UNIVERSIDADE DE SÃO PAULO

FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE DE RIBEIRÃO PRETO

DEPARTAMENTO DE ECONOMIA

PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA – ÁREA: ECONOMIA APLICADA

DIEGO MENEZES DOS SANTOS PINHEIRO

Diga-me com quem andas: efeito de pares sobre lócus de controle

ORIENTADOR: PROF. DR. DANIEL DOMINGUES DOS SANTOS

RIBEIRÃO PRETO

2019

Prof. Dr. Vahan Agopyan

Reitor da Universidade de São Paulo

Prof. Dr. André Lucirton Costa

Diretor da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto

Prof. Dr. Sérgio Kannebley Junior

Chefe do Departamento de Economia

Prof. Dr. Sergio Naruhiko Sakurai

Coordenador do Programa de Pós-Graduação em Economia – Área: Economia Aplicada

DIEGO MENEZES DOS SANTOS PINHEIRO

Diga-me com quem andas: efeito de pares sobre lócus de controle

Dissertação apresentada ao Programa de Pós-Graduação em Economia – Área: Economia Aplicada da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Mestre em Ciências. Versão Corrigida. A original encontra-se disponível na FEA-RP/USP.

ORIENTADOR: PROF. DR. DANIEL DOMINGUES DOS SANTOS

RIBEIRÃO PRETO

2019

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Pinheiro, Diego Menezes dos Santos
Diga-me com quem andas: efeito de pares sobre lócus de controle. Ribeirão Preto, 2019.
51 p. : il. ; 30 cm

Dissertação de Mestrado, apresentada ao Programa de Pós-Graduação em Economia da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo. Área de concentração: Economia Aplicada.
Orientador: Santos, Daniel Domingues dos.

1. Habilidades não-cognitivas. 2. Lócus de controle. 3. Ensino Médio - Brasil. 4. Interação social. 5. Efeito de pares.

Agradecimentos

Ao meu orientador Daniel Santos, por me inserir no tema fascinante que são as habilidades socioemocionais e me ajudar a definir a história que eu conto aqui.

Aos membros da Comissão Examinadora, por todas as valiosas contribuições.

À Capes, pelo apoio financeiro.

Aos técnicos e funcionários do Serviço de Pós-Graduação da FEA-RP/USP, pelas orientações e mensagens de apoio às vésperas de praticamente todos os prazos.

À Carolina Assis e Vinícius Princiotti do Laboratório de Estudos e Pesquisas em Economia Social (Lepes), por tão gentilmente me ajudarem a entender o funcionamento das análises psicométricas na prática e a desvendar as nuances da base de dados utilizada neste trabalho.

À Clarissa Gandour, Tomás do Valle, João Pedro Vieira, Rhayana Holz e Helena Arruda do Climate Policy Initiative (CPI), pelo imensurável suporte à conclusão deste estudo, ainda que ele não pertença às suas áreas de atuação na Economia. O interesse genuíno demonstrado por cada um de vocês em aprender sobre habilidades não-cognitivas e tudo o que esse tema envolve, para então serem capazes de me auxiliar, foi memorável e me deixou ainda mais grato por pertencer à essa grande família que é o CPI.

Aos meus colegas de turma e amigos Felipe Costa, Laura Ogando, Marcos Paulo Cambrainha e Raytza Yoshimura, pelo companheirismo em todas as horas, conhecimento dividido, preocupação e, principalmente, por terem sido minha família em Ribeirão Preto.

Às minhas amigas e irmãs Isabela Cid, Luana Miranda, Mariana Pinheiro e Sthefany Esposito, por se fazerem sempre presentes, ainda que à 800km de distância, e transformarem meus medos e ansiedades em motivação e esperança.

Aos meus pais, por sonharem esse sonho junto comigo, e terem me ajudado a torná-lo realidade.

À Deus, por absolutamente tudo.

PINHEIRO, D. M. S. **Diga-me com quem andas: efeito de pares sobre lócus de controle.** 2019. 51p. Dissertação (Mestrado) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2019.

Resumo

Esta dissertação reavalia o papel que a interação entre os pares tem na determinação do lócus de controle, uma medida de quão pessoalmente responsável uma pessoa se sente com relação ao que acontece na sua vida. O estudo se concentra no contexto escolar, utilizando os dados de alunos majoritariamente no Ensino Médio do sistema de ensino público e privado de Sertãozinho, um município brasileiro perto da cidade de São Paulo. De forma análoga ao trabalho que produziu a primeira análise dessa relação, a estratégia empírica utiliza variáveis instrumentais baseadas na estrutura de painel dos dados, porém de modo menos restritivo. Diferente de tal estudo, os resultados desta avaliação apresentam evidência de efeito de pares endógeno significativo sobre o lócus de controle. O efeito é positivo e se mantém robusto à inclusão de controles adicionais.

Palavras-chave: Efeito de Pares; Lócus de Controle; Variável Instrumental.

PINHEIRO, D. M. S. **Birds of a feather flock together: peer effects on locus of control.** 2019. 51p. Dissertation (Master Degree) – School of Economics, Business and Accounting of Ribeirão Preto, University of São Paulo, Ribeirão Preto, 2019.

Abstract

This dissertation re-evaluates the role of peer interaction in the determination of the locus of control, a measure of how personally responsible an individual feels about what happens in their life. The essay focuses on the school context, using data of students mostly in the High School of the public and private schooling system of Sertãozinho, a Brazilian municipality near the city of São Paulo. Similar to the research that produced the first analysis of this relationship, the empirical strategy uses instrumental variables based on the panel structure of the data, but in a less restrictive way. Unlike that study, the results of this evaluation show evidence of significant endogenous peer effect on the locus of control. The effect is positive and remains robust to the inclusion of additional controls.

Keywords: Peer Effects; Locus of Control; Instrumental Variable.

Conteúdo

1. Introdução 11

2. Estratégia Empírica 14

3. Dados 17
   3.1 Lócus de Controle 18
   3.2 Controles 21
   3.3 Estatísticas Descritivas 22

4. Resultados 27
   4.1 Principais Resultados: Efeito de Pares 27
   4.2 Robustez: Controles Adicionais 31

5. Conclusão 37

Bibliografia 39

Apêndice A 43
   A.1 Análise Fatorial 43
   A.2 Análise Psicométrica 44
      A.2.1 Lócus de Controle de Tel Aviv 44
   A.3 Questionários 48
      A.3.1 Lócus de Controle de Tel Aviv, 2012 48
      A.3.2 Lócus de Controle de Tel Aviv, 2017 50

Lista de Figuras

Figura 1 *Scree plots* dos Questionários de Lócus de Controle 20

Figura 2 Histograma do Escore de Lócus de Controle 20

Figura 3 Proporção dos pares de 2017 que não conheciam *i* em 2012 28

Lista de Tabelas

Tabela 1 Estatísticas Descritivas 23
Tabela 2 Média do Fator de Lócus de Controle por Grupo 24
Tabela 3 Média do Escore de Lócus de Controle por Grupo 24
Tabela 4 Média do Fator de Lócus de Controle por Quartil dos Pares 25
Tabela 5 Média do Escore de Lócus de Controle por Quartil dos Pares 26
Tabela 6 Primeiro-Estágio 29
Tabela 7 Segundo-Estágio 31
Tabela 8 Primeiro-Estágio – Robustez, Controles Adicionais 33
Tabela 9 Segundo-Estágio – Robustez, Controles Adicionais 36
Tabela 10 *Loadings* dos itens do Lócus de Controle de Tel Aviv, 2012 46
Tabela 11 *Loadings* dos itens do Lócus de Controle de Tel Aviv, 2017 47

1. Introdução

Desenvolvimentos recentes nos campos da educação, da economia, da psicologia e da neurociência mostram que as chamadas habilidades socioemocionais são tão importantes quanto as medidas cognitivas [1]na previsão de vários resultados da vida. Como exemplo, há evidências de que alunos que são mais abertos a novas experiências estéticas, culturais e intelectuais e que são mais organizados, esforçados e responsáveis possuem melhor desempenho escolar (Poropat, 2009). Há evidências também de que essas mesmas qualidades estão ligadas de forma positiva à salários e outras medidas de resultado do mercado de trabalho (Almlund et al., 2011).

Uma característica importante dessas habilidades não-cognitivas é se originar da interação entre predisposições biológicas e fatores socioambientais. (Primi et al., 2016). No entanto, o grau de influência de cada um desses agentes no processo de determinação das habilidades varia de acordo com a habilidade em formação. Por exemplo, tomemos dois dos domínios socioemocionais estudados por John et al. (2008), extroversão e amabilidade. Os autores dizem que o domínio da extroversão *"implica uma abordagem energética em relação ao mundo social"*, enquanto que o domínio da amabilidade *"contrasta uma orientação pró-social e comunitária em relação aos outros com antagonismo"*. De acordo com essas definições, é esperado que o ambiente social desempenhe um papel fundamental na construção da extroversão de um indivíduo, ao passo que sua amabilidade seja de responsabilidade maior dos fatores biológicos.

Uma vez que fator biológico é uma variável exógena, e, portanto, fora do escopo de políticas públicas, entender melhor a interação entre ambiente social e habilidades não-cognitivas é crucial para o desenho de políticas que visam utilizar este mecanismo como forma de aprimorar essas habilidades. No caso de crianças e adolescentes, por exemplo, cujo meio social está majoritariamente concentrado na escola, qualquer alteração na composição desse meio (por exemplo, rearranjo de turmas) pode modificar algumas dessas capacidades.

Este trabalho busca contribuir com a análise de como essa interação pode impactar uma habilidade socioemocional em específico, o lócus de controle.

O lócus de controle é uma medida de quão pessoalmente responsável uma pessoa se sente com relação ao que acontece na sua vida (Rotter, 1966). Indivíduos que acreditam

[1] Aqui, este termo está associado às habilidades de leitura, interpretação de texto, resolução de problemas e memorização.

que suas ações desempenham um papel importante na determinação do curso de suas vidas são ditos ter um lócus interno, ao passo que as pessoas com lócus externo normalmente verão o destino e a sorte como os principais responsáveis pela condução de suas vidas. Alguns estudos já foram capazes de demonstrar que esse construto é essencialmente interpessoal, no sentido de que é fortemente influenciado pelo meio social (ver Strauser et al., 2002 e Schneewind, 1995). Em paralelo a isso, um conjunto de evidências empíricas associam um lócus mais interno a alguns resultados positivos relevantes, como maior desempenho acadêmico (Findley & Cooper, 1983), maior produtividade no trabalho (Judge & Bono, 2001) e menor chance de cometer crimes (Urzua, 2008). Portanto, entender a dinâmica de formação desse construto através da interação com o ambiente social possui implicações que podem contribuir para o sucesso em variás áreas da vida da população.

A forma de interação social representada neste estudo é a de efeito de pares. Esta abordagem é muito utilizada em análises cuja unidade de observação são estudantes (Sacerdote, 2011), como é o meu caso.

Eu uso um banco de dados longitudinal auto reportado por estudantes do sistema de ensino público e privado de Sertãozinho, um município brasileiro perto da cidade de São Paulo. Em sua última coleta de informações, estes alunos encontravam-se majoritariamente no Ensino Médio. Esta base de dados me fornece um questionário de lócus de controle, a partir do qual eu calculo a estimativa dessa variável, e também informações sobre a composição das classes e das redes de amizades que me permitem identificar pares.

Eu utilizo o arcabouço tradicional dos estudos de efeito de pares, o modelo *linear-in-means*. Este modelo pressupõe que indivíduos pertencem a um grupo cujos membros influenciam e são influenciados por cada um deles, porém não possuem ligação com quaisquer indivíduos fora daquele grupo. Além disso, ele assume também que todos os indivíduos dentro do grupo são igualmente influentes, de modo que a média do grupo - excluindo-se o indivíduo *i* - é a variável relevante na função de produção de habilidades.

Num trabalho pioneiro sobre interações sociais, Manski (1993) destaca as dificuldades de identificar corretamente os parâmetros deste modelo. O autor definiu como “Problema da Reflexão” o fato de que tanto o comportamento médio de um grupo exerce influência sobre o comportamento de um indivíduo que pertence àquele grupo quanto o inverso. Sabemos que isso acarreta em viés de simultaneidade no estimador do efeito de pares.

Para contornar essa questão, eu utilizo a estratégia de identificação proposta por Hanushek et al. (2003): instrumentalizar o lócus de controle corrente dos pares por seu lócus passado. Essa abordagem pressupõe a existência de correlação serial do lócus de controle intra-indivíduos.

Uma primeira tentativa de explicar como o efeito de pares pode influenciar o lócus de controle foi feita em Chikitani et al. (2015). Contudo, os autores não encontram evidência de que esse efeito seja significativo, o que contraria a característica interpessoal desse construto. Eles utilizam a abordagem de Bramoullé et al. (2009) que implica em uma restrição a mais do que a de Hanushek et al. (2003). Assim, através dessa estratégia de identificação menos restritiva este trabalho pretende reavaliar essa relação.

Diferente de tal estudo, os resultados desta avaliação mostram evidência de efeito de pares endógeno significativamente positivo sobre o lócus de controle. Isso nos diz que, em média, quanto mais interno/externo for o lócus de controle dos pares do indivíduo *i*, mais interno/externo será o lócus desse indivíduo. Esta relação se mantém estável à inclusão de controles adicionais para outras habilidades socioemocionais.

O trabalho prossegue da seguinte forma: Seção 2 detalha a estratégia de identificação; Seção 3 apresenta a base de dados, os procedimentos de construção das variáveis e as estatísticas descritivas; Seção 4 reporta e discute os resultados; e Seção 5 conclui com implicações de políticas públicas.

2. Estratégia Empírica

Para identificar a minha relação de interesse – de que modo os pares de um indivíduo afetam seu lócus de controle – eu utilizo o típico modelo *linear-in-means* de efeito de pares. Esse modelo descreve a variável de resultado y_{it} do indivíduo *i* no período corrente *t* como uma função linear das suas características correntes e do resultado e características médias correntes dos seus pares. Seja P_{it} o grupo dos pares do agente *i* no período corrente *t*, com $i \notin P_{it}$, e $n_{it} = \sum_{j \in N} 1\{j \in P_{it}\}$ o número de pessoas neste grupo, em que *N* é a amostra completa. O modelo pode ser escrito, para cada *i*, da seguinte forma:

$$y_{it} = \alpha_1 + \beta_1 \frac{\sum_{j \in P_{it}} y_{jt}}{n_{it}} + X'_{it}\gamma_1 + \frac{\sum_{j \in P_{it}} X'_{jt}}{n_{it}} \delta_1 + \varepsilon_{it} \qquad (2\text{-}1)$$

As variáveis y_{it} e y_{jt} representam o lócus de controle corrente dos indivíduos e dos pares desses indivíduos, respectivamente. Da mesma forma, as variáveis X_{it} e X_{jt} sintetizam o vetor de características correntes de cada um desses agentes. ε_{it} é o termo de erro idiossincrático. O parâmetro β_1 nos dá o efeito endógeno de pares (*"a propensão de um indivíduo se comportar de alguma forma varia com o comportamento do grupo"*), enquanto δ_1 é o efeito contextual de pares (*"a propensão de um indivíduo se comportar de alguma forma varia com as características exógenas do grupo"*), conforme a denominação de Manski (1993).

Este arcabouço pressupõe duas hipóteses importantes:

i. indivíduos pertencem a um grupo cujos membros influenciam e são influenciados pelos outros, porém não possuem links com qualquer outra pessoa fora desse grupo.

ii. todos os indivíduos dentro de um grupo são igualmente influentes, de modo que a média do grupo - excluindo-se o indivíduo *i* - é a variável relevante na função de produção de y_{it}.

Sob a hipótese (i), o termo que mede o efeito endógeno dos pares se torna problemático porque $j \in P_{it}$ implicará então que $i \in P_{jt}$. Sabemos que isso acarreta em viés de simultaneidade no estimador do parâmetro β_1. Num trabalho pioneiro sobre interações sociais, Manski (1993) classifica esse obstáculo como "Problema da Reflexão". Esta seção propõe uma estratégia de variáveis instrumentais para identificar o efeito causal do lócus de controle dos pares de um indivíduo sobre o seu lócus de controle.

A estrutura em formato de painel da base de dados utilizada neste trabalho me permite replicar a mesma abordagem de variáveis instrumentais empregada em Hanushek et al. (2003): instrumentalizar o lócus de controle corrente dos pares por seu lócus passado.

A validade desse instrumento reside (i) no fato dele ser não-correlacionado com o termo de erro ε_{it} na equação 2-1, condicional em todas as variáveis observáveis incluídas nesta equação (restrição de exclusão) e (ii) de haver correlação serial do lócus de controle intra-indivíduos (restrição de inclusão). A condição (i) poderia ser violada somente se indivíduos que são pares no período corrente já fossem par um do outro desde o período passado. Eu contorno esta situação usando como instrumento somente o lócus de controle passado de quem é par do indivíduo *i* no período corrente e não era no período passado.

Dessa forma, eu argumento que o único canal através do qual o lócus de controle passado dos pares se relaciona com o lócus de controle corrente do indivíduo *i* é via lócus de controle corrente dos seus pares.

Eu demonstro a validade da condição (ii) através da estimação da equação de primeiro-estágio:

$$\frac{\sum_{j\in P_{it}} y_{jt}}{n_{it}} = \mu + \rho \frac{\sum_{j\in P_i^*} y_{jt-1}}{n_i^*} + \upsilon_{it} \qquad (2\text{-}2)$$

onde a variável dependente é o lócus de controle médio corrente dos pares correntes do indivíduo *i* e o regressor é o lócus de controle médio passado dos indivíduos que são pares correntes do indivíduo *i* mas não eram seus pares no passado. O parâmetro P_i^* representa o grupo dos pares do agente *i* no período corrente *t* que não eram seus pares no período passado *t-1*, com $i \notin P_i^*$, e $n_i^* = \sum_{j\in N} 1\{j \in P_i^*\}$ o número de pessoas neste grupo.

A equação a ser estimada no segundo-estágio é dada por:

$$y_{it} = \alpha_2 + \beta_2 \frac{\sum_{j\in P_{it}} y_{jt}}{n_{it}} + X'_{it}\gamma_2 + \frac{\sum_{j\in P_{it}} X'_{jt}}{n_{it}} \delta_2 + \xi_{it} \qquad (2\text{-}3)$$

onde o lócus de controle médio corrente dos pares encontra-se instrumentalizado pelo seu valor médio passado.

Caso seja significativo, $\beta_2 > 0$ indica que quanto mais interno/externo for o lócus de controle dos pares do indivíduo *i*, mais interno/externo será o seu próprio lócus, enquanto que se $\beta_2 < 0$ quanto mais interno/externo for o lócus de controle dos pares do indivíduo *i*, mais externo/interno será o lócus desse indivíduo. Quando significativas, as estimativas de efeito de pares endógeno tendem a ser positivas (Hamilton et al., 2003; Mas & Moretti, 2009; Sacerdote, 2001). Embora o papel do efeito de pares sobre a determinação do lócus de controle – ou de habilidades não-cognitivas em geral – seja amplamente não explorado,

Bursztyn & Jensen (2015) mostram que alunos do Ensino Médio são muito empenhados em se comportar conforme as normas sociais dos colegas de turma – ou pelo menos sinalizam que agem dessa forma –, o que também implica em esperar que o coeficiente estimado de efeito de pares endógeno seja positivo. Eu realizo um teste de robustez para explorar a sensibilidade deste coeficiente à inclusão de demais habilidades socioemocionais como controles.

3. Dados

A análise empírica deste trabalho é inteiramente baseada em um conjunto de dados longitudinais auto reportado por estudantes do sistema de ensino público e privado de Sertãozinho, um município brasileiro perto da cidade de São Paulo. Esses dados são de origem privada e foram gentilmente cedidos pelo Laboratório de Estudos e Pesquisas em Economia Social (Lepes), que foi responsável pelas duas últimas fases de coleta dessas informações.

Sertãozinho é um município da região nordeste do estado de São Paulo com população estimada de 124.453 habitantes (IBGE, 2018). Trata-se de uma região relativamente rica – renda per capita de R$42.168 em 2016 contra R$30.407 do Brasil neste mesmo ano (IBGE, 2016) – e, mais importante, possui um sistema público de educação muito mais avançado do que o padrão nacional: em 2017, 65% dos alunos da rede pública desta região possuíam conhecimento apropriado[2] de matemática no 5º ano do ensino fundamental, de acordo com o Sistema de Avaliação da Educação Básica (Saeb), enquanto que esse mesmo valor para o país e para o estado de São Paulo foi 26% e 41%, respectivamente.

A primeira fase de coleta de dados foi realizada em 2008. Reuniu-se informações de alunos do 2º ano do Ensino Fundamental de 22 escolas em Sertãozinho que se voluntariaram a participar. Desse total, 14 escolas eram públicas e 8 eram privadas. 2087 alunos levaram para casa um formulário de perguntas para ser respondido pelos pais. Essencialmente, este questionário abordava as características pessoais, socioeconômicas e familiares dos estudantes.

Quatro anos depois realizou-se a segunda coleta. Como o objetivo era entrevistar as mesmas crianças entrevistadas em 2008, reuniu-se informações de todas as escolas públicas de Sertãozinho que possuíam alunos matriculados no 6º ano do Ensino Fundamental além das escolas privadas também com alunos nesta fase que se dispuseram a participar. Ademais, devido aos casos de repetência, esta coleta também abrangeu as escolas que possuíam alunos de 2008 matriculados no 4º ou 5º ano do Ensino Fundamental. No total, a amostra de 2012 conta com a participação de 31 escolas de Sertãozinho, sendo 17 municipais, 8 estaduais e 6 privadas. O total de entrevistados contabiliza 5219 alunos. Desse montante, 1705 estudantes pertencem a ambas as coletas. Esse valor representa aproximadamente 82% da amostra de 2008.

[2] Desempenho maior ou igual a 250 na prova de matemática do Saeb, cujo resultado varia de 0 a 500.

Nesta etapa, apesar dos pais também terem contribuído com informações, os próprios alunos foram responsáveis por auto reportar suas características pessoais, socioeconômicas e familiares. Além dessas informações que são comuns à ambas as coletas, a base de dados de 2012 conta também com questionários de personalidade. Os alunos responderam às versões adaptadas para a idade e para o Brasil de três instrumentos psicológicos: o Lócus de Controle de Tel Aviv (Milgram & Milgram, 1975), o *Big Five Inventory* (BFI; John & Srivastava, 1999) e o *Social Skills Rating System*, de Gresham & Elliott (2007), sumarizado em Crosby (2011).

Em 2017 foi feita a terceira e até então última coleta. Novamente, a pesquisa foi realizada de modo censitário com o objetivo de facilitar a logística de aplicação nas séries que concentravam a maior parte dos alunos originários das coletas de 2008 e 2012. Dessa forma, foram entrevistadas todas as escolas que possuíam alunos matriculados no 9º ano do Ensino Fundamental e 1º e 2º ano do Ensino Médio de Sertãozinho. Além disso, os demais alunos da coorte que estavam matriculados em outras séries também participaram da coleta, sendo colocados junto a outras turmas no momento da aplicação dos questionários. No total, a amostra de 2017 conta com a participação de 24 escolas de Sertãozinho, sendo 6 municipais, 9 estaduais, 6 privadas e 1 federal. O total de entrevistados contabiliza 4148 alunos. Desse montante, 1065 estudantes pertencem às três coletas, 2008, 2012 e 2017 (51% da amostra de 2008) e 3141 estudantes pertencem às duas últimas coletas, 2012 e 2017 (60% da amostra de 2012).

Nesta etapa, os alunos também reportaram suas próprias características pessoais, socioeconômicas e familiares e participaram dos mesmos questionários de personalidade aplicados em 2012. Além da composição de turmas, que é presente nas três bases de dados, a base de 2017 dispõe também de uma segunda maneira de inferir ligações entre os estudantes. Foi pedido que eles listassem seus cinco melhores amigos na escola e também em Sertãozinho. Na Seção 3.3 falamos sobre como a inclusão dessa pergunta no questionário dos alunos contribui com a minha análise.

3.1 Lócus de Controle

O lócus de controle é um traço de personalidade desenvolvido por Rotter (1966). Este construto mede quão pessoalmente responsável uma pessoa se sente com relação ao que acontece na sua vida. Indivíduos que acreditam que suas ações desempenham um papel

importante na determinação do curso de suas vidas são ditos ter um lócus interno, ao passo que as pessoas com lócus externo normalmente verão o destino e a sorte como os principais responsáveis pela condução de suas vidas. O senso comum espera que indivíduos internos depositem mais esforço para cumprir seus objetivos. Com efeito, um conjunto de evidências empíricas associam um lócus mais interno a alguns resultados positivos, como maior desempenho acadêmico (Findley & Cooper, 1983), maior produtividade no trabalho (Judge & Bono, 2001), menor chance de cometer crimes (Urzua, 2008) e maior cuidado com a saúde (Wallston & Wallston, 1978). Contudo, ainda que as evidências tipicamente associem um lócus interno a resultados positivos, esta relação não é monotônica. Há indícios de que esse aspecto do lócus de controle também está ligado a estresse e depressão (Benassi et al., 1988). De fato, existem situações na vida que estão além do nosso controle, e pessoas que não percebem ou não aceitam isso estão mais propensas a sofrer desses males.

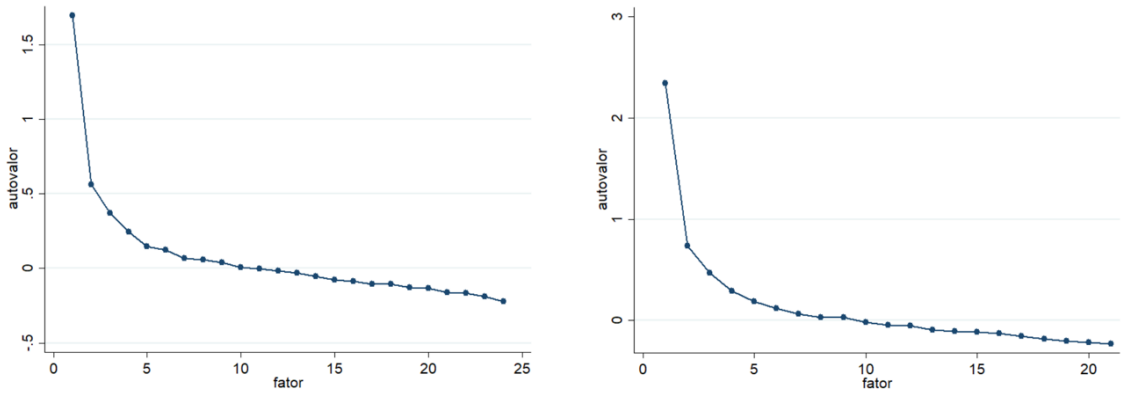
A construção dessa variável é baseada no questionário de Lócus de Controle de Tel Aviv que foi aplicado aos alunos durante a coleta de dados de 2012 e 2017[3]. Devido os alunos estarem mais maduros na última pesquisa realizada (idade média de 16 anos), as perguntas do questionário de 2017 não são as mesmas do questionário de 2012. Entretanto, a estrutura do questionário permaneceu a mesma: o estudante se depara com uma situação hipotética do cotidiano e deve escolher entre duas opções de resposta. Uma das opções classifica aquela resposta como do tipo interna e a outra do tipo externa. O questionário de 2012 conta com 24 questões e o de 2017 com 21. Através da análise fatorial eu sou capaz de resumir as informações contidas nas respostas dessas questões em uma escala de dimensão menor, conhecida como fator[4]. De fato, conforme mostram os *scree plots* na Figura 1, toda a informação contida nos questionários de 2012 e 2017 pode ser razoavelmente descrita por um único fator para cada um desses questionários.

Uma vez que os *loadings* associados a cada um desses fatores tipicamente possuem o sinal esperado (ver Apêndice A.2), eu os utilizo como uma medida do lócus de controle.

[3] O Apêndice A.3 apresenta os questionários completos.
[4] Ver Apêndice A.1 para entender os mecanismos envolvidos nesta técnica.

Figura 1: *Scree plots* dos Questionários de Lócus de Controle

(a) 2012 (b) 2017

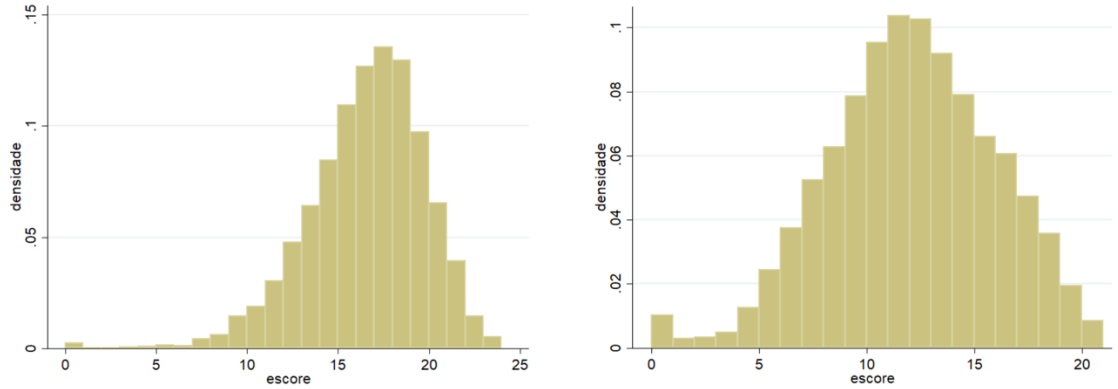


Notas: A figura apresenta os autovalores associados a cada fator extraído da análise fatorial exercida às respostas do questionário de Lócus de Controle de Tel Aviv aplicado em 2012 (Figura 1.a) e em 2017 (Figura 1.b).

Além desse cálculo, eu também uso a fórmula pioneira de construção desse domínio socioemocional, proposta por Rotter (1966). Essa medida alternativa é baseada na soma das respostas de cada indivíduo, em que cada resposta do tipo interno vale 1 ponto. Isto significa que para cada indivíduo esse escore varia entre 0 e 24 para o questionário de 2012 e 0 e 21 para o questionário de 2017. A distribuição empírica desse escore é apresentada na Figura 2.

Figura 2: Histograma do Escore de Lócus de Controle

(a) 2012 (b) 2017



Notas: A figura apresenta a densidade do lócus de controle calculado como a soma das respostas de cada indivíduo ao questionário de Lócus de Controle de Tel Aviv, em que cada resposta do tipo interno vale 1 ponto. Esse escore é calculado com as respostas ao questionário aplicado em 2012, que possui 24 questões (Figura 2.a), e em 2017, que possui 21 questões (Figura 2.b).

3.2 Controles

É importante ressaltar que a escolha do grupo de covariadas X desta análise requer grande cautela, uma vez que elas também cumprem o papel de variáveis instrumentais na equação 2-2 da Seção 2 (ver Seção 4.1). Variáveis que são irrelevantes na determinação do lócus de um indivíduo provavelmente serão instrumentos fracos.

Embora o lócus de controle pareça não diferir entre gêneros na fase adulta (De Brabander & Boone, 1990), meninas adolescentes tendem a apresentar um lócus mais interno do que meninos nesta mesma fase (Manger & Eikeland, 2000). A idade também é um fator determinante deste construto, uma vez que as pessoas se tornam mais internas conforme o seu avanço, atingindo o pico durante a meia idade (Lefcourt, 1982; Heckhausen & Schulz, 1995). Com relação ao ambiente familiar, Strauser et al. (2002) argumenta que pais que trabalham cumprem um papel importante no quanto os filhos valorizam o esforço, resultando em filhos mais internos. Além disso, Schneewind (1995) ressalta que ter irmãos torna as crianças mais externas. Baixo status socioeconômico também está associado com lócus mais externo (Freed & Tompson, 2011), porém os canais através dos quais isto acontece ainda se encontram em debate. Por último, etnia e cultura também importam: Berry et al. (2002) argumenta que pessoas negras são mais externas do que as brancas nos Estados Unidos, mesmo após controlar para o status socioeconômico; Uba (2003) conclui que sociedades mais individualistas promovem lócus mais interno ao comparar o lócus de controle de chineses que vivem na América e chineses que vivem em Hong Kong – estes últimos são tipicamente mais externos.

A escolha de *dummies* para gênero e raça parece então ser óbvia. O mesmo vale para a idade dos alunos. Entretanto, a minha medida de idade apresentou variação muito baixa entre os indivíduos, de modo que eu decidi não incluí-la no conjunto de regressores. Para representar a influência do ambiente familiar, eu utilizo uma *dummy* que indica se o estudante mora com algum irmão. Embora o conjunto de questionários me forneça a informação sobre a atividade praticada pelos pais, que me permitiria dizer se eles são desempregados ou não, o número de observações faltantes para essa questão é extremamente alto, e, portanto, usá-la reduziria significativamente o tamanho da amostra. Por último, como *proxy* para o status socioeconômico eu utilizo uma *dummy* que indica se o pai do estudante completou o Ensino Médio.

Todas essas informações são provenientes do conjunto de questionários auto reportado pelos alunos, com exceção da *dummy* “Pai completou o Ensino Médio” que reúne informações reportadas tanto pelos alunos quanto pelos pais. Essa variável, assim como as outras, é reportada por ambos os indivíduos nas coletas de 2012 e 2017. No entanto, no seu caso em específico observa-se uma grande divergência entre o que é reportado por cada uma das partes. Eu adotei o critério de considerar a resposta dos pais como a correta. Entretanto, como a taxa de participação dos pais nas coletas de dados é muito inferior à dos alunos, eu considero a resposta destes últimos nos casos em que falta informação dos pais para essa questão.

3.3 Estatísticas Descritivas

A abordagem de variáveis instrumentais utilizada como estratégia de identificação neste trabalho consiste em instrumentalizar o lócus de controle corrente dos pares por seu lócus passado. A viabilidade dessa estratégia reside na possibilidade de construção desse lócus para um mesmo conjunto de indivíduos em dois períodos do tempo. Vimos na Seção 3.1 que isso foi feito utilizando a resposta de alunos do sistema de ensino público e privado de Sertãozinho ao questionário de Lócus de Controle de Tel Aviv em 2012 e 2017.

Eu utilizo duas definições de pares. A primeira é baseada na informação disponível de composição das classes. Existe uma vasta literatura de efeito de pares que compartilha dessa mesma definição (Hoxby, 2000; Boozer & Cacciola, 2001; Hanushek et al., 2003; Betts & Zau, 2004; Vigdor & Nechyba, 2007; Carman & Zhang, 2012; Burke & Sass, 2013; Chikitani et al., 2015). Sob a hipótese de alocação aleatória dos alunos nas turmas, eu argumento que não existe auto-seleção nesses grupos. A segunda definição é baseada na rede de amizades dos alunos. Na última coleta de dados em 2017 foi pedido que eles listassem seus cinco melhores amigos na escola e também em Sertãozinho. Como eu sou capaz de identificar parte desses amigos da escola na minha base de dados, eu uso esta informação como uma forma alternativa de inferir ligações entre os estudantes. Entretanto, neste caso a auto-seleção é predominante entre os grupos. Portanto, os resultados gerados a partir desta abordagem não serão interpretados como efeito causal.

Do total de 3141 alunos que eu consigo encontrar nas bases de 2012 e 2017, somente 1124 (36%) não possui nenhuma informação faltante que seja necessária para a minha

análise empírica seguindo a definição de pares que usa colegas da turma. Quando uso a lista de melhores amigos da escola, esse número é ainda menor, 385 (12%). Essa queda nas observações em função da mudança na definição dos pares é esperada devido à maior dificuldade de identificação dos amigos na base de dados. A listagem desses amigos foi feita nominalmente, e na maioria dos casos com imperfeições. A Tabela 1 apresenta as médias e desvios-padrão das variáveis usadas na análise empírica para ambas essas amostras.

Tabela 1: Estatísticas Descritivas

Variável	Amigos		Turma	
	Média	Desv. Pad.	Média	Desv. Pad.
Indivíduos *i*'s, 2017				
Lócus de Controle Fator	0,03	0,84	0,04	0,87
Lócus de Controle Escore	12,05	3,53	12,05	3,68
Menino	0,48	0,50	0,48	0,50
Branco	0,41	0,49	0,42	0,49
Mora com irmãos	0,89	0,32	0,84	0,37
Pai completou o Ensino Médio	0,32	0,47	0,29	0,45
Pares, 2017				
Lócus de Controle Fator	0,10	0,76	0,01	0,49
Lócus de Controle Escore	12,31	3,20	11,91	2,06
Menino	0,49	0,47	0,51	0,29
Branco	0,45	0,46	0,40	0,28
Mora com irmãos	0,89	0,29	0,85	0,20
Pai completou o Ensino Médio	0,34	0,43	0,33	0,28
Pares, 2012				
Lócus de Controle Fator	0,26	0,58	0,15	0,40
Lócus de Controle Escore	17,28	2,32	16,99	1,41
Menino	0,49	0,47	0,51	0,29
Branco	0,44	0,46	0,39	0,28
Mora com irmãos	0,90	0,28	0,90	0,15
Pai completou o Ensino Médio	0,26	0,41	0,23	0,25
Observações	385		1124	

Notas: A tabela apresenta média e desvio-padrão para as variáveis usadas na análise empírica. As variáveis correspondentes aos pares estão em formato de média, conforme especificação do modelo *linear-in-means* (ver Seção 2). As variáveis Menino, Branco, Mora com irmãos e Pai completou o Ensino Médio são *dummies.*

Vemos que apesar da grande diferença no número de observações, ambas as amostras possuem estatísticas similares para as variáveis de interesse.

A Tabela 2 apresenta a média e o desvio-padrão do fator de lócus de controle para cada grupo das minhas covariadas, também utilizando ambas as amostras. Nesta análise, eu utilizo somente informações dos indivíduos *i*'s em 2017. Primeiramente vemos que mais uma vez os valores encontrados em ambas as amostras divergem muito pouco. Também, vemos que as *dummies* de "Menino" e "Mora com irmãos" vão contra o que diz a literatura de psicologia sobre essas variáveis (ver Seção 3.2): na minha amostra, meninos e indivíduos que vivem com os irmãos são em média mais internos.

Tabela 2: Média do Fator de Lócus de Controle por Grupo

Variável	Amigos		Turma	
	d = 0	d = 1	d = 0	d = 1
Menino	-0,06	0,13	-0,06	0,15
	(0,87)	(0,80)	(0,87)	(0,85)
Branco	0,03	0,04	0,03	0,05
	(0,84)	(0,85)	(0,86)	(0,88)
Mora com irmãos	-0,02	0,04	-0,04	0,05
	(0,94)	(0,83)	(0,86)	(0,87)
Pai completou o Ensino Médio	-0,04	0,19	-0,01	0,16
	(0,85)	(0,80)	(0,87)	(0,85)
Observações	385		1124	

Notas: A tabela apresenta a média do lócus de controle extraído da análise fatorial, por grupo. Eu utilizo somente informações dos indivíduos *i*'s em 2017. Todas as variáveis são *dummies*. As colunas "d=1" indicam que a *dummy* está ativa, e "d=0" o caso contrário. Desvios-padrão em parênteses.

A Tabela 3 mostra que esses resultados se mantêm quando eu uso o escore de lócus de controle no lugar do fator.

Tabela 3: Média do Escore de Lócus de Controle por Grupo

Variável	Amigos		Turma	
	d = 0	d = 1	d = 0	d = 1
Menino	11,71	12,42	11,60	12,55
	(3,57)	(3,45)	(3,61)	(3,69)
Branco	12,04	12,06	12,01	12,12
	(3,51)	(3,56)	(3,66)	(3,71)
Mora com irmãos	11,79	12,08	11,81	12,10
	(4,05)	(3,46)	(3,59)	(3,70)
Pai completou o Ensino Médio	11,79	12,60	11,86	12,54
	(3,57)	(3,37)	(3,70)	(3,58)
Observações	385		1124	

Notas: A tabela apresenta a média do lócus de controle escore por grupo. Eu utilizo somente informações dos indivíduos *i*'s em 2017. Todas as variáveis são *dummies.* As colunas "d=1" indicam que a *dummy* está ativa, e "d=0" o caso contrário. Desvios-padrão em parênteses.

A Tabela 4 apresenta a média e o desvio-padrão do fator de lócus de controle dos indivíduos *i*'s para cada quartil do fator de lócus de controle médio dos pares. Eu utilizo informações dos indivíduos *i*'s e dos pares somente em 2017. Em geral, podemos afirmar que a média do lócus dos indivíduos aumenta de acordo com a ordem dos quartis dos lócus médios dos pares. Essa média possui variação da ordem de 0,29 na amostra de amigos e 0,20 na amostra de turma, entre o 1º e o 4º quartil. Esses números representam 34% e 23%, respectivamente, do desvio-padrão do lócus do total dos indivíduos nessas amostras em 2017.

Tabela 4: Média do Fator de Lócus de Controle por Quartil dos Pares

Variável	Amigos		Turma	
	Média	Desv. Pad.	Média	Desv. Pad.
1º quartil	-0,16	0,86	-0,02	0,89
2º quartil	0,10	0,82	0,00	0,90
3º quartil	0,07	0,85	0,00	0,85
4º quartil	0,13	0,81	0,18	0,83
total	**0,03**	**0,84**	**0,04**	**0,87**
Observações	385		1124	

Notas: A tabela apresenta a média e o desvio-padrão do lócus de controle fator dos indivíduos *i*'s, por quartil do lócus de controle fator médio dos pares. Eu utilizo informações dos indivíduos *i*'s e dos pares somente em 2017.

A Tabela 5 reporta essa mesma análise, usando o escore de lócus de controle no lugar do fator. Eu encontro a mesma relação positiva entre a média do lócus dos indivíduos e a ordem dos quartis dos lócus médios dos pares. Contudo, ela ocorre em menor magnitude na amostra de amigos. A variação da média do lócus dos indivíduos, entre o 1º e o 4º quartil, como porcentagem do desvio-padrão do lócus do total dos indivíduos nessa amostra é da ordem de 17% – metade do valor encontrado com o lócus no formato fator utilizando essa mesma amostra. Na amostra de turma essa medida permanece estável, em torno de 24%.

Tais tendências, ainda que sugestivas de efeito de pares endógeno positivo, não são conclusivas devido à presença de endogeneidade – simultaneidade, ou "Problema da Reflexão", considerando a amostra de turma e simultaneidade mais auto-seleção

considerando a amostra de amigos (ver Seção 2 e segundo parágrafo desta seção). A seção seguinte apresenta a versão causal dessa relação.

Tabela 5: Média do Escore de Lócus de Controle por Quartil dos Pares

Variável	Amigos		Turma	
	Média	Desv. Pad.	Média	Desv. Pad.
1º quartil	11,65	3,53	11,63	3,79
2º quartil	11,73	3,68	11,89	3,65
3º quartil	12,59	3,39	12,20	3,78
4º quartil	12,23	3,47	12,52	3,43
total	**12,05**	**3,53**	**12,05**	**3,68**
Observações	385		1124	

Notas: A tabela apresenta a média e o desvio-padrão do lócus de controle escore dos indivíduos *i*'s, por quartil do lócus de controle escore médio dos pares. Eu utilizo informações dos indivíduos *i*'s e dos pares somente em 2017.

4. Resultados

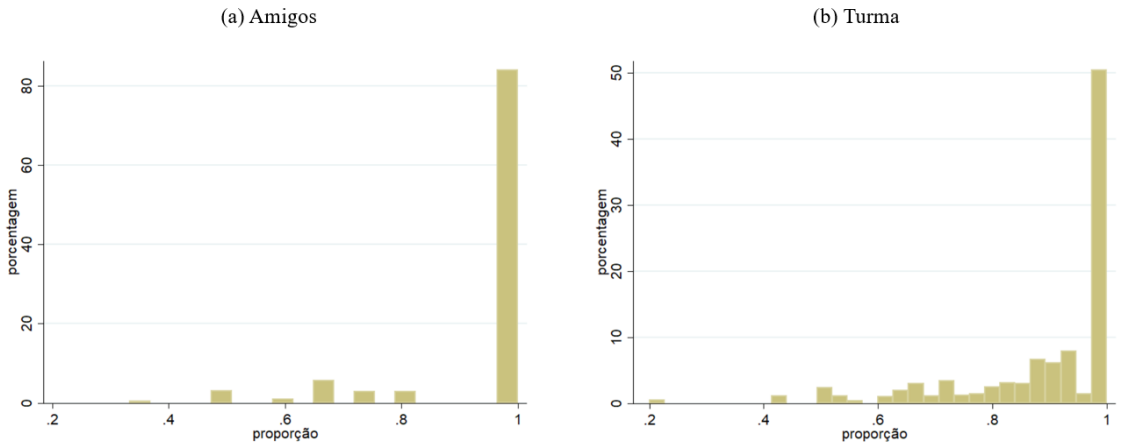
Esta seção se inicia testando empiricamente a existência de efeito de pares endógeno sobre lócus de controle. Em seguida, é feita a análise de robustez que usa demais habilidades não-cognitivas como controles adicionais. Em todas as análises desta seção as medidas de lócus de controle encontram-se padronizadas.

4.1 Principais Resultados: Efeito de Pares

A estimação do parâmetro β_2 de efeito de pares endógeno na equação de segundo-estágio 2-3 (Seção 2) passa antes pela etapa de instrumentalização do lócus de controle corrente dos pares por seu lócus passado. Isto implica em estimar a equação de primeiro-estágio 2-2 (Seção 2). Antes disso, relembre da Seção 2 que a validade da restrição de exclusão desse instrumento depende dos indivíduos que são pares no período corrente não serem par um do outro no período passado. Como *proxy* para "ser par um do outro no período passado" eu utilizo a informação de composição de turma de 2012. Assim, eu uso como instrumento somente o lócus de controle passado de quem é par do indivíduo *i* em 2017 e não era da mesma turma que ele em 2012.

Se apenas uma quantidade pequena de pares satisfazem essa restrição, significa que a variabilidade do meu instrumento é muito baixa. Neste caso, a estimação da equação 2-2 exibiria um coeficiente não-significativo para ρ, o que poderia prejudicar a estratégia de identificação de β_2. A Figura 3 resume a proporção de pares que cumprem essa restrição. Vemos que a maior parte dos indivíduos *i*'s não pertenciam à mesma turma que a maioria de seus pares em 2012. Este resultado é valido tanto para a amostra que segue a definição de pares como melhores amigos da escola (Figura 3.a) quanto a que usa os colegas de turma (Figura 3.b) – mais de 80% dos indivíduos *i*'s não conheciam 80 à 100% dos seus pares em 2012, tanto na amostra de amigos quanto na de colegas de turma. Assim, eu espero encontrar um valor significativo de ρ na estimação da equação 2-2 utilizando ambas as amostras.

Figura 3: Proporção dos pares de 2017 que não conheciam *i* em 2012



Notas: A figura apresenta a porcentagem dos indivíduos *i*'s de acordo com a proporção dos pares de 2017 que não eram da mesma turma que eles em 2012. Essa informação é demonstrada utilizando a amostra que segue a definição de pares como melhores amigos da escola (Figura 3.a) e a que usa os colegas de turma (Figura 3.b).

A estimação da equação de primeiro-estágio 2-2 é reportada na Tabela 6. Além do lócus de controle médio de 2012 dos pares, eu incluo como instrumentos adicionais as características médias de 2012 desses pares. A coluna 1 reporta os resultados utilizando a amostra de melhores amigos e o lócus de controle extraído da análise fatorial. Na coluna 2 eu utilizo a medida alternativa desta variável, seu formato escore. As colunas 3 e 4 reportam os resultados utilizando a amostra de colegas da turma e seguem as mesmas especificações das colunas 1 e 2, respectivamente.

Como esperado, obtemos um valor estatisticamente relevante de ρ tanto para a amostra de melhores amigos quanto para de turma. Em termos de significância este coeficiente também é similar entre os modelos que utilizam o lócus de controle na sua forma fator e os que o utilizam na sua forma de escore. Os resultados indicam que o lócus de controle médio de 2012 dos pares tem relação positiva significativa com o seu lócus de controle médio de 2017.

Note que as estatísticas-F de todas as especificações geram preocupação sobre a força dos meus instrumentos, uma vez que $F < 10$ sugere que os instrumentos são fracos (Stock et al., 2002). Este resultado já era esperado, pois todos os instrumentos se encontram em formato de média, o que reduz parte da sua variabilidade. Para contornar esta situação, eu recorro ao teste de Anderson & Rubin (1949) que produz inferência robusta à presença de instrumentos fracos para avaliar a significância do regressor endógeno no segundo-estágio.

Tabela 6: Primeiro-Estágio

	Amigos		Turma	
	(1)	(2)	(3)	(4)
	LC Fator	LC Escore	LC Fator	LC Escore
Lócus de Controle dos Pares, em 2012	0,118***	0,126**	0,194***	0,179**
	(2,87)	(2,38)	(3,06)	(2,10)
Características dos Pares, em 2012				
Menino	-0,061	0,005	-0,045	0,225
	(-0,18)	(0,01)	(-0,08)	(0,41)
Branco	-0,246	-0,240	1,002***	1,166***
	(-1,31)	(-1,20)	(3,04)	(3,78)
Mora com irmãos	-0,340	-0,433*	0,064	0,211
	(-1,52)	(-1,91)	(0,10)	(0,32)
Pai completou o Ensino Médio	0,403**	0,350**	0,363	0,215
	(2,57)	(2,21)	(1,09)	(0,66)
Observações	385	385	1124	1124
Estatística-F	2,477	5,705	3,332	3,019

Notas: A tabela reporta os resultados estimados da equação 2-2 de primeiro-estágio (Seção 2). Todas as variáveis, incluindo as dependentes, dizem respeito aos pares e, portanto, estão em formato de média. Além do lócus de controle médio de 2012 dos pares, eu incluo como instrumentos adicionais as características médias de 2012 desses pares. As variáveis dependentes representam o lócus de controle médio de 2017 dos pares, nas versões fator e escore. Estatísticas-t em parênteses. Erros padrão calculados com cluster de escola. Os coeficientes dos controles do segundo-estágio e da constante foram omitidos por simplificação visual. Significância: *** p<0.01, ** p<0.05, * p<0.10.

A Tabela 7 reporta os coeficientes estimados para o efeito de pares endógeno sobre o lócus de controle. Quando estimada por um simples MQO (colunas 1 a 4), essa relação aparece de forma significativa apenas no modelo que utiliza a amostra de colegas de turma e o lócus de controle escore (coluna 4). Contudo, sabemos que estes resultados estão incorretos (ver discussão na Seção 2). De fato, a magnitude e a significância dos coeficientes estimados usando MQO em 2 estágios (colunas 5 a 8) indicam que os resultados encontrados no MQO simples estão viesados. Vemos que os modelos que utilizam a amostra de amigos (colunas 5 e 6) mantiveram seus coeficientes estatisticamente não significativos. Porém, mesmo com a simultaneidade controlada pelo processo de estimação em dois estágios, sabemos que estes modelos permanecem endógenos devido à auto-seleção presente no processo de formação dos pares (ver discussão na Seção 3.3). Sua inclusão neste estudo é apenas para servir de análise comparativa aos resultados dos modelos que utilizam a amostra de colegas de turma (colunas 7 e 8). Os resultados do MQO em 2 estágios para estes últimos mostram que existe efeito de pares endógeno positivo sobre o lócus de controle. Tudo o mais constante, o aumento de 1 desvio-padrão do lócus de controle médio dos pares aumenta o lócus de controle do indivíduo *i* em 0,16 desvios-padrão, quando medido em fator, e 0,26 desvios-

padrão, quando medido em escore, em média. Vimos que o modelo de turma também apresenta valor significativo deste coeficiente quando estimado por MQO simples (coluna 4). Quando comparamos este valor com o seu respectivo estimado por MQO em 2 estágios (coluna 8), temos que não só a significância aumenta – de 10% para 5% – mas também a magnitude – de 0,06 desvios-padrão para 0,26. Logo, ele encontrava-se subestimado.

Essa tabela também mostra que, apesar dos instrumentos fracos, não há divergência entre o que as estatísticas-t de inferência padrão e as estatísticas-AR de inferência robusta à instrumentos fracos dizem sobre a significância do coeficiente da variável de efeito de pares endógeno no segundo-estágio. Ambas apontam para a irrelevância desse parâmetro estimado com a amostra de amigos (colunas 5 e 6) e para a significância do mesmo estimado com a amostra de colegas de turma (colunas 7 e 8).

Esse resultado é consistente com a literatura de habilidades socioemocionais, que possui evidências de que o lócus de controle é uma habilidade fortemente influenciada pelo ambiente social (Strauser et al., 2002 e Schneewind, 1995). Essa relação encontrada também está de acordo com os estudos de efeito de pares: quando significativas, as estimativas de efeito de pares endógeno tendem a ser positivas (Hamilton et al., 2003; Mas & Moretti, 2009; Sacerdote, 2001). Apesar disso, a evidência de efeito de pares sobre a determinação do lócus de controle obtida neste trabalho diverge do que foi encontrado em Chikitani et al. (2015), que realiza o primeiro esforço de estimar esse impacto e não encontra evidências de que ele seja significativo. Esses autores também usam a base de dados dos alunos de Sertãozinho e o modelo empírico *linear-in-means* de efeito de pares, utilizando a informação de composição de turma para identificar os pares. Porém, no ano em que eles produziram o estudo, a coleta de dados de 2017 não havia sido feita. Relembre da Seção 3 que os questionários de personalidade, dentre eles o de Lócus de Controle de Tel Aviv, só foram introduzidos a partir da coleta de dados de 2012. Dessa forma, a estratégia de identificação proposta em Hanushek et al. (2003), utilizada neste trabalho, para lidar com o "Problema da Reflexão" subjacente ao modelo *linear-in-means* não encontrava-se viável para esses autores. Como alternativa, eles usam a abordagem de Bramoullé et al. (2009) e utlizam o lócus corrente de um terceiro indivíduo como instrumento para o lócus corrente dos pares atuais do indivíduo *i*, dado que esse terceiro indivíduo era par desses pares no passado e não é mais no período corrente. Contudo, vê-se que essa abordagem implica em uma restrição a mais do que a de Hanushek et al. (2003): além de supôr que o lócus de controle possui correlação serial intra-indivíduos,

ela também pressupõe a existência de efeito de pares endógeno no período passado, entre o terceiro indivíduo e os pares atuais do indivíduo *i*. Dado o conjunto de evidências que apontam para o resultado de efeito de pares endógeno positivo, mais a característica interpessoal do lócus de controle, a relação encontrada aqui oferece indício de que essa suposição adicional da abordagem de Bramoullé et al. (2009) foi forte demais ao ser utilizada no contexto da base de dados dos estudantes de Sertãozinho com informações apenas de 2008 e 2012 e do modelo *linear-in-means* que por si só já impõe uma série de limitações, conforme discutido na Seção 2.

Tabela 7: Segundo-Estágio

	Amigos		Turma		Amigos		Turma	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	LC Fator	LC Escore	LC Fator	LC Escore	LC Fator	LC Escore	LC Fator	LC Escore
Efeito de Pares Endógeno								
Lócus de Controle dos Pares	0,099	0,064	0,062	0,064*	0,035	0,031	0,162*	0,255**
	(1,65)	(1,36)	(1,68)	(1,79)	(0,13)	(0,10)	(1,84)	(2,18)
Características Individuais								
Menino	0,216	0,165	0,233***	0,252***	0,230	0,173	0,226***	0,240***
	(1,45)	(1,16)	(4,69)	(5,32)	(1,38)	(0,98)	(4,72)	(5,25)
Branco	0,008	0,005	0,019	0,020	0,009	0,004	0,008	-0,002
	(0,10)	(0,05)	(0,47)	(0,45)	(0,12)	(0,05)	(0,19)	(-0,04)
Mora com irmãos	0,041	0,070	0,089	0,062	0,052	0,075	0,083	0,047
	(0,29)	(0,45)	(1,47)	(0,99)	(0,40)	(0,49)	(1,44)	(0,80)
Pai completou o Ensino Médio	0,265**	0,217**	0,189***	0,183***	0,268***	0,219***	0,188***	0,187***
	(2,67)	(2,43)	(3,55)	(3,27)	(2,95)	(2,65)	(3,54)	(3,24)
Efeito de Pares Contextual								
Menino	-0,051	0,042	0,152	0,142	-0,042	0,047	0,025	-0,095
	(-0,35)	(0,30)	(1,43)	(1,27)	(-0,31)	(0,34)	(0,13)	(-0,56)
Branco	-0,134	-0,172	0,126	0,126	-0,137	-0,174*	0,197	0,241**
	(-1,09)	(-1,54)	(1,23)	(1,31)	(-1,21)	(-1,69)	(1,31)	(2,19)
Mora com irmãos	-0,072	-0,032	0,104	0,080	-0,087	-0,039	0,108	0,077
	(-0,34)	(-0,15)	(0,66)	(0,47)	(-0,39)	(-0,19)	(0,75)	(0,47)
Pai completou o Ensino Médio	-0,035	-0,021	-0,158	-0,159	-0,019	-0,012	-0,084	0,008
	(-0,29)	(-0,17)	(-1,43)	(-1,41)	(-0,14)	(-0,09)	(-0,60)	(0,06)
Observações	385	385	1124	1124	385	385	1124	1124
R^2	0,043	0,032	0,034	0,035	0,039	0,031	0,026	0,006
AR - Chi2					5,69	5,66	13,23	14,58
Prob > AR					0,338	0,341	0,021	0,012

Notas: A tabela reporta os resultados estimados da equação 2-3 de segundo-estágio (Seção 2). As colunas 1 a 4 apresentam os coeficientes estimados por MQO simples. As colunas 5 a 8 apresentam os coeficientes estimados por MQO em 2 estágios usando lócus de controle e características médias de 2012 dos pares como instrumentos para o lócus de controle médio de 2017 dos pares. As variáveis que dizem respeito aos pares estão em formato de média. Estatísticas-t em parênteses. Erros padrão calculados com cluster de escola. O coeficiente da constante foi omitido por simplificação visual. Significância: *** p<0.01, ** p<0.05, * p<0.10.

4.2 Robustez: Controles Adicionais

Nas últimas décadas, manifestou-se entre os psicólogos um consenso de que a maneira mais eficaz de analisar a personalidade humana consiste em observá-la em cinco dimensões, conhecidas como os Cinco Grandes Fatores, ou *Big Five:* Abertura a Novas

Experiências, Extroversão, Amabilidade, Conscienciosidade e Estabilidade Emocional, ou Neuroticismo (Santos & Primi, 2014). Para alguns autores (como exemplo, Almlund et al., 2011), o lócus de controle pode ser um atributo relacionado a esses cinco grandes domínios de personalidade, dada a maneira como se relaciona estatisticamente com medidas destes domínios. Assim, eu adiciono esses cinco traços de personalidade na equação 2-3 de segundo estágio, tanto para os indivíduos quanto para os pares. Como essas variáveis serão automaticamente incluídas como controles adicionais na equação 2-2 de primeiro-estágio, devido à forma como a técnica de estimação por MQO em 2 estágios opera, eu utilizo somente os seus valores de 2012 por motivos de exogeneidade. Além do questionário de Lócus de Controle de Tel Aviv, o BFI também foi aplicado aos alunos durante as coletas de dados de 2012 e 2017 (ver Seção 3). Esse instrumento foi desenhado para medir esses cinco grandes domínios utilizando 44 questões numa escala *likert*[5] de 5 pontos. Usando Teoria Clássica dos Testes[6] e correção por aquiescência[7] eu sou capaz de resumir as informações desse questionário em um escore para cada indivíduo em cada domínio. Eu utilizo então esses escores, calculados com as respostas de 2012, como regressores adicionais na equação 2-3 de segundo-estágio e, por conseguinte, na equação 2-2 de primeiro-estágio.

A Tabela 8 reporta os resultados do primeiro-estágio. Vemos que o coeficiente estimado do lócus de controle de 2012 dos pares permanece altamente estável em todas as especificações em termos de significância e magnitude com relação aos modelos sem os *Big Five* (ver Tabela 6). Assim, novamente temos que o lócus de controle médio de 2012 dos pares tem relação positiva significativa com o seu lócus de controle médio de 2017. Dentre os outros domínios de 2012 dos pares, que foram acrescentados, destaca-se o domínio da Estabilidade Emocional que apresentou valores significativamente positivos em todas as especificações. Além disso, chama-se atenção para o fato de nenhum dos cinco grandes domínios referentes aos indivíduos *i*'s terem apresentado coeficiente estatisticamente significativo. Isto corrobora com a validade da restrição de exclusão, que depende essencialmente dos indivíduos não conhecerem seus pares em 2012.

---

[5] Escala de pontos que operacionaliza o grau de concordância com afirmações sobre características pessoais em questionários de autorrelato usados em pesquisa em várias áreas (psicologia, educação, medicina e outras). O nome se refere ao seu inventor, o psicólogo Rensis Likert.

[6] Cada questão do BFI é associada a um dos cinco domínios. Como exemplo, o domínio da Amabilidade possui 9 itens. A Teoria Clássica dos Testes calcula o escore de cada um desses domínios para cada indivíduo simplesmente através da média das respostas individuais aos seus respectivos itens.

[7] Tendência dos entrevistados de concordar ou discordar em demasia ao responder às perguntas do questionário.

Novamente, as estatísticas-F de significância conjunta dos instrumentos indicam que eles são fracos. Assim, a significância do coeficiente de efeito de pares endógeno no segundo-estágio será interpretada à luz da estatística-AR de inferência robusta à instrumentos fracos.

Tabela 8: Primeiro-Estágio – Robustez, Controles Adicionais

	Amigos		Turma	
	(1)	(2)	(3)	(4)
	LC Fator	LC Escore	LC Fator	LC Escore
Lócus de Controle dos Pares, em 2012	0,116*	0,134**	0,196***	0,175**
	(1,93)	(2,15)	(4,01)	(2,07)
Características dos Pares, em 2012				
Menino	0,328	0,468	-0,351	-0,118
	(0,43)	(0,82)	(-0,62)	(-0,21)
Branco	-0,313*	-0,306	0,557**	0,644***
	(-1,72)	(-1,64)	(2,47)	(2,83)
Mora com irmãos	-0,169	-0,291	-0,229	-0,047
	(-0,61)	(-1,08)	(-0,42)	(-0,09)
Pai completou o Ensino Médio	0,297	0,254	0,353	0,193
	(1,51)	(1,48)	(1,32)	(0,77)
Big Five Individuais, em 2012				
Amabilidade	-0,031	-0,013	0,013	0,007
	(-0,28)	(-0,11)	(0,27)	(0,14)
Conscienciosidade	-0,035	-0,072	-0,042	-0,061
	(-0,31)	(-0,55)	(-0,73)	(-1,06)
Extroversão	-0,180**	-0,165**	0,038	0,029
	(-2,24)	(-2,10)	(0,99)	(0,77)
Estabilidade Emocional	-0,029	-0,011	0,065	0,069
	(-0,25)	(-0,09)	(1,29)	(1,54)
Abertura a Novas Experiências	-0,037	-0,019	-0,022	-0,021
	(-0,34)	(-0,18)	(-0,51)	(-0,43)
Big Five dos Pares, em 2012				
Amabilidade	-0,007	0,009	0,306	0,352*
	(-0,07)	(0,08)	(1,45)	(1,80)
Conscienciosidade	0,047	0,154*	0,199	0,048
	(0,45)	(1,66)	(0,75)	(0,15)
Extroversão	0,157	0,138	0,374**	0,285
	(1,25)	(1,06)	(2,10)	(1,36)
Estabilidade Emocional	0,202**	0,147*	0,524*	0,729**
	(2,04)	(1,70)	(1,94)	(2,56)
Abertura a Novas Experiências	0,159	0,158*	-0,286*	-0,123
	(1,40)	(1,85)	(-1,90)	(-0,88)
Observações	328	328	955	955
Estatística-F	0,977	1,698	4,582	3,185

Notas: A tabela reporta os resultados estimados da equação 2-2 de primeiro-estágio (Seção 2). Todas as variáveis que dizem respeito aos pares estão em formato de média. Além do lócus de controle médio de 2012 dos pares, eu incluo como instrumentos adicionais as características médias de 2012 desses pares. Além disso, são incluídos controles adicionais para os *Big Five* individuais e dos pares, também calculados em 2012. As variáveis dependentes representam o lócus de controle médio de 2017 dos pares, nas versões

fator e escore. Estatísticas-t em parênteses. Erros padrão calculados com cluster de escola. Os coeficientes dos controles usuais do segundo-estágio e da constante foram omitidos por simplificação visual. Significância: *** p<0.01, ** p<0.05, * p<0.10.

A Tabela 9 reporta os coeficientes estimados para o efeito de pares endógeno sobre o lócus de controle, agora levando em consideração os *Big Five*. Antes de mais nada, destaca-se a divergência, em todas as especificações, entre a significância do coeficiente de efeito de pares endógeno expressa pela estatística-t e pela estatística-AR. As estatísticas-t de inferência usual apontam para a irrelevância desse parâmetro em todos os modelos, ao passo que as estatísticas-AR de inferência robusta à instrumentos fracos dizem o total oposto. Conforme discutido no parágrafo acima, este segundo resultado será levado em conta. Sob a ótica desses resultados, temos agora coeficiente estatisticamente significativo de efeito de pares endógeno nos modelos que utilizam a amostra de amigos (colunas 1 e 2). Relativamente aos modelos que usam a amostra de amigos para estimar esse parâmetro por MQO em 2 estágios sem controlar para os *Big Five* (Tabela 7, colunas 5 e 6), essa nova estimativa apresenta aumento da ordem de 0,18 desvios-padrão no modelo que usa o lócus de controle fator e de 0,15 desvios-padrão no modelo que usa o lócus de controle escore. Os modelos que utilizam a amostra de turma permanecem estáveis com relação à relevância do coeficiente de efeito de pares endógeno (colunas 3 e 4). Contudo, observa-se uma queda considerável em sua magnitude – da ordem de 0,11 desvios-padrão – em relação à sua estimação sem controlar para os *Big Five* (Tabela 7, colunas 7 e 8).

A nova configuração de todas essas estimativas para o coeficiente de efeito de pares endógeno exibe valores desse parâmetro muito superiores nos modelos que definem os pares de acordo com os melhores amigos da escola, em comparação aos modelos que desenham esta relação através da composição da classe. Essa relação é plausível e sua justificativa se baseia na auto-seleção subjacente à formação dos pares na base dos melhores amigos (ver Seção 3.3). Espera-se que uma pessoa, ao escolher seus melhores amigos, leve em consideração os indivíduos com os quais possui comportamento e preferências em comum. Em função disso, o coeficiente de efeito de pares endógeno estimado com as informações dessa amostra, se significativo, deveria ser superior ao valor encontrado para este parâmetro utilizando a amostra de turma, em que supõe-se atribuição aleatória dos pares.

Essa instabilidade observada na magnitude do coeficiente que pretende estimar o efeito causal da relação entre o lócus de controle dos pares de um certo indivíduo *i* e o

lócus desse indivíduo é consequência da importante ligação entre o lócus de controle e os *Big Five*, conforme evidências do trabalho de Almlund et al. (2011). De fato, os domínios da Extroversão, Estabilidade Emocional e Abertura a Novas Experiências referentes aos indivíduos *i*'s apresentam coeficientes estatisticamente significativos em todas as especificações na Tabela 9. Entretanto, tal flutuação não modifica o significado principal desse coeficiente: novamente, temos evidência de efeito de pares endógeno positivo sobre lócus de controle.

Tabela 9: Segundo-Estágio – Robustez, Controles Adicionais

	Amigos		Turma	
	(1)	(2)	(3)	(4)
	LC Fator	LC Escore	LC Fator	LC Escore
Efeito de Pares Endógeno				
Lócus de Controle dos Pares	0,216	0,176	0,046	0,145
	(0,67)	(0,54)	(0,31)	(1,00)
Características Individuais				
Menino	0,332*	0,278	0,340***	0,364***
	(1,91)	(1,51)	(6,12)	(6,69)
Branco	-0,019	-0,024	-0,005	-0,014
	(-0,22)	(-0,24)	(-0,10)	(-0,23)
Mora com irmãos	-0,045	-0,016	0,072	0,047
	(-0,27)	(-0,10)	(0,93)	(0,63)
Pai completou o Ensino Médio	0,219**	0,156*	0,163***	0,159***
	(2,12)	(1,70)	(2,92)	(2,76)
Efeito de Pares Contextual				
Menino	-0,045	0,054	0,051	-0,054
	(-0,25)	(0,34)	(0,29)	(-0,31)
Branco	-0,115	-0,161	-0,063	-0,005
	(-1,07)	(-1,49)	(-0,51)	(-0,05)
Mora com irmãos	-0,045	0,018	0,221**	0,186
	(-0,21)	(0,09)	(1,97)	(1,29)
Pai completou o Ensino Médio	-0,075	-0,057	-0,118	-0,027
	(-0,59)	(-0,47)	(-0,74)	(-0,19)
Big Five Individuais, em 2012				
Amabilidade	-0,020	0,010	0,033	0,065*
	(-0,19)	(0,09)	(0,77)	(1,67)
Conscienciosidade	0,128	0,096	0,212***	0,188***
	(1,03)	(0,83)	(3,10)	(2,86)
Extroversão	0,292***	0,279***	0,137***	0,137***
	(2,62)	(2,60)	(3,06)	(2,90)
Estabilidade Emocional	0,147*	0,170**	0,126**	0,112**
	(1,84)	(2,21)	(2,35)	(2,57)
Abertura a Novas Experiências	0,178*	0,173**	0,079*	0,095**
	(1,96)	(2,08)	(1,79)	(2,21)
Big Five dos Pares, em 2012				
Amabilidade	0,120	0,112	0,006	-0,027
	(1,32)	(1,26)	(0,08)	(-0,29)
Conscienciosidade	-0,282**	-0,327***	-0,209	-0,221*
	(-1,99)	(-2,60)	(-1,62)	(-1,84)
Extroversão	-0,047	-0,022	0,022	-0,037
	(-0,44)	(-0,18)	(0,22)	(-0,41)
Estabilidade Emocional	-0,035	-0,018	-0,070	-0,129
	(-0,28)	(-0,14)	(-0,50)	(-1,20)
Abertura a Novas Experiências	0,104	0,144	0,105	0,141**
	(1,17)	(1,39)	(1,22)	(2,24)
Observações	328	328	955	955
R^2	0,103	0,099	0,086	0,085
AR - Chi2	20,63	21,33	14,21	16,25
Prob > AR	0,001	0,001	0,014	0,003

Notas: A tabela reporta os resultados estimados da equação 2-3 de segundo-estágio (Seção 2). Todos os coeficientes são estimados por MQO em 2 estágios usando lócus de controle e características médias de 2012 dos pares como instrumentos para o lócus de controle médio de 2017 dos pares. As variáveis que dizem respeito aos pares estão em formato de média. São incluídos controles adicionais para os *Big Five* de 2012 individuais e dos pares. Estatísticas-t em parênteses. Erros padrão calculados com cluster de escola. O coeficiente da constante foi omitido por simplificação visual. Significância: *** p<0.01, ** p<0.05, * p<0.10.

5. Conclusão

Desenvolvimentos recentes nos campos da educação, da economia, da psicologia e da neurociência mostram que as chamadas habilidades socioemocionais, são tão importantes quanto as medidas cognitivas na previsão de vários resultados da vida. Uma característica importante dessas habilidades não-cognitivas é se originar da interação entre predisposições biológicas e fatores socioambientais (Primi et al., 2016). Uma vez que fator biológico é uma variável exógena, e, portanto, fora do escopo de políticas públicas, entender melhor a interação entre ambiente social e habilidades não-cognitivas é crucial para o desenho de políticas que visam utilizar este mecanismo como forma de aprimorar essas habilidades. Este trabalho busca contribuir com a análise de como essa interação pode impactar uma habilidade socioemocional em específico, o lócus de controle, que mede quão pessoalmente responsável uma pessoa se sente com relação ao que acontece na sua vida.

A forma de interação social representada neste estudo é a de efeito de pares e o meu modelo empírico é o típico *linear-in-means* de análises sobre esse efeito. As questões econométricas inerentes a este arcabouço são resolvidas usando a metodologia proposta em Hanushek et al. (2003).

Uma primeira tentativa de explicar como o efeito de pares pode influenciar o lócus de controle foi feita em Chikitani et al. (2015). Contudo, os autores não encontram evidência de que esse efeito seja significativo, o que contraria a característica interpessoal desse construto. Eles utilizam a abordagem de Bramoullé et al. (2009) que implica em uma restrição a mais do que a de Hanushek et al. (2003).

Diferente de tal estudo, os resultados desta avaliação mostram evidência de efeito de pares endógeno significativamente positivo sobre o lócus de controle. Este resultado, que se manteve robusto à inclusão de controles adicionais para outros domínios socioemocionais, é consistente com a literatura de habilidades não-cognitivas, que possui evidências de que o lócus de controle é uma habilidade fortemente influenciada pelo ambiente social (Strauser et al., 2002 e Schneewind, 1995). Essa relação encontrada também está de acordo com os estudos de efeito de pares que mostram que, quando significativas, as estimativas de efeito de pares endógeno tendem a ser positivas (Hamilton et al., 2003; Mas & Moretti, 2009; Sacerdote, 2001).

Esta análise fornece importantes implicações políticas que podem potencialmente contribuir para o sucesso em variás áreas da vida da população. Em particular, este trabalho contribui com as políticas de desenho ótimo de turma. Com relação ao lócus de controle em específico, vimos que os colegas de turma desempenham um importante papel na formação dessa característica. Assim, a composição da classe é um dos meios pelos quais a escola pode aprimorar essa habilidade. Esse estudo também colabora com o debate atual sobre as diferentes formas de ensino. Deve-se ter em mente que essa oportunidade que as escolas tradicionais têm de aperfeiçoar certas características de um indivíduo através da sua interação social não é capaz de ser reproduzida nos ensinos a distância ou domiciliar, que vêm sendo discutidos recentemente.

Por último, este trabalho deixa em aberto a oportunidade de ser expandido à outras habilidades socioemocionais. Em Chikitani et al. (2015) os autores constroem uma tentativa de obter a relação causal entre efeito de pares e os *Big Five*, mas são impedidos pela baixa consistência dos dados. A amostra utilizada neste estudo é capaz de contornar essa situação por ser composta por indivíduos cuja idade média é 16 anos [8]– 5 anos a mais que a idade média dos indivíduos da amostra de Chikitani et al. (2015).

[8] Sabe-se que o grau de compreensão dos questionários de personalidade e, portanto, a qualidade das suas respostas, é função positiva da idade dos entrevistados (Soto et al., 2008).

Bibliografia

Almlund, M., Duckworth, A. L., Heckman, J. J., & Kautz, T. D. (2011). Personality psychology & economics. In E. Hanushek, S. Machin, & L. Woessman (Eds.), *Handbook of the Economics of Education* (Vol. 4, pp. 1–181). Amsterdam, The Netherlands: Elsevier.

Anderson, T. W., & Rubin, H. (1949). Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations. *The Annals of Mathematical Statistics,* 20(1):46-63.

Barrick, M. R., Stewart, G. L., Neubert, M. J., & Mount, M. K. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of applied psychology,* 83(3):377.

Benassi, V. A., Sweeney, P. D., & Dufour, C. L. (1988). Is there a relation between locus of control orientation and depression? *Journal of abnormal psychology,* 97(3):357.

Benet-Martínez, V., & John, O. P. (1998). Los cinco grandes across cultures and ethnic groups: Multitrait-multimethod analyses of the big five in spanish and english. *Journal of personality and social psychology,* 75(3):729.

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2002). *Cross-cultural psychology: Research and applications.* Cambridge University Press.

Betts, J. R., & Zau, A. (2004). Peer groups and academic achievement: Panel evidence from administrative data. *Unpublished manuscript.*

Boozer, M. A., & Cacciola, S. E. (2001). Inside the black box of project star: Estimation of peer effects. *Yale Economic Growth Center Discussion Paper,* (832).

Burke, M. A., & Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of Labor Economics,* 31(1):51–82.

Bramoullé, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of econometrics,* 150(1):41–55.

Bursztyn, L., & Jensen, R. (2015). How does peer pressure affect educational investments? *The Quarterly Journal of Economics,* 130(3), 1329-1367.

Carman, K. G., & Zhang, L. (2012). Classroom peer effects and academic achievement: Evidence from a Chinese middle school. *China Economic Review,* 23(2), 223-237.

Chikitani, M., Ponczek, V., & Pinto, C. (2015). Peer Effects on Locus of Control. *Anais do 37º Encontro Brasileiro de Econometria.*

Crosby, J. W. Test review: F. M. Gresham & S. N. Elliot (2011). Social Skills

Improvement System Rating Scales. Mineapollis, MN: NCS Pearson, 2008. *Journal of Psychoeducational Assesment.* 29, pp. 292-296.

De Brabander, B., & Boone, C. (1990). Sex differences in perceived locus of control. *The Journal of social psychology,* 130(2):271–272.

Duckworth, A. L., & Seligman, M. E. (2005). Self-discipline outdoes iq in predicting academic performance of adolescents. *Psychological science,* 16(12):939–944.

Findley, M. J., & Cooper, H. M. (1983). Locus of control and academic achievement: A literature review. *Journal of personality and social psychology,* 44(2):419.

Freed, R. D., & Tompson, M. C. (2011). Predictors of parental locus of control in mothers of pre-and early adolescents. *Journal of Clinical Child & Adolescent Psychology,* 40(1):100– 110.

Hamilton, B. H., Nickerson, J. A., & Owan, H. (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of political Economy,* 111(3):465–497.

Hampson, S. E., & Friedman, H. S. (2008). Personality and health: A lifespan perspective.

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of applied econometrics,* 18(5):527–544.

Heckhausen, J., & Schulz, R. (1995). A life-span theory of control. *Psychological review,* 102(2):284.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics,* 24(3):411–482.

Hoxby, C. (2000). *Peer effects in the classroom: Learning from gender and race variation* (No. w7867). National Bureau of Economic Research.

IBGE (2016). Produto Interno Bruto a preços correntes e Produto Interno Bruto per capita segundo as Grandes Regiões, Unidades da Federação e os Municípios – 2016. Consultado em Abril de 2019.

IBGE (2018). Estimativas da população residente nos municípios brasileiros com data referência em 1º de julho de 2018. Consultado em Abril de 2019.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big-Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of Personality: Theory and research* (pp. 114–158). New York, NY: Guilford Press.

John, O. P., & Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality: Theory and research* (Vol. 2, pp. 102–138). New York: Guilford Press.

Judge, T. A., & Bono, J. E. (2001). Relationship of core self-evaluations traitsself-esteem, generalized self-efficacy, locus of control, and emotional stability with job satisfaction and job performance: A meta-analysis. *Journal of applied Psychology,* 86(1):80.

Kline, P. (2013). *Handbook of Psychological Testing.* Routledge.

Lee, L.-f. (2007). Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics,* 140(2):333–374.

Lefcourt, H. M. (1982). Locus of Control: Current Trends in Theory and Research. *Psychology Press.*

Manger, T., & Eikeland, O.-J. (2000). On the relationship between locus of control, level of ability and gender. *Scandinavian Journal of psychology,* 41(3):225–229.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies,* 60(3):531–542.

Mas, A., & Moretti, E. (2009). Peers at work. *American Economic Review,* 99(1):112–145.

Milgram, N. A., & Milgram, R. M. (1975). Dimensions of locus of control in children. *Psychological Reports,* 37(2):523–538.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences,* page 201010076.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: big five correlates of gpa and sat scores. *Journal of personality and social psychology,* 93(1):116.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin,* 135, 322–338. doi: 10.1037/a0014996.

Primi, R., Santos, D., John, O. P., & De Fruyt, F. (2016). Development of an inventory assessing social and emotional skills in Brazilian youth. *European Journal of Psychological Assessment.*

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied,* 80(1):1.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly journal of economics,* 116(2):681–704.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they

and how much do we know thus far? In *Handbook of the Economics of Education* (Vol. 3, pp. 249-277). Elsevier.

Santos, D., & Primi, R. (2014). Desenvolvimento socioemocional e aprendizado escolar: uma proposta de mensuração para apoiar políticas públicas. *Relatório sobre resultados preliminares do projeto de medição de competências socioemocionais no Rio de Janeiro. São Paulo: OCDE, SEEDUC, Instituto Ayrton Senna*.

Schneewind, K. A. (1995). Impact of family processes on control beliefs. *Self-efficacy in changing societies,* pages 114–148.

Soto, C. J., John, O. P., Gosling, S. D., and Potter, J. (2008). The developmental psychometrics of big five self-reports: acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology,* 94(4):718.

Spearman, C. (1904). "general intelligence," objectively determined and measured. *The American Journal of Psychology,* 15(2):201–292.

Stock, J. H., & Wright, J. H. (2000). GMM with Weak Identification. *Econometrica,* 68(5):1055-1096.

Stock, J. H., Wright, J. H., & Yogo, M. (2002). A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business and Economic Statistics,* 20(4):518-529.

Strauser, D. R., Ketz, K., & Keim, J. (2002). The relationship between self-efficacy, locus of control and work personality. (self-efficacy and locus of control). *The journal of rehabilitation,* 68(1):20.

Uba, L. (2003). *Asian Americans: Personality patterns, identity, and mental health.* Guilford Press.

Urzua, S. (2008). Racial labor market gaps the role of abilities and schooling choices. *Journal of Human Resources,* 43(4):919–971.

Vigdor, J. L., & Nechyba, T. S. (2007). Peer effects in North Carolina public schools. In: Woessmann, L., Peterson, P. E. (Eds.), Schools and the Equal Opportunity Problem. *The MIT Press,* pp. 73-101.

Wallston, B. S., & Wallston, K. A. (1978). Locus of control and health: a review of the literature. *Health Education & Behavior,* 6(1):107–117.

Apêndice A

A.1 Análise Fatorial

A análise fatorial é uma técnica estatística cujo objetivo principal é reduzir a dimensão de um conjunto de dados enquanto retém o máximo possível da informação original. Esse instrumento também é muito útil quando as variáveis de interesse são difíceis de avaliar diretamente, e toda a informação disponível sobre elas consiste em medidas ruidosas que possuem variações em comum. Nesse caso, a análise fatorial é capaz de separar o ruído da variação sistemática e resumi-la em uma ou mais variáveis latentes. Os fatores representam essas dimensões latentes que resumem ou explicam o conjunto de medidas observadas, e os coeficientes associados a eles são chamados de *loadings*. Esse mecanismo pressupõe válida a hipótese de que tais medidas são combinações lineares dos fatores e um termo de erro.

O uso dessa técnica em psicologia tem início na pesquisa de Charles Spearman sobre inteligência (Spearman, 1904), comumente conhecida como *g theory*. Ao observar que os escores de diferentes testes de habilidade mental – habilidades em matemática, verbais, raciocínio lógico, entre outras – estavam correlacionados positivamente, Spearman postulou a hipótese de que todos aqueles resultados estavam medindo conjuntamente e de forma implícita a inteligência de cada pessoa. Através da análise fatorial ele utilizou todos aqueles escores para extrair um fator comum de inteligência que ele denominou *g*. Além de seu uso na avaliação empírica dos construtos teóricos, a análise fatorial também foi empregada como base para o desenvolvimento da teoria do Modelo dos Cinco Fatores (John et al., 2008), também conhecida como *Big Five Personality traits.*

Nesse estudo, eu utilizo a análise fatorial para extrair fatores que poderiam ser interpretados como uma variável de interesse não observável, o lócus de controle. Uma deficiência específica dessa análise é que ela não produz soluções únicas. Como sua meta é maximizar a variância explicada, dado um conjunto de fatores, qualquer transformação unitária de uma solução também é uma solução. Um procedimento padrão, chamado de rotação varimax, consiste em usar uma transformação desses fatores que os diferencie ao máximo, de modo que um item que possui seu *loading* forte em um fator tenha ele fraco nos outros. Dessa forma, os fatores passam a ser interpretados de acordo com os itens que eles estão associados.

Eu recorro à literatura de psicologia para avaliar o número de fatores esperados e sua interpretação. Através dessa referência eu analiso a qualidade da informação que posso extrair dos questionários psicológicos em meus dados. Como exemplo, se um questionário foi projetado para produzir um certo número de fatores, um afastamento desse padrão é tido como evidência de que os dados podem ter alguns problemas.

Como cada fator adicional explica uma parte menor da variação nos dados, eu determino o número de fatores que serão extraídos com base no momento em que o fator marginal não é mais importante. No entanto, a própria análise fatorial não fornece nenhuma orientação para avaliar quais fatores de fato importam. Assim, vários métodos foram propostos. Aqui, eu comparo os autovalores associados a cada fator e decido o número de fatores no momento em que a diferença entre dois autovalores subsequentes fica muito pequena. Essa abordagem é conhecida como *scree plot test*, porque é usualmente apresentada através de um gráfico dos fatores e seus respectivos autovalores. Como essa análise é sujeita a arbitrariedades, eu também utilizo o critério de Kaiser que sugere manter apenas os fatores que possuem autovalor acima de um.

A.2 Análise Psicométrica

A.2.1 Lócus de Controle de Tel Aviv

O Lócus de Controle de Tel Aviv é um questionário desenhado para avaliar em que medida indivíduos atribuem situações correntemente vividas a decisões e atitudes por eles tomadas no passado (lócus interno), ou ao acaso, sorte ou ações e decisões tomadas por terceiros (lócus externo). Esse questionário possui uma estrutura muito simples, contendo itens com apenas 2 alternativas de resposta: uma do tipo interna e outra do tipo externa. Ele foi aplicado aos alunos do sistema de ensino público e privado de Sertãozinho durante a coleta de dados de 2012 e 2017. Devido os alunos estarem mais maduros na última pesquisa realizada (idade média de 16 anos), os itens do questionário de 2017 não são os mesmos do questionário de 2012. Além disso, o questionário de 2012 conta com 24 questões enquanto que o de 2017 com 21. Conforme mostrado na Figura 1 (Seção 3.1), um único fator é capaz de explicar grande parte da variância presente em cada um desses questionários. Assim, temos 1 fator de lócus de controle que resume o questionário de

2012 e 1 fator que resume o questionário de 2017. Esse resultado é consistente com a literatura de psicologia, que interpreta este único fator de lócus de controle como uma medida de quão internos os indivíduos são.

As Tabelas 10 e 11 apresentam os *loadings* e seus sinais esperados para cada item dos questionários de 2012 e 2017, respectivamente. Uma vez que cada resposta do tipo interno vale 1 ponto, espera-se que todos os *loadings* sejam positivos. Essa condição é totalmente atendida para o fator extraído do questionário de 2017 (Tabela 11). Contudo, 3 dos 24 *loadings* do fator calculado a partir do questionário de 2012 não possuem o sinal antecipado (*loadings* 11, 12 e 22 na Tabela 10). Vale ressaltar, entretanto, que esses três parâmetros estão dentre os 5 cujo valor absoluto é menor do que 0,1. Assim, eu argumento que a inconsistência desses valores não compromete a minha medida de lócus de controle extraída da análise fatorial.

Essa imprecisão observada no fator de lócus de controle de 2012 pode ser explicada pelo *alpha de Cronbach.* Esse indicador é uma medida de confiabilidade dos dados que são utilizados na análise fatorial. A regra de bolso usual sugere que *alpha* deve ser maior do que 0,70 (Kline, 2013). Para o questionário de Lócus de Controle de Tel Aviv de 2012 temos *alpha* igual à 0,52 enquanto que para o questionário de 2017 esse valor é de 0,69, muito mais próximo da regra de bolso.

Tabela 10: *Loadings* dos itens do Lócus de Controle de Tel Aviv, 2012

Item	Loading	Sinal Esperado
1	0,0257	+
2	0,3325	+
3	0,3807	+
4	0,3263	+
5	0,2258	+
6	0,2615	+
7	0,1916	+
8	0,2537	+
9	0,1901	+
10	0,4208	+
11	-0,0157	+
12	-0,0386	+
13	0,1030	+
14	0,4773	+
15	0,1791	+
16	0,1378	+
17	0,3668	+
18	0,2932	+
19	0,2576	+
20	0,1593	+
21	0,4449	+
22	-0,0869	+
23	0,2953	+
24	0,0267	+

Notas: A tabela apresenta os *loadings* associados ao fator de lócus de controle extraído da análise fatorial exercida às respostas do questionário de Lócus de Controle de Tel Aviv aplicado em 2012.

Tabela 11: *Loadings* dos itens do Lócus de Controle de Tel Aviv, 2017

Item	Loading	Sinal Esperado
1	0,2301	+
2	0,4230	+
3	0,4875	+
4	0,3690	+
5	0,2833	+
6	0,1679	+
7	0,3446	+
8	0,4177	+
9	0,1587	+
10	0,2958	+
11	0,2390	+
12	0,1312	+
13	0,1144	+
14	0,3498	+
15	0,5023	+
16	0,2541	+
17	0,4182	+
18	0,2912	+
19	0,3882	+
20	0,1400	+
21	0,5159	+

Notas: A tabela apresenta os *loadings* associados ao fator de lócus de controle extraído da análise fatorial exercida às respostas do questionário de Lócus de Controle de Tel Aviv aplicado em 2017.

A.3 Questionários

Esta seção apresenta os questionários completos de Lócus de Controle de Tel Aviv que foram aplicados em 2012 e 2017. Os itens em negrito são as respostas do tipo interno.

A.3.1 Lócus de Controle de Tel Aviv, 2012

1. Suponha que você mudou de bairro, fez amizade com algumas crianças novas e se tornou o chefe do grupo. Isso aconteceu porque: a) você encontrou o grupo certo. **b) você tem as qualidades de chefe.**

2. Quando um de seus irmãos ou amigos aceita uma de suas sugestões é porque: **a) você explicou a razão de sua sugestão e o convenceu de que você está certo.** b) ele fica cansado de discutir com você.

3. Suponha que você se saiu bem na escola em uma matéria que geralmente você acha difícil. Isto foi porque: a) você teve sorte e alguém o ajudou. **b) você se esforçou mais do que de costume.**

4. Se o seu professor acha bom o seu trabalho na escola é porque: a) ele gosta de você. **b) seu trabalho na escola é bom mesmo.**

5. Imagine que você queira assistir a um filme e seus amigos concordem em ir com você. Eles concordaram porque: a) não havia outro filme que valia à pena ver. **b) você os convenceu de que valia à pena assistir àquele filme.**

6. Suponha que seus pais deram uma bronca em você. Foi porque: a) eles estavam nervosos naquele momento. **b) você fez alguma coisa errada.**

7. Quando você tira nota baixa numa prova é porque: **a) você não estava preparado para a prova.** b) a prova era muito difícil.

8. Quando seus pais brigam com você é porque: a) eles estão de mau humor. **b) geralmente a culpa é sua.**

9. Suponha que você mude para outro bairro e não consiga fazer novos amigos de sua idade. Isto acontece porque: a) você não teve sorte, pois eles não deixam ninguém entrar no grupo. **b) você não se esforçou bastante para fazer novos amigos.**

10. Se você não passa de ano é porque: **a) seu trabalho na escola não foi suficientemente bom.** b) o professor tem alguma coisa contra você.

11. Quando você briga com um de seus irmãos é porque: a) geralmente a culpa é dele. **b) geralmente a culpa é sua.**

12. Quando seus irmãos não concordam em brincar daquilo que você quer, é porque: a) eles nunca querem brincar da mesma coisa que você. **b) você não aceita brincar daquilo que eles querem.**

13. Suponha que você não tenha muitos amigos. Isto é porque: **a) você não tem muita facilidade em fazer amigos.** b) sempre existem crianças que não gostam de você.

14. O professor elogiou seu comportamento durante um passeio da escola. Isto foi porque: **a) seu comportamento no passeio mereceu mesmo ser elogiado.** b) seu professor estava de bom humor.

15. Depois de discutir com você por muito tempo, seus pais lhe permitiram fazer uma viagem com seus amigos. Isto foi porque: **a) você os convenceu.** b) eles ficaram cansados de discutir com você.

16. Imagine que você encontrou um grupo de crianças na praia. Aí você sugeriu que todos brincassem de um certo jogo e todos aceitaram a sugestão. Isto aconteceu porque: **a) você conseguiu convencê-los.** b) eles já queriam brincar daquele jogo.

17. Suponha que seus pais digam que você está indo bem na escola. Isto é porque: a) eles gostam de qualquer coisa que você faça. **b) seu trabalho na escola é bom.**

18. Quando você esquece alguma coisa que foi dita em aula, isto provavelmente acontece porque: **a) você não fez bastante esforço para lembrar.** b) o professor não explicou muito bem.

19. Se você é muito querido entre os amigos é por causa: a) das pessoas com quem você anda e das coisas que você tem. **b) do esforço que você faz para que os amigos gostem de você.**

20. Quando você discute com seus amigos é porque: **a) você não desiste de suas ideias.** b) seus amigos são cabeças duras.

21. Quando você se sai muito bem em uma prova na escola, isso acontece porque: a) você teve sorte. **b) você estudou muito para a prova.**

22. Quando você não chega a um acordo com outras crianças é porque: **a) você não aceita a opinião delas.** b) as outras crianças geralmente não aceitam as suas opiniões.

23. Imagine que seu professor não tenha uma boa opinião a seu respeito. Isto é porque: **a) você não se saiu bem na escola no ano passado.** b) ele não gosta de você.

24. Se seu irmão, sua irmã ou seu melhor amigo concordaram em dar alguma coisa deles a você é porque: a) ele ou ela não estão precisando daquela coisa. **b) você soube pedir.**

A.3.2 Lócus de Controle de Tel Aviv, 2017

1. Você acredita que a maioria dos problemas se resolverão sozinhos se você apenas não lidar com eles? a) sim. **b) não.**

2. Você é frequentemente acusado por coisas que não foram sua culpa? a) sim. **b) não.**

3. Você sente que muitas vezes não vale a pena se esforçar porque, de qualquer modo, as coisas nunca dão certo mesmo? a) sim. **b) não.**

4. Você sente que os pais escutam o que os filhos têm a dizer? **a) sim.** b) não.

5. Quando você recebe uma punição, você costuma achar injusta? a) sim. **b) não.**

6. Você acha difícil fazer um amigo mudar de opinião? a) sim. **b) não.**

7. Você sente que é quase impossível mudar a opinião que seus pais têm sobre as coisas? a) sim. **b) não.**

8. Quando faz alguma coisa errada, você sente que existe muito pouco que você pode fazer para consertar? a) sim. **b) não.**

9. Você acredita que as pessoas boas nos esportes já nasceram assim? a) sim. **b) não.**

10. Você sente que a melhor maneira de lidar com os problemas é apenas não pensar neles? a) sim. **b) não.**

11. Quando algum colega decide te bater, você sente que pouco pode fazer para impedi-lo? a) sim. **b) não.**

12. Você já sentiu pessoas serem más contigo sem motivo algum? a) sim. **b) não.**

13. Você acredita que pode mudar o que acontecerá com você amanhã pelo que faz hoje? **a) sim.** b) não.

14. Você acredita que quando coisas ruins vão acontecer, elas simplesmente acontecem, não importa o que você tente fazer para impedi-las? a) sim. **b) não.**

15. Você sente que não adianta tentar fazer as coisas do seu jeito em casa? a) sim. **b) não.**

16. Quando algum colega decide ser seu inimigo, você acredita que pode fazer pouca coisa para mudar a situação? a) sim. **b) não.**

17. Você sente que tem pouca influência em casa para decidir sobre em que canal deve estar a TV ou sobre o que comer? a) sim. **b) não.**

18. Quando alguém de sua idade não gosta de você, acredita que há pouco o que fazer para mudar a situação? a) sim. **b) não.**

19. Você sente que é inútil se esforçar na escola porque a maioria dos outros alunos é mais inteligente que você? a) sim. **b) não.**

20. Você é o tipo de pessoa que acredita que planejar o futuro faz as coisas darem mais certo? **a) sim.** b) não.

21. Você sente que tem pouca influência nas decisões do cotidiano em sua casa? a) sim. **b) não.**