

Raposo, Isabel Pessoa de Arruda; Gonçalves, Michela Barreto Camboim

Article

Peer effects and educational achievement: Evidence of causal effects using age at school entry as exogenous variation for Peer quality

EconomiA

Provided in Cooperation with:

The Brazilian Association of Postgraduate Programs in Economics (ANPEC), Rio de Janeiro

Suggested Citation: Raposo, Isabel Pessoa de Arruda; Gonçalves, Michela Barreto Camboim (2020) : Peer effects and educational achievement: Evidence of causal effects using age at school entry as exogenous variation for Peer quality, EconomiA, ISSN 1517-7580, Elsevier, Amsterdam, Vol. 21, Iss. 1, pp. 18-37, <https://doi.org/10.1016/j.econ.2020.04.003>

This Version is available at:

<https://hdl.handle.net/10419/266947>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer effects and educational achievement: evidence of causal effects using age at school entry as exogenous variation for Peer quality

Isabel Pessoa de Arruda Raposo*, Michela Barreto Camboim Gonçalves

Researcher at the Fundação Joaquim Nabuco – Ministry of Education, Rua Dois Irmãos, 92 - Ed. Anexo Anízio Teixeira - Apipucos – Recife/PE CEP: 52071-440, Brazil

Received 12 March 2018; received in revised form 5 June 2019; accepted 23 April 2020

Available online 27 May 2020

Abstract

This paper evaluates the diffusion of peer effects on academic achievement of 4th grade students in the Brazilian public school system. Using data from *Prova Brasil* 2013, the identification strategy builds on the use of an IV approach, in which the instruments for peers' performance are the proportions of classmates born in the second semester of the year (and alternatively, in each quarter). The idea behind the instruments is that compulsory school enrolment laws generate variation in the child's age at school entry, which, in turn, make the date of birth within the year an important determinant of educational achievement and, at the same time, plausibly exogenous to the quality of the student's peers. The results demonstrate that classrooms with higher proportions of peers born in the 2nd semester (*started school at a relatively older age*) tend to perform better, on average, than those that concentrate children born in the 1st semester, even after the inclusion of a wide range of control variables. For the math and Portuguese language evaluations, a one standard deviation increase in the classmates' test scores improves individual achievement by 30% of a SD.

Keywords: Peer effects; Instrumental variables; School enrolment laws

JEL Classification: I20; I21; C13

© 2020 The Authors. Production and hosting by Elsevier B.V. on behalf of National Association of Postgraduate Centers in Economics, ANPEC. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

One of the prominent issues within the economics of education is devoted to understanding the role of peers in educational outcomes¹. The behavioral influence received from friends in the social interaction process might affect educational results not only during the schooling period, but also latter in life, having an effect on standards of educational attainment to employment decisions. If the individual academic performance increases as the average class performance increases, then peer effects act as a social multiplier providing, as a result, implications for educational policies (Manski, 1993). These policies can guide group formation, and can be applied to school unit organization or to the whole educational system design.

* Corresponding author.

E-mail address: i.raposo@hotmail.com (I.P.d.A. Raposo).

¹ The sense of the term *peers* in this article refers to all members belonging to the same classroom.

Despite the fact that the role of peers in educational outcomes has already been vastly exploited in the international literature, in Brazil such literature is still scarce with few exceptions, such as the works of [Koppensteiner \(2018\)](#); [Raposo \(2015\)](#); [Firpo et al. \(2015\)](#); [Oliveira \(2015\)](#), and [Pinto \(2008\)](#). The use of instrumental variables is the most frequently reported method in the literature², and the identification hypothesis is based on the use of an appropriate instrument, which determines peers' educational performance, but which is nonetheless exogenous to any unobserved factor related to the potential individual outcome. For instance, [Firpo et al. \(2015\)](#) and [Pinto \(2008\)](#) build their empirical approach using the allocation rule of the students into classrooms and the teacher's wage as instruments to the quality of peers. In France, [Goux and Maurin \(2007\)](#) use spatially represented data on households in France and attributes the birth month of neighboring adolescents as a tool for peer influence. Their idea is that the date of birth is an important determinant of children's initial school performance and is exogenous to the quality of the neighborhood where they live.

In the line of [Goux and Maurin \(2007\)](#) work, our study provides additional evidence for the Brazilian context and offers a new contribution for the literature since it develops the use of instruments not yet tested for this country. Specifically, our study seeks to answer the following question: "what are the impacts of peer effects on 4th grade students' academic outcomes from Brazilian public schools?" In order to address to the endogenous influence of classmates, we will instrument the quality of the student body with the proportions of peers born in the second semester of the year (and alternatively, with the proportion of births in each quarter), as proposed by the work of [Goux and Maurin \(2007\)](#)³.

The intuition behind such instrument is based on the relative age literature⁴, which demonstrates that date of birth within the year is an important determinant of educational achievement and, at the same time, it is plausibly exogenous to the quality of the student's peers. The idea is that educational achievement is related to age at school entry because of the age variation induced by compulsory school enrolment laws. In many countries there are enrolment laws that define the age at school entry based on a specified cut-off date, which induce some children to enter at a relatively older (or younger) age. For instance, in the US and other European countries enrolment laws require students to enter school in the fall of the year in which they turn six. In Chile and other Latin American countries this cut-off is the first semester of the year. In Brazil for the year 2009, when most of the students from our study sample entered the school, there was a Federal law establishing that a child should enter in the 1st year if he/ she turned six years old until the beginning of the school year, which usually occurred in March, those born after that should wait until the next year ([BRASIL, 2008](#)⁵). From 27 Federal States, only six of them had different resolutions regarding the cut-off age for school entry, defined by their local boards of Education (see appendix [Table A2](#)).

For the empirical estimations, we use data from the *Prova Brasil*⁶ exams for the year of 2013 and find that classrooms with higher proportions of peers born in the 2nd semester tend to perform better than those that concentrate children

² Usually the empirical identification strategies rely either on the use of (i) exogenous instruments for peer's educational outcome ([Sund, 2009](#); [Ding and Lehrer, 2007](#); [Vigdor and Nechyba, 2007](#); [Hanushek et al., 2003](#); [Evans et al., 1992](#); [Case and Katz, 1991](#)); (ii) experimental or quasi-experimental data ([Oosterbeek and Van Ewijk, 2014](#); [Eisenkopf et al., 2011](#); [Duflo et al., 2008](#); [Zimmerman, 2003](#); [Sacerdote, 2001](#)) and (iii) mostly recently on the architecture of social networks (e.g. [Patacchini and Venanzoni, 2014](#); [Badev, 2013](#); [Patacchini et al., 2011](#); [Mele, 2010](#); [Calvó-Armengol et al., 2009](#); [Bramoullé et al., 2009](#); [Ballester et al., 2006](#)).

³ Similar IV approach was firstly proposed by [Angrist and Krueger \(1991; 1992\)](#) and [Angrist and Imbens \(1995\)](#). [Angrist and Krueger \(1991; 1992\)](#) show that students' quarter of birth interacts with compulsory enrolment laws in the US and age at school entry to generate variation in years of schooling. They show that educational attainment is related to age at school entry because children who enter at an older age (born in early quarters) are permitted to drop out after having completing less schooling than children who enter at a younger age. [Angrist and Imbens \(1995\)](#) based on such evidence also use student's quarter of birth as an instrument for schooling in a wage equation.

⁴ As pointed out by [Goux and Maurin \(2007\)](#), there is a literature on the influence of date of birth within the year (called the relative age effect) going back a few decades ([Barnsley et al., 1985](#); [Allen and Barnsley, 1993](#)). It is shown that relative age has an impact on achievement in competitive activities (Hockey, Soccer), on achievement at school, on emotional development and even on the probability of committing suicide.

⁵ According to legal guidance and standards set by the Brazilian National Board of Education / Ministry of Education in Resolution CNE/CEB 4/2008, to enter the first year of elementary education, the child must be six (6) years old by the beginning of the school year. Some States had different resolutions defined by their local boards of Education (see appendix [Table A2](#)). In the year of 2010, a new resolution explicitly defined the cut-off date on March, 31st at the Resolution CNE/CEB 6/2010, according to the following articles: Article III – To enter the first year of elementary education, the child must be six (6) years old by March 31st before they are enrolled. Article IV – Children that reach the age of six (6) years after the date set in Article III must be enrolled at pre-school.

⁶ *Prova Brasil* is a census exam that evaluates students' performances in Portuguese and mathematics on elementary education (4th and 8th grades) from all urban public schools in Brazil. It is the Brazilian National System to Evaluate Basic Education from the Brazilian Ministry of Education.

born in the 1st semester. For the math and Portuguese language evaluations, a one standard deviation increase in the peers' test scores improves individual achievement by 30% of a standard deviation.

The rest of the paper is organized into four additional sections. Section two describes *Prova Brasil* dataset and offers a descriptive analysis from the sample variables use. The third one presents the IV identification strategies using information on peer's date of birth to estimate peer effects on academic performance. Section four provides the estimated results, and some robustness tests to verify the validity of the instruments. Finally, the last one brings together the main findings of the article.

2. The dataset and descriptive statistics of the variables

This paper uses data from the Brazilian National System to Evaluate Basic Education also known as *Prova Brasil*, for the year 2013. Created in 2005 by the *National Institute of Educational Research Anísio Teixeira* from the Brazilian Ministry of Education (MEC/Inep), the *Prova Brasil* evaluation consists of a census assessment involving students from 4th and 8th grades of elementary schools in Brazil, and offers a nationwide measurement of the quality of education. Schools with at least twenty students enrolled in the assessed grades are eligible to participate in the assessment and their pupils are submitted to two types of exams: mathematics and Portuguese language. At the end of the tests, the student answers a questionnaire containing questions about date of birth, age and sex, as well as, scholar history, studying habits and family socioeconomic background. Interviewers also administer two other questionnaires for the school Principals and teachers, in which they collect a large set of information on internal and external aspects of the school environment.

For the present study we consider only the sample of 4th grade students from public schools that had at least ten students being evaluated in the *Prova Brasil* exams of 2013. The analysis could be conducted only after this year, in order to create a sample composed of children who entered school when the Federal law of 2008 was valid. As a consequence, each student in our dataset who was 10 years old in the beginning of the school year in 2013, entered in the first grade when he or she was 6 year old, as required by the Federal Law of 2008, which started to be valid in 2009. On average, children born after March start school relatively older. For our sample study, the average age at school entry in the 1st semester is 6.3 years old, while in the 2nd semester, it is 6.8. This difference is statistically significant at less than 1% level. We excluded all the schools located in the six states that did not follow the Resolution CNE/CEB 4/2008 (BRASIL, 2008) in the year of 2009 (the school entry year of the sample students), since for those units the students born in the second semester are not necessarily older than their peers, which breaks down the intensity status of the instrument.

After the exclusion of individuals with inadequate or missing information, the final sample consists of about 1 million students from 4th grade and 47,902 schools distributed over the Brazilian states and the Federal District. Table 1 presents the descriptive statistics of the variables selected for the estimated models of this paper. The description of each variable is provided on the appendix section, Table A1. The scale of the *Prova Brasil* evaluations ranges from 0 to 500 and indicates in which position the student of a given school unit is located. Such position reveals the skills the students have already acquired in the learning process, or those that are still being developed. For the 4th grade, the minimum desired level for mathematics should be 225, while in Portuguese language this inferior limit should be 200. In the case of our sample, as presented in Table 1, the 4th graders achieve in average below the minimum level for both the math and Portuguese language evaluations, with average test score of 213 in mathematics, and 197 in Portuguese language.

3. Identification using information on peer's date of birth

The central question of this paper is whether student's educational achievement is affected by the performance of other peers from the same classroom. The identification strategy builds on the use of an instrumental variable which determines peers' performance at school, but which is nonetheless exogenous to any unobserved factor related to the potential individual outcome. Specifically, the instruments being considered here are the proportions of peers in the student's classroom who were born in the second semester of the year, as proposed by the work of Goux and Maurin (2007). These authors use the distribution of neighbors' dates of birth as instrumental variable to identify the effect of neighbors' educational advancement on an adolescent's performance at school. Their IV estimates suggest that a one standard deviation increase in the proportion of neighbors who have already been held back a grade at age 15 increases

Table 1
Descriptive statistics of the variables.

	N	Average	Standard Deviation
<i>Dependent Variables (Y)</i>			
Math grade at <i>Prova Brasil</i> exam (4 th grade)	1,053,506	207.87	50.00
Portuguese Language grade at <i>Prova Brasil</i> exam (4 th grade)	1,053,506	192.52	48.57
<i>Instruments (Z)</i>			
Proportion of peers born in the 2 nd semester	1,506,897	0.48	0.12
<i>Peers' Educational Achievement (\bar{Y})</i>			
Peers math grade (4 th grade)	1,053,506	206.25	30.05
Peers Portuguese language grade (4 th grade)	1,053,506	190.94	27.54
<i>Students' Characteristics (X)</i>			
Semester II	1,506,897	0.49	0.50
Age	1,036,496	4.02	1.12
Male sex	1,020,405	0.50	0.50
White	884,898	0.31	0.46
Black	884,898	0.12	0.33
Student lives with the mother	1,040,255	0.89	0.31
Educational level (mother)	675,279	3.79	1.50
Literacy (mother)	1,048,732	0.94	0.23
Student lives with the father	1,033,909	0.63	0.48
Educational level (father)	563,820	3.70	1.59
Literacy (father)	1,031,954	0.91	0.29
Preschool	1,021,837	1.90	0.90
Grade retention	1,506,897	1.39	0.64
Student has already abandoned school	1,506,897	1.10	0.36
<i>Schools' Characteristics (W)</i>			
Number of enrolled students	1,506,897	699.48	432.52
Day-shift classes	1,506,897	0.55	0.50
Night-shift classes	1,506,897	0.00	0.01
Federal schools	1,506,897	0.00	0.03
Municipal schools	1,506,897	0.75	0.44
Urbana	1,506,897	0.90	0.31

Source: MEC/ Inep (2013).

an adolescent's probability of grade repetition between the age of 15 and 16 by about 10–15 percentage points (i.e. about 20% of a SD).

Following this approach, one simple way to identify the influence of scholar peer effects, in our study, is to test whether student's performance at school is affected by the distribution of dates of birth within the year of other peers from the same class. As already mentioned, the intuition behind such instrument is based on the compulsory school enrolment laws that generate variation in the child's age at school entry, which, in turn, make the date of birth within the year an important determinant of educational achievement and, at the same time, plausibly exogenous to the quality of the student's peers⁷.

Based on this argument, we propose to estimate the following system of equations based on a Two-Step Generalized Method of Moments (GMM-IV) procedure:

⁷ In the related literature, there are studies that focuses on the potential linkage between birth events and socio-cultural aspects (Pinedo and Bermudez, 2015), which is particularly problematic for the purpose of our identification strategy, since birth seasonality could rather be reflecting family background. However, this has not been empirically observed. For instance, Lam and Miron (1987) find that the seasonal pattern of children's births is unrelated to the wealth and marital status of their parents for several countries' populations. In Brazil, Moreira (2013) does not find statistically significant differences of birth patterns among populations from different geographic regions, neither among socio-demographics characteristics of the mother, such as, age, marital status, education and occupation. According to this author, these evidences suggest that Brazilian conceptions are rather related to daily luminosity (photoperiod). As a result, the period of birth, in itself, should not be correlated with educational outcome, suggesting that this variable is plausibly exogenous for the purposes of this paper.

$$Y_{-icj}^- = \alpha_1 + \beta_1 Z_{cj} + \sum_{k=1}^K \theta_1^k X_{icj}^k + \sum_{s=1}^S \delta_1^s W_j^s + v_{icj} \quad (1)$$

$$Y_{icj} = \alpha_2 + \tau_2 \hat{Y}_{-icj}^- + \sum_{k=1}^K \theta_2^k X_{icj}^k + \sum_{s=1}^S \delta_2^s W_j^s + (\varepsilon_{icj} + \tau_2 \hat{v}_{icj}) \quad (2)$$

In the first step, Eq. (1), we regress the average peer's grade $Y_{-icj}^- = \frac{1}{n_{cj}} \sum_{-i=1}^{n_{cj}} y_{-icj}$, excluding student i , at the classroom c and school j ⁸, on the excluded instrument Z_{cj} composed by the proportion of 4th graders born in the second semester of the year at the classroom level, and on other predetermined variables of student i (X) and school j (W), as already described in Table 1. In the second step, Eq. (2), the instrumented average grade (\hat{Y}_{-icj}^-) is included to capture the effect of peer's performance on the individual outcome of interest Y_{icj} . There are two target outcomes being considered in this study: the individual achievement on math and on Portuguese language at the *Prova Brasil* exams of 2013. The estimation procedure is based on the GMM-IV technique and it is robust to a variety of violations of the assumption of independent and identically distributed errors (iid).

In the IV approach there are two indispensable requirements necessary for the instruments to be considered valid, the independence and monotonicity assumptions, which cannot be directly checked because it involves unobservable variables⁹. Nevertheless, there are some testable implications that can be used as robustness tests and will be further discussed in the section of results.

4. Results

The first results presented in Table 2 focus on the reduced form effect and analyse the student's performance on math and Portuguese language as a function of own date of birth, and peer's date of birth. This reduced form analysis does not distinguish endogenous peer's behavior from exogenous social effects. Two models are estimated, Model 1 considers the semester as the period of birth and in Model 2 the periods of birth are divided in quarters. In both cases the estimates reveal that the student's grades, either on math or Portuguese language, are higher if their classmates were born in the second semester rather than the first (Model 1) or if they were born in the later quarters rather than the first (Model 2). For both models, the maturity of peers (proportions) presents a stronger positive correlation with the student's grade than own maturity (dummies), such results were also found by Goux and Maurin (2007).

Table 3 presents the first stage results. Again two types of models are estimated, considering periods of birth as semester (Model 1) or quarters (Model 2). Columns (1) and (3) provide specifications without student and school characteristics, for peers' average grades on math and Portuguese language, respectively, while columns (2) and (4), include all the control variables. We decided to include the states dummies even in the simplest specification, in order to control for regional specificities that might be related to student performance regardless of the instrument. For all the specifications and models, the estimates show a strong and significant correlation of the instruments with the potentially endogenous variable (peers' achievement). In Model 1, classes with high proportions of peers born in the 2nd semester (or later quarters, Model 2) tend to perform better, in average, than those that concentrate children born in the 1st semester (or 1st quarter, Model 1), even after the inclusion of the student period of birth and a wide range of individual socio-economic and school features. Estimates do not change much between the models with and without the control variables. The first stage results capture the relative age effect on the quality of peers, induced here by the exogenous variation of the age at school entry defined by enrolment laws.

The 2nd stage IV and OLS results are exhibited in Table 4, again for Model 1 (semester) and Model 2 (quarters). Basically, OLS and IV estimates are similar in magnitude, although the IV standards errors are much higher. Focusing first on the IV estimates for Model 1 and math evaluation, the positive and significant 2nd stage IV coefficient ($\hat{\beta} = 0.79$) and its corresponding standardized coefficient ($\hat{b} = 0.30$)¹⁰, indicates that a one standard deviation increase in the peers' grade improves individual achievement by 30% of a SD. For the Portuguese language exam, this effect represents an

⁸ Where n_{cj} is the total number of 4th graders students at classroom c and school j minus 1, the excluded student i .

⁹ Suppose a specific classroom has a high proportion of students born in the second semester, and denote its average performance as Y_{cj1}^- . The non-observed counterfactual Y_{cj0}^- for that same unit would provide its average performance had that class concentrated more students born in the first semester. However, only one of Y_{cj1}^- (or Y_{cj0}^-) is observed.

¹⁰ The standardized coefficient is calculated as follows: $\hat{b} = \hat{\beta} * [sd(\bar{Y})/sd(Y)]$.

Table 2
Reduced form effect of classmates' date of birth on the student's performance at the 4th grade.

Model 1	Math	Portuguese Language
Proportion of peers born in the 2 nd semester	4.94*** (0.60)	4.35*** (0.59)
Student born in the 2 nd semester (dummy)	2.91*** (0.15)	2.87*** (0.14)
Additional controls ^a	Yes	Yes
State dummies	Yes	Yes
R ²	0.25	0.26
N ^o of observations	380,712	
Model 2	Math	Portuguese Language
Proportion of peers born the 2 nd quarter	8.09*** (0.87)	5.08*** (0.84)
Proportion of peers born the 3 rd quarter	9.89*** (0.86)	7.89*** (0.84)
Proportion of peers born the 4 th quarter	8.00*** (0.87)	5.86*** (0.85)
Student born in the 2 nd quarter (dummy)	2.48*** (0.20)	2.39*** (0.19)
Student born in the 3 rd quarter (dummy)	5.20*** (0.21)	5.05*** (0.20)
Student born in the 4 th quarter (dummy)	3.23*** (0.20)	3.20*** (0.20)
Additional controls ^a	Yes	Yes
State dummies	Yes	Yes
R ²	0.25	0.26
N ^o of observations	380,712	

Source: Original compilation based on MEC/ Inep (2013).

^a Additional controls include student i ($X_{i,c,j}$) and school j (W_j) characteristics described in Table 1. Robust standard errors in between parentheses. ***, **, * represent the statistical significance at the 1%, 5% and 10% levels, respectively.

increase of 29% SD of the student grade. These results are robust when the instruments consider different peers' period of birth, semester or quarters. For Model 2, the corresponding standardized coefficient produces an impact of 32% SD on individual math grade, very similar to Model 1, while for Portuguese language this impact is a bit smaller, 25% SD of the student grade. The results also remain robust to a different sample of *Prova Brasil*. Estimations for the exams of 2015 are presented in appendix Tables A5 and A6 and even though some results are a little sensitive regarding their magnitudes, the direction of the effect of each variable is robust to the estimation with a different year of the exam (as the signs of the coefficients don't change).

Note that we find a stronger peer effects than the study of Goux and Maurin (2007), whose work found an impact of 20% SD on the probability of being held back a grade. Possibly the endogenous peer effects operate more sharply on the student performance in isolated disciplines, such as Portuguese language or mathematics, than on other outcomes, such as grade retention that capture the student's overall performance. Also other studies that use student test scores as dependent variables, find a stronger peer influence than Goux & Maurin. For instance, for the Brazilian investigations, Koppensteiner (2018) find a peer impact of 40% SD of student grade, Firpo et al. (2015) also find 40%, Raposo (2015) encounters a smaller standardized effect of 6% and Oliveira (2015) provides estimates ranging from 15% to 18% SD. More mixed evidences are found for the international literature, for example Vardardottir (2013) find a peer impact of 32% SD of the student grade, Patacchini et al. (2011) find 10% SD, Calvo-Armengol et al. (2009) find 7%, Duflo et al. (2008) 14% SD, Hanushek et al. (2003) 35% SD and Sacerdote (2001) 5% SD¹¹.

Returning to Table 4, when we compare IV with OLS estimates, we find that for both types of exams, the OLS coefficients seem to overestimate the influence of peers on individual achievement. In the OLS estimations, the corresponding standardized coefficient produces an impact of 42% SD on individual math grade (for both Model 1 and 2),

¹¹ Comparisons to previous studies should be established with caution because of different methods used, different outcomes and mainly because of heterogeneities in the peer reference groups.

Table 3

1st stage IV estimates: correlations between peers' educational achievement and the proportion of classmates by date of birth as the excluded instruments.

<i>Model 1</i>	Math		Portuguese Language		
	(1)	(2)	(3)	(4)	
Proportion of peers born in the 2 nd semester	6.73***	(0.30)	6.50*** (0.39)	5.95*** (0.28)	5.66*** (0.36)
Student born in the 2 nd semester (<i>dummy</i>)	–		0.57*** (0.07)	–	0.53*** (0.07)
Additional controls ^a	No		Yes	No	Yes
State dummies	Yes		Yes	Yes	Yes
R ²	0.40		0.48	0.38	0.45
H ₀ : “instruments are jointly valid”					
Sargan-Hansen J test	Equation exactly identified		Equation exactly identified		
H ₀ : “equation is weakly identified”					
Kleibergen-Paap Wald rk F statistic	1150.75 ⁺⁺		418.69 ⁺⁺	1056.61 ⁺⁺	376.40 ⁺⁺
N ^o of observations	1,053,506		378,034	1,053,506	378,034
<i>Model 2</i>	(1)	(2)	(3)	(4)	
Proportion of peers born the 2 nd quarter	11.91***	(0.43)	9.87*** (0.56)	9.90*** (0.40)	7.85*** (0.52)
Proportion of peers born the 3 rd quarter	15.14***	(0.43)	13.54*** (0.56)	13.53*** (0.40)	11.89*** (0.51)
Proportion of peers born the 4 th quarter	10.12***	(0.44)	9.24*** (0.57)	8.17*** (0.41)	7.17*** (0.52)
Student born in the 2 nd quarter (<i>dummy</i>)	–		0.64*** (0.11)	–	0.55*** (0.09)
Student born in the 3 rd quarter (<i>dummy</i>)	–		1.03*** (0.11)	–	1.00*** (0.10)
Student born in the 4 th quarter (<i>dummy</i>)	–		0.79*** (0.11)	–	0.67*** (0.10)
Additional controls ^a	No		Yes	No	Yes
State dummies	Yes		Yes	Yes	Yes
R ²	0.28		0.35	0.24	0.33
H ₀ : “instruments are jointly valid”					
Sargan-Hansen J test	3.02	(p = 0.22)	2.32 (p = 0.31)	2.75 (p = 0.25)	3.02 (p = 0.22)
H ₀ : “equation is weakly identified”					
Kleibergen-Paap Wald rk F statistic	1078.57 ⁺⁺⁺		326.82 ⁺⁺⁺	968.67 ⁺⁺⁺	284.17 ⁺⁺⁺
N ^o of observations	1,053,506		378,034	1,053,506	378,034

Source: Original compilation based on MEC/ Inep (2013).

^a Additional controls include student i ($X_{i,c,j}$) and school j (W_j) characteristics described in Table 1. Standard errors in between parentheses. Estimates efficient for arbitrary heteroskedasticity and autocorrelation. Statistics robust to heteroskedasticity and autocorrelation. ***, **, * represent the statistical significance at the 1%, 5% and 10% levels, respectively. ++ Test statistics to be compared with Stock-Yogo (2005) weak ID test critical values: 16.38, 8.96, 6.66 and 5.53 corresponding to 10%, 15%, 20% and 25% maximal IV size, respectively. H₀ is not accepted in all the cases. +++ Test statistics to be compared with Stock-Yogo (2005) weak ID test critical values: 22.30, 12.83, 9.54 and 7.80 corresponding to 10%, 15%, 20% and 25% maximal IV size, respectively, as well as, 13.91, 9.08, 6.46, and 5.39 corresponding to 5%, 10%, 20% and 30% maximal IV relative bias, respectively. H₀ is not accepted in all the cases. The estimated coefficients of all variables are presented in Tables A3–A6 from the appendix section.

while for the Portuguese language exams this effect represents an increase of 37% SD of the student test score (again for both Models). This finding is somehow expected once endogenous group selection is likely to lead to upward bias in the OLS coefficient, given that members of the same group tend to share similar unobserved attributes.

4.1. Overidentification, weak identification and other robustness tests

According to Nichols (2006) the standard errors on IV estimates are likely to be larger than OLS estimates, and much larger if the excluded instrumental variables are only weakly correlated with the endogenous regressors. Young

Table 4
 OLS and 2nd stage IV estimates: the effect of peers' educational achievement on student's performance at the 4th grade.

<i>Model 1</i>	Math		Portuguese Language	
	OLS	2 nd stage	OLS	2 nd stage
Peers average math grade	0.70*** (0.003)	0.79*** (0.09)	–	–
Peers average Port. Lang. grade	–	–	0.66*** (0.003)	0.77*** (0.10)
Student born in the 2 nd semester (<i>dummy</i>)	2.49*** (0.14)	2.47*** (0.15)	2.51*** (0.14)	2.45*** (0.15)
Additional controls ^a	Yes	Yes	Yes	Yes
State dummies	Yes	Yes	Yes	Yes
R ²	0.35	0.35	0.33	0.33
N ^o of observations	378,034	378,034	378,034	378,034
<i>Model 2</i>	OLS	2 nd stage	OLS	2 nd stage
Peers average math grade	0.70*** (0.003)	0.78*** (0.06)	–	–
Peers average Port. Lang. grade	–	–	0.66*** (0.003)	0.66*** (0.07)
Student born in the 2 nd quarter (<i>dummy</i>)	2.03*** (0.19)	2.00*** (0.19)	2.03*** (0.19)	2.02*** (0.19)
Student born in the 3 rd quarter (<i>dummy</i>)	4.46*** (0.19)	4.41*** (0.20)	4.04*** (0.19)	4.40*** (0.20)
Student born in the 4 th quarter (<i>dummy</i>)	2.66*** (0.19)	2.62*** (0.20)	2.75*** (0.19)	2.75*** (0.20)
Additional controls ^a	Yes	Yes	Yes	Yes
State dummies	Yes	Yes	Yes	Yes
R ²	0.35	0.35	0.33	0.33
N ^o of observations	378,034	378,034	378,034	378,034

Source: Original compilation based on MEC/ Inep (2013).

^a Additional controls include student i ($X_{i,c,j}$) and school j (W_j) characteristics described in Table 1. The 1st stage estimates also include these additional controls. Standard errors in between parentheses. Estimates efficient for arbitrary heteroskedasticity and autocorrelation. Statistics robust to heteroskedasticity and autocorrelation. ***, **, * represent the statistical significance at the 1%, 5% and 10% levels, respectively. The estimated coefficients of all variables are presented in Tables A3 to A6 from the appendix section.

(2018) in a review of 31 papers published in the journals of the American Economic Association, shows how non-iid errors weaken 1st stage relations, raising the relative bias of 2SLS and generating mean squared error that is larger than biased OLS in almost all published papers. Non-iid errors also increase the probability of spuriously large test statistics when the instruments are irrelevant, particularly in highly leveraged regressions and particularly in joint tests of coefficients, i.e. 1st stage F tests. As a consequence, weak instrument pre-tests based upon F-statistics are found to be largely uninformative of both size and bias.

In order to avoid wrong inference, in our study all the estimated models are based on the IV-GMM technique, robust to heteroskedasticity and autocorrelation in the data process. Thus, all the tests implemented are consistent with violations on the iid error assumption.

Referring again to Table 3, we first present the overidentification tests to check the validity of our instruments for peers' academic achievement. The Sargent-Hansen J statistics test the null hypothesis that "instruments are jointly valid", i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. The J statistic is consistent in the presence of heteroskedasticity and autocorrelation and it could only be calculated for Model 2, since it has more than one excluded instruments. As can be verified from Table 3 (Model 2 with control variables), the null hypothesis of the validity of our instruments cannot be rejected at 1% level, for both evaluations: math (p-value=0.31 > 0.01) and Portuguese language (p-value = 0.22 > 0.01).

Table 5
Placebo regressions of the instruments with various dependent variables.

	Dependent variables						
	Class particip. index	Teacher education	Teacher job contract	Principal education	Principal elected	Physical aspect of classrooms	Library
<i>Model 1</i>							
Proportion of peers born in the 2 nd semester	0.009 (0.007)	0.01 (0.01)	0.03 (0.02)	0.07 (0.007)	−0.01 (0.01)	−0.04 (0.03)	−0.04 (0.04)
Additional controls ^a	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.29	0.05	0.03	0.04	0.20	0.03	0.20
N ^o of obs	381,825	379,116	379,116	379,116	375,439	371,030	367,256
<i>Model 2</i>							
Proportion of peers born in the 2 nd quarter	0.005 (0.01)	−0.03 (0.02)	−0.01 (0.02)	−0.007 (0.01)	0.07*** (0.002)	−0.05 (0.04)	−0.08 (0.06)
Proportion of peers born in the 3 rd quarter	0.005 (0.01)	−0.05 (0.01)	0.01 (0.02)	−0.007 (0.01)	0.02 (0.02)	−0.07* (0.04)	−0.07 (0.06)
Proportion of peers born in the 4 th quarter	0.02* (0.01)	0.02 (0.02)	0.04 (0.02)	0.01 (0.01)	0.02 (0.02)	−0.07* (0.04)	−0.08 (0.06)
Additional controls ^a	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.29	0.05	0.06	0.04	0.20	0.03	0.20
N ^o of obs	381,825	379,116	379,116	379,116	375,439	371,030	367,256

Source: Original compilation based on MEC/ Inep (2013).

^a Additional controls include student i ($X_{i,c,j}$) and school j (W_j) characteristics described in Table 1. Robust standard errors in between parentheses. ***, **, * represent the statistical significance at the 1%, 5% and 10% levels, respectively. Class participation index is a ratio between the number of students in the class taking the *Prova Brasil* exams and the total number of students at school taking these tests in the 4th grade; Teacher education is a dummy variable which is equal to 1 if the teacher has only high school; Teacher job contract is a dummy equal to 1 if the teacher has a temporary contract with the school; Principal education is a dummy equal to 1 if the Principal has only high school; Principal elected is a dummy equal to 1 if the school Principal has been elected to the job position; Physical aspect of classrooms represents the state of conservation of the class being classified as good, regular, bad or inexistent; Library indicates the presence (or not) of a library in the school and its state of conservation being classified as good, regular, bad or inexistent.

In addition, we provide a test of weak instruments under the null hypothesis that “the equation is weakly identified”, meaning the excluded instruments are irrelevant or uncorrelated with the endogenous regressors. The test of weak instruments based on the Kleibergen-Paap rk statistic together with critical values provided by Stock and Yogo (2005) is the appropriated test when the iid assumptions are dropped, because they restrict size and bias of IV tests to OLS standards. The results of these tests are exhibit in Table 3 and for both models H_0 is not accepted.

The tests presented so far demonstrate the instruments are not weakly correlated with the endogenous peers’ performance and they are jointly valid, a plausible indication that in our empirical strategy the independence assumption is being respected. Another indirect way to check for this independence assumption is to estimate placebo regressions in order to verify if the instruments are not related to other school or class characteristics that also influence the student test scores. Table 5 presents the results for these placebo tests. For Model 1, none of the variables tested were significantly related to the instrument defined as the proportion of peers born in the 2nd semester, while for Model 2 some of the three proportions were significant, but in none of the regressions they were jointly significant. Indeed, in appendix Table A7 a balancing test on the observable attributes of school and students demonstrate that there are no apparent discrepancies between classrooms with higher proportions of 4th graders born in the 1st (or 2nd) semester. Although this is not a direct way to test for the independence assumption, this balancing exercise on observables, together with the placebo regressions, overidentification check and the tests of weak instruments offer a reliable way to defend the exogeneity of the instruments.

Table 6

IV regressions for the effect of peers' educational achievement on student's performance over the restricted (complier) and unrestricted sample.

Independent variables	Dep. var.: Student math grade (2 nd stage)		Dep. var: Student Port. Lang. grade (2 nd stage)	
	Unrestricted	Restricted (compliers)	Unrestricted	Restricted (compliers)
<i>Model 1</i>				
Peers average math grade	0.74*** (0.09)	0.53*** (0.13)	–	–
Peers average Port. Lang. grade	–	–	0.77*** (0.10)	0.59*** (0.15)
Born 2 nd semester	2.47*** (0.15)	1.16*** (0.20)	2.45*** (0.15)	1.48*** (0.20)
Additional controls ^a	Yes	Yes	Yes	Yes
State dummies	Yes	Yes	Yes	Yes
R ²	0.35	0.31	0.33	0.30
H ₀ = "Equal coefficients"	$F_{stat.} = 0.18 \not\geq F_{(tab)} = 1.16$		$F_{stat.} = 0.06 \not\geq F_{(tab)} = 1.16$	
N ^o of observations	378,034	221,661	378,034	221,661
<i>Model 2</i>				
Peers average math grade	0.78*** (0.06)	0.71*** (0.08)	–	–
Peers average Port. Lang. grade	–	–	0.66*** (0.07)	0.61*** (0.09)
Born in the 2 nd quarter	2.00*** (0.19)	1.06** (0.42)	2.02*** (0.19)	2.02*** (0.41)
Born in the 3 rd quarter	4.41*** (0.20)	2.39*** (0.39)	4.40*** (0.20)	3.41*** (0.39)
Born in the 4 th quarter	2.62*** (0.20)	1.34*** (0.31)	2.75*** (0.20)	2.21*** (0.31)
Additional controls ^a	Yes	Yes	Yes	Yes
State dummies	Yes	Yes	Yes	Yes
R ²	0.35	0.32	0.33	0.30
H ₀ = "Equal coefficients"	$F_{stat.} = 0.07 \not\geq F_{(tab)} = 1.16$		$F_{stat.} = 0.07 \not\geq F_{(tab)} = 1.16$	
N ^o of observations	378,034	221,661	378,034	221,661

Source: Original compilation based on MEC/ Inep (2013).

^a Additional controls include student i ($X_{i,c,j}$) and school j (W_j) characteristics described in Table 1. Standard errors in between parentheses. Estimates efficient for arbitrary heteroskedasticity and autocorrelation. Statistics robust to heteroskedasticity and autocorrelation. ***, **, * represent the statistical significance at the 1%, 5% and 10% levels, respectively.

Finally, we should check if our IV strategy is in accordance with the monotonicity assumption¹². This requirement cannot be directly checked, because again it involves unobservable variables, only one of \bar{Y}_{c1} (or \bar{Y}_{c0}) is observed.

¹² The *monotonicity assumption* demands the treatment status (the peers' average performance) to be indeed induced by the instrument status (proportion of 4th graders born by semester). In the present paper, monotonicity means that classrooms with higher proportion of students born in the second semester should perform at least as well (or better) as they would have performed had they concentrated more students born in the first semester. As a consequence, the intended effect of our instrument is a local average treatment effect (LATE) identified only over the subpopulation of compliers. Here the compliers are the students who turned 6 years old by the beginning of the academic year and enrolled in the 1st grade, as well as those who had to wait one more year to start school if they turned 6 after this period. Notice that this compliance rule, due to the monotonicity assumption, could be violated if students (or their parents) manage to defy the legal age entrance at school. Suppose for instance parents share the beliefs it is better to enter school at a relatively older age and a child who turn 6 years old in March would be held on one more year before start the 1st grade. This same child would begin the 1st grade relatively older breaking down the instrument status, in which children born in the first semester should be relatively younger. This sort of manipulation turns down the compliance rule and the IV approach is no longer valid.

However, we propose a test that consists on restricting the estimated models over the subpopulation of compliers, in which we identify only students who enrolled in the 1st grade when they turned 6 years old by March, as well as those who had to wait one more year to start school if they turned 6 after this period. The subsample of compliers is restricted to 221,661 individuals, representing 59% of our sample¹³. Table 6 presents the results for both models and type of exams. For all the cases, the peer effects coefficients reduce in the restricted sample, however they are not statistically different from those of the complete sample, as illustrated by the test F. This kind of exercise provides an indirect way of proving that the treatment intensity (peers' average performance) responds robustly to the instruments and, as such, we assume there is no reason to expect schools to respond differently to the instrument status.

5. Final considerations

The present paper provides new evidence for the role of peer effects in the Brazilian public school system. It aims to answer to the question: “what are the impacts of peer effects on 4th grade students' academic outcomes from Brazilian public schools?” Using data from *Prova Brasil* 2013, our identification strategy is based on the IV approach and contributes to the Brazilian literature by adopting the use of instruments that have never been tested for this country. The quality of the student body is instrumented here with the proportions of peers born in the second semester of the year, as proposed by the work of Goux and Maurin (2007). The idea behind these instruments is that educational achievement is related to age at school entry because of the age variation induced by compulsory school enrolment laws. In Brazil, children born in the second semester start school at a relatively older age, because of the Federal laws (BRASIL, 2008, 2010).

The results demonstrate that classrooms with higher proportions of peers born in the 2nd semester tend to perform better, in average, than those that concentrate children born in the 1st semester, even after the inclusion of a wide range of control variables. For the math and Portuguese language evaluations, a one standard deviation increase in the peers' test scores improves individual achievement by 30% of a standard deviation. Tests for weak instruments, overidentification and other robustness exercises were implemented and favored the strength of the instruments.

To better understand the role of peer effects in the school environment is a key issue to improve the development of educational policies. Further research is needed to help understanding the mechanisms through which children influence each other, in order to better prescribe classroom interventions which takes advantage of the peer effects dissemination.

¹³ The high number of students not compliers (about 40%) demonstrates that the enforcement and punishment by the government to ensure compliance with Resolution CNE / CEB 6/2010 have not being effective. Possibly, some schools and parents were awaiting the annulment of this law because of the innumerable judicial processes that several states had requested to question the constitutionality of the cutoff entry age. Until this issue was not resolved in the Federal Supreme Court, the states have found legal certainty not to comply with this resolution. However, in 2018 the Brazilian Federal Supreme Court finally determined the constitutionality of Resolution CNE / CEB 6/2010, canceling other state level resolutions that presented different cutoff entry age of this federal resolution (*Parecer* CNE / CEB N^o 2/2018).

Appendix

Table A1
Definition of the variables.

Dependent Variables (Y)	Definition of Variables
Math grade at Prova Brasil exam (4 th grade)	Grade of student <i>i</i> on math test transformed to the scale with mean = 250 and standard deviation = 50
Portuguese Language grade at Prova Brasil exam (4 th grade)	Grade of student <i>i</i> on Portuguese language test transformed to the scale with mean = 250 and standard deviation = 50
Instruments (Z)	
Proportion of peers born in the 2 nd semester	Proportion of peers in the student classroom born in the second semester of the year (4 th graders)
Peers' Educational Achievement (\bar{Y})	
Peers math grade (4 th grade)	Peers average math grade in the student classroom
Peers Portuguese language grade (4 th grade)	Peers average Portuguese language grade in the student classroom
Students' Characteristics (X)	
Semester II	Dummy equals 1 if students were born in the months from July to December
Age	Age of the student, takes values: 1 = 8 years old; 2 = 9 years old; 3 = 10 years old; 4 = 11 years old; 5 = 12 years old; 6 = 13 years old; 7 = 14 years old and 8 = 15 years old or more
Male sex	Dummy equals 1 if student is male
White	Dummy equals 1 if students declare themselves as white
Black	Dummy equals 1 if students declare themselves as black
Student lives with the mother	Dummy equals 1 if student lives with the mother
Educational level (mother)	Educational attainment, takes values: 1 = has never studied; 2 = 4 th grade incomplete; 3 = has finish 4 th grade, but hasn't finish 8 th grade; 4 = has finish 8 th grade, but hasn't finish high school; 5 = has completed high school; 6 = College degree
Literacy (mother)	Dummy equals 1 if student's mother knows how to read
Student lives with the father	Dummy equals 1 if student lives with the father
Educational level (father)	Similar to Educational level of mother
Literacy (father)	Dummy equals 1 if student's father knows how to read
Preschool	Age at which student began to frequent school, takes values: 1 = between 0 and 3 years old; 2 = between 4 and 5 years old; 3 = between 6 and 7 years old; 4 = for 8 years old or more
Grade retention	Variable takes values: 1 = has never repeated; 2 = has repeated one grade; 3 = has repeated two grades or more
Student has already abandoned school	Variable takes values: 1 = has never abandoned; 2 = has abandoned one year; 3 = has abandoned two years or more
Schools' Characteristics (W)	
Number of enrolled students	Total number of enrolled students at the school <i>j</i>
Day-shift classes	Classes take place between 7am to 12am
Night-shift classes	Classes take place between 6 pm to 10pm
Federal schools	The school is under the responsibility of Federal Government
Municipal schools	The school is under the responsibility of Municipal Government
Urbana	The school is located in urban area

Source: MEC/ Inep (2013).

Table A2
Brazilian laws regarding minimum age for school entry in 2009.

State	1 st grade entry age	Birthday cutoff	Statute
Rio Grande do Sul, Santa Catarina, Rio de Janeiro, Minas Gerais, Mato Grosso, Mato Grosso do Sul, Goiás, Distrito Federal, Amazonas, Acre, Amapá, Roraima, Rondônia, Tocantins, Pará, Bahia, Alagoas, Pernambuco, Paraíba, Rio Grande do Norte, Maranhão	6	Beginning of school year	CNE/CEB N ^o : 4/2008 – MEC ¹
São Paulo	6	31-December	DELIBERATION CEE N ^o 73/2008 - SÃO PAULO ²
Espírito Santo	6	1-March	RESOLUTION CEE N ^o 1790/2008 – EDUCATIONAL BOARD OF ESPÍRITO SANTO ³
Paraná	6	31-December	LAW N ^o 16.049 OF FEBRUARY, 19, 2009 ⁴
Sergipe	6	30-April	ORDINANCE N ^o 5957/2008 – GOVERNMENT OF SERGIPE ⁵
Ceará	6	30-April	RESOLUTION N ^o 0410/2006 - EDUCATIONAL BOARD OF CEARÁ ⁶
Piauí	6	30-June	RESOLUTION CEE/PI N ^o 141/2007 ⁷

Source:

¹ BRASIL. Ministério da Educação. Resolution CNE/CEB no. 4, February 20th 2008. Published at *Diário Oficial da União*, Brasília, June 10th 2008. <http://portal.mec.gov.br/cne/arquivos/pdf/2008/pceb004_08.pdf>. Accessed June, 20, 2016.

² DELIBERATION CEE N^o 73/2008. <http://desumare.edunet.sp.gov.br/Supervisao/09/deliberacoes/deliberacao_CEE_7308.pdf>. Accessed June, 20, 2016.

⁵ ORDINANCE N^o 5957/2008. <http://seed.se.gov.br/arquivos/Portaria_Diretrizes_Escolas_Publicas_Estaduais_2009.pdf>. Accessed June, 20, 2016.

⁶ RESOLUTION N^o 0410/2006 - EDUCATIONAL BOARD OF CEARÁ. <http://portal.mec.gov.br/arquivos/pdf/acs_resolucao410.2006.pdf> Accessed June, 20, 2016.

Table A3

IV regressions for the effect of peers' educational achievement on student's performance: sample of *Prova Brasil* 2013 – Model 1.

	Math		Portuguese language	
	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)
<i>Model 1</i>				
Peers average math grade	–	0.74*** (0.09)	–	–
Peers average Port. Lang. grade	–	–	–	0.77*** (0.10)
Proportion of peers born in the 2 nd semester	6.50*** (0.39)	–	5.66*** (0.36)	–
Born 2 nd semester	0.57*** (0.07)	2.47*** (0.15)	0.53*** (0.07)	2.45*** (0.15)
Age	–1.12*** (0.04)	–2.46*** (0.12)	–1.17*** (0.04)	–2.20*** (0.14)
Male sex	–0.49*** (0.07)	3.79*** (0.14)	–0.51*** (0.07)	–10.21*** (0.14)
White	1.33*** (0.08)	–0.37* (0.19)	1.33*** (0.08)	0.46** (0.20)

Table A3 (Continued)

	Math		Portuguese language	
	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)
Black	−2.76*** (0.11)	−8.18*** (0.32)	−2.27*** (0.11)	−6.55*** (0.31)
Student lives with the mother	1.59*** (0.12)	2.20*** (0.27)	1.52*** (0.11)	1.91*** (0.27)
Educational level (mother)	1.54*** (0.03)	2.80*** (0.15)	1.59*** (0.03)	2.83*** (0.17)
Literacy (mother)	1.07*** (0.18)	0.77** (0.34)	0.89*** (0.16)	−0.34 (0.32)
Student lives with the father	1.54*** (0.08)	1.73*** (0.20)	1.36*** (0.07)	1.46*** (0.20)
Educational level (father)	0.77*** (0.03)	0.82*** (0.09)	0.90*** (0.03)	1.04*** (0.10)
Literacy (father)	1.56*** (0.15)	1.77*** (0.30)	1.50*** (0.13)	1.77*** (0.29)
Preschool	−0.50*** (0.04)	−0.81*** (0.09)	−0.48*** (0.04)	−0.86*** (0.09)
Grade retention	−1.89*** (0.07)	−10.63*** (0.21)	−1.97*** (0.06)	−10.94*** (0.23)
Student has already abandoned school	−3.47*** (0.11)	−8.54*** (0.37)	−3.10*** (0.10)	−7.82*** (0.37)
Day-shift classes	0.26*** (0.07)	−0.33** (0.14)	−0.05 (0.07)	−0.37** (0.13)
Night-shift classes	0.18 (2.40)	0.75 (4.46)	−2.97 (1.95)	−2.71 (4.11)
Federal schools	37.11*** (0.92)	4.39 (3.72)	37.26*** (0.87)	3.09 (4.15)
Municipal schools	−3.37*** (0.10)	0.22 (0.34)	−3.95*** (0.09)	0.52 (0.43)
Urbana	11.45*** (0.13)	−0.35 (1.03)	11.24*** (0.11)	−0.94 (1.15)
Number of enrolled students	−0.00*** (0.00)	−0.00*** (0.00)	−0.00*** (0.00)	−0.00** (0.00)
Constant	177.99*** (0.40)	72.84*** (15.92)	163.78*** (0.36)	69.15*** (16.74)
State dummies	Yes	Yes	Yes	Yes
R ²	0.48	0.35	0.45	0.33
N ^o of observations	378,034	378,034	378,034	378,034

Source: Original compilation based on MEC/ Inep (2013).

Standard errors in between parentheses. Estimates efficient for arbitrary heteroskedasticity and autocorrelation. Statistics robust to heteroskedasticity and autocorrelation. ***, **, * represent the statistical significance at the 1%, 5% and 10% levels, respectively.

Table A4

IV regressions for the effect of peers' educational achievement on student's performance: sample of *Prova Brasil* 2013 – Model 2.

	Math		Portuguese language	
	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)
<i>Model 2</i>				
Peers average math grade	–	0.78*** (0.06)	–	–
Peers average Port. Lang. grade	–	–	–	0.66*** (0.07)
Proportion of peers born in the 2 nd quarter	9.87*** (0.56)	–	7.85*** (0.52)	–
Proportion of peers born in the 3 rd quarter	13.54*** (0.56)	–	11.89*** (0.51)	–
Proportion of peers born in the 4 th quarter	9.24*** (0.57)	–	7.17*** (0.52)	–
Born in the 2 nd quarter	0.64*** (0.11)	2.00*** (0.19)	0.55*** (0.09)	2.02*** (0.19)
Born in the 3 rd quarter	1.03*** (0.11)	4.41*** (0.20)	1.00*** (0.10)	4.40*** (0.20)
Born in the 4 th quarter	0.79*** (0.11)	2.62*** (0.20)	0.67*** (0.10)	2.75*** (0.20)
Age	–1.16*** (0.04)	–2.61*** (0.10)	–1.21*** (0.04)	–2.48*** (0.11)
Male sex	–0.48*** (0.07)	3.83*** (0.14)	–0.50*** (0.07)	–10.23*** (0.14)
White	1.33*** (0.08)	–0.38** (0.17)	1.33*** (0.08)	0.60*** (0.18)
Black	–2.75*** (0.11)	–8.13*** (0.26)	–2.27*** (0.11)	–6.77*** (0.25)
Student lives with the mother	1.58*** (0.12)	2.17*** (0.25)	1.51*** (0.11)	2.05*** (0.25)
Educational level (mother)	1.54*** (0.03)	2.77*** (0.10)	1.59*** (0.03)	2.99*** (0.12)
Literacy (mother)	1.05*** (0.18)	0.73** (0.33)	0.87*** (0.16)	–0.28 (0.31)
Student lives with the father	1.54*** (0.08)	1.71*** (0.17)	1.36*** (0.07)	1.60*** (0.17)
Educational level (father)	0.76*** (0.03)	0.81*** (0.07)	0.89*** (0.03)	1.13*** (0.08)
Literacy (father)	1.54*** (0.15)	1.73*** (0.28)	1.48*** (0.13)	1.91*** (0.27)
Preschool	–0.50*** (0.04)	–0.80*** (0.08)	–0.48*** (0.04)	–0.91*** (0.08)
Grade retention	–1.85*** (0.07)	–10.46*** (0.17)	–1.93*** (0.06)	–11.00*** (0.18)
Student has already abandoned school	–3.45*** (0.11)	–8.46*** (0.28)	–3.09*** (0.10)	–8.10*** (0.28)
Day-shift classes	0.26*** (0.07)	–0.34** (0.14)	–0.05 (0.07)	–0.38** (0.13)
Night-shift classes	0.59 (2.36)	0.85 (4.47)	–2.64 (1.95)	–2.91 (4.07)
Federal schools	37.36*** (0.92)	4.10 (2.78)	37.49*** (0.87)	7.11** (3.03)
Municipal schools	–3.36*** (0.10)	0.26 (0.26)	–3.94*** (0.09)	0.12 (0.32)

Table A4 (Continued)

	Math		Portuguese language	
	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)
Urbana	11.40*** (0.13)	−0.50 (0.69)	11.20*** (0.11)	0.20 (0.77)
Number of enrolled students	−0.00*** (0.00)	−0.00*** (0.00)	−0.00*** (0.00)	−0.00*** (0.00)
Constant	172.96*** (0.46)	70.22*** (10.39)	159.79*** (0.42)	85.83*** (11.02)
State dummies	Yes	Yes	Yes	Yes
R ²	0.35	0.35	0.33	0.33
N ^o of observations	378,034	378,034	378,034	378,034

Source: Original compilation based on MEC/ Inep (2013).

Standard errors in between parentheses. Estimates efficient for arbitrary heteroskedasticity and autocorrelation. Statistics robust to heteroskedasticity and autocorrelation. ***, **, * represent the statistical significance at the 1%, 5% and 10% levels, respectively.

Table A5

IV regressions for the effect of peers' educational achievement on student's performance: sample of *Prova Brasil 2015* – Model 1.

	Math		Portuguese language	
	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)
<i>Model 1</i>				
Peers average math grade	–	0.89*** (0.22)	–	–
Peers average Port. Lang. grade	–	–	–	0.69*** (0.16)
Proportion of peers born in the 2 nd semester	3.57*** (0.51)	–	5.22*** (0.52)	–
Born 2 nd semester	0.25** (0.10)	1.66*** (0.20)	0.36*** (0.10)	1.94*** (0.21)
Age	−1.01*** (0.05)	−1.33*** (0.25)	−1.25*** (0.06)	−1.42*** (0.23)
Male sex	−0.28*** (0.10)	4.53*** (0.20)	−0.33*** (0.10)	−8.83*** (0.21)
White	1.25*** (0.11)	−0.16 (0.35)	1.14*** (0.11)	0.10 (0.29)
Black	−3.01*** (0.16)	−8.02*** (0.74)	−2.64*** (0.16)	−8.21*** (0.54)
Student lives with the mother	1.31*** (0.16)	1.92*** (0.43)	1.26*** (0.16)	2.09*** (0.40)
Educational level (mother)	1.25*** (0.04)	2.45*** (0.29)	1.39*** (0.04)	2.80*** (0.24)
Literacy (mother)	0.17 (0.25)	−0.24 (0.48)	0.74** (0.26)	−0.34 (0.53)
Student lives with the father	1.68*** (0.11)	1.37*** (0.43)	1.51*** (0.11)	1.37*** (0.33)
Educational level (father)	0.66*** (0.04)	0.76*** (0.16)	0.87*** (0.04)	1.41*** (0.16)
Literacy (father)	1.02*** (0.20)	1.75*** (0.44)	1.38*** (0.21)	2.55*** (0.46)
Preschool	−0.79*** (0.05)	−1.06*** (0.21)	−0.73*** (0.05)	−0.94*** (0.16)
Grade retention	−1.71*** (0.10)	−10.46*** (0.43)	−1.75*** (0.10)	−11.63*** (0.35)

Table A5 (Continued)

	Math		Portuguese language	
	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)
Student has already abandoned school	−2.02*** (0.14)	−4.45*** (0.53)	−2.19*** (0.15)	−7.00*** (0.47)
Day-shift classes	−2.13*** (0.10)	−0.03 (0.51)	−2.46*** (0.10)	−0.73 (0.45)
Night-shift classes	−1.19 (2.64)	−5.14 (5.05)	−5.00** (1.97)	3.55 (8.11)
Federal schools	33.14*** (0.66)	−2.37 (7.70)	31.28*** (0.86)	4.00 (5.62)
Municipal schools	−1.44*** (0.12)	0.63 (0.40)	−2.60*** (0.12)	0.04 (0.49)
Urbana	4.83*** (0.16)	−1.54 (1.13)	7.08*** (0.17)	0.17 (1.20)
Number of enrolled students	0.03*** (0.00)	−0.01 (0.01)	0.03*** (0.00)	−0.00 (0.01)
Constant	190.53*** (0.53)	36.14 (42.99)	175.84*** (0.55)	79.06*** (2910)
State dummies	Yes	Yes	Yes	Yes
R ²	0.41	0.31	0.42	0.30
N ^o of observations	164,423	164,423	164,423	164,423

Source: Original compilation based on MEC/ Inep (2015).

Standard errors in between parentheses. Estimates efficient for arbitrary heteroskedasticity and autocorrelation. Statistics robust to heteroskedasticity and autocorrelation. ***, **, * represent the statistical significance at the 1%, 5% and 10% levels, respectively.

Table A6

IV regressions for the effect of peers' educational achievement on student's performance: sample of *Prova Brasil* 2015 – Model 2.

	Math		Portuguese language	
	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)
<i>Model 2</i>				
Peers average math grade	−	0.68*** (0.15)	−	−
Peers average Port. Lang. grade	−	−	−	0.64*** (0.11)
Proportion of peers born in the 2 nd quarter	4.39*** (0.73)	−	5.48*** (0.73)	−
Proportion of peers born in the 3 rd quarter	7.44*** (0.72)	−	10.47*** (0.73)	−
Proportion of peers born in the 4 th quarter	3.73*** (0.73)	−	4.96*** (0.75)	−
Born in the 2 nd quarter	0.54*** (0.14)	2.89*** (0.28)	0.73*** (0.14)	2.91*** (0.29)
Born in the 3 rd quarter	0.77*** (0.14)	4.20*** (0.29)	1.08*** (0.14)	4.54*** (0.31)
Born in the 4 th quarter	0.28*** (0.14)	2.20*** (0.27)	0.38** (0.14)	2.34*** (0.28)
Age	−1.08*** (0.06)	−1.89*** (0.19)	−1.35*** (0.06)	−1.83*** (0.19)

Table A6 (Continued)

	Math		Portuguese language	
	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)	1 st stage: peers' average grade (Dep. var.)	2 nd stage: student grade (Dep. var.)
Male sex	−0.27** (0.10)	4.51*** (0.19)	−0.31*** (0.10)	−8.80*** (0.20)
White	1.25*** (0.11)	0.11 (0.28)	1.14*** (0.11)	0.15 (0.26)
Black	−3.00*** (0.16)	−8.66*** (0.54)	−2.63*** (0.16)	−8.32*** (0.44)
Student lives with the mother	1.30*** (0.16)	2.18*** (0.37)	1.24*** (0.16)	2.12*** (0.37)
Educational level (mother)	1.25*** (0.04)	2.71*** (0.20)	1.39*** (0.04)	2.85*** (0.18)
Literacy (mother)	0.15 (0.25)	−0.24 (0.48)	0.72** (0.26)	−0.35 (0.52)
Student lives with the father	1.67*** (0.11)	1.72*** (0.32)	1.51*** (0.11)	1.42*** (0.28)
Educational level (father)	0.66*** (0.04)	0.90*** (0.12)	0.87*** (0.04)	1.45*** (0.13)
Literacy (father)	1.00*** (0.20)	1.95*** (0.41)	1.36*** (0.21)	2.59*** (0.43)
Preschool	−0.79*** (0.05)	−1.21*** (0.16)	−0.73*** (0.05)	−0.96*** (0.14)
Grade retention	−1.64*** (0.10)	−10.49*** (0.31)	−1.65*** (0.10)	−11.37*** (0.28)
Student has already abandoned school	−2.00*** (0.14)	−4.82*** (0.41)	−2.17*** (0.15)	−7.03*** (0.39)
Day-shift classes	2.15*** (0.10)	−0.51 (0.37)	−2.48*** (0.10)	−0.86** (0.34)
Night-shift classes	−0.56 (2.58)	−4.89 (4.99)	−4.20** (1.95)	3.81 (8.06)
Federal schools	33.10*** (0.66)	4.83 (5.33)	31.24*** (0.86)	5.46 (4.24)
Municipal schools	−1.44*** (0.12)	0.33 (0.31)	−2.59*** (0.12)	−0.07 (0.38)
Urbana	4.82*** (0.16)	−0.47 (0.78)	7.08*** (0.17)	0.52 (0.87)
Number of enrolled students	0.03*** (0.00)	0.00 (0.00)	0.03*** (0.00)	0.00 (0.00)
Constant	188.37*** (0.60)	77.60** (28.47)	173.15*** (0.63)	87.08*** (20.21)
State dummies	Yes	Yes	Yes	Yes
R^2	0.41	0.32	0.43	0.30
N ^o of observations	164,423	164,423	164,423	164,423

Source: Original compilation based on MEC/ Inep (2015).

Standard errors in between parentheses. Estimates efficient for arbitrary heteroskedasticity and autocorrelation. Statistics robust to heteroskedasticity and autocorrelation. ***, **, * represent the statistical significance at the 1%, 5% and 10 % levels, respectively.

Table A7

Summary statistics of individual and school characteristics by schools with higher proportions of peers born by semester.

Individual and school characteristics	Higher proportion of peers born in the 1 st semester		Higher proportion of peers born in the 2 nd semester ^a	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	3.97	1.11	4.07	1.11
Proportion of peers born in the 2 nd semester in the classroom	0.40	0.07	0.58	0.07
Total number of enrolled students in the school	697.47	424.72	697.95	434.79
Male	0.53	0.50	0.53	0.50
White	0.32	0.47	0.32	0.47
Black	0.12	0.33	0.12	0.32
Student lives with the mother	0.91	0.29	0.91	0.29
Educational level (mother)	3.89	1.49	3.87	1.49
Literacy (mother)	0.95	0.22	0.95	0.22
Student lives with the father	0.70	0.46	0.70	0.46
Educational level (father)	3.81	1.56	3.76	1.55
Literacy (father)	0.92	0.27	0.92	0.27
Preschool	1.86	0.86	1.86	0.86
Grade retention	1.37	0.62	1.37	0.63
Student has already abandoned school	1.10	0.35	1.10	0.35
Day-shift classes	0.56	0.50	0.55	0.50
Night-shift classes	0.00	0.02	0.00	0.01
Federal school	0.00	0.03	0.00	0.04
Municipal school	0.74	0.44	0.73	0.45
Urban area	0.90	0.30	0.89	0.31
N ^o of observations	195,247		182,787	

Source: Original compilation based on MEC/ Inep (2013).

^a It is a dummy variable that takes the value 1 when the proportion of 4th graders born in the 2nd semester, in the student's classroom, is at least as high as the proportion of 4th graders born in the 1st semester; and the value 0 otherwise.

References

- Allen, J., Barnsley, R., 1993. Streams and tiers: the interaction of ability, maturity and training in system with age-dependent recursive selection. *J. Hum. Resour.* 28 (3), 649–659.
- Angrist, J.D., Imbens, G.W., 1995. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *J. Am. Stat. Assoc.* 90 (430), 431–442.
- Angrist, J.D., Krueger, A.B., 1991. Does Compulsory School Enrolment Affect Schooling and Earnings? *Q. J. Econ.* 106, 979–1014.
- Angrist, J.D., Krueger, A.B., 1992. The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples. *J. Am. Stat. Assoc.* 87 (418), 328–336.
- Badev, A., 13-05 2013. *Discrete Games in Endogenous Networks: Theory and Policy*. PSC Working Paper Series, PSC.
- Ballester, C., Calvó-Armengol, A., Zenou, Y., 2006. Who's who in networks. wanted: the key player. *Econometrica* 74 (5), 1403–1417.
- Barnsley, R.H., Thompson, A.H., Barnsley, P.E., 1985. Hockey success and birth date: the relative age effect. *Canadian Association for Health, Physical Education, and Recreation Journal* 51, 23–28.
- Bramoullé, Y., Djebbari, H., Fortin, B., 2009. Identification of peer effects through social networks. *J. Econom.* 150, 41–55.
- Brasil, 2008. Ministério Da Educação. Resolution CNE/CEB no. 4, Dated February 20th 2008. Defines the Operational Guidelines About the Initial Three Years of Elementary Education. *Diário Oficial Da União, Brasília, June 10th*, Available on: <http://portal.mec.gov.br/cne/arquivos/pdf/2008/pceb004_08.pdf>. Link last accessed on: June 20th 2016.
- Brasil, 2010. Ministério Da Educação. Resolution CNE/CEB no. 6, Dated October 20th 2010. Defines the Operational Guidelines for Enrolling Students Into Elementary and Childhood Education. *Diário Oficial Da União, Brasília, October 21st*, Available on: <http://portal.mec.gov.br/index.php?option=com_content&id=14906&Itemid=866>. Link last accessed on: June 20th 2016.
- Calvó-Armengol, A., Patacchini, E., Zenou, Y., 2009. Peer effects and social networks in education. *Rev. Econ. Stud.* 76 (4), 1239–1267.
- Case, A.C., Katz, L.F., 1991. The Company You Keep: the Effects of Family and Neighborhood on Disadvantaged Youths. NBER Working Paper., pp. 3705.
- Ding, W., Lehrer, S.F., 2007. Do peers affect student achievement in china's secondary schools? *Rev. Econ. Stat.* 89 (2), 300–312.
- Duflo, E., Dupas, P., Kremer, M., 2008. Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence From a Randomized Evaluation in Kenya. NBER Working Paper., pp. 14475.
- Eisenkopf, G., Hessami, Z., Fischbacher, U., Heinrich, U., 2011. Academic Performance and Single-sex Schooling: Evidence From a Natural Experiment in Switzerland. CESifo Working Paper: Economics of Education., pp. 3592.

- Evans, W.N., Oates, W.E., Schwab, R.M., 1992. Measuring peer group effects: a study of teenage behavior. *J. Polit. Econ.* 100 (5), 966–991.
- Firpo, S., Jales, H., Pinto, C., 2015. Measuring peer effects in the Brazilian school system. *Appl. Econ.* 47 (32).
- Goux, D., Maurin, E., 2007. Close neighbours matters: neighbourhood effects on early performance at school. *Econ. J.* 117 (523), 1193–1215.
- Hanushek, E.A., Kain, J.F., Markman, J.M., Rivkin, S.G., 2003. Does peer ability affect student achievement? *J. Appl. Econom.* 18 (5), 527–544.
- Koppensteiner, M.F., 2018. Relative age, class Assignment and academic performance: evidence from Brazilian primary schools. *Scand. J. Econ.* 120 (1), 296–325.
- Lam, D., Miron, J., 1987. The Seasonality of Births in Human Populations. Research Report 87-114, University of Michigan, Population Studies Center.
- Manski, C.F., 1993. Identification of endogenous social effects: the reflection problem. *Rev. Econ. Stud.* 60 (3), 531–542.
- Mele, A., 2010. A Structural Model of Segregation in Social Networks. Working Papers 10-16, NET Institute. [dataset] Ministério Da Educação – Mec / Instituto Nacional De Estudos E Pesquisas Educacionais Anísio Teixeira – Inep. Saeb/ Prova Brasil – 2015. Available on: <<http://provaabril.inep.gov.br/web/guest/microdados>> Link Last Accessed on: November 22, 2018.
- Moreira, M.M., 2013. Sazonalidade Dos Nascimentos No Brasil - 2000-2010. XXIX Congresso Latinoamericano De Sociologia, 2013, Chile. *Anais Do XXIX Congresso Latinoamericano De Sociologia*.
- Nichols, A., 2006. Weak instruments: An overview and new techniques. Stata 5th North American Meeting Presentation, 2006. Available on: <<https://www.stata.com/meeting/5nasug/wiv.pdf>>. Link last accessed on: November 22, 2018.
- Oliveira, V.R., 2015. O impacto da formação das turmas sobre o desempenho dos alunos: evidências para as escolas públicas brasileiras. In: de Aquino Fonseca Neto, Fernando, Barrantes Hidalgo, Álvaro (Eds.), *III Encontro Pernambucano De Economia Políticas Para O Desenvolvimento Estadual*, 1, 1 ed. Recife: Villalux, pp. 99–116.
- Oosterbeek, H., Van Ewijk, R., 2014. Gender peer effects in university: evidence from a randomized experiment. *Econ. Educ. Rev.* 38, 51–63.
- Patacchini, E., Venanzoni, G., 2014. Peer effects in the demand for housing quality. *J. Urban Econ.* 83, 6–17.
- Patacchini, E., Rainone, E., Zenou, Y., 2011. Dynamic Aspects of Teenage Friendships and Educational Attainment. CEPR Discussion Paper., pp. 8223.
- Pinedo, W.J.I., Bermudez, B.C., 2015. Month of Birth and Socioeconomic Outcomes of Adults: Evidence From Brazil, *Anais Do XLIII Encontro Nacional De Economia [Proceedings of the 43rd Brazilian Economics Meeting]*, ANPEC - Associação Nacional Dos Centros De Pósgraduação Em Economia [Brazilian Association of Graduate Programs in Economics].
- Pinto, C.C.X., 2008. Semiparametric Estimation of Peer Effects. Tese de Doutorado - University of California, Berkeley.
- Raposo, I.P.A., 2015. O Papel Da Rede De Amizades E Da Formação Aleatória De Turmas Por Faixa Etária Sobre O Desempenho. Tese De Doutorado – Universidade Federal De Pernambuco.
- Sacerdote, B., 2001. Peer effects with random assignment: results for Dartmouth rommates. *The Quarterly Journal of Economics* 116 (2), 681–704.
- Stock, J.H., Yogo, M., 2005. Testing for weak instruments in linear IV regression. Ch. 5. In: Stock, J.H., Andrews, D.W.K. (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*. Cambridge University Press.
- Sund, K., 2009. Estimating peer effects in Swedish high school using school, teacher, and student fixed effects. *Econ. Educ. Rev.* 28, 329–336.
- Vardardottir, A., 2013. Peer effects and academic achievement: a regression discontinuity approach. *Econ. Educ. Rev.* 36 (C), 108–121.
- Vigdor, J.L., Nechyba, T.J., 2007. Peer effects in North Carolina public schools. In: Woessmann, L., Peterson, P.E. (Eds.), *Schools and the Equal Opportunity Problem*. MIT Press, pp. 73–102.
- Young, A., Available on: <<http://personal.lse.ac.uk/YoungA/>>. Link last accessed on: November 22, 2018. Consistency Without Inference: Instrumental Variables in Practical Application. London School of Economics, Working Paper.
- Zimmerman, D.J., 2003. Peer effects in academic outcomes: evidence from a natural experiment. *Rev. Econ. Stat.* 85 (1), 9–23.