

Submeter à área de Economia Regional e Agrícola

**A SAÚDE DOS AMIGOS DE SALA DE AULA INTERFERE NO
DESEMPENHO ESCOLAR NO ALUNO?**

Isabel Pessoa de Arruda Raposo

Pesquisadora da Fundação Joaquim Nabuco – Fundaj/MEC

Doutora em Economia pelo PIMES/UFPE

E-mail: i_raposo@hotmail.com

Endereço Profissional: Rua Dois Irmãos, 92 - Ed. Anexo Anízio Teixeira - Apipucos – Recife/PE CEP: 52071-440

Telefone: (81) 996330181

Michela Barreto Camboim Gonçalves

Pesquisadora da Fundação Joaquim Nabuco – Fundaj/MEC

Doutora em Economia pelo PIMES/UFPE

E-mail: michelabcg@hotmail.com

Endereço Profissional: Rua Dois Irmãos, 92 - Ed. Anexo Anízio Teixeira - Apipucos – Recife/PE CEP: 52071-440

Telefone: 999170335 Fax: (81) 30736499

A SAÚDE DOS AMIGOS DE SALA DE AULA INTERFERE NO DESEMPENHO ESCOLAR NO ALUNO?

JEL: I12, I15, I25.

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 (5ª série) 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 se baseiam: (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 mostraram que os alunos apresentam melhor desempenho acadêmico quando se relacionam com amigos de sala de aula que possuem um bom estado geral de saúde. E mais, a seleção dos alunos ao grupo de amigos na classe tende a ser compensatória, ou seja, alunos com algum tipo de doença tendem a fazer amizades com colegas saudáveis.

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. Furthermore, the sorting of students into groups tend to be compensatory, meaning that unhealthy individuals tend to be friends with health classroom mates.

Key words: Health, School achievement, Friendship network, Peer effects.

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 *et al.* (2011) argumenta 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 lado, 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 a que o professor gaste mais tempo de aula disciplinando a turma em vez de promover o aprendizado das matérias. Sendo assim, este 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 o 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 dos 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. Primeiramente, 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 se associarem 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, independentemente 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 *et al.* (2009) ressaltam que como 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.

Introduzida esta discussão, cabe então apresentar a pergunta principal do artigo: a saúde do grupo de amigos de sala de aula se relaciona em alguma medida com o rendimento escolar do aluno? Motivado por este questionamento, o presente 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, já que utiliza uma base de dados inédita da Fundação Joaquim Nabuco [Fundaj/ MEC, (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 (2014), Mele (2010), Calvo-Armengol *et al.* (2009), Bramoullé *et al.* (2009) e Ballester *et al.* (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 se propõe 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 indivíduos 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 *et al.* (2004), os estudantes são motivados a agir de acordo com o grupo e a adotar comportamentos similares daqueles 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 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)¹, 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 relacionando 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 partir 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

¹ 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.

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 espacialmente distribuídas pelas 18 microrregiões da cidade do Recife². 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 1.855 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 não puderam ser corretamente emparelhadas ou o estudante não nomeou nenhum amigo de sua classe. Outras perdas (cerca de 15%) se deveram à eliminação de observações com informações *missings* para o conjunto das variáveis incluídas no modelo (21). 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 (2014) com 5,4% e Mele (2010) com 5,5%. Apenas Calvó-Armengol *et al.* (2009) com 55% e Bramoullé *et al.* (2009) com 61% trabalharam com um percentual da amostra original semelhante ao deste artigo.

A Tabela 1 apresenta a definição e estatísticas descritivas das variáveis utilizadas nas estimações deste artigo. 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 de 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

² 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 (Prefeitura do Recife *et al.*, 2005).

tipo de doença³. Dentre aqueles apontados como doentes pelos seus responsáveis, 18% dos estudantes são acometidos por doenças físicas (tais como asma, alergias, diabetes, doenças cardíacas, etc.) e apenas 1% por doenças de cunho emocional como depressão, ansiedade, pânico, traumas psicológicos, dentre outras⁴. 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 uma avaliação subjetiva, esta última variável está sendo incorporada nas estimativas 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 ¼ deles tenham 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.

³ 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.

⁴ Para a classificação das doenças emocionais os autores consideraram o item [13] e as respostas do item [14] especificadas como “o aluno é nervoso”, “tem problemas emocionais”, “é hiperativo”, “tem medos, transtornos psicológicos”, “tem Transtorno de Déficit de Atenção e Hiperatividade” e “sofre traumas de bullying”.

Tabela 1
Definição e estatísticas descritivas das variáveis

		continua
Definição das variáveis		Desvio-padrão
		Média
<i>Variável dependente</i>		
Nota de matemática no final do ano $y_{i,t}$	Nota do aluno i na prova de matemática realizada pela Pesquisa Fundaj (2013) ao final do ano letivo. A escala da nota varia de 0 a 100.	41,15 15,66
<i>Características de saúde</i>		
Saúde do aluno	<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	0,78 0,41
Doença física	<i>Dummy</i> igual 1 se responsável declara que o aluno possui algum tipo de doença física e 0 se o responsável declara que o aluno não possui doenças físicas	0,18 0,38
Doença emocional	<i>Dummy</i> igual 1 se responsável declara que o aluno possui algum tipo de doença emocional e 0 se o responsável declara que o aluno não possui doenças emocionais	0,01 0,10
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	0,64 0,48
<i>Características individuais (X)</i>		
Nota de matemática no início do ano y_i^0	Nota do aluno i na prova de matemática realizada pela Pesquisa Fundaj (2013) no início do ano letivo. A escala da nota varia de 0 a 100.	41,92 15,88
Sexo masculino	<i>Dummy</i> igual a 1 se o aluno é do sexo masculino	0,45 0,50
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	0,19 0,39
Idade	Idade do aluno	11,19 0,87
Novato	<i>Dummy</i> igual 1 se o aluno estuda a menos de 1 ano na escola pesquisada	0,73 0,45
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	2,57 1,52
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	1,49 0,66
Frequente casa do amigo	Nº de amigos da turma que o aluno costuma frequentar a residência	1,06 1,25
Religiosidade	Aluno responde à questão “Você costuma ir à igreja/culto”: 1=sempre ou quase sempre, 2=às vezes, 3=nunca ou quase nunca	1,75 0,70
Clubes, academias	<i>Dummy</i> igual a 1 se o aluno frequenta algum clube, centro desportivo ou academia de ginástica no seu bairro	0,23 0,42
Sexo masculino (pais ou responsável)	<i>Dummy</i> igual a 1 se um dos pais ou responsável pelo aluno é do sexo masculino	0,13 0,34
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	0,17 0,38

Tabela 1
Definição e estatísticas descritivas das variáveis

Idade (pais ou responsável)	Idade de um dos pais ou responsável pelo aluno	38,48	continuação 8,16
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	8,85	3,44
Situação conjugal (pais ou responsável)	<i>Dummy</i> igual 1 para os pais/ responsáveis casados, com união estável ou concubinato	0,55	0,50
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)	0,82	0,38
Beneficiário de programa social (pais ou responsável)	<i>Dummy</i> igual 1 se pais/ responsáveis recebem algum auxílio do Governo	0,62	0,49
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	1,20	0,51
<i>Característica dos pares (GX)</i>		Valores médios de todas as variáveis de controle dos estudantes do grupo de amigos diretos do aluno <i>i</i>	
<i>Características da rede/ turma</i>			
Tamanho da rede	Número de alunos na rede/ turma	13,25	5,59

Nº de observações: 1.855 alunos

Nº de redes/ turmas: 139 turmas

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 intra-grupo. 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 deste 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):

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

$$i = 1, \dots, 1.855; \quad k = 1, \dots, 139$$

Em que λ 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 η_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 na Tabela 1.

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

Em notação matricial (1) denota-se por:

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

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

Em que G é construída de maneira a formar uma matriz diagonal em bloco, onde a matriz de interação de cada sala de aula g_k , $k = 1, 2, \dots, 139$, forma um bloco específico⁵. Como resultado, os alunos que pertencem a uma determinada rede ou turma g_k não se conectam a estudantes integrantes de outra rede 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 correlacionados com o erro da regressão. Logo, o requerimento da exogeneidade das matrizes de relacionamentos G e $SAUDE$ 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 (η_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); Calvó-Armengol *et al.*, (2009); Bramoullé *et al.*, (2009) e Lee, (2007). Esta mesma estratégia, contudo, não pode ser adotada para o controle do efeito fixo individual, já 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 trás embutido em sua estatística aqueles componentes não observáveis (como esforço, habilidade, etc.) invariantes ao longo do ano letivo⁶. Nesse sentido, y_i^0 representa uma *proxy* para o *background* educacional inicial do estudante.

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 Como ilustração considere duas matrizes de relacionamentos para duas turmas hipotéticas cada uma com três alunos, g_1 e g_2 . A matriz conjunta diagonal G , denota-se por:

$$g_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad g_2 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad G = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

⁶ Considere dois modelos com estrutura de regressão para as notas do início e final do ano de um aluno i :

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

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 $\beta_2 = \theta \beta_1$ isto implica que os efeitos das variáveis X e u mudam a uma mesma taxa constante θ entre $t-1$ e t . Sob tais condições, a inclusão de y_i^0 no modelo empírico (2) permite o controle desta condição inicial fixa para cada aluno.

5 RESULTADOS

Os resultados das estimações do modelo empírico (3) são apresentados nesta seção. Uma variedade de modelos é testada utilizando a rede indireta de amigos⁷, ou seja, o caso em que G é simétrica. Estimativas com a matriz não-simétrica também foram operacionalizadas, porém em virtude da grande quantidade de zeros, problemas computacionais inviabilizaram as estimações⁸.

Na Tabela 2 encontram-se reportados os resultados estimados para diferentes especificações do modelo (3). Em cada especificação - colunas (1) a (3) - as estimativas são operacionalizadas por meio de ordem crescente de inclusão de regressores. O coeficiente positivo e significativo λ 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 as 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 (3)
Variável dependente: nota no final do ano

	(1)	(2)	(3)
Saúde do amigo (λ_j)	0,761	0,761	0,924
<i>(Estatística de teste)</i>	(3,513)	(3,505)	(3,760)
Características individuais	Sim	Sim	Sim
Características dos amigos	Sim	Sim	Sim
Saúde do aluno	Não	Sim	Sim
Saúde percebida do responsável	Não	Sim	Sim
Efeito fixo da rede	Não	Não	Sim
R^2 ajustado	0,294	0,293	0,345
Nº de observações	1.855	1.855	1.855
Nº de redes	139	139	139

Fonte: elaboração própria com base na Pesquisa Fundaj 2013.

Notas: Os resultados completos das estimações incluindo todas as variáveis de controle podem ser solicitados aos autores

⁷ No caso das redes indiretas (G é simétrica) $g_{ij} = g_{ji} = 1$, porém se a rede for considerada de maneira direta então $g_{ij} = 1$ e $g_{ji} = 0$.

⁸ Alguns estudos demonstram que os resultados do efeito dos pares não se alteram quanto à simetria da matriz G . Vejam-se os estudos de Patacchini e Venanzoni (2014) e Calvó-Armengol *et al.* (2009).

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 se relaciona 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.

Sendo assim, a estratégia de identificação proposta neste artigo depende da observância de $E(\varepsilon_{i,k} | \mathbf{x}_k, g_{ij,k}, \mathbf{SAUDE}_k, \eta_k) = 0$. Patacchini e Venanzoni (2014) testam esta suposição baseando-se no trabalho de Goldsmith-Pinkham e Imbens (2013). Considere-se novamente o modelo empírico (2):

$$y_{i,k} = \mu y_i^0 + \sum_{j=1}^N \lambda g_{ij,k} \mathbf{SAUDE}_{j,k} + \theta_{i,k}(\mathbf{x}) + \underbrace{\zeta_{i,k} + \eta_k + \varepsilon_{i,k}}_{e_{i,k}} \quad (2')$$

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 e não observáveis:

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

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 consiste em checar se há correlação significativa entre os resíduos da equação (2') e a probabilidade de formar amizade. Sendo assim, é possível substituir $|\zeta_{i,k} - \zeta_{j,k}|$ em (4) por $|\widehat{e}_{i,k} - \widehat{e}_{j,k}|$ de (2') 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 turma η_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, 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.

Tabela 3
Teste de robustez – Estimativas para equação (4)

	OLS
Diferença entre os resíduos ($\widehat{\phi}$) sem controle de efeito fixo (<i>p-value</i>)	0,00001 (0,0529)
Diferença entre os resíduos ($\widehat{\phi}$) 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 $n = 1.855$, o que gera um total de observações para a estimação de (4) de $[n*(n-1)]/2 = 1.719.585$ observações. As variáveis de controle são as mesmas incluídas no modelo (3).

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 foram estimados três modelos autorregressivos espaciais para as variáveis de saúde do aluno (*dummy* assumindo 1 quando o aluno possui algum tipo de doença e 0 no caso contrário), doenças físicas (*dummy* assumindo 1 quando o aluno possui alguma doença física e 0 quando não possui doenças) e doenças emocionais (*dummy* assumindo 1 quando o aluno possui alguma doença emocional e 0 quando não possui doenças), conforme a equação que se segue.

$$Z_{i,k} = \mu y_i^0 + \sum_{j=1}^n \rho g_{ij,k} Z_{j,k} + \theta_{i,k}(\mathbf{x}) + \phi_z \zeta_{i,k} + \phi_\eta \eta_k + \varepsilon_{i,k} \quad (5)$$

Em que Z corresponde aos três *outcomes* de saúde testados. 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. Entretanto, quando a correlação é testada separadamente para as doenças físicas e emocionais, o que se constata é que há evidências de correlações espaciais inversas e significantes, ou seja, a seleção dos alunos ao seu grupo de amigos ocorre de maneira compensatória. Estudantes que possuem algum tipo de doença, física ou emocional, tendem a fazer amizades com outros que não sejam acometidos por tais enfermidades.

Tabela 4
Teste de robustez – Estimativas para equação (5)

Estimativas ρ para as variáveis de saúde	OLS
Saúde do aluno	-0,010
(<i>p-value</i>)	(0,196)
Doença física	-0,065
(<i>p-value</i>)	(0,000)
Doença emocional	-0,073
(<i>p-value</i>)	(0,000)

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 relação entre saúde e desempenho educacional, via utilização de rede de amizades. Dispondo-se dos dados de uma pesquisa da Fundação Joaquim Nabuco – Fundaj no ano de 2013, a estratégia de identificação do efeito da saúde do amigo do aluno sobre a sua performance escolar se baseia 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. E mais, a seleção dos alunos ao grupo de amigos na classe tende a ser compensatória, isto é, alunos com algum tipo de doença tendem a fazer amizades com colegas saudáveis. 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 adoece 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. **Discrete games in endogenous networks: theory and policy**. Disponível em: <http://www.antonbadev.com/papers/dscr_games_endog_networks.pdf>. Acesso em: 4 jul. 2014.
- BALLESTER, C., CALVÓ-ARMENGOL, A. AND ZENOU, Y. **Who's who in networks. wanted: the key player**. *Econometrica*, v. 74(5), p. 1403–1417, September, 2006.
- BERNDT, T. J.; PERRY, T. B. **Children's perceptions of friendships as supportive relationships**. *Developmental Psychology*, 22, 640–648, 1986.
- BRAMOULLÉ, Y., DJEBBARI, H. AND FORTIN, B. **Identification of peer effects through social networks**. *Journal of Econometrics*, v. 150, p. 41-55, 2009.
- CALVÓ-ARMENGOL, A., PATAACCHINI, E. AND ZENOU, Y. **Peer effects and social networks in education**. *The Review of Economic Studies*, v. 76(4), p. 1239-1267, Oct., 2009.
- CHRISTAKIS, N.; FOWLER, J. **The spread of obesity in a large social network over 32 years**. *The New England Journal of Medicine* 357, 370-379, 2007.
- COHEN-COLE, E.; FLETCHER, J. M. **Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic**. *Journal of Health Economics* 27, 1382-1387, 2008.
- CURRIE, J., **Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development**. *Journal of Economic Literature* 47(1), 87–122. 2009.
- DEGIORGI, G., PELLIZZARI, M. AND WOOLSTON, W. G. **Class size and class heterogeneity**. NBER Working Paper 16405, September 2010.
- DING, W.; LEHRER, F. **Do Peers Affect Student Achievement in China's Secondary Schools?** *The Review of Economic Statistics*, vol.89, n.º2, pp.300-312, 2007.
- EPSTEIN, J. L. **The influence of friends on achievement and affective outcomes**. In J. L. Epstein & N. Karweit (Eds.), *Friends in school* (pp.177–200). New York: Academic Press, 1983.
- FLETCHER, J. M. **Social interactions and smoking: evidence using multiple student cohorts, instrumental variables, and school fixed effects**. *Health Economics*, 19: 466-484, 2010.
- _____. **Peer influences on adolescent alcohol consumption: evidence using an instrumental variables/fixed effect approach**. *Journal of Population Economics*, 25: 1265-1286, 2012.
- FUNDAÇÃO JOAQUIM NABUCO - FUNDAJ. **Coordenação de Estudos Econômicos e Populacionais. Acompanhamento longitudinal do desempenho escolar de alunos da rede pública de ensino fundamental do Recife**. Recife, 2013.
- GOLDSMITH-PINKHAM, P., AND IMBENS, G.W. **Social networks and the identification of peer effects**. *Journal of Business and Economic Statistics*, v. 31, p. 253–264. 2013.
- HARRIS, J. R. **Where is the child's environment? A group socialization theory of development**. *Psychological Review*, v.102(3), P.458-486, 1995.
- LEE, L. **Identification and estimation of econometric models with group interactions, contextual factors and fixed effects**. *Journal of Econometrics*, v. 140, p. 333–374, 2007.
- MANSKI, C. F. **Identification of endogenous social effects: the reflection problem**. *The Review of Economic Studies*, v. 60(3), p. 531-542, Jul., 1993.

- MELE, A. **A structural model of segregation in social networks.** Working Papers 10-16, NET Institute, 2010.
- PATACCHINI, E.; RAINONE, E.; ZENOU, Y. **Dynamic aspects of teenage friendships and educational attainment,** CEPR Discussion Paper 8223, 2011.
- PATACCHINI, E.; VENANZONI, G. **Peer effects in the demand for housing quality.** Journal of Urban Economics, v. 83, p. 6–17, 2014.
- PINTO, C. C. X. **Semiparametric estimation of peer effects.** 2008. Tese de Doutorado - University of California, Berkeley.
- RAPOSO, I. P. A. **O papel da rede de amizades e da formação aleatória de turmas por faixa etária sobre o desempenho.** 2015. Tese de Doutorado – Universidade Federal de Pernambuco.
- SCHNEEWEIS, N. AND WINTER-EBMER, R. **Peer effects in Austrian schools.** Empirical Economics, v. 32(2), p. 387-409, 2007.
- WENTZEL, K. R.; CALDWELL, K. A.; MCNAMARA, C. B. **Friendships in Middle School: Influences on Motivation and School Adjustment.** Journal of Educational Psychology, vol.96, n°2, 195-203, 2004.
- ZIMMER, R. **A new twist in the educational tracking debate.** Economics of Education Review, v. 22, p. 307-315, 2003.