (0,382)
d=-0,6	[0,039; 0,052]	-2,777 (0,249)	-1,642 (0,223)	-0,454 (0,257)	0,938 (0,253)	2,208 (0,292)	3,963 (0,305)	5,684 (0,359)
d=-0,7	[0,016; 0,02]	-3,547 (0,279)	-2,1 (0,253)	-0,774 (0,276)	0,791 (0,293)	2,279 (0,301)	4,029 (0,335)	5,92 (0,316)

Fonte: Elaboração própria com base nos dados da Fundaj (2003).

Nota: Desvio-padrão obtido por *bootstrap* entre parênteses. Resultados obtidos a partir do pacote SENSATT do Stata disponibilizado por Nannicini (2007).

Um segundo exercício de simulação proposto por Ichino et al (2006) é baseado na construção de “killer” *confounder*. Ou seja, em vez de “imitar” o comportamento das variáveis observáveis binárias, procura-se obter um conjunto de parâmetros  $p_{ij}$ , tal que, se  $U$  fosse observado, o ATT seria conduzido para zero (Nannicini, 2007). A observação dos valores associados aos *treatment-effect* e *outcome effect* deste conjunto de parâmetros permitiria avaliar, pois, a plausibilidade dessa configuração particular de parâmetros. Para tanto, é necessário fixar valores pré-determinados de probabilidades. Seja  $d = p_1 - p_0 > 0$ , isto é, a probabilidade de estar no grupo dos alunos moradores da favela maior que a probabilidade de estar no grupo dos não moradores da favela, e seja  $s = p_{01} - p_{00} < 0$ , isto é, a probabilidade de um resultado negativo maior do que a probabilidade de um resultado positivo na variável de interesse quando o grupo é composto pelos alunos não tratados.

Os resultados estimados para o ATT (entradas da tabela) associados a este novo conjunto de simulações são apresentados na tabela 8, a seguir, para cada par de situações quanto a valores de  $s$  e  $d$  considerado. Note-se que, como levado a efeito por Ichino et al (2006), ao longo de cada linha,  $d$  é mantido fixo enquanto  $s$  aumenta. Ademais, o valor pré-determinado de  $d$  está associado ao intervalo de



variação do *outcome-effect* estimado. De modo análogo,  $s$  é associado ao intervalo de variação do *treatment-effect* em cada coluna.

Como se pode observar a partir dos números da Tabela 8, a obtenção de valores insignificantes para o ATT exigem valores bastante reduzidos para o *outcome-effect* ou bastante elevados para o *treatment-effect*, sugerindo a pouca plausibilidade da existência e tais *cofounders*. Por exemplo, para conduzir o efeito da favela para zero quando  $s = 0,4$  e  $d = -0,2$ , o impacto da variável omitida sobre a probabilidade relativa de ser tratado (morar em favela), ou seja, o *treatment-effect* precisa ser de valor 4,752. Isto é, a condição associada a tal variável omitida deve elevar a probabilidade relativa de ser tratado por um fator de maior que 4,5 vezes e a probabilidade relativa de se ter uma valor da nota maior que a média de quase 2 vezes. Dada a extensão do conjunto de informações utilizado nesta pesquisa, parece pouco plausível a existência de um tal fator. Ressalta-se, ainda, que os valores extremos correspondentes ao *outcome-effect* e o *treatment-effect* da tabela 8 diferem em muito das variáveis reais utilizadas na análise anterior.

### 5.3 Robustez dos resultados

Dois conjuntos de *checks* de robustez são considerados nesta subseção. O primeiro conjunto de testes de robustez busca verificar a possibilidade da sobre-estimação do efeito negativo da favela no desempenho escolar a partir da consideração de grupos específicos de famílias de alunos, mas com mesmo estimador até aqui empregado (*propensity score matching*). O segundo conjunto de *checks* utiliza diferentes estimadores para obter o impacto da condição de morador da favela sobre o desempenho escolar, mas mantém o conjunto total inicial de alunos.

No primeiro caso, busca-se verificar se existem fatores não observáveis associados a condições específicas associadas às condições de estudo, aos critérios de escolha do local de moradia ou à qualidade da escola que poderiam afetar os resultados, uma vez que os alunos favelados são economicamente mais desfavorecidos e frequentam uma unidade escolar com infraestrutura mais precária. Neste sentido, são obtidas estimativas para o ATT limitando-se, de forma alternada, a amostra aos estudantes que não possuem computador com internet em casa, aos alunos cujos responsáveis escolheram o local de moradia por outros motivos que não a condição financeira e, finalmente, aos alunos que frequentam uma escola com dificuldade de contratar profissionais de apoio pedagógico. Os resultados encontrados a partir o algoritmo *kernel matching* estão reportados na tabela 9, a seguir.

**Tabela 9: Influência da favela sobre o desempenho escolar - Análise de robustez**

	Coefficiente	Bootstrap Std. Err.	z	P>z	[95% Conf. Interval]	
	Alunos que não possuem computador com internet					
Kernel	-1,368083	0,7649822	-1,79	0,074	-2,86742	0,131254
	Escolheu o bairro por motivos diferentes de “condição financeira”					
Kernel	-1,000415	0,5453977	-1,83	0,067	-2,06937	0,068544
	Estuda em uma escola com carência de profissionais de apoio pedagógico					
Kernel	-1,804244	0,9281395	-1,94	0,052	-3,62336	0,014876

Fonte: Elaboração própria com base nos dados da Fundaj (2003).

Nota: Desvio-padrão obtido por *Bootstrap* e calculado após 100 replicações.

Como se pode observar, os coeficientes estimados para as três amostras restritas foram negativos e significantes (ao menos ao nível de 10%), corroborando com as evidências anteriores de que existe um efeito favela. Além disso, todos os três ATT's estimados a partir destas amostras restritas não são estatisticamente diferentes daquele encontrado anteriormente para amostra completa (-1,22).

O segundo teste de robustez consiste na obtenção de estimativas para a influência da moradia na favela sobre o desempenho escolar através dos métodos de reponderação (IPW e IPWRA) e da estratégia proposta por Kline (2011) a partir da decomposição se Oaxaca-Blinder. Como exposto na metodologia, os estimadores IPWRA e aquele associado à proposta de Kline (2011) são duplamente robustos, portanto, a confirmação dos resultados a partir dos mesmos reforça a credibilidade dos resultados até aqui obtidos. O novo conjunto de resultados desses procedimentos é apresentado a partir da Tabela 10.

**Tabela 10: Influência da favela sobre o desempenho escolar - Análise de robustez: métodos de reponderação e Oaxaca-Blinder**

	Coef.	Bootstrap Std. Err.	z	P>z	[95% Conf. Interval]	
IPW	-1,027	0,529	-1,940	0,052	-2,064	0,009
IPWRA	-1,019	0,489	-2,080	0,037	-1,976	-0,061
Oaxaca-Blinder	-1,230	0,589	2,090	0,037	-2,385	-0,075
Controles						
Aluno	sim					
Responsáveis	sim					
Escola	sim					

Fonte: Elaboração própria com base nos dados da Fundaj (2003).

Nota: Desvio-padrão obtido por *Bootstrap* e calculado após 100 replicações.

Como se pode aprender a partir dos valores (ATT) estimados da Tabela 10, os coeficientes estimados são todos negativos, estatisticamente significantes e com magnitudes semelhantes ao que já obtido nos exercícios econométricos anteriores. Tais evidências sugerem, mais uma vez, que um estudante que mora na favela tende a ter um aproveitamento escolar menor do que outro com características socioeconômicas semelhantes, mas que mora fora das áreas urbanas mais segregadas.

## 6 Considerações finais

Como mostraram recentemente Oliveira e Silveira Neto (2015), Recife é reconhecidamente uma cidade macrossegredada e figura entre aquelas com a maior proporção de moradores residentes em favelas no Brasil (IBGE, 2011). Em que medida tal segregação espacial de indivíduos condiciona seus resultados sociais é uma questão praticamente inexplorada nas pesquisas sociais aplicadas. Com pretensão de iniciar o preenchimento dessa lacuna, o presente estudo teve como objetivo principal, pois, analisar especificamente a influência da favela sobre o desempenho escolar dos alunos da rede pública de ensino da Cidade do Recife. Apesar de certas características conhecidas, muito pouco é conhecido sobre a potencial influência que a moradia localizada na favela tem sobre a trajetória individual das crianças ou adolescentes.

A principal dificuldade empírica deriva do fato dos alunos não serem aleatoriamente distribuídos ao longo do território urbano, onde as favelas tendem a abrigar os estudantes com perfil socioeconômico menos favorável ao estudo. Nesse sentido, uma simples comparação entre o desempenho médio obtido por alunos moradores e não moradores da favela seria uma análise ingênua, já que podemos atribuir ao local de moradia os efeitos decorrentes das diferenças no perfil dos alunos. Para contornar esse problema, exploramos a base de dados da Fundaj (2013) que se destaca por informar, de maneira inédita, os endereços dos alunos da rede pública de ensino da Cidade do Recife. O georreferenciamento dos endereços, combinado com o *shapefile* dos aglomerados subnormais disponibilizado pelo IBGE (2010), permitiu a identificação dos alunos entre moradores das favelas recifenses.

Como estratégia empírica, adotamos a técnica de *propensity score matching* (PSM) e análise de sensibilidade proposta por Ichino et al (2008). Adicionalmente, utilizamos dois métodos de reponderação e o coeficiente de Oaxaca-Blinder com o intuito de verificar se o efeito tratamento (isto é, morar na favela) é sensível a diferentes ponderações e ao método utilizado na estimação.

Os resultados encontrados, que são robustos a diversas especificações, sugerem que as áreas urbanas mais segregadas impacta de forma negativa o desempenho médio dos estudantes em matemática. Em média, 50% da diferença das médias incondicionais da nota de matemática entre os dois grupos de alunos pode ser atribuída às favelas. Isto é, a média incondicional dos alunos favelados é 2,3 pontos a menos que a dos demais alunos no teste de matemática, em uma escala que pode variar de 0 a 100 pontos. Após controlarmos pelas características observáveis, essa diferença diminui para aproximadamente 1,2 pontos.

Tais resultados estão alinhados com as evidências nacionais e internacionais que apontaram para os efeitos adversos de se crescer nas áreas urbanas mais segregadas, o que pode dificultar a mobilidade

social dos mais pobres (Wilson, 1987; Jencks e Mayer, 1990; Cutler, 1991; Torres et al, 2004 e Alves *et al.*, 2008).

Ressaltamos que embora os resultados encontrados nesse estudo sejam interessantes no sentido de ser a primeira evidência empírica do efeito causal da favela sobre o aprendizado do aluno na Cidade do Recife, não conseguimos distinguir por qual mecanismo essa influência operaria. Essa distinção é extremamente necessária para o desenho de políticas públicas que procurem atenuar, de modo eficaz, os efeitos negativos de se morar na favela. Nesse sentido, novas pesquisas devem ser feitas para cobrir essa lacuna e, possivelmente, com dados longitudinais dos alunos e do local de moradia para o melhor controle das características não observáveis.

## Referências

- AINSWORTH, J. W. Why does it take a village-the mediation of neighborhood effects on educational achievement. *Soc. F.*, v. 81, p. 117, 2002.
- ALVES, F., FRANCO, C., RIBEIRO, L. C. Q. Segregação Residencial e Desigualdade Escolar no Rio de Janeiro. In: RIBEIRO, L.C. Q e KAZTMAN, R. *A cidade contra a Escola: Segregação urbana e desigualdades educacionais em grandes cidades da América Latina*. Rio de Janeiro: Letra Capital. 2008
- ANDERSON, E. Neighborhood effects on teenage pregnancy. *The urban underclass*, v. 375, 1991.
- ANGRIST, J. D.; PISCHKE, J.S.. **Mostly harmless econometrics: An empiricist's companion**. Princeton university press, 2008.
- BAYER, P.; ROSS, S. L.; TOPA, G.. Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy*, v. 116, n. 6, p. 1150-1196, 2008.
- BOLSTER, A.; BURGESS, S.; JOHNSTON, R.; JONES, K.; PROPPER, C; SARKER, R. Neighbourhoods, households and income dynamics: a semi-parametric investigation of neighbourhood effects, *Journal of Economic Geography*, vol. 7(1), pp. 1–38, 2007.
- BURGOS, M. B.; PAIVA, A.. A Escola e a Favela. **Rio de Janeiro: editora PUC-Rio**, 2009.
- BURSIK JR, R. J; GRASMICK, H. G. **Neighborhoods & crime**. Lexington Books, 1999.
- CALIENDO, M.; KOPEINIG, S.. Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, v. 22, n. 1, p. 31-72, 2008.
- CASE, A. C.; KATZ, L. F. **The company you keep: The effects of family and neighborhood on disadvantaged youths**. National Bureau of Economic Research, 1991.
- CAVALCANTI, H.; SOUZA, M. A. A. ; MIRANDA, L. ; NEVES, N. . Tipologia e Caracterização dos Assentamentos Precários: Região Metropolitana do Recife (RMR). 2010. (Relatório de pesquisa).
- CHETTY, R.; HENDREN, N.; KATZ, L. F. The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. **National Bureau of Economic Research**, 2015.
- CUTLER, D. M.; GLAESER, E. L. Are ghettos good or bad?. *The Quarterly Journal of Economics*, v. 112, n. 3, p. 827-872, 1997.
- CIRA, D. A. Urban upgrading in latin america and the caribbean. **En Breve, World Bank**, n. 3, 2002.
- COSTA, F. Guetos ou favelas? *Romanica Olomucensia* **25.1**, p.37–45, 2013.
- CRANE, J.. The epidemic theory of ghettos and neighborhood effects on dropping out and teenage childbearing. *American journal of Sociology*, v. 96, n. 5, p. 1226-1259, 1991.
- DEHEJIA, Rajeev H.; WAHBA, Sadek. Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, v. 84, n. 1, p. 151-161, 2002.
- DURLAUF, S. N. Neighborhood effects. *Handbook of regional and urban economics*, v. 4, p. 2173-2242, 2004.
- ELLEN, I. G.; TURNER, M. A.. Does neighborhood matter? Assessing recent evidence. *Housing policy debate*, v. 8, n. 4, p. 833-866, 1997.
- ELLIOT, D. S.; WILSON, W. J.; HUIZINGA, D.; SAMPSON, R. J., ELLIOTT, A.; RANKIN, B. The effects of neighborhood disadvantage on adolescent development. *Journal of Research in Crime and Delinquency*, v. 33, n. 4, p. 389-426, 1996.
- ESTPHAL, E.. Urban Slums, pacification and discrimination: Afield experiment in Rio de Janeiro's labor market. **Massachusetts, Estados Unidos**, 2014.
- EVANS, W. N.; OATES, W. E.; SCHWAB, R. M. Measuring peer group effects: A study of teenage behavior. *Journal of Political Economy*, v. 100, n. 5, p. 966-991, 1992.
- FORTIN, N.; LEMIEUX, T.; FIRPO, S.. Decomposition methods in economics. *Handbook of labor economics*, v. 4, p. 1-102, 2011.
- GARCÍA CRUZ, G. A.; NICODEMO, C.. Job Search Channels, Neighborhood Effects and Wages Inequality in Developing Countries: The Colombian Case. 2013.
- GIBBONS, S.; SILVA, O.; WEINHARDT, . Everybody needs good neighbours? Evidence from students' outcomes in England. *The Economic Journal*, v. 123, n. 571, p. 831-874, 2013.
- GOUX, D.; MAURIN, E.. Close neighbours matter: Neighbourhood effects on early performance at school. *The Economic Journal*, v. 117, n. 523, p. 1193-1215, 2007.
- HARDING, D. J. Counterfactual Models of Neighborhood Effects: The Effect of Neighborhood Poverty on Dropping Out and Teenage Pregnancy 1. *American Journal of Sociology*, v. 109, n. 3, p. 676-719, 2003.
- HELMERS, Christian; PATNAM, Manasa. Does the rotten child spoil his companion? Spatial peer effects among children in rural India. *Quantitative Economics*, v. 5, n. 1, p. 67-121, 2014.
- HIRANO, K.; IMBENS, G. W.; RIDDER, G.. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, v. 71, n. 4, p. 1161-1189, 2003.

IBGE. Censo Demográfico 2010: aglomerados subnormais – primeiros resultados. Rio de Janeiro: IBGE, 2011. Disponível em: [http://www.ibge.gov.br/english/estatistica/populacao/censo2010/aglomerados\\_subnormais/agsn2010.pdf](http://www.ibge.gov.br/english/estatistica/populacao/censo2010/aglomerados_subnormais/agsn2010.pdf).

ICHINO, A.; MEALLI, F.; NANNICINI, T.. From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity?. **Journal of Applied Econometrics**, v. 23, n. 3, p. 305-327, 2008.

IMBENS, G. W.; WOOLDRIDGE, J. M. Recent developments in the econometrics of program evaluation. **Journal of economic literature**, v. 47, n. 1, p. 5-86, 2009.

JENCKS, C.; MAYER, S. E. The social consequences of growing up in a poor neighborhood. In: **Inner-city poverty in the United States**, v. 111, p. 186, 1990.

KLINE, P.. Blinder-oaxaca as a Reweighting Estimator. **American Economic Review: Papers and Proceedings**, 101, pp. 532-537, 2011.

KLING, J. R.; LUDWIG, J.; KATZ, L. F. Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment. **The Quarterly Journal of Economics**, v. 120, n. 1, p. 87-130, 2005.

LUDWIG, J., LADD, H.F., DUNCAN, G.J.: **Urban Poverty and Educational Outcomes**. Brookings-Wharton Pap. Urban Aff. 147–201, 2001.

MAYER, S. E.; JENCKS, C.. Growing up in poor neighborhoods: How much does it matter?. **Science**, v. 243, n. 4897, p. 1441-1446, 1989.

MENEZES-FILHO, N. A. Os determinantes do desempenho escolar do Brasil. IFB, 2007

MONTEIRO, Joana; ROCHA, Rudi. Drug battles and school achievement: evidence from Rio de Janeiro's favelas. **Review of Economics and Statistics**, n. 0, 2012.

NANNICINI, T. Simulation-based sensitivity analysis for matching estimators. **Stata Journal, Citeseer**, v. 7, n. 3, p. 334, 2007.

NETA, M.. **Geografia e Literatura: decifrando a paisagem dos mocambos do Recife**, 116 f. 2005. Dissertação (Mestrado em Geografia)–Universidade Federal do Rio de Janeiro, Rio de Janeiro.

OLIVEIRA, T. G.; SILVEIRA NETO, R. M.. Segregação residencial na Cidade do Recife: Um estudo da sua configuração. **Revista Brasileira de Estudos Regionais e Urbanos**, v. 9, n. 1, p. 71-92, 2016.

PASTERNAK, S.; D'OTTAVIANO, C.. Favelas no Brasil e em São Paulo: avanços nas análises a partir da Leitura Territorial do Censo de 2010. **Cadernos Metr pole.**, v. 18, n. 35, p. 75-100, 2016.

PAIVA, A. R. Cidadania e formas de solidariedade social na favela. . In: BURGOS, M. B.; PAIVA, A.. A Escola e a Favela. **Rio de Janeiro: editora PUC-Rio**, 2009.

PERO, V.; CARDOSO, A.; ELIAS, P.. Discriminação no mercado de trabalho: o caso dos moradores de favelas cariocas. **Coleção Estudos Cariocas**, v. 5, 2005.

PLOTNICK, R. D.; HOFFMAN, S. D. The effect of neighborhood characteristics on young adult outcomes: Alternative estimates. **Social Science Quarterly**, p. 1-18, 1999.

RAPOSO, I. P. A, MENEZES, T. A; CARVALHO, R. , N BREGA, R. Z. **Difusão do efeito dos pares** dentro da rede de amigos de sala de aula. In: Anais XIII Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos, 2015.

RIBEIRO, E. Impactos educacionais nas Unidades de Polícia Pacificadora: Explorando os efeitos sobre os fluxos docentes. **Revista Intratextos** , v. 4, p. 27-52, 2013a.

RIBEIRO, E. Vizinhança, violência urbana e educação no Rio de Janeiro: Efeitos territoriais e resultados escolares. BIB. **Revista Brasileira de Informação Bibliográfica em Ciências Sociais** , v. 1, p. 5-182, 2013b.

RIBEIRO, Luiz Cesar de Queiroz; KOSLINSKI, Mariane Campelo. **Fronteiras Urbanas e Oportunidades Educacionais: o caso do Rio de Janeiro**. 33ª Reuni  o Anual da ANPOCS, 2009.

ROCHA, L. P.; PESSOA, M.; MACHADO, D. C.. Discriminação espacial no mercado de trabalho: o caso das favelas do Rio de Janeiro. **Revista Brasileira de Estudos Regionais e Urbanos**, v. 7, n. 1, p. 38-57, 2015.

RODRIGUES, R.I. Moradia Prec ria e Viol ncia na Cidade de S o Paulo. Rio de Janeiro: Ipea, 2006 (texto para discuss o).

ROSENBAUM, J. E. Changing the geography of opportunity by expanding residential choice: Lessons from the Gautreaux program. **Housing Policy Debate**, v. 6, n. 1, p. 231-269, 1995.

ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. **Biometrika**, p. 41-55, 1983

SANT'ANNA, M. J. G.. O papel do territ rio na configura  o das oportunidades educativas: efeito escola e efeito vizinhan a. **Cidade: olhares e trajet rias. Rio de Janeiro: Garamond**, p. 167-192, 2009.

SHARKEY, P.; FABER, J. W. Where, when, why, and for whom do residential contexts matter? Moving away from the dichotomous understanding of neighborhood effects. **Annual Review of Sociology**, v. 40, p. 559-579, 2014.

S ŁCZYŃSKI, T.. The Oaxaca–Blinder unexplained component as a treatment effects estimator. **Oxford Bulletin of Economics and Statistics**, v. 77, n. 4, p. 588-604, 2015.

SOARES, J. F.; RIGOTTI, J. I. R.; ANDRADE, L. T. As desigualdades socioespaciais e o efeito das escolas p blicas de Belo Horizonte. In: RIBEIRO, L.C. Q e KAZTMAN, R. **A cidade contra a Escola: Segrega  o urbana e desigualdades educacionais em grandes cidades da Am rica Latina**. Rio de Janeiro: Letra Capital. 2008

SOUSA, A. **Do mocambo   favela: Recife, 1920-1990**. Editora Universitaria, 2003.

VARTANIAN, T. P. Adolescent neighborhood effects on labor market and economic outcomes. **Social Service Review**, v. 73, n. 2, p. 142-167, 1999.

VASCONCELLOS, P.; ROCHA, L. F. Intera  o social e evas o escolar nas favelas do Rio de Janeiro–um problema de identifica  o. In: **Anais do XXXIV Encontro Nacional de Economia**. ANPEC-Associa  o Nacional dos Centros de P s gradua  o em, 2006.

WACQUANT, L.. What is a ghetto? Building a sociological concept. **Revista de Sociologia e Pol tica**, n. 23, p. 155-164, 2004.

WILSON, W. J. The truly disadvantaged: the inner city, the underclass and public policy. Chicago. **University of Chicago**, 1987.