

# A Saúde dos Amigos de Sala de Aula Interfere no Desempenho Escolar do Aluno?

## Isabel Pessoa de Arruda Raposo

Pesquisadora – Fundação Joaquim Nabuco (Fundaj/MEC)

Endereço: Rua Dois Irmãos, 92 – Ed. Anexo Anízio Teixeira 3900 – Recife/PE

CEP: 52071-440 – E-mail: [i\\_raposo@hotmail.com](mailto:i_raposo@hotmail.com) – <https://orcid.org/0000-0001-7304-546X>

## Michela Barreto Camboim Gonçalves

Pesquisadora – Fundação Joaquim Nabuco (Fundaj/MEC)

Endereço: Rua Dois Irmãos, 92 – Ed. Anexo Anízio Teixeira 3900 – Recife/PE

CEP: 52071-440 – E-mail: [michelabcg@hotmail.com](mailto:michelabcg@hotmail.com) – <https://orcid.org/0000-0003-3536-3971>

Recebido: 02/12/2016. Aceite: 25/01/2018.

## Resumo

Este artigo tem como objetivo investigar se existe alguma relação entre a saúde dos amigos e o desempenho acadêmico de alunos do 6º ano de escolas públicas da cidade do Recife. As estimações empíricas utilizam uma pesquisa da Fundação Joaquim Nabuco (Fundaj 2013) que traz informações cruciais para as estratégias de identificação propostas. Essas estratégias baseiam-se: (i) na arquitetura de redes de amizades em sala de aula composta por grupos de referência heterogêneos, (ii) na utilização de uma avaliação longitudinal do desempenho em duas provas de matemática, (iii) no controle do efeito fixo da rede e (iv) em estimações de modelos espaciais utilizando a matriz de amizades da turma. Os resultados demonstram que os alunos apresentam melhor desempenho acadêmico quando se relacionam com amigos de sala de aula que possuem um bom estado geral de saúde. Um aumento de 1 desvio padrão (d.p.) na proporção de amigos saudáveis, leva a um acréscimo de aproximadamente 9% do d.p. da nota média de matemática do aluno ao final do ano letivo.

## Palavras-Chave

Saúde. Desempenho escolar. Rede de amizades. Efeito de pares.

## Abstract

This paper evaluates the influence of classmate's health on individual academic achievement of six graders students from public schools in the city of Recife/Pernambuco. The empirical estimations use a new educational dataset originated from a Brazilian public institution (Fundação Joaquim Nabuco 2013), which provides crucial information for the identification strategies. Such strategies exploit: (i) the architecture of friendship networks in the classrooms, (ii) a longitudinal evaluation of school achievement in two math tests applied at the beginning and end of school year, (iii) the control of network fixed effects and (iii) spatial models estimation using friendship network matrix within classroom. The results show students perform better when their classmates' friends exhibit a good health status. An increase of one standard deviation (sd) in the ratio of healthy peers improves by 9% sds the student's grade.



Esta obra está licenciada com uma Licença Creative Commons Atribuição-Não Comercial 4.0 Internacional.

**Keywords**

Health. School achievement. Friendship network. Peer effects.

**JEL Classification**

I20. I10. C21.

**1. Introdução**

O papel dos pares na função de produção educacional é uma das proeminentes questões que a economia da educação se dedica a entender. A literatura sobre a influência dos pares na proficiência do aluno é vasta e inclui diversas áreas do conhecimento. Em um importante trabalho da psicologia, Harris (1995) demonstra que os amigos são fonte de interação, motivação e aspiração dentro do processo de aprendizagem e são relativamente mais importantes para o desenvolvimento de adolescentes do que os pais ou o ambiente domiciliar. Ela argumenta que na busca pela integração ao grupo de pares, as crianças podem mudar radicalmente sua personalidade quando longe da supervisão dos pais. Comportamentos desviantes daquilo esperado pelo grupo em geral não são aceitos e tendem a ser excluídos.

Patacchini, Rainone e Zenou (2011) argumentam que, por um lado, o aluno pode se beneficiar das externalidades (*spillovers*) do conhecimento geradas pelos debates e questionamentos de outros colegas de turma e, por outro, tem potencial para influenciar diretamente os padrões de desempenho de sua sala de aula. Um aluno mais indisciplinado, por exemplo, pode perturbar a concentração dos demais colegas, levando um professor a gastar mais tempo de aula disciplinando a turma ao invés de promover o aprendizado de disciplinas. Sendo assim, esse aluno influencia e também é influenciado diretamente pelas características e comportamento de seus pares. Essa influência de pares também pode se difundir por meio de um efeito imitação, em que os alunos teriam motivações individuais para apresentar um comportamento/desempenho equivalente ao do grupo no qual está inserido.

Se o efeito de pares sobre o desempenho acadêmico é suficientemente forte, então é cabível a intervenção de políticas públicas educacionais. Essas políticas podem guiar desde a formação da turma, passando pela organização da unidade escolar até a concepção do sistema educacional como

um todo. Será que misturar estudantes com desempenhos diferentes é melhor do que segregá-los? Escolas separadas por gênero elevam a performance acadêmica? Alunos de diferentes idades devem ser alocados para mesma turma ou a convivência entre mais novos e mais velhos é benéfica? A criança deve estudar no mesmo bairro de moradia? Misturar alunos com necessidades especiais aos demais estudantes melhora o resultado dos primeiros sem comprometer o desempenho desses últimos? Questionamentos desta natureza dão suporte à elaboração de políticas educacionais e são de interesse não somente dos gestores escolares, mas também dos pais, alunos e toda sociedade.

Nesse contexto, uma questão pouco explorada dentro da literatura de efeito de pares diz respeito à relação entre a saúde dos amigos e o desempenho escolar. É sabido que a saúde do aluno afeta a sua performance acadêmica (Currie 2009), mas será que a saúde do amigo interfere na sua nota? É possível que amigos doentes possam atrapalhar os resultados escolares individuais. Seja por necessitarem de mais atenção para receber a matéria atrasada em virtude de faltas, seja porque ao estarem atrasados em relação ao grupo, os amigos doentes podem reduzir as externalidades do conhecimento, quando deixam de participar de debates e questionamentos do conteúdo abordado em sala de aula ou no estudo em grupo. No entanto, na literatura econômica existem hipóteses alternativas que podem potencialmente explicar a relação entre saúde dos amigos e desempenho escolar sem que existam efeitos de redes sociais. Primeiro, indivíduos podem escolher seus amigos baseados em fatores associados ao seu próprio status de saúde, o que em economia é chamado de homofilia (propensão de pessoas com traços semelhantes associarem-se entre si). Segundo, indivíduos podem ajustar o seu comportamento quando expostos a influências comuns do grupo. Esse tipo de efeito é chamado de influências de contexto ou exógenas. Por exemplo, a escola pode estar inserida numa comunidade mais pobre e carente de infraestrutura social, como saneamento e coleta de lixo, o que pode aumentar a propensão a doenças no seu corpo discente, independente dos laços de amizade que possam existir entre os alunos.

O grande desafio empírico é, portanto, conseguir identificar separadamente o efeito endógeno dos pares dos demais efeitos de contexto e homofilia. É muito difícil distinguir se o desempenho do estudante é influenciado por atributos do seu grupo ou se simplesmente resulta de suas próprias características individuais.

Além disso, não é fácil obter informações que permitam identificar a interação dos pares dentro de um espaço relevante de convivência, o que faz com que a definição dos grupos seja arbitrária e, em geral, em níveis muito agregados. Calvo-Armengol, Patacchini e Zenou (2009) ressaltam que como a influência dos amigos é geralmente tratada como uma externalidade intragrupo, na medida em que esses grupos são definidos de maneira imprecisa, erros de medida vão permear toda a análise de identificação do efeito dos pares.

O objetivo principal desse artigo consiste em investigar em que medida a saúde do grupo de amigos de sala de aula se relaciona com o rendimento escolar do aluno. O trabalho procura contribuir para o debate, apresentando uma abordagem que explica a difusão da influência da saúde dos pares sobre o desempenho escolar a partir da estrutura da rede de amizades do aluno. Trata-se de uma investigação bastante singular na literatura brasileira, uma vez que utiliza uma base de dados inédita da Fundação Joaquim Nabuco (Fundaj 2013), que faz um levantamento detalhado dos grupos de amigos em sala de aula e de aspectos relativos à saúde dos estudantes. Nesse exercício, a estratégia de identificação explora a arquitetura dessas redes sociais para separar os efeitos exógenos dos pares dos demais efeitos de contexto, tal como estudado em Patacchini e Venanzoni (2014), Badev (2013), Mele (2010), Calvo-Armengol, Patacchini e Zenou (2009), Bramoullé, Djebbari e Fortin (2009) e Ballester, Calvo-Armengol e Zenou (2006).

Além desta introdução, o trabalho se desenvolve em mais seis seções. A seção dois faz uma revisão das evidências já publicadas sobre o tema. A terceira seção discute sobre a base de dados e variáveis utilizadas nas estimações. A quarta apresenta o modelo empírico e as estratégias de identificação. Os resultados são apresentados na quinta seção e, por fim, são tecidas as considerações finais do artigo.

## 2. Revisão da Literatura

Esta seção propõe-se a apresentar algumas evidências publicadas no campo da economia da educação dedicadas ao estudo do efeito dos pares sobre desempenho escolar, sociabilidade e a influência comportamental de in-

divíduos no ambiente escolar. Essas evidências procuram identificar os mecanismos pelos quais os padrões de comportamento do aluno e seu grupo estão associados.

Berndt e Perry (1986) argumentam que adolescentes tendem a imitar, por observação, o comportamento de seus amigos devido, principalmente, à forte ligação emocional existente no relacionamento de amizade. Os autores complementam que quanto mais afetiva for a relação de amizade e quanto maior a frequência de interação entre eles, mais influência o grupo de amizade proporciona sobre o comportamento individual. Para Wentzel, Caldwell e Mcnamara (2004), os estudantes são motivados a agir de acordo com o grupo e a adotar comportamentos similares aos de amigos com os quais têm vínculo emocional forte. Por exemplo, os estudantes são geralmente mais sociáveis quando tiverem amigos sociáveis.

Em relação à influência dos amigos sobre o desempenho individual, diversos autores encontram efeitos positivos e não lineares dos pares sobre o desempenho escolar como em Zimmer (2003), Schneeweis e Winter-Ebmer (2007), Pinto (2008), De Giorgi, Pellizzari e Woolston (2010). Epstein (1983), por exemplo, identifica que estudantes com bom desempenho no início do período analisado, cujos amigos tinham desempenho inferior, reduziram sua performance. Por outro lado, estudantes que inicialmente apresentaram desempenho inferior a seus amigos, melhoraram suas notas no final do período analisado. O autor ainda argumenta que a influência dos amigos de alto desempenho parece ser maior que a influência dos amigos de desempenho inferior. Raposo (2015), usando dados da rede de amizades em sala de aula, encontra não somente um efeito positivo dos amigos diretos sobre o resultado acadêmico individual, mas também identifica que alunos mais centrais, em termos de suas centralidades de Katz-Bonacich, possuem maior seu desempenho acadêmico.

Na área de saúde, o trabalho de Christakis e Fowler (2007) encontra efeito endógeno positivo e significativo entre ser obeso e ter amigos obesos. Em outras palavras, “obesidade seria contagiosa”, ou seja, ter amigos obesos pode fazer com que a pessoa se torne obesa também, principalmente pelo efeito de imitação de costumes. Por exemplo, aumentando a tolerância individual ao peso, alterando hábitos individuais de ingestão de comidas mais calóricas, tabagismo e redução do interesse em atividades físicas. Em contrapartida, Cohen-Cole e Fletcher (2008) adicionaram variáveis de contexto e controles para características não observáveis ao modelo de

Christakis e Fowler (2007), e encontram um efeito de contágio da obesidade significativamente reduzido e estatisticamente nulo.

Fletcher (2010), a partir de uma arquitetura de redes sociais, estimou se a decisão individual de fumar é influenciada pelas decisões dos colegas de sala de aula. Utilizando dados do National Longitudinal Study of Adolescent (Add Health),<sup>1</sup> encontrou que aumentando em 10% a proporção de amigos de sala de aula que fumam, aumenta em três pontos percentuais a probabilidade do estudante fumar. Em outro estudo parecido, Fletcher (2012) encontrou que aumentando a proporção de colegas de sala de aula que consomem álcool em 10%, aumenta a probabilidade individual de ingestão alcoólica em cinco pontos percentuais.

O que se observa da discussão até aqui apresentada é uma profusão de evidências publicadas, que relaciona comportamento individual àquele do grupo de pares, mas ainda há uma grande lacuna no que diz respeito à influência do status de saúde dos amigos sobre o resultado cognitivo individual. A motivação maior deste artigo é preencher, pelo menos em parte, tal vazio na literatura.

### 3. A Base de Dados e as Estatísticas Descritivas das Variáveis dos Modelos

Este artigo utiliza uma base de dados inédita proveniente de uma pesquisa realizada pela Fundação Joaquim Nabuco – Fundaj - no ano de 2013, com uma amostra representativa de alunos do 6º ano (5ª série) de escolas públicas da cidade do Recife/Pernambuco. Intitulada *Acompanhamento Longitudinal do Desempenho Escolar de Alunos da Rede Pública de Ensino Fundamental do Recife*, a pesquisa avaliou o desempenho do aluno a par-

<sup>1</sup> A *Add Health* é uma pesquisa longitudinal junto a adolescentes de escolas americanas, que tem como um dos principais destaques o levantamento da rede de amizades desses estudantes. A pesquisa possui uma amostra de 90.118 alunos distribuídos de forma representativa em 132 escolas públicas e privadas dos Estados Unidos. Desde 1994, a pesquisa já coletou cinco painéis (1994, 1995, 1996, 2001 e 2008) junto a alunos do 7º ano até o final do ensino médio. Uma subamostra desses estudantes (cerca de 20.000) responde a um questionário bem detalhado que contém informações sobre a rede de amizades. Nesta os alunos indicam até dez melhores amigos, sendo cinco do sexo feminino e cinco do sexo masculino.

tir de duas provas de matemática (elaboradas pela Fundaj e aplicadas ao início e final do ano letivo) e coletou uma série de informações relacionadas a aspectos internos e externos à escola, por meio de quatro tipos de questionários (um para o aluno, outro para o principal responsável por sua vida acadêmica, outro para o professor de matemática e outro para o diretor da escola). Dentre as informações aferidas, o principal destaque da Pesquisa foi o levantamento da rede de amizades do aluno dentro da sala de aula, um tipo de dado inédito nas pesquisas quantitativas já conduzidas no Brasil e crucial para identificação da influência dos pares no processo de aprendizagem. Nessas perguntas, os alunos listavam até cinco melhores amigos e informavam se eram colegas de sala, se estudavam com os amigos indicados, se frequentava a casa do amigo ou se conversou com ele sobre algum problema. Adicionalmente, para esses amigos listados pelo aluno, os pais diziam se os conheciam, assim como suas famílias, e que tipo de influência os mesmos seriam para seu filho. Assim, para cada amigo de sala de aula, dispõe-se não somente da nota deste colega nas duas avaliações de matemática, como também de todas as demais informações levantadas pelos questionários. A Pesquisa levantou ainda a percepção do responsável sobre diversos aspectos relacionados à saúde do estudante, disponibilizando assim informações relevantes para investigar em que medida o estado geral de saúde dos amigos pode afetar o desempenho acadêmico do aluno.

Ao todo, a Pesquisa (Fundaj 2013) entrevistou 4.191 alunos, 3.670 pais ou responsáveis, 120 diretores e 131 professores de 120 escolas especialmente distribuídas pelas 18 microrregiões da cidade do Recife.<sup>2</sup> Em algumas escolas com maior número de matrículas no 6º ano foram sorteadas duas turmas, e não somente uma, razão pela qual a quantidade total de turmas selecionadas para amostra foi de 146.

Excluindo indivíduos com informações inadequadas ou *missing*, a amostra final é constituída por 139 redes/turmas e 1.854 alunos. Esta redução de 56% em relação ao tamanho original da amostra (4.196 alunos e 146 turmas) se deve, em grande parte, ao próprio processo de construção das redes de amigos: para 48% dos alunos entrevistados, suas amizades citadas

<sup>2</sup> Cada Região Político-Administrativa da cidade do Recife é dividida em três microrregiões “visando à definição das intervenções municipais em nível local e articulação com a população” e compostas por um ou mais dos 94 Bairros estabelecidos pelo Decreto Municipal 14.452, de 26 de outubro de 1988, para levantamento de informações para o IBGE e para o Sistema de Informações e Planejamento do Recife. As 18 Microrregiões correspondem à divisão das Regiões Político-Administrativas, que foi idealizada em 1995 pela Secretaria de Políticas Sociais, para organizar as reuniões do Orçamento Participativo inicialmente limitadas às associações e aos seus representantes (Recife and PNUD 2005).

não puderam ser corretamente emparelhadas ou o estudante não nomeou nenhum amigo de sua classe. Outras perdas (cerca de 15%) foram devidas à eliminação de observações com informações *missings* para o conjunto das variáveis incluídas no modelo a ser testado. Esta larga diminuição no tamanho da amostra é comum quando se trabalha com informações de redes de amizades. Diversos trabalhos que utilizam dados de uma pesquisa americana que também faz o levantamento da rede de amigos do aluno, a *Add Health*, enfrentam perdas ainda maiores. Por exemplo, Patacchini e Venanzoni (2014) trabalham com apenas 19% da amostra saturada da *Add Health*; Badev (2013) com 5,4% e Mele (2010) com 5,5%. Apenas Calvo-Armengol, Patacchini e Zenou (2009) com 55% e Bramoullé, Djebbari e Fortin (2009) com 61% trabalharam com um percentual da amostra original semelhante ao deste artigo.

A Tabela 1 apresenta as estatísticas descritivas das variáveis utilizadas nas estimações deste artigo. As definições e unidades de medidas dessas variáveis se encontram no Quadro A1 do apêndice. O desempenho em matemática foi avaliado no início e final do ano letivo e em média os alunos da amostra utilizada tiveram um percentual de acerto de cerca de 40% nas duas provas, não apresentando variação significativa no período. A nota média inicial foi ligeiramente superior a final, 41,92 e 41,15, mas também apresentou uma dispersão mais elevada. A fração de estudantes do sexo masculino é 0,45, e dos que se declaram brancos, 0,19. Os alunos possuem em média 11 anos de idade, o que é esperado para alunos do 6º ano do ensino fundamental. Há um elevado percentual de novatos, cerca de 70% da amostra, o que pode ser uma decorrência da municipalização do ensino fundamental, levando à migração de alunos de escolas estaduais para municipais.

No que diz respeito às características de saúde do aluno e do seu responsável, todas as avaliações foram feitas por parte do responsável pela vida escolar do estudante. De acordo com o declarado pelos responsáveis, um percentual de 78% dos alunos não possui nenhum tipo de doença.<sup>3</sup> A avaliação do estado geral da saúde do responsável indica que em média 64% deles conceituam sua saúde como boa ou muito boa. Por se tratar de

<sup>3</sup> Especificamente a pesquisa da Fundaj (2013) faz a seguinte pergunta ao responsável pelo aluno: "O aluno possui algum tipo de doença?" (múltipla escolha). As respostas admitem as seguintes opções: [01] Não; [02] Sim, asma/doenças respiratórias; [03] Sim, alergias; [04] Sim, dores de cabeça; [05] Sim, diabetes; [06] Sim, hipertensão; [07] Sim, doença cardíaca; [08] Sim, doença de pele; [09] Sim, dores musculares; [10] Sim, doença no estômago/intestino; [11] Sim, doença psiquiátrica; [12] Sim, incontinência urinária; [13] Sim, ansiedade/ depressão; [14] Sim, outra (especificar); [99] NS/NR.

uma avaliação subjetiva, esta última variável está sendo incorporada nas estimações com o objetivo de ponderar para o grau individual de otimismo (ou pessimismo) quanto ao juízo crítico que faz tanto de sua saúde, como também da saúde do aluno pelo qual é responsável.

A maioria dos alunos (72%) afirma que estuda as matérias da escola no mínimo três dias por semana. Mais da metade (60%) declara que o professor de matemática sempre o elogia quando tira nota boa ou faz a tarefa. A participação dos alunos da amostra em missas ou cultos é expressiva; 85% declaram frequentar a igreja às vezes ou sempre. Porém, menos de 1/4 deles tem o hábito de frequentar clubes ou academias de sua vizinhança. Um percentual de 58% dos alunos frequenta a casa de pelo menos 1 amigo da sala.

No que diz respeito às demais características do principal responsável pela vida escolar do aluno, 82% deles são pais naturais ou adotivos, 87% é do sexo feminino, 17% se declaram brancos e possuem em média 38 anos de idade. A escolaridade média desses responsáveis é de quase nove anos de estudos, o que corresponde ao ensino fundamental completo e a grande maioria (85%) costuma conferir o boletim do aluno. Um percentual de 45% desses responsáveis não é casado, nem se encontra em nenhum tipo de união estável. Além disso, 62% deles recebem algum tipo de benefício social do Governo.

**Tabela1 - Estatísticas descritivas das variáveis utilizadas nos modelos**

	<b>Média</b>	<b>Desvio-padrão</b>
Nota de matemática no final do ano $y_{i,t}$	41,15	15,66
<b>Características individuais (<math>X</math>)</b>		
Nota de matemática no início do ano	41,92	15,88
Saúde	0,78	0,41
Sexo masculino	0,45	0,5
Raça branca	0,19	0,39
Idade	11,19	0,87
Novato	0,73	0,45
Dedicação ao estudo	2,57	1,52
Elogio do professor	1,49	0,66
Frequenta casa do amigo	1,06	1,25
Religiosidade	1,75	0,7
Clubes, academias	0,23	0,42
Saúde percebida do responsável	0,64	0,48
Sexo masculino (pais ou responsável)	0,13	0,34
Raça branca (pais ou responsável)	0,17	0,38
Idade (pais ou responsável)	38,48	8,16
Escolaridade (pais ou responsável)	8,85	3,44
Situação conjugal (pais ou responsável)	0,55	0,5
Parentesco (pais ou responsável) com o aluno	0,82	0,38
Beneficiário de programa social (pais ou responsável)	0,62	0,49
Boletim (pais ou responsável)	1,2	0,51
<b>Característica dos pares (GX) - Valores médios ou agregados das variáveis de controle do grupo de amigos diretos do aluno <math>i</math></b>		
Saúde - amigos (agregado)	1,96	1,45
Nota de matemática no início do ano - amigos	42,33	12,2
Sexo masculino - amigos	0,43	0,45
Raça branca - amigos	0,19	0,28
Religiosidade - amigos	1,76	0,49
Novato - amigos	0,72	0,39
Clubes, academias - amigos	0,23	0,29
Dedicação ao estudo - amigos	2,61	1,11
Elogio do professor - amigos	1,49	0,46
Escolaridade (pais ou responsável) - amigos	8,87	2,63
Sexo masculino (pais ou responsável) - amigos	0,14	0,25
Raça branca (pais ou responsável) - amigos	0,18	0,28
Idade (pais ou responsável) - amigos	38,44	5,97
Situação conjugal (pais ou responsável) - amigos	0,53	0,36
Parentesco (pais ou responsável) com o aluno - amigos	0,83	0,26
Beneficiário de programa social (pais ou responsável) - amigos	0,6	0,36
Boletim (pais ou responsável) - amigos	1,2	0,37
<b>Características da rede/ turma</b>		
Nº de redes/ turmas: 139 turmas		
Nº de alunos por rede/ turma	13,25	5,59
Nº de observações: 1.854 alunos		

Fonte: elaboração própria com base na Pesquisa (Fundaj 2013).

#### 4. Modelo Empírico e Estratégias de Estimação

Nesta seção apresenta-se o modelo que embasa as estimações empíricas do artigo, descrevendo-se em que medida a base de dados utilizada ajuda a minimizar os diversos problemas de endogeneidade relacionados à identificação do efeito da saúde dos pares sobre o desempenho escolar individual. A equação empírica do presente artigo é uma extensão do modelo *linear-in-means* de interação social proposto por Manski (1993), em que o *outcome* individual é uma função da média do *outcome* do grupo ao qual o indivíduo pertence. O presente modelo, contudo, avança em relação ao de Manski (1993) na medida em que permite para heterogeneidade intragrupo. Nos estudos mais tradicionais de *peer effects*, os indivíduos interagem em grupos homogêneos formados, por exemplo, por todos os alunos da turma, sem distinguir se o indivíduo interage em grupos específicos dentro desse grande grupo. Como a base de dados utilizada neste artigo dispõe da rede de relacionamentos do aluno dentro de sua turma, é possível identificar os grupos de referência de cada aluno, o qual é formado pelos amigos indicados por ele. Esses grupos não somente podem ser distintos entre si, como também podem se sobrepor, já que é possível que dois alunos tenham um ou mais amigos em comum e é justamente esta arquitetura de redes diretas de amizades composta por grupos de referência heterogêneos que permite a identificação do efeito dos pares.

Seja  $y_i$  uma variável que expressa o desempenho acadêmico do aluno  $i$ , como por exemplo, a sua nota em testes escolares e  $SAUDE_j$  que corresponde a uma *dummy* assumindo o valor 1 quando o amigo  $j$  não possui nenhum tipo de doença relatada pelo seu responsável e 0 para o caso em ele é acometido por alguma enfermidade. A influência da saúde dos pares sobre o desempenho individual pode ser denotada por um modelo do tipo SLX (*Spatial Lag of X*) com *lag* espacial incidindo sobre o vetor **SAUDE** de dimensão  $n \times 1$ , conforme a Equação (1):

$$y_{i,k} = \mu y_i^0 + \sum_{j=1}^n \lambda g_{ij,k} SAUDE_{j,k} + \theta_{i,k}(\mathbf{x}) + \phi_\zeta \zeta_{i,k} + \phi_\eta \eta_k + \varepsilon_{i,k} \quad (1)$$

$$i = 1, \dots, 1.854; \quad k = 1, \dots, 139$$

Em que  $\lambda$  é o coeficiente que captura o efeito da saúde dos amigos do aluno sobre o seu rendimento escolar;  $y_i^0 \geq 0$  representa o *background* educacional inicial do aluno, podendo ser interpretado como uma condição inata para aprendizagem e  $g_{ij}$  denotam as conexões dentro da rede. Existe

uma conexão ativa, ou uma relação de amizade, dentro da rede quando  $g_{ij} = 1$  e não ativa quando  $g_{ij} = 0$ , por convenção  $g_{ii} = 0$ . As variáveis  $\zeta_{i,k}$  e  $\eta_k$  representam as heterogeneidades não observáveis individuais e da turma, respectivamente. O componente  $\theta_i(\mathbf{x})$  introduz a heterogeneidade exógena que captura as diferenças observáveis entre os indivíduos. Exemplos seriam sexo, raça, idade, *background* familiar, além de algumas características dos amigos diretos, tais como média da escolaridade dos pais, da composição sociodemográfica da turma, dentre outros. O componente  $\theta_i(\mathbf{x})$  se expressa conforme a Equação (2) e a definição de suas variáveis apresenta-se no Quadro A1 do apêndice.

$$\theta_i(\mathbf{x}) = \sum_{m=1}^M \beta_m x_i^m + \frac{1}{g_i} \sum_{m=1}^M \sum_{j=1}^n \gamma_m g_{ij} x_j^m \quad (2)$$

Em notação matricial (1) denota-se por:

$$\mathbf{y} = \mu \mathbf{y}^0 + \lambda \mathbf{GSAUDE} + \theta(\mathbf{x}) + \phi_\eta \boldsymbol{\eta} + \phi_\zeta \boldsymbol{\zeta} + \boldsymbol{\varepsilon} \quad (3)$$

Onde:  $\theta(\mathbf{x}) = \beta \mathbf{X} + \gamma \mathbf{GX}$

Em que  $\mathbf{G}$  é construída de maneira a formar uma matriz diagonal em bloco, onde a matriz de interação de cada sala de aula  $\mathbf{g}_k$ ,  $k = 1, 2, \dots, 139$ , forma um bloco específico.<sup>4</sup> Como resultado, os alunos que pertencem a uma determinada rede ou turma  $\mathbf{g}_k$  não se conectam a estudantes integrantes de outra rede  $\mathbf{g}'_k$  e o número total de alunos corresponde, portanto, a soma deles em cada rede, assim:  $n = \sum_{k=1}^{139} n_k$ .

Estimativas de mínimos quadrados ordinários (*Ordinary Least Squares* - OLS) são conduzidas na estimação do modelo empírico (3). As propriedades ótimas dos estimadores OLS, contudo, só são obtidas se, e somente se, o vetor saúde dos amigos e a matriz de amizades não estiverem corre-

<sup>4</sup> Como ilustração considere duas matrizes de relacionamentos para duas turmas hipotéticas cada uma com três alunos,  $\mathbf{g}_1$  e  $\mathbf{g}_2$ . A matriz conjunta diagonal  $\mathbf{G}$ , denota-se por:

$$\mathbf{g}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \mathbf{g}_2 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \mathbf{G} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

lacionados com o erro da regressão. Logo, o requerimento da exogeneidade das matrizes de relacionamentos  $\mathbf{G}$  e  $\mathbf{SAUDE}$  é uma condição necessária para obtenção de estimativas consistentes.

As estratégias de identificação propostas neste artigo se basearão nas hipóteses de que, uma vez condicionando-se o erro da regressão aos efeitos fixos da rede ( $\eta_k$ ) e individual ( $\zeta_{i,k}$ ), esses requerimentos são atendidos. *Dummies* por turma/rede é a estratégia adotada para o controle de efeito fixo da rede. Este é o procedimento usualmente adotado na literatura, conforme visto em Patacchini e Venanzoni (2014), Calvo-Armengol, Patacchini e Zenou (2009), Bramoullé, Djebbari e Fortin (2009) e Lee (2007). Esta mesma estratégia, contudo, não pode ser adotada para o controle do efeito fixo individual, uma vez que não se dispõe de dados em painel. No entanto, como a variável de *outcome* é observada em dois períodos, início e final do ano letivo, a estratégia de controle do efeito fixo individual se baseará em duas suposições, devidamente justificadas mais adiante: (i) uso da nota inicial do aluno e (ii) suposição de que quando as redes são suficientemente pequenas, o próprio controle de efeito fixo da turma é uma boa aproximação para também capturar os atributos individuais não observados.

No que diz respeito à primeira suposição, Ding e Lehrer (2007) argumentam que a nota inicial do aluno seria uma estatística suficiente para capturar uma variedade de influências que podem confundir a análise e incluem todo o histórico observável e não observável do *background* familiar, escolar e da comunidade do aluno. Os autores assumem, por hipótese, que a nota inicial do estudante obedece a um processo de Markov e, sendo assim, os fatores observáveis e não observáveis anteriores a  $t-1$  se ajustam a uma mesma taxa, de forma que nenhuma dessas características deixaria de ser representada por  $y_{i,t-1}$ . A suposição desses autores permite, portanto, que a nota inicial do aluno funcione como uma espécie de efeito fixo do estudante, já que traz embutido em sua estatística aqueles componentes não observáveis (como esforço, habilidade, etc.) invariantes ao longo do ano letivo.<sup>5</sup> Nesse sentido,  $y_i^0$  representa uma *proxy* para o *background* educacional inicial do estudante.

<sup>5</sup> Considere dois modelos com estrutura de regressão para as notas do início e final do ano de um aluno  $i$ :

$$y_i^0 \equiv y_{i1} = \underbrace{\alpha_1 X_{i1}}_{\text{Observável}} + \underbrace{\beta_1 u_i + \varepsilon_{i1}}_{\text{Não-observável}} \quad e \quad y_{i2} = \underbrace{\alpha_2 X_{i2}}_{\text{Observável}} + \underbrace{\beta_2 u_i + \varepsilon_{i2}}_{\text{Não-observável}} .$$

onde  $u_i$  é um componente não observável invariante no ano letivo. De acordo com Boardman e Murnane (1979), se  $\alpha_2 = \theta \alpha_1$  e  $\beta_2 = \theta \beta_1$  e isto implica que os efeitos das variáveis  $X$  e  $u$  mudam a uma mesma taxa constante  $\theta$  entre  $t-1$  e  $t$ . Sob tais condições, a inclusão de no modelo empírico (1) permite o controle desta condição inicial fixa para cada aluno.

Com o objetivo de testar a validade das estratégias de identificação aqui propostas, dois testes de robustez são operacionalizados na próxima seção de resultados para checar a exogeneidade da matriz **GSAUDE** face ao controle de efeitos fixos das redes (*dummies* por turma) e do indivíduo (nota inicial).

## 5. Resultados

Os resultados das estimações do modelo empírico (1) são apresentados nesta seção. Uma variedade de modelos é testada utilizando a rede indireta de amigos,<sup>6</sup> ou seja, o caso em que  $\mathbf{G}$  é simétrica. Estimativas com a matriz não simétrica também foram operacionalizadas, porém, em virtude da grande quantidade de zeros, problemas computacionais inviabilizaram as estimações.<sup>7</sup>

Na Tabela 2 encontram-se reportados os resultados estimados para diferentes especificações da Equação (1). Em cada especificação - colunas (1) a (3) - as estimativas são operacionalizadas por meio de ordem crescente de inclusão de regressores. O coeficiente  $\lambda$  representa o quanto do desempenho médio de matemática de um determinado aluno está sendo impactado pela mudança na proporção de seus amigos de turma que se encontra saudável. Retomando a expressão da Equação (1), o efeito marginal deve ser interpretado como a mudança esperada no desempenho do aluno  $i$  a partir de um incremento marginal na fração de seus amigos saudáveis. Como se trata de um somatório, temos então uma proporção agregada dos amigos saudáveis, cujo valor médio para o total da amostra é de 1,96, conforme apresentado na Tabela 1. Traduzindo esse efeito em termos de desvio padrão, o coeficiente padronizado de 0,086 do modelo mais completo (coluna 3 da Tabela 2) indica que um aumento de 1 desvio padrão (d.p.) na proporção de amigos saudáveis, levará a um acréscimo de aproximadamente 9% do d.p. da nota média de matemática do aluno ao final do ano letivo. Em termos relativos, a saúde dos amigos destaca-se como o segundo preditor mais importante do desempenho do aluno ao final do ano letivo, ficando atrás apenas do próprio desempenho no início

<sup>6</sup> No caso das redes indiretas ( $\mathbf{G}$  é simétrica)  $g_{ij} = g_{ji} = 1$ , porém, se a rede for considerada de maneira direta, então  $g_{ij} = 1$  e  $g_{ji} = 0$ .

<sup>7</sup> Alguns estudos demonstram que os resultados do efeito dos pares não se alteram quanto à simetria da matriz  $\mathbf{G}$ . Vejam-se os estudos de Patacchini e Venanzoni (2014) e Calvo-Armengol, Patacchini e Zenou (2009).

do ano. A influência da saúde dos amigos é superior, inclusive, ao efeito da nota de matemática do início do ano letivo desses amigos, cujo impacto é de 6% para os modelos (1) e (2) e nulo para o modelo (3) em que há um controle para o efeito fixo das turmas. Quando se observa o impacto da saúde do próprio aluno, não se verifica significância estatística em nenhum dos modelos estimados, possivelmente esse efeito está sendo absorvido pelas demais variáveis de controles do indivíduo, sobretudo sua nota inicial que, como já argumentado, funciona como uma espécie de “efeito fixo” do estudante e carrega a influência de toda a sua história anterior.

As demais variáveis de controle utilizadas nas três especificações também repercutem alguns achados comumente observados na literatura em que a escolaridade dos pais amplia a performance do estudante; as famílias beneficiárias de programas sociais, tais como o Bolsa-Família, possuem filhos com piores resultados escolares; alunos mais velhos possuem desempenho mais baixo, e aqueles cujos pais checam seus boletins escolares tendem a apresentar maiores notas.

Para observar a robustez do efeito pretendido frente a diferentes especificações, foram estimados três modelos distintos - colunas (1) a (3) da Tabela 2. O coeficiente positivo e significativo  $\lambda$  demonstra que o resultado acadêmico individual está diretamente correlacionado com o estado de saúde de seu grupo de amigos de sala de aula, em qualquer das especificações adotadas. A força desta correlação não se altera entre os modelos (1) e (2), isto é, quando se passa a controlar para a saúde do próprio aluno e também do seu responsável. Entretanto, quando se incluem as *dummies* de efeito fixo da rede, o coeficiente cresce e torna-se mais preciso, revelando como o controle dos atributos comuns ao grupo (contexto) ajuda a capturar as heterogeneidades não observáveis presentes na análise, as quais, se deixadas de fora, podem enviesar os coeficientes.

Em outras palavras, os resultados encontrados indicam que quando os amigos de sala de aula do estudante gozam de uma boa saúde, isto pode melhorar o rendimento do aluno em questão. É possível que quando os colegas adoecem com maior regularidade, o desempenho do aluno seja prejudicado em razão da redução das externalidades do conhecimento que deixam de ser geradas pelos debates e questionamentos dos estudantes que faltam às aulas por estarem doentes. Esse é um mecanismo possível que explica como a saúde do amigo pode interferir nos resultados escolares individuais. Mas podem existir outros. Por exemplo, amigos doentes podem interferir na concentração, motivação e dedicação ao estudo.

**Tabela 2 - Estimativas para Equação (1) - Variável dependente é a nota no final do ano**

	1	2	3
<b>Saúde dos amigos (<math>\lambda</math>)</b>	0,071***	0,071***	0,086***
<b>[Estatística de teste]</b>	[3,513]	[3,505]	[3,759]
	0,447***	0,448***	0,411***
Nota de matemática no início do ano	[21,438]	[21,421]	[18,748]
	0,011	0,011	0,014
Sexo masculino	[0,305]	[0,305]	[0,415]
	-0,040**	-0,040**	-0,031
Raça branca	[-1,936]	[-1,936]	[-1,516]
	-0,066***	-0,066***	-0,087***
Idade	[-3,170]	[-3,164]	[-3,841]
	0,003	0,003	0,014
Novato	[0,123]	[0,133]	[0,506]
	-0,008	-0,008	0,004
Dedicação ao estudo	[-0,394]	[-0,397]	[0,178]
	-0,027	-0,027	-0,032
Elogio do professor	[-1,332]	[-1,325]	[-1,572]
	-0,031	-0,031	-0,028
Frequenta casa do amigo	[-1,532]	[-1,517]	[-1,325]
	-0,012	-0,012	-0,011
Clubes, academias	[-0,591]	[-0,588]	[-0,570]
	-0,002	-0,002	0,006
Religiosidade	[-0,108]	[-0,107]	[0,285]
	0,039*	0,039*	0,021
Escolaridade (pais ou responsável)	[1,865]	[1,848]	[0,968]
	0,023	0,023	0,002
Sexo masculino (pais ou responsável)	[1,099]	[1,086]	[0,100]
	0,02	0,02	0,027
Raça branca (pais ou responsável)	[0,987]	[0,987]	[1,314]
	0,040*	0,041*	0,03
Idade (pais ou responsável)	[1,839]	[1,825]	[1,301]
	0,028	0,028	0,014
Situação conjugal (pais ou responsável)	[1,377]	[1,378]	[0,713]
	0,019	0,019	0,011
Parentesco (pais ou responsável) com o aluno	[0,899]	[0,894]	[0,509]
	-0,036*	-0,036*	-0,022
Beneficiário de programa social (pais ou responsável)	[-1,685]	[-1,674]	[-1,006]
	-0,051**	-0,051**	-0,041*
Boletim (pais ou responsável)	[-2,552]	[-2,540]	[-2,014]
	0,059***	0,058***	0,016
Nota de matemática dos amigos no início do ano	[2,771]	[2,747]	[0,640]
		-0,002	0,013
Saúde do aluno		[-0,100]	[0,650]
		0,004	0,008
Saúde percebida do responsável		[0,177]	[0,388]
Demais características dos amigos	Sim	Sim	Sim
Efeito fixo da rede	Não	Não	Sim
R2 ajustado	0,294	0,293	0,345
Nº de observações	1.854	1.854	1.854
Nº de redes	139	139	139

Fonte: elaboração própria com base na Pesquisa (Fundaj, 2013).

Notas: Essa tabela apresenta a estimação de mínimos quadrados ordinários da Equação (1), com ordem crescente de inclusão de regressores e coeficientes padronizados. \*\*\*, \*\*, \* representam as significâncias estatísticas a menos de 1%, 5% e 10%, respectivamente. As estimativas completas incluindo as demais características dos amigos encontram-se nas Tabelas A1 a A3 do apêndice.

### 5.1. Testes de Robustez para as Estratégias de Identificação

Existem duas questões centrais que podem colocar em xeque a estratégia de identificação dos modelos apresentados na seção anterior. A primeira diz respeito ao fato de que os indivíduos costumam escolher seus amigos por homofilia, fazendo amizades com outro que tenha um status de saúde semelhante ao seu. A segunda questão relaciona-se às influências comuns do contexto ao qual o grupo como um todo está inserido. Por exemplo, a escola pode estar inserida numa comunidade mais pobre e carente de infraestrutura social, como saneamento e coleta de lixo, o que pode aumentar a propensão a doenças no seu corpo discente, independentemente dos laços de amizade que possam existir entre os alunos. Se essas questões não forem devidamente incorporadas ao modelo, então a suposição de exogeneidade do termo de erro não fica garantida.

O modelo de identificação descrito pela Equação (1) baseia-se nos pressupostos de que a grande lista de controles inserida no modelo é capaz de capturar as características que conduzem à classificação de indivíduos em grupos (idade, gênero, raça, educação e ocupação dos pais, etc.) e que quaisquer características não observadas remanescentes são capturadas ao nível da rede. Os efeitos fixos de rede atuam, portanto, como um remédio para o viés de autosseleção, em uma mesma rede, de indivíduos com características similares não observadas. O pressuposto subjacente é que tais características não observadas são comuns a todos indivíduos dentro de cada rede, o que é uma suposição razoável em nosso estudo de caso em que as redes são bastante pequenas.

No entanto, essa estratégia falha se houver atributos não observáveis no nível do aluno que afetem tanto o desempenho escolar individual quanto a probabilidade de formar links de amizade,  $\zeta_{i,k}$  ou seja, se a variável não observada for não degenerada e  $\phi_\zeta \neq 0$  da Eq.(1). Patacchini e Venanzoni (2014) testam este tipo de endogeneidade baseando-se no trabalho de Goldsmith-Pinkham e Imbens (2013). Considere-se novamente o modelo empírico (1):

$$y_{i,k} = \mu y_i^0 + \sum_{j=1}^N \lambda g_{ij,k} SAUDE_{j,k} + \theta_{i,k}(\mathbf{x}) + \underbrace{\phi_\zeta \zeta_{i,k} + \phi_\eta \eta_k + \varepsilon_{i,k}}_{e_{i,k}} \quad (1')$$

Para entender o problema, suponha agora que exista um modelo de formação da rede de amizades, em que as variáveis que explicam a ligação entre dois estudantes  $i$  e  $j$  pertencentes a uma rede  $k$  ( $g_{ij,k}$ ), são as distâncias entre eles em termos das características observáveis  $X$  e não observáveis  $\zeta$ :

$$g_{ij,k} = \alpha + \sum_{m=1}^M \psi_m |x_{i,k}^m - x_{j,k}^m| + \varphi_\zeta |\zeta_{i,k} - \zeta_{j,k}| + \varphi_\eta \eta_k + u_{ij,k} \quad (4)$$

onde a variável dependente é uma *dummy* que assume o valor 1, se existe uma conexão entre  $i$  e  $j$ , e 0 no caso contrário. Um teste para a presença de heterogeneidades não observáveis ao nível do aluno, que afetem tanto o seu desempenho quanto a formação da rede de amigos, consiste em checar se há correlação significativa entre os resíduos da Equação (1') e a probabilidade de formar amizade. Sob a hipótese de endogeneidade da matriz  $\mathbf{G}$ , quando dois indivíduos  $i$  e  $j$  são amigos entre si ( $g_{ij}=1$ ), então as distâncias entre suas características observadas tendem a ser relativamente pequenas em virtude da homofilia. Conforme demonstrado por Goldsmith-Pinkham e Imbens (2013), temos que se  $E[|\zeta_i - \zeta_j|/|X_i - X_j| = x, g_{ij} = 1]$  é decrescente em  $x$ , então  $E[|\varepsilon_i - \varepsilon_j|/|X_i - X_j| = x, g_{ij} = 1]$  também será decrescente em  $x$ , ou seja, as distâncias dos atributos não observados  $\xi$  entre os estudantes  $i$  e  $j$  seguem o mesmo padrão daqueles atributos observados  $X$  e, como um resultado, é possível se utilizar dos resíduos estimados para detectar a influência de heterogeneidades não observadas. É factível, então, substituir  $|\zeta_{i,k} - \zeta_{j,k}|$  em (4) por  $|\widehat{\varepsilon}_{i,k} - \widehat{\varepsilon}_{j,k}|$  de (1') e estimar o modelo (4). Evidência de exogeneidade para a matriz  $\mathbf{G}$  seria encontrar  $\widehat{\phi} = 0$ . Os resultados deste teste estão reportados na Tabela 3, e o que se constata é que, quando não há controle do efeito fixo da rede/turma  $\eta_k$ , existe uma correlação significativa entre a probabilidade de formar amizade e as similaridades não observáveis entre os pares. No entanto, quando *dummies* por turma são inseridas no modelo, esta correlação significativa desaparece, como se vê na segunda linha da Tabela 3. Logo, condicionando-se a um amplo conjunto de controles, as características dos pares e aos efeitos fixos das redes, não há evidências de outros atributos individuais não observáveis que possam enviesar os resultados aqui encontrados.

Tabela 3 -Teste de robustez – Estimativas para Equação (4)

	OLS
Diferença entre os resíduos ( $\hat{\phi}$ ) sem controle de efeito fixo ( <i>p-value</i> )	0,00001 -0,0529
Diferença entre os resíduos ( $\hat{\phi}$ ) com controle de efeito fixo ( <i>p-value</i> )	0,00001 -0,504

Fonte: elaboração própria com base na Pesquisa (Fundaj 2013).

Notas: As observações incluem todas as combinações  $ij$  entre os pares da amostra utilizada com  $n = 1.854$ , o que gera um total de observações para a estimação de (4) de  $[n*(n-1)/2 = 1.717.731 \text{ observações}]$ . As variáveis de controle são as mesmas incluídas na Equação (1).

Um segundo teste de robustez procura identificar a existência de endogeneidade em relação ao estado de saúde do aluno e seu amigo, uma vez que é plausível supor que os alunos apresentem uma propensão a se agrupar por homofilia, inclusive nos aspectos relacionados à saúde. Nesse sentido, foi estimado um modelo autorregressivo espacial para a variável do estado de saúde do aluno, conforme equação que se segue.

$$SAUDE_{i,k} = \delta y_i^0 + \sum_{j=1}^n \rho g_{ij,k} SAUDE_{j,k} + \theta_{i,k}(x) + \tau_{\zeta} \zeta_{i,k} + \tau_{\eta} \eta_k + \epsilon_{i,k} \quad (5)$$

Os resultados se encontram reportados na Tabela 4, e como se observa, não há existência de correlação espacial estatisticamente significativa para a variável “saúde do aluno”, mostrando que não há uma tendência sistemática de alunos saudáveis se agruparem com amigos saudáveis, nem dos não saudáveis se associarem a colegas também não saudáveis.

Tabela 4- Teste de robustez – Estimativas para Equação (5)

Estimativas $\rho$ para as variáveis de saúde	SAR
Saúde do aluno ( <i>p-value</i> )	-0,009 0,18

Fonte: elaboração própria com base na Pesquisa (Fundaj 2013).

## 6. Considerações Finais

O objetivo maior deste artigo é o de contribuir para o debate que trata da relação entre saúde e desempenho educacional, via utilização de rede de amizades. Dispondo-se dos dados de uma pesquisa da Fundação Joaquim Nabuco no ano de 2013, a estratégia de identificação do efeito da saúde do amigo do aluno sobre a sua performance escolar baseia-se na arquitetura de redes diretas de amizades composta por grupos de referência heterogêneos, na utilização de uma avaliação longitudinal do desempenho em duas provas de matemática e no controle do efeito fixo da rede.

Os resultados mostram que a saúde do amigo é muito importante para a explicação dos resultados escolares. Ou seja, ter amigo saudável afeta a nota do estudante de forma direta. Em outras palavras, quando o estudante se relaciona com outros colegas saudáveis, ele tende a exibir um melhor desempenho acadêmico. Os resultados também revelam que quando se controla o viés de variáveis omitidas (como por exemplo, diferentes maneiras de cuidar da própria saúde) o impacto da saúde do amigo sobre o desempenho é ainda maior. Os testes de robustez revelam que, condicionando-se a um amplo conjunto de controles, às características dos pares e aos efeitos fixos das redes, não há evidências de outros atributos individuais não observáveis que possam enviesar os resultados aqui encontrados. Evidências dessa natureza podem auxiliar na elaboração de critérios para organização dos alunos em grupos. Se o aluno saudável é capaz de melhorar o desempenho do amigo que adoce com maior frequência, então deve haver uma distribuição heterogênea e equilibrada dos estudantes de acordo com o seu status de saúde. O desafio é encontrar uma composição de pares que eleve o aprendizado agregado.

A formulação de uma política pública que atenda a essas novas evidências empíricas deve ter como ponto de partida a compreensão de que somente o acesso e a permanência na escola não garantem um futuro com melhor qualidade de vida à geração de estudantes. É preciso também assegurar-lhes aprendizagem de qualidade. Destaca-se que muito ainda pode ser feito no intuito de entender os fatores condicionantes/determinantes do aprendizado/desempenho dos alunos. Sobretudo no que tange ao desenho das bases de dados que possibilitem estudos mais acurados com a inclusão de outras questões que levem em consideração que a evolução dos alunos é multidimensional e que aprendizado também envolve o domínio de competências de natureza física, afetiva e comportamental.

## Referências

- Badev, A. 2013. "Discrete Games in Endogenous Networks: Theory and Policy". [https://repository.upenn.edu/cgi/viewcontent.cgi?referer=https://www.google.com.br/&httpsredir=1&article=1044&context=psc\\_working\\_papers](https://repository.upenn.edu/cgi/viewcontent.cgi?referer=https://www.google.com.br/&httpsredir=1&article=1044&context=psc_working_papers). Acesso em: 27/2/2018.
- Ballester, Corolario, Antoni Calvó-Armengol e Yves Zenou. 2006. "Who's Who in Networks. Wanted: The Key Player". *Econometrica* 74(5): 1403–1417, September.
- Berndt, Thomas J., e Perry, T. Bridgett. 1986. "Children's Perceptions of Friendships as Supportive Relationships". *Developmental Psychology* 22(5): 640–648, September.
- Boardman, Anthony E., e Richard J Murnane. 1979. "Using Panel Data to Improve Estimates of the Determinants of Educational Achievement". *Sociology of Education* 52(2): 113-121, April.
- Bramoullé, Yann, Habiba Djebbari, e Bernard Fortin. 2009. "Identification of Peer Effects through Social Networks". *Journal of Econometrics* 150: 41-55.
- Calvó-Armengol, Antoni, Eleonora Patacchini, e Yves Zenou. 2009. "Peer Effects and Social Networks in Education". *The Review of Economic Studies* 76(4): 1239-1267, October.
- Christakis, Nicholas A. e James H. Fowler. 2007. "The Spread of Obesity in a Large Social Network Over 32 Years". *The New England Journal of Medicine* 357: 370-379, July.
- Cohen-Cole, Ethan and Jason M. Fletcher. 2008. "Is Obesity Contagious? Social Networks vs. Environmental Factors in the Obesity Epidemic". *Journal of Health Economics* 27(5): 1382-1387, September.
- Currie, Janet. 2009. "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development". *Journal of Economic Literature* 47(1): 87–122.
- De Giorgi, Giacomo D., Michele Pellizzari, and William Gui Woolston. 2010. "Class Size and Class Heterogeneity". NBER Working Paper 16405, September. <http://www.nber.org/papers/w16405>.
- Ding, Weili, e Steven F. Lehrer. 2007. "Do Peers Affect Student Achievement in China's Secondary Schools?" *The Review of Economics Statistics* 89(2): 300-312, May.
- Epstein, J. L. 1983. "The influence of friends on achievement and affective outcomes." In *Friends in School: Patterns of Selection and Influence in Secondary Schools*, edited by J. L. Epstein & N. Karweit, (New York: Academic Press), 177–200.
- Fletcher, Jason. M. 2010. "Social Interactions and Smoking: Evidence Using Multiple Student Cohorts, Instrumental Variables, and School Fixed Effects." *Health Economics* 19(4): 466-484, April.
- \_\_\_\_\_. 2012. "Peer Influences on Adolescent Alcohol Consumption: Evidence Using an Instrumental Variables/Fixed Effect Approach." *Journal of Population Economics* 25(4): 1265-1286.
- Fundação Joaquim Nabuco - Fundaj. Coordenação de Estudos Econômicos e Populacionais. 2013. "Acompanhamento Longitudinal do Desempenho Escolar de Alunos da Rede Pública de Ensino Fundamental do Recife". Recife.
- Goldsmith-Pinkham, Paul, and Guido W. Imbens. 2013. "Social Networks and the Identification of Peer Effects". *Journal of Business and Economic Statistics* 31(3): 253–264.
- Harris, Judith. R. 1995. "Where is The Child's Environment? A Group Socialization Theory of Development." *Psychological Review* 102(3): 458-486.
- Lee, Lung-Fei. 2007. "Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors and Fixed Effects." *Journal of Econometrics* 140(2): 333–374.
- Manski, Charles. F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* 60(3): p. 531-542, July.
- Mele, Angelo. 2010. "A Structural Model of Segregation in Social Networks." NET Institute Working Papers 10-16. [http://www.netinst.org/NET\\_Working\\_Papers.html](http://www.netinst.org/NET_Working_Papers.html)

- Patacchini, Eleonora, Edoardo Rainone, e Yves Zenou. 2011. "Dynamic Aspects of Teenage Friendships and Educational Attainment." CEPR Discussion Paper No. DP8223, February. <https://ssrn.com/abstract=1758432>
- Patacchini, Eleonora, e Giuseppe Venanzoni. 2014. "Peer Effects in the Demand for Housing Quality." *Journal of Urban Economics* 83: 6–17.
- Pinto, Cristine. C. X. 2008. "Semiparametric Estimation of Peer Effects." PhD dissertation, University of California, Berkeley.
- Recife, Prefeitura, e Pnud. 2005. "Metodologia de Divisão do Território do Recife Adotada no Atlas Municipal Do Desenvolvimento Humano." In: *Atlas do Desenvolvimento Humano no Recife*.
- Raposo, Isabel. P. A. 2015. "O Papel da Rede de Amizades e da Formação Aleatória de Turmas por Faixa Etária sobre o Desempenho". Tese de Doutorado, Universidade Federal de Pernambuco.
- Schneeweis, Nicole, e Winter-Ebmer, Rudolf. 2007. "Peer Effects in Austrian Schools." *Empirical Economics* 32(2): 387-409.
- Wentzel, Kathryn. R.; Carolyn M. Barry, e Kathryn A. Caldwell. 2004. "Friendships in Middle School: Influences on Motivation and School Adjustment." *Journal of Educational Psychology* 96(2): 195-203.
- Zimmer, Ron. 2003. "A New Twist in the Educational Tracking Debate." *Economics of Education Review*, 22(3): 307-315.

## Apêndice

<b>Quadro A1</b>	
<b>Definições das variáveis utilizadas nos modelos</b>	
Nota de matemática no final do ano $y_{i,t}$	Nota do aluno $i$ na prova de matemática realizada pela Pesquisa Fundaj (2013) ao <b>final</b> do ano letivo. A escala da nota varia de 0 a 100.
<i>Características individuais (X)</i>	
Nota de matemática no início do ano	Nota do aluno $i$ na prova de matemática realizada pela Pesquisa Fundaj (2013) no <b>início</b> do ano letivo. A escala da nota varia de 0 a 100.
Saúde	<i>Dummy</i> igual 1 se o responsável declara que o aluno não possui nenhum tipo de doença e 0 se o responsável declara que o aluno possui alguma doença
Sexo masculino	<i>Dummy</i> igual a 1 se o aluno é do sexo masculino
Raça branca	<i>Dummy</i> igual 1 para os alunos que se declaram brancos e 0 os que se declaram negros, pardos, indígenas ou amarelos
Idade	Idade do aluno
Novato	<i>Dummy</i> igual 1 se o aluno estuda a menos de 1 ano na escola pesquisada
Dedicação ao estudo	Aluno responde à questão "Com que frequência você estuda as matérias da escola": 1=todos os dias da semana, 2=apenas nos dias que tem aula, 3=3 dias por semana, 4=menos de 3 dias por semana, 5=apenas quando tem prova, 6=nunca ou quase nunca
Elogio do professor	Aluno responde à questão "O(A) prof(a) de matemática elogia ou dá parabéns quando você tira boa nota ou faz a tarefa bem feita": 1=sempre ou quase sempre, 2=às vezes, 3=nunca ou quase nunca
Frequenta casa do amigo	Nº de amigos da turma que o aluno costuma frequentar a residência
Religiosidade	Aluno responde à questão "Você costuma ir à igreja/culto": 1=sempre ou quase sempre, 2=às vezes, 3=nunca ou quase nunca
Clubes, academias	<i>Dummy</i> igual a 1 se o aluno frequenta algum clube, centro desportivo ou academia de ginástica no seu bairro
Saúde percebida do responsável	<i>Dummy</i> igual 1 se o responsável considera que o seu estado geral de saúde é bom ou muito bom e 0 se o responsável considera sua saúde regular ou ruim
Sexo masculino (pais ou responsável)	<i>Dummy</i> igual a 1 se um dos pais ou responsável pelo aluno é do sexo masculino
Raça branca (pais ou responsável)	<i>Dummy</i> igual 1 para os pais/ responsáveis que se declaram brancos e 0 os que se declaram negros, pardos, indígenas ou amarelos
Idade (pais ou responsável)	Idade de um dos pais ou responsável pelo aluno
Escolaridade (pais ou responsável)	Pais/ responsáveis respondem à questão: "Qual a série mais elevada concluída com aprovação": 1=1º ano (alfabetização), ..., 9=9º ano (antiga 8ª série); 10=1º ano do ensino médio, ..., 12=3º ano do ensino médio; 13=1º ano universidade, ..., 18=6º ano universidade
Situação conjugal (pais ou responsável)	<i>Dummy</i> igual 1 para os pais/ responsáveis casados, com união estável ou concubinato
Parentesco (pais ou responsável) com o aluno	<i>Dummy</i> igual 1 para pais naturais ou adotivos e 0 para os demais casos (avós, tios, irmãos, padrasto/madrasta, outros)

<b>Quadro A1</b> <b>Definições das variáveis utilizadas nos modelos (Continuação)</b>	
Beneficiário de programa social (pais ou responsável)	<i>Dummy</i> igual 1 se pais/ responsáveis recebem algum auxílio do Governo
Boletim (pais ou responsável)	Pais/responsáveis respondem à questão “Você confere o boletim do aluno”: 1=sempre ou quase sempre, 2=às vezes, 3=nunca ou quase nunca
<i>Característica dos pares (GX)</i>	
Saúde - amigos (agregado)	Valor agregado (somatório) da <i>dummy</i> “Saúde do aluno” dos amigos do estudante <i>i</i>
Nota de matemática no início do ano - amigos	Média da “Nota de matemática no início do ano” dos amigos do estudante <i>i</i>
Sexo masculino - amigos	Média da <i>dummy</i> “Sexo masculino” dos amigos do estudante <i>i</i>
Raça branca - amigos	Média da <i>dummy</i> “Raça branca” dos amigos do estudante <i>i</i>
Religiosidade - amigos	Média da variável “Religiosidade” dos amigos do estudante <i>i</i>
Novato – amigos	Média da <i>dummy</i> “Novato” dos amigos do estudante <i>i</i>
Clubes, academias - amigos	Média da <i>dummy</i> “Clubes, academias” dos amigos do estudante <i>i</i>
Dedicação ao estudo - amigos	Média da variável “Dedicação ao estudo” dos amigos do estudante <i>i</i>
Elogio do professor - amigos	Média da variável “Elogio do professor” dos amigos do estudante <i>i</i>
Escolaridade (pais ou responsável) - amigos	Média da variável “Escolaridade” dos amigos do estudante <i>i</i>
Sexo masculino (pais ou responsável) - amigos	Média da <i>dummy</i> “Sexo masculino (pais ou responsável)” dos amigos do estudante <i>i</i>
Raça branca (pais ou responsável) - amigos	Média da <i>dummy</i> “Raça branca (pais ou responsável)” dos amigos do estudante <i>i</i>
Idade (pais ou responsável) - amigos	Média da variável “Idade (pais ou responsável)” dos amigos do estudante <i>i</i>
Situação conjugal (pais ou responsável) – amigos	Média da <i>dummy</i> “Situação conjugal (pais ou responsável)” dos amigos do estudante <i>i</i>
Parentesco (pais ou responsável) com o aluno – amigos	Média da <i>dummy</i> “Parentesco (pais ou responsável)” dos amigos do estudante <i>i</i>
Beneficiário de programa social (pais ou responsável) – amigos	Média da <i>dummy</i> “Beneficiário de programa social (pais ou responsável)” dos amigos do estudante <i>i</i>
Boletim (pais ou responsável) - amigos	Média da variável “Boletim (pais ou responsável)” dos amigos do estudante <i>i</i>

**Tabela A1 - Estimativas completas para o modelo 1 da Tabela 2**

	Estimate	Standardized	Std. Error	t value	Pr(> t )
Saúde dos amigos	0,761	0,071	0,217	3,513	0
Nota de matemática no início do ano	0,441	0,447	0,021	21,438	< 2e-16
Sexo masculino	0,337	0,011	1,107	0,305	0,761
Raça branca	-1,579	-0,04	0,815	-1,936	0,053
Idade	-1,193	-0,066	0,376	-3,17	0,002
Novato	0,112	0,003	0,904	0,123	0,902
Dedicação ao estudo	-0,082	-0,008	0,209	-0,394	0,694
Elogio do professor	-0,636	-0,027	0,477	-1,332	0,183
Frequente casa do amigo	-0,388	-0,031	0,253	-1,532	0,126
Clubes, academias	-0,442	-0,012	0,748	-0,591	0,555
Religiosidade	-0,049	-0,002	0,45	-0,108	0,914
Escolaridade (pais ou responsável)	0,18	0,039	0,096	1,865	0,062
Sexo masculino (pais ou responsável)	1,043	0,023	0,949	1,099	0,272
Raça branca (pais ou responsável)	0,826	0,02	0,836	0,987	0,324
Idade (pais ou responsável)	0,077	0,04	0,042	1,839	0,066
Situação conjugal (pais ou responsável)	0,868	0,028	0,63	1,377	0,169
Parentesco (pais ou responsável) com o aluno	0,798	0,019	0,887	0,899	0,369
Beneficiário de programa social (pais ou responsável)	-1,167	-0,036	0,693	-1,685	0,092
Boletim (pais ou responsável)	-1,557	-0,051	0,61	-2,552	0,011
Nota de matemática dos amigos no início do ano	0,075	0,059	0,027	2,771	0,006
Sexo masculino - média	0,453	0,013	1,244	0,364	0,716
Raça branca - média	-1,736	-0,031	1,158	-1,499	0,134
Religiosidade - média	-0,423	-0,013	0,644	-0,657	0,512
Novato - média	0,262	0,006	1,017	0,257	0,797
Clubes, academias - média	-0,302	-0,006	1,103	-0,274	0,784
Dedicação ao estudo - média	0,12	0,008	0,293	0,41	0,682
Elogio do professor - média	-0,769	-0,023	0,693	-1,109	0,267
Escolaridade (pais ou responsável) - média	0,292	0,049	0,127	2,303	0,021
Sexo masculino (pais ou responsável) - média	0,937	0,015	1,332	0,704	0,482
Raça branca (pais ou responsável) - média	0,69	0,012	1,13	0,611	0,541
Idade (pais ou responsável) - média	0,16	0,061	0,057	2,801	0,005
Situação conjugal (pais ou responsável) - média	0,131	0,003	0,876	0,15	0,881
Parentesco (pais ou responsável) com o aluno - média	1,154	0,02	1,22	0,946	0,344
Beneficiário de programa social (pais ou responsável) - média	-1,467	-0,034	0,95	-1,545	0,123
Boletim (pais ou responsável) - média	-0,81	-0,019	0,842	-0,961	0,336
Intercepto	23,193	0	6,509	3,563	0

$R^2$  ajustado: 0,294  
 Nº de observações: 1.854  
 Nº de redes: 139

Fonte: elaboração própria com base na Pesquisa Fundaj 2013.

**Tabela A2 - Estimativas completas para o modelo 2 da Tabela 2**

	Estimate	Standardized	Std. Error	t value	Pr(> t )
Saúde dos amigos	0,761	0,071	0,217	3,505	0
Nota de matemática no início do ano	0,441	0,448	0,021	21,421	< 2e-16
Sexo masculino	0,337	0,011	1,108	0,305	0,761
Raça branca	-1,58	-0,04	0,816	-1,936	0,053
Idade	-1,192	-0,066	0,377	-3,164	0,002
Novato	0,121	0,003	0,905	0,133	0,894
Dedicação ao estudo	-0,083	-0,008	0,209	-0,397	0,691
Elogio do professor	-0,633	-0,027	0,478	-1,325	0,185
Frequenta casa do amigo	-0,385	-0,031	0,254	-1,517	0,129
Clubes, academias	-0,44	-0,012	0,749	-0,588	0,557
Religiosidade	-0,048	-0,002	0,451	-0,107	0,915
Escolaridade (pais ou responsável)	0,179	0,039	0,097	1,848	0,065
Sexo masculino (pais ou responsável)	1,036	0,023	0,954	1,086	0,278
Raça branca (pais ou responsável)	0,826	0,02	0,837	0,987	0,324
Idade (pais ou responsável)	0,078	0,041	0,043	1,825	0,068
Situação conjugal (pais ou responsável)	0,868	0,028	0,63	1,378	0,168
Parentesco (pais ou responsável) com o aluno	0,794	0,019	0,888	0,894	0,371
Beneficiário de programa social (pais ou responsável)	-1,162	-0,036	0,694	-1,674	0,094
Boletim (pais ou responsável)	-1,552	-0,051	0,611	-2,54	0,011
Nota de matemática dos amigos no início do ano	0,075	0,058	0,027	2,747	0,006
Saúde do aluno	-0,076	-0,002	0,768	-0,1	0,921
Saúde percebida do responsável	0,121	0,004	0,684	0,177	0,86
Sexo masculino - média	0,446	0,013	1,246	0,358	0,721
Raça branca - média	-1,738	-0,031	1,159	-1,5	0,134
Religiosidade - média	-0,424	-0,013	0,645	-0,658	0,511
Novato - média	0,255	0,006	1,018	0,251	0,802
Clubes, academias - média	-0,297	-0,005	1,104	-0,269	0,788
Dedicação ao estudo - média	0,121	0,009	0,294	0,413	0,679
Elogio do professor - média	-0,773	-0,023	0,694	-1,114	0,265
Escolaridade (pais ou responsável) - média	0,292	0,049	0,127	2,3	0,022
Sexo masculino (pais ou responsável) - média	0,946	0,015	1,333	0,709	0,478
Raça branca (pais ou responsável) - média	0,695	0,012	1,133	0,613	0,54
Idade (pais ou responsável) - média	0,16	0,061	0,057	2,8	0,005
Situação conjugal (pais ou responsável) - média	0,124	0,003	0,878	0,141	0,888
Parentesco (pais ou responsável) com o aluno - média	1,153	0,02	1,221	0,944	0,345
Beneficiário de programa social (pais ou responsável) - média	-1,461	-0,034	0,951	-1,536	0,125
Boletim (pais ou responsável) - média	-0,807	-0,019	0,843	-0,957	0,339
Intercepto	23,102	0	6,59	3,505	0

$R^2$  ajustado: 0,293  
 Nº de observações: 1.854  
 Nº de redes: 139

Fonte: elaboração própria com base na Pesquisa Fundaj 2013.

**Tabela A3 - Estimativas completas para o modelo 3 da Tabela 2**

	Estimate	Standardized	Std. Error	t value	Pr(> t )
Saúde dos amigos	0,924	0,086	0,246	3,759	0
Nota de matemática no início do ano	0,406	0,411	0,022	18,748	< 2e-16
Sexo masculino	0,456	0,014	1,099	0,415	0,678
Raça branca	-1,236	-0,031	0,816	-1,516	0,13
Idade	-1,576	-0,087	0,41	-3,841	0
Novato	0,501	0,014	0,991	0,506	0,613
Dedicação ao estudo	0,038	0,004	0,213	0,178	0,859
Elogio do professor	-0,763	-0,032	0,486	-1,572	0,116
Frequente casa do amigo	-0,347	-0,028	0,262	-1,325	0,185
Clubes, academias	-0,427	-0,011	0,749	-0,57	0,569
Religiosidade	0,128	0,006	0,45	0,285	0,775
Escolaridade (pais ou responsável)	0,095	0,021	0,099	0,968	0,333
Sexo masculino (pais ou responsável)	0,096	0,002	0,959	0,1	0,92
Raça branca (pais ou responsável)	1,112	0,027	0,846	1,314	0,189
Idade (pais ou responsável)	0,057	0,03	0,044	1,301	0,194
Situação conjugal (pais ou responsável)	0,455	0,014	0,639	0,713	0,476
Parentesco (pais ou responsável) com o aluno	0,453	0,011	0,891	0,509	0,611
Beneficiário de programa social (pais ou responsável)	-0,7	-0,022	0,696	-1,006	0,315
Boletim (pais ou responsável)	-1,248	-0,041	0,62	-2,014	0,044
Nota de matemática dos amigos no início do ano	0,02	0,016	0,032	0,64	0,523
Saúde do aluno	0,503	0,013	0,774	0,65	0,516
Saúde percebida do responsável	0,269	0,008	0,692	0,388	0,698
Sexo masculino - média	0,698	0,02	1,251	0,558	0,577
Raça branca - média	-1,225	-0,022	1,24	-0,988	0,323
Religiosidade - média	-0,327	-0,01	0,676	-0,484	0,629
Novato - média	-0,255	-0,006	1,37	-0,186	0,852
Clubes, academias - média	-0,906	-0,017	1,172	-0,773	0,44
Dedicação ao estudo - média	0,294	0,021	0,319	0,921	0,357
Elogio do professor - média	-0,792	-0,023	0,757	-1,046	0,296
Escolaridade (pais ou responsável) - média	0,089	0,015	0,141	0,627	0,531
Sexo masculino (pais ou responsável) - média	0,032	0,001	1,385	0,023	0,981
Raça branca (pais ou responsável) - média	0,865	0,015	1,206	0,717	0,473
Idade (pais ou responsável) - média	0,115	0,044	0,061	1,873	0,061
Situação conjugal (pais ou responsável) - média	-0,172	-0,004	0,951	-0,181	0,856
Parentesco (pais ou responsável) com o aluno - média	-0,006	0	1,295	-0,004	0,996
Beneficiário de programa social (pais ou responsável) - média	-0,884	-0,021	1,018	-0,868	0,385
Boletim (pais ou responsável) - média	-0,128	-0,003	0,922	-0,139	0,889
Intercepto	31,185	0	8,06	3,869	0

 $R^2$  ajustado: 0,345

Nº de observações: 1.854

Nº de redes: 139

Fonte: elaboração própria com base na Pesquisa Fundaj 2013.