CAM
Centre for Applied Microeconometrics

Department of Economics
University of Copenhagen

http://www.econ.ku.dk/CAM/

Working paper no. 2014-03

Should I Stay or Should I Go: Peer Effects in Absenteeism

Mette Ejrnæs, Anders Holm and Daniel Le Maire

Should I Stay or Should I Go: Peer Effects in Absenteeism*

Mette Ejrnæs†, Anders Holm‡ and Daniel le Maire§

University of Copenhagen

October 11, 2014

Abstract

We study peer effects in absenteeism in compulsory school using unique Danish data that track students across schools and allow us to separate unauthorized absence from absence due to e.g. illness. Our identification strategy uses lagged peer absenteeism for new students in the class and student fixed effects to identify peer effects. We find significant peer effects in absenteeism. However, the peer effects are heterogeneous. For example, entry of a boy with a high level of absenteeism into a class only tends to increase the absence of the incumbent boys, but the entry of a high absenteeism girl tends to increase the absence of both incumbent boys and girls. Further, peer effects also seem heterogeneous across own absenteeism such that the entry of a peer with a high level of absence has larger effects for students who already have a high level of absenteeism.

JEL codes: I21, I31, C31, C33

Key words: absenteeism, peer effects, compulsory school.

1 Introduction

The increase in student absenteeism has raised public concern in the US as well as many other places, see e.g. Harris (2013) and Jackson et al. (2012). The effect of absenteeism on students' achievements has also

*We thank Fane Groes, Anna Folke Larsen, Nikolaj Harmon, Arthur Lewbel and seminar participants at CAM, the Copenhagen Education Network and KORA for valuable comments. Furthermore, we thank the municipalities of Copenhagen and Frederiksberg for access to data and Mette Friis-Hansen for efficient research assistance. We are grateful for financial support from the Danish Strategic Research Council.

†Department of Economics, University of Copenhagen. email: mette.ejrnes@econ.ku.dk.

‡Department of Sociology, University of Copenhagen. email: ah@soc.ku.dk.

§Department of Economics, University of Copenhagen. email: daniel.le.maire@econ.ku.dk

spurred interest among researchers. Romer (1993) is one early example of interest in the effects of absenteeism among college students. Other examples are Durden and Ellis (1995) and Dobkin et al. (2010). They all find negative effects of absenteeism. However, while the magnitude and the effects of student absenteeism on cognitive outcomes are well established, the causes of absenteeism are less well understood. Knowing the mechanisms behind student absenteeism are obviously important in order to design policies that aim at bringing down student absenteeism. This paper analyzes whether absenteeism is in part driven by peer effects. Our contribution is fourfold. First, we are the first to examine peer effects in absenteeism in compulsory schools. Second, we use unique detailed data on absenteeism and a novel approach to provide evidence of peer effects in absenteeism. Third, we show that peer effects are heterogeneous across students, and fourth, we provide indicative evidence for - at least part of - the mechanisms behind the peer effects.

Our identification strategy builds on the approach in Lavy et al. (2012), who use lagged peer achievements and student fixed effects to identify peer effects. Our identification strategy relies on the idea that we can observe whether students respond to a new classmate's absence. To break the reflexion problem, we measure the unauthorized absence of an incoming student before he or she enters the class, where the absence could not have been influenced by future classmates. Furthermore, we use fixed effects regressions at student-class level to control for time invariant characteristics and non-random allocation of students to classes. To test for time-varying confounders, we also run two placebo tests in order to rule out selection on trends in absenteeism in the receiving class. Our identification strategy uses changes in peer composition that is more narrow than in other studies of peer effects which use differences in class composition across grades and years (Hoxby, 2000), or across subjects and years (Lavy et al., 2012). By only using the variation in peer composition that comes from a new student entering the class we can specifically address when and how our estimator will be biased.

Our empirical analyses are based on a large and detailed Danish dataset from schools in the metropolitan area of Copenhagen. The data track students across schools and classes and, thereby, allow us to follow students who change schools, which is crucial for our identification strategy. Our results show evidence of peer effects in absenteeism in compulsory schools. Furthermore, we find heterogeneous peer effects. Peer effects are strongest for students who already had a high level of absenteeism. Also, we find that girls only respond to the absenteeism of other girls, while boys react to the absenteeism of students of both genders. The heterogeneous effects provide additional evidence on the underlying mechanisms driving peer effects. Given our detailed data, we can also examine whether students are absent on the same days, and we find evidence supporting the hypothesis that the peer effects in absenteeism are partly driven by the fact that students enjoy being absent more if they are jointly absent with their classmates.

Our study relates to two branches of the literature on student behavior: the literature on absenteeism and the literature on peer effects. The reason why increasing student absenteeism has raised concerns is that higher absenteeism would be expected to hamper human capital formation. This intuitive relationship has found solid empirical support as several researchers have provided a strong link between school attendance in general and cognitive outcomes. Romer (1993) and Durden and Ellis (1995) show that truancy affects achievement for students in an economics course. Gottfried (2011) finds, using a sibling fixed effects approach, that school absence affects academic achievement. Stanca (2006) finds that attending class versus not attending has a causal effect on learning outcomes for university students. These results are in accordance with the literature examining the the effect of official number of school days.[1] In sum, there is ample evidence that time spent in school, whether it is generated by absence or changes of school length, has a causal effect on test scores and academic achievement. Thus, affecting how much time students spend in school, e.g. through peers, may have an important causal effect on human capital formation.

Economists have increasingly focused attention on school peer effects, see Epple and Romano (2011), Sacerdote (2011) and many others. Also, Card and Giuliano (2013) study peer effects on absenteeism among adolescents as one of many outcomes. Recent literature on peer effects in school puts emphasis on non-linear effects of peers. For instance, Lavy et al. (2012) find that being with students from the bottom of the ability distribution lowers own outcomes. Burke and Sass (2013) find no linear peer effects but peer effects in different parts of the ability distribution of students.

The remainder of this paper is organized as follows. In the next section, we review some potential ex-

[1] Büttner and Thomsen (2010) use a reform that cut length of schooling to show that time spent in school has a positive impact on learning outcomes. Fitzpatrick et al. (2011) exploit quasi-randomness in the timing of assessment dates to find that an additional year of schooling results in large gains in test scores. Wills (2014) finds that fewer days at school due to a teacher strike has a large and significant impact on students' test scores. Webbink and Gerristen (2013) find that length of schooling, as caused by varying school age entry rules across countries, affects PISA test scores. Carlson et al. (2012) find that the variation in days in school generated by random variation in test taking affects test scores. Pischke (2007) finds that changes in school year length has a significant effect on test scores in Germany.

planations for why peer absenteeism might affect own absenteeism. Section three explains briefly the Danish educational system, section four explains our identification strategy and our econometric method. Section five illustrates the data we use to provide evidence of our hypothesis, section six shows our empirical results, while section seven offers some concluding remarks.

2 Peer effects and absenteeism

In this section, we review different explanations for peer effects in absenteeism relevant for our study. First, students may derive value from joint leisure from absence in the sense that being absent from teaching is more valuable when spent with one or more classmates. Hence, the arrival of a student with high absenteeism increases the value to time spent outside teaching. If students also derive utility from effort, e.g. through higher grades, one would expect that students who derive less value from teaching are most influenced by the arrival of a new student with a high level of absenteeism. Conversely, a new student with a low level of absenteeism should not affect peer absenteeism as this student does not improve the utility of spending (more) time outside class. This is in line with Cicala et al. (2011) who argue that when agents are sorted into new peers, their behavior depends on the placement in the new peer distribution, and Fruehwirth (2012) who finds that peer effects primarily exist within reference groups (race in her case) and are not class wide.

Second, Akerlof and Kranton (2002) introduce a model where the utility of a student's effort in school in part depends on student identity in the sense that the student derives utility from identifying herself with different stereotypes - in the case of Akerlof and Kranton, these are school leading crowd, burnouts and nerds. Each stereotype is associated with a certain amount of effort. The identification with each of the stereotypes influences optimal student effort. As absenteeism is the reverse of student effort, the model of Akerlof and Kranton is relevant when considering peer effects of absenteeism. Changes in the distribution of student absenteeism e.g. by the arrival of a new student, may affect the utility of a particular student identifying with a certain stereotype. If the new student has a high level of absenteeism, this may increase the utility of similar students of identifying with the burnout or the low effort stereotype. This implies that students with a high level of absenteeism are more affected than students with a low level of absenteeism. Conversely, if the new student has a low level of absenteeism, this may increase the utility of relatively low absenteeism students to identify with the nerd or leading crowd stereotype but it may not affect those student with a high level of absenteeism who are already identifying themselves with the burnout identity. As with joint leisure, this mechanism is also supported by the findings of Fruehwirth (2012) as the utility of identifying with a stereotype depends on that stereotype being present among ones peers.

Third, if disruptive behavior and student absence is positively correlated, a new student with a high level of absenteeism may also imply a negative impact on learning opportunities for other students. This in turn implies that the returns for the other students of showing up in school decrease, which again may induce some students to increase their absenteeism. This is in line with Carrell and Hoekstra (2010) and Kristoffersen et al. (2014) who find that having a disruptive student in the class decreases test score results.

Finally, if punishment for absenteeism is more difficult to enforce when several students are absent, the arrival of a new student with a high level of absenteeism lowers the average dis-utility from being punished for being absent for the other students in the class. This increases the incentive for being absent for all students in the class. Moreover, if students also derive utility from being present in class (e.g. from learning), this effect will be relatively more predominant for students with lower utility from learning.

In sum, we have provided four different but non-exclusive arguments as to why we should expect spillover effects from absenteeism. In particular, we have provided arguments as to why we should expect significant spillover effects from the arrival of a new student with a high level of absenteeism and why we might expect lower or no spillover effects from the arrival of a new student with a low level of absenteeism.

3 Danish compulsory school system

Compulsory schooling in Denmark covers nine years, usually starting at age six and ending in the 9th, or after a voluntary 10th grade, at age 15 or 16. The compulsory school is comprehensive and all students are taught the same curriculum. Some students might be taken out for special education based on their cognitive ability or other characteristics but they are subject to the same curriculum as students in ordinary classes.[2] Upon completing compulsory school almost all students take an exit exam in the core subjects in compulsory schools (Danish, Science, Math, English, and French or German). Since 1993, the Danish compulsory school system has been a completely de–tracked system from grade 0 to grade 9. Students enter kindergarten class (0th grade)

[2] Students in public schools (folkeskolen) that receive special education amount to 8.7 percent of all students (Sourcė: "Specialundervisning i folkeskolen: veje til bedre organisering", Undervisningsministeriet, June 2010).

and leave after final exams in 9th grade. Students are expected to stay with the same class throughout the entire time in compulsory school except if they change residence or if they, for other reasons, want to change schools. All subjects are taught to the students within the same class.[3] Thus, they are exposed to the same peers in all subjects throughout their entire time in compulsory school except if they change class or if new students are allocated into their class because they change residence or move – individually – between schools. Also the social life in school such as school trips and parties etc. are organized around the classes. Therefore, it is more meaningful to define classmates as the relevant peers rather than students of the same grade but in different classes. This definition is especially meaningful in the lower grades (less than 7th grade).

In the Danish compulsory school system, it is common that the same teachers are following the class for a couple of years, e.g. often first to third grade are taught by the same set of teachers, then fourth to sixth grade are taught by a new set of teachers and finally seventh to nineth grade by a third set of teachers. Furthermore, one of the teachers will be assigned as the primary teacher of the class (often the teacher who teaches Danish) and this teacher will be responsible for the well-being of the students. It is this teacher or the headmaster of the school who will take action in relation to absenteeism of students.

In Denmark, absenteeism has recently received increasing attention among policy makers and administrators in the compulsory school system and most schools have now implemented strict rules of how to act if students are absent. Further, schools are now expected to report absenteeism to the national educational statistical agency (through the municipalities) and reporting follows a specified set of rules of how absenteeism is recorded into unauthorized absence and absence due to illness . We return to details of the recording of absence in the data section.

4 Empirical approach

It is well recognized that identifying peer effects is not a simple task (see Manski (1993) and Sacerdote (2011) for a thorough description of identification problems in models with peer effects). In a standard linear-in-mean

[3] There are a few optional courses (like German, French, Arts and sports classes) in 7th- 9th grade that might be taught together with students from other classes.

model, peer effects can be described as

$$Y_{it} = \beta_0 + \beta_1 \bar{Y}_{it(-i)} + \beta_2 X_{it} + \beta_3 \bar{X}_{it(-i)} + \varepsilon_{it}, \tag{1}$$

where Y_{it} is the outcome of student i at time t, X_{it} contains individual specific characteristics, and $\bar{Y}_{it(-i)}$ ($\bar{X}_{it(-i)}$) is the average of absenteeism (characteristics) of the peers (in our case the classmates). In this model, a student's outcome is affected by the mean outcome of the classmates and the mean characteristics of the classmates. The direct effect of peers' outcome is measured by β_1 and called the endogenous peer effect; it is assumed that $0 \leq \beta_1 < 1$. The effect of peers' characteristics is measured by β_3 and called the contextual effect.

Using OLS to estimate the parameters in equation (1) is problematic for at least three reasons. First, because the outcome of student i is affected by student j and vice versa (the reflection problem), which implies that $\bar{Y}_{it(-i)}$ is an endogenous variable. Second, the selection into peer groups (here classes) may be based on unobserved characteristics e.g. parental characteristics. Third, in this specification it is not easy to disentangle the endogenous effect of the outcome of other peers β_1 from the contextual effect of the characteristics of peers β_3.

In this study, we will address the first two problems but we will not be able to separate endogenous peer effects from contextual effects. We overcome the reflection problem by measuring absenteeism of a future classmate before he or she has entered the class. We use the common way to overcome the endogenous formation of peer groups by using student fixed effects and assume that, conditional on individual effects, the selection into classes is random (see Hoxby 2000, Lavy et al. 2012).

To explain our identification strategy further, we illustrate it in a simple model. We assume that we observe absenteeism in two periods ($T = 2$) for students in the class. For simplicity, it is also assumed that there are n students in the class in period 1 and that one new student enters the class in period 2. Finally, we assume that there were no new students in period 1. The unauthorized absence for student i who is in the same class both in period 1 and 2 (we label them stayers) is measured by Y_{it}. The unauthorized absence of the new student is measured by Z_{it}. The measure Z_{i1} is measured the period before entering a new class. In our model, we assume that all individual characteristics are time invariant and, therefore, absorbed in an individual fixed effect μ_i.[4] This assumption is, of course, a strong assumption but given that in our empirical analyses we only observe students for four years, most of the student characteristics e.g. parental background information will

[4] This means that $\beta_2 X_{it} = \mu_i$.

be constant.[5] As mentioned earlier, we are not able to disentangle the contextual effects from the endogenous effects and thus we can only identify $\beta_1 + \beta_3$. To simplify the notation, we assume that $\beta_3 = 0$.[6] By using equation (1) and imposing the assumptions described above, our model is given by

$$Y_{it} = \beta_0 + \beta_1 \bar{Y}_{it(-i)} + \mu_i + \varepsilon_{it}. \tag{2}$$

In the first period, the average absenteeism in a class is given by[7]

$$\bar{Y}_1 = \frac{\beta_0}{1-\beta_1} + \frac{\bar{\mu}}{1-\beta_1} + \frac{\bar{\varepsilon}_1}{1-\beta_1}.$$

Assume a new student $i = n+1$ enters the class. The individual fixed effect of the new student is μ_{n+1}. The average absence of the stayers in the class (only students who are in the class in both periods) can be found as:

$$\bar{Y}_2 = \frac{\beta_0}{1-\beta_1} + \frac{\bar{\mu}_i}{1-\beta_1} + \frac{\beta_1}{(1-\beta_1)(n+\beta_1)}(\mu_{n+1} - \bar{\mu}) + \frac{n\bar{\varepsilon}_2 + \beta_1 \varepsilon_{n+1,2}}{(1-\beta_1)(n+\beta_1)}. \tag{3}$$

We can then investigate how the average absence of stayers changes when a new student enters the class

$$\bar{Y}_2 - \bar{Y}_1 = \frac{\beta_1}{(1-\beta_1)(n+\beta_1)}(\mu_{n+1} - \bar{\mu}) + \frac{n\bar{\varepsilon}_2 + \beta_1 \varepsilon_{n+1,2}}{(1-\beta_1)(n+\beta_1)} - \frac{\bar{\varepsilon}_1}{1-\beta_1}. \tag{4}$$

From the expression above we can see that if positive peer effects exist ($\beta_1 > 0$), the change in the average absence of stayers will depend on the fixed effect of the new student. If the new student has a higher individual effect than the new classmates i.e. $\mu_{n+1} > \bar{\mu}$, then the unauthorized absence of stayers will increase. We will use this observation to test empirically for peer effects.

In the empirical analyses, we will use the unauthorized absence of the student before he or she moves into the class, that is $Z_{n+1,1}$, as a proxy for the student's fixed effect: μ_{n+1}. The idea is that since $Z_{n+1,1}$ is measured in the period before entering the new class, it cannot be affected by the new classmates.

We propose to use the following empirical model, where $Z_{n+1,1}$ is used as a proxy for μ_{n+1}:

$$\begin{aligned} Y_{i1} &= \alpha_0 + \mu_i + \upsilon_{i1} \\ Y_{i2} &= \alpha_0 + \alpha_1 Z_{n+1,1} + \mu_i + \upsilon_{i2}. \end{aligned}$$

The model can be interpreted as a reduced form model. The estimation is done by a fixed effect estimator. In Appendix A.2, we have derived the relation between the parameters in the empirical model and the parameters

[5] In the empirical analyses we control for class grade and can allow that the absenteeism changes over class grades.

[6] In all the calculations in the following paragraphs β_1 can be replaced by $\beta_1 + \beta_3$.

[7] The details about the average outcome of stayers are given in Appendix A.1.

in equation (2). We can show that the fixed effect estimator converges in probability to

$$\underset{C\to\infty}{\text{plim}}\hat{\alpha}_1 = \frac{\beta_1}{(1-\beta_1)(n+\beta_1)} \frac{\left(E(\mu_{n+1}^2) - E(\bar{\mu}\mu_{n+1})\right)}{E(\mu_{n+1}^2) + \sigma_m^2},$$

where C is the number of classes and σ_m^2 is the variance of measurement error in $Z_{n+1,1}$. $E(\bar{\mu}_c\mu_{n+1,c})$ is the covariance between the individual effect of the new student and the average of the individual effects of new classmates.[8] This covariance indicates whether there exists positive or negative sorting of new students into classes. If $E(\bar{\mu}\mu_{n+1}) > 0$, it indicates positive sorting, such that students with high absenteeism are more likely to enter classes where the other students also have high absenteeism. From the expression above, we see that if $\underset{C\to\infty}{\text{plim}}\hat{\alpha}_1 > 0$ then $\beta_1 > 0$ and peer effects exists.[9] Therefore, if we can find significant and positive impacts of the lagged absenteeism of the incoming students on their new classmates' absenteeism, we will take this as evidence that peer effects exist.

It is not possible directly to obtain an estimate of β_1 because $Z_{n+1,1}$ is only a proxy for the true individual effect, μ_{n+1}, and there might be sorting of incoming students and classes. However, if there is either no sorting or positive sorting, we can show that the estimate of α_1 is likely to be a downward biased estimate of $\frac{\beta_1}{(1-\beta_1)(n+\beta_1)}$. We also see that α_1 will be smaller for classes with a large class size compared to small classes, which intuitively is because the new student has a relatively smaller impact on the mean in a large class.

To get an idea of the sorting of new students, we can use the OLS estimator. In Appendix A.2, we show that the asymptotic difference between OLS and fixed effect depends on the term $E(\bar{\mu}\mu_{n+1})$. If the OLS estimator of α_1 is larger than the fixed effect estimator, it suggests positive sorting of new students. In this case, we can calculate a lower bound on the estimate of β_1.

In the empirical model, we assume that $Z_{n+1,1}$ is uncorrelated with $\Delta\upsilon_{i2} = \upsilon_{i2} - \upsilon_{i1}$. A threat to this assumption is if new students are sorted into new classes on the basis of their own absenteeism and the time-trend of absenteeism in the receiving class. We address this potential problem by employing two different placebo tests.

In the empirical analyses, we exploit that the estimation data set covers four years and use the variation both in the unauthorized absence of the incoming students but also the fact that not all classes receive new students every year. Therefore we include a dummy for whether new students enter the class $c(i)$, $D_{c(i)t}$.

[8] Notice that $E(\bar{\mu}_c\mu_{n+1,c}) = \text{cov}(\bar{\mu}_c, \mu_{n+1,c})$ since $E(\bar{\mu}_c) = 0$.

[9] Notice, that $\frac{(E(\mu_{n+1}^2) - E(\bar{\mu}\mu_{n+1}))}{E(\mu_{n+1}^2) + \sigma_m^2} < 1$ if positive sorting or no sorting.

Furthermore, there might also be classes that receive more than one student and in this case we use the average unauthorized absence of the new students one year before they enter the class $\bar{Z}_{ct-1}$. Furthermore, we include characteristics of the class and school, e.g. class size and level of grade; $S_{c(i),t}$. The empirical specification is given by:

$$Y_{it} = \alpha_0 + \alpha_1(\bar{Z}_{c(i),t-1} \times D_{c(i),t}) + \alpha_2 D_{c(i),t} + \alpha_3 S_{c(i),t} + \mu_i + \upsilon_{it}. \quad (5)$$

This specification contains the linear-in-mean model in equation (2) as a special case. In the linear-in-mean model, we see that whether a new student has a positive or negative impact on the classmates depends on level of absenteeism of the new student, μ_{n+1}, compared to the new classmates $\bar{\mu}$, see equation (3). If the new student has the same level of absenteeism, $(\mu_{n+1} - \bar{\mu}) = 0$, then there will be no impact. In our specification (equation (5)) we can obtain the linear-in-mean model by imposing that $\alpha_2 = -\alpha_1\bar{\mu}$.[10] In the empirical analyses, we do not impose the linear-in-mean assumption, but estimate α_2. Thus, we can investigate whether the linear-in-mean model holds; that is whether $\alpha_2 < 0$. Our specification also allows for other types of peer effects; e.g. if $\alpha_2 \geq 0$ then it implies that even new students with no absenteeism do not affect the classmates by lowering their absenteeism. The specification also allows for the possibility that new students entering the class in itself affects the students' outcome.

An alternative approach is to use an IV estimation of the Manski model (equation (1)). In our context, we could use $\bar{Z}_{c(i),t-1}$ as an instrument for $Y_{c(i),t}$. However, Angrist (2014) argues strongly against such an IV approach for estimating peer effects as the instrument $\bar{Z}_{c(i),t-1}$ is a weak instrument. Hence, we refrain from using the IV approach. However, what we do when we estimate equation (5) is exactly the same as estimating a first-stage equation for the IV estimation.

Finally, we could also have chosen to use variation in average outcome of classmates induced by students leaving the class. However, we consider the decision to leave the class as an endogenous decision and, therefore, do not use this variation. Students who leave the class are not defined as stayers and we will not include them in the main analyses. However, we will use the leavers to perform a placebo test later.

[10] Our model would then be $Y_{it} = \alpha_0 + \alpha_1 D_{c(i),t}(\bar{Z}_{c(i),t-1} - \bar{\mu}_i) + \alpha_3 S_{c(i),t} + \mu_i + \upsilon_{it}$

5 Data

5.1 Source of data

The data used in this paper stem from the registration of absenteeism in schools at the metropolitan area of Copenhagen. The data contain information on schools in the Copenhagen and Frederiksberg municipalities. Focus on school absenteeism and demands on school reporting on absenteeism has increased over the years. Schools are required to record student absenteeism and to clearly distinguish between four different categories of absence: illness, authorized absence, unauthorized absence (full day) and unauthorized absence (part day). The uniqueness of these data is that the absence is registered using a unique personal identifier. This means that if a student moves from one school to another school within the metropolitan area we can follow the student.

The absenteeism is registered on a daily basis and the reason for absenteeism is recorded. We have data for the period August 2007 to June 2012. In our main analyses we will not consider absence due to illness, but to validate our data we also examine absence due to illness. To give an overview of the data, we have aggregated the days of absence into a weekly measure for each of the four different types of absence. In Figure 1, we show the total number of students who, on average, were absent in a given week in the period August 2008 to June 2012.[11] First, we notice that illness is the most common reason for absence. The figure shows a clear seasonal pattern in the days of *absence due to illness*. The days of absence almost doubles during the winter compared to the summer months. The data also reveal a very high number of students who were absent in week 45 in 2009, which coincides with a flu epidemic in Denmark in that particular week.[12] Similarly, we can explain the spikes for the other years as outbreak of flu epidemics. If we turn to the second most common reason for absence, *unauthorized absence (part day)*, we also see a seasonal pattern with more students not being present in some of the lessons during the winter, although it is much less pronounced. The third reason for absence is *authorized absence*. In this case, the parents have ensured that the student is granted permission from the school to be absent. Examples of authorized absence are visits to doctors and holidays outside the official holiday periods.

[11] Notice, for vacation weeks the absence is zero. The graphs are only made for 1st-7th grade because we then have a balanced sample.

[12] See http://www.ssi.dk/Aktuelt/Nyhedsbreve/INFLUENZA-NYT/2009-2010/INFLUENZA-NYT%20-%20uge%205%20-%202010.aspx

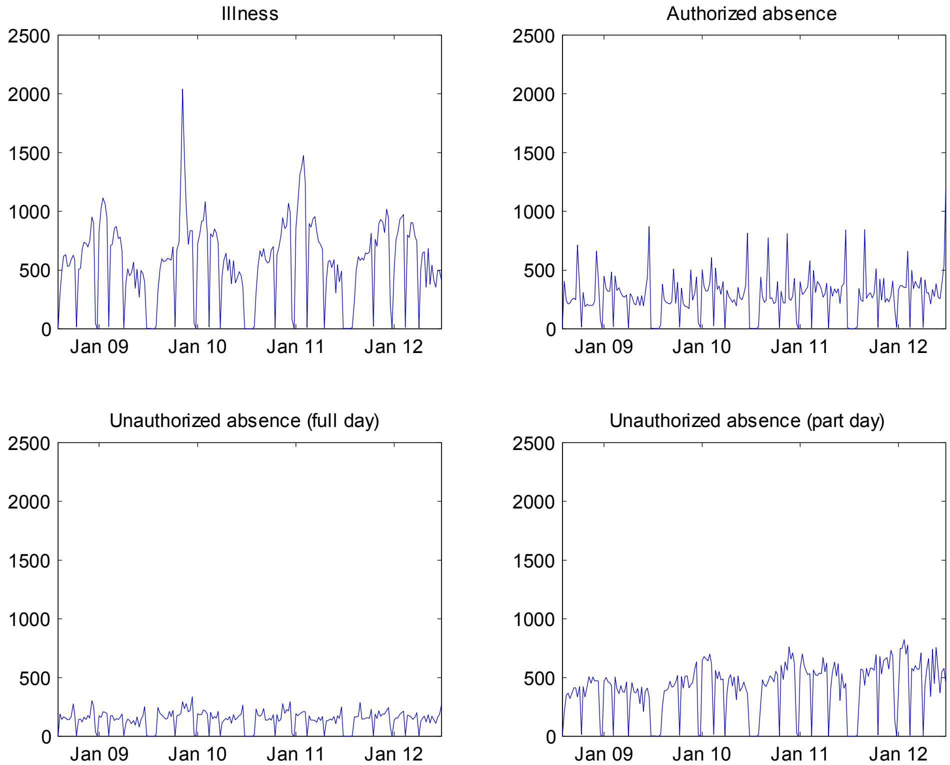

Figure 1: Different types of absence

With respect to the latter, notice the clear pattern that students are more likely to be absent for legal reasons just before or after the official holidays. Finally, we have *unauthorized absence (full day)* which is lower than the other measures and fairly constant over the period and over the school year.

In sum, the overall picture of the data suggests that the data are accurately registered. We can validate the data for absence due to illness by comparing the data to official statistics for outbreaks of the flu and we see a high degree of co-movement. This indicates that these data are accurately measured and we see no reason why there should be systematic mis-reporting in the three other measures of absence given that they are registered in the same system and by the same teachers.

5.2 Sample selection and definition of unauthorized absence

Our sample contains data from five school years, 2007/08 – 2011/12, covering grades 0 to 9. Before converting data from student-class spells, we have de-selected spells for students with special needs who are taught in special classes. The total number of student-year observations for Copenhagen and Frederiksberg municipalities is 212,597 observations for 58,106 students and 173 schools.

We delete schools not reporting absence to the municipality each quarter during our sample period. It is almost exclusively independent and private schools which do not report absence, and observations for these 109 (typically small) schools are deleted. We only delete seven of the public schools and all seven schools did not report absence for the school year 2007/08 due to closures and school mergers.[13] Deleting all observations for schools not reporting absence in just a single quarter leaves us with $150,160$ observations for $43,495$ students.

It turns out that absence of the older cohorts of students, 7th-9th grade in 2007/08, is poorly recorded so we discard observations for grade $7-9$ in 2007/08, grade $8-9$ in 2008/09, and grade 9 in 2009/10. Hence, the sample is reduced to $143,822$ observations for $41,087$ students.

The first observation for each individual is used to construct lagged absence. Hence, in selecting the estimation sample, we sequentially drop the first school year 2007/08 ($21,106$ observations), observations for grade 0 ($15,874$ observations) and remaining observations where we do not observe the student in the previous year ($5,288$ observations). This leaves us with $101,554$ observations for $34,610$ students.

We remove observations with class sizes, respectively, below the bottom and above the top 0.25 percentile. Sometimes two classes are merged, but the primary objective in this paper is to focus on classes receiving only a few students. With class mergers, the short-run interaction between the new students and incumbent students is likely to be different and since classes merged together are usually from the same school we restrict attention to classes which receive three students or less. However, as a robustness check we show that we obtain similar, albeit slightly less significant, results using all student changes. Restricting attention to classes receiving less than four students, reduces the sample to $97,832$ observations for $34,272$ students.

In our estimations, we focus on the effect on the absence of the students not changing class in between two school years when one or more students enter the class. We have $7,867$ observations for class changers and $89,965$ observations for class stayers. We restrict attention to class changes happening in the summer between two school years to avoid the problem of students attending the new class for part of the previous year. For this, we use the spell dimension of the data set to require that the class changer's spell in the previous class runs until the end of that school year and that the spell for the new class does not begin before the subsequent school year begins. This implies that we effectively only consider $3,117$ class changes for $2,984$ students. For

[13] School mergers only took place in the summer of 2008 during our sample period and schools participating in the school mergers and continuing to exist reported absence in all quarters, we observe.

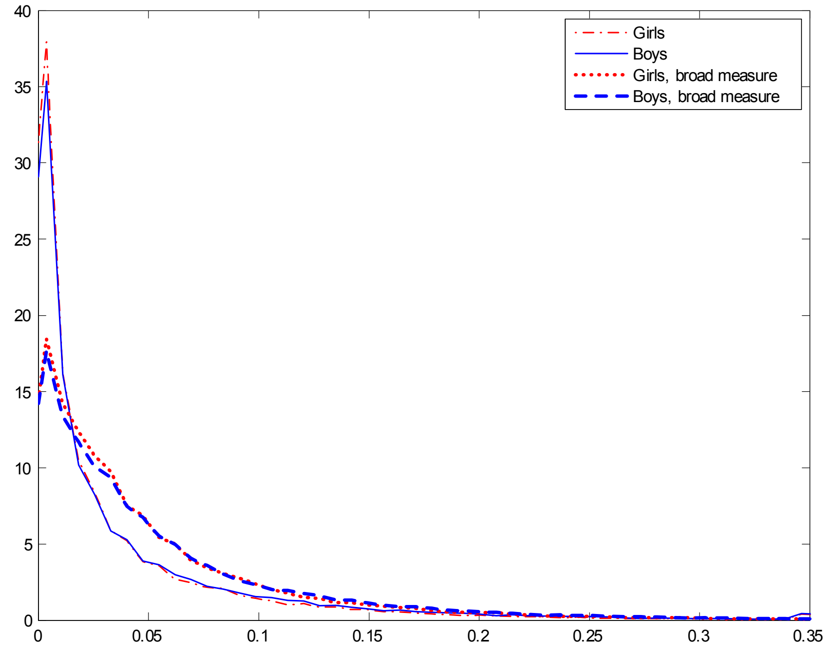

Figure 2: The distribution of measures of absence for girls and boys

the class stayers, we also require that they do not change class within the first month of the new school year. Thereby, we reduce the sample of stayers to 89,188 observations for 33,256 students allocated to 1,630 different classes in 62 schools.

In our analyses, we define our measure of absenteeism based on the number of days the student was *unauthorized absent* full or part day. We have chosen this measure because it clearly captures absence due to violations of the rules of the school. In the robustness analyses, we also use a different definition of absenteeism where we include authorized absence. Based on the number of days of absence and the number of school days, we calculate the fraction of days during a school year the student has been unauthorized absent.[14] A normal school year is around 200 days.[15] On average, we find that students are unauthorized absent (entirely or partly) about 7 days per year, which is equivalent to about 3.6 percent of the year. In Figure 2, we show the distribution of absence for boys and girls using both our measure of unauthorized absence and the broader measure including authorized absence. The graph shows that a high fraction of students have very little unauthorized absence.

[14] We censor the data by dropping the top 0.25 percent of the observation. The 99.75 percentile for our measure of unauthorized absence is 0.415.

[15] The number of school days varies a little from year to year (depending on holidays) and the grade of the student.

5.3 Descriptive statistics

In Table 1, descriptive statistics are reported for our sample. First, it is noticed that the average unauthorized absence for students is 3.6 percent. Using the broader measure of absence including authorized absence but still excluding absence due to illness, the average increases to 5.3 percent.

Besides the registration of absence, the data contain information on the students and the classes. For the students, we know the grade and gender. Furthermore, we have unique identifiers for the students, classes and schools. Based on these identifiers we can calculate the class and school size. In Table 1, we show descriptive statistics for our estimation sample: the stayers, and for comparison, we also shows descriptives on the entrance students. Our estimation sample covers grades 1 to 9 and contain 89,188 observations. Table 1 shows that the average grade is around 4. About half of our sample is boys. The average class size is around 21.5. There seem to be no systematic differences between stayers and entrance students.

When looking at the average unauthorized absence for different grades (the figure is not shown), it turns out that the unauthorized absence is almost constant for 1st to 5th grade and is around 3 percent. From 5th grade unauthorized absence starts to increase up to on average 5.8 percent in grade 9.

If we compare unauthorized absence for classes and schools of different sizes, we find that unauthorized absence is higher in small sized classes. For classes below the median class size of 22 students, the average absence is 3.8 percent, while for classes with 23 and more students, the average is 3.0 percent. This could be due to selection away from classes with bad teachers or bad peers. In our empirical analysis, we take into account such selection effects by eliminating student and class fixed effects and we further address the issue in two placebo tests.

6 Results

We begin by presenting the estimation results based on equation (5). In Table 2, we present both OLS estimations (columns 1-2), student fixed effects estimations (column 3-4) and student-class fixed effects estimations (column 5-6).[16] For each type of estimation we present results from two different specifications: a simple spec-

[16] In this specification, we only use variation for stayers. We have fixed effects for each combination of class and student for stayers.

	No. of obs.	Mean	Std.	Min	Max
Stayers					
Student absence. in pct.	89,188	0.036	0.058	0	0.415
Student absence in pct. incl. authorized absence	89,188	0.053	0.065	0	0.455
Class size	89,188	21.54	3.43	11	30
Class grade	89,188	4.38	2.40	1	9
Years in class (truncated)	89,188	2.05	1.03	1	4
School year	89,188	2009.66	1.090	2008	2011
Boy	89,188	0.502	0.500	0	1
Entrance students					
Student absence. in pct.	3,117	0.039	0.066	0	0.415
Student absence in pct. incl. authorized absence	3,117	0.055	0.072	0	0.455
Class size	3,117	22.44	4.11	11	30
Class grade	3,117	4.58	2.47	1	9
Years in class (truncated)	3,117	0	0	0	0
School year	3,117	2009.52	1.11	2008	2011
Boy	3,117	0.507	0.50	0	1

Table 1: Descriptive statistics of the sample

ification and a specification with class characteristics. If we compare the estimates of the absenteeism of new students, α_1, from equation (5), we see that they are all significant at the 1 percent level, but OLS estimates are substantially higher ($0.088 - 0.114$) than the fixed effect estimates ($0.022 - 0.046$). The comparison reveals that OLS seems to over estimate the peer effects. As discussed in section 4, the higher OLS estimates indicate positive sorting of new students into new classes. In other words, students changing classes share the same unobserved characteristics as the students in their new class. Controlling for student-class fixed effects, or student fixed effects lowers the estimate substantially. When including class characteristics, the estimates are further reduced; the estimate of α_1 varies between $0.022 - 0.024$. Our preferred specification is the specification in column 6 with student-class fixed effect and school and class characteristics and in the following we will concentrate the discussion on this baseline specification.

6.1 Baseline specification

Table 2 suggests that there exist peer effects in absenteeism which are significantly different from zero. In the preferred specification with student-class fixed effects, the estimate of α_1 is 0.022. This implies that if a new student enters a class and he or she was unauthorized absent 9.4 percent, which is one standard deviation above the mean, in the year before the move, the absenteeism in the class goes up by about 0.30 percentage points on average. This is equivalent to 0.5 extra days of unauthorized absence. This should be compared to a mean of 3.6 or 7.2 days with unauthorized absence. Since we find evidence for positive sorting, we can calculate a lower bound estimate of β_1. The lower bound estimate of β_1 is estimated to about 0.33,[17] corresponding to a 0.0187 increase in absenteeism from a one standard deviation increase in peer absenteeism. Our estimate is in the range of the estimated peer effects which are found in the literature (see an overview of empirical estimates of β_1 in Sacerdote (2011) and Gibbons and Telhaj (2008)).

Another interesting result is that the coefficient on the dummy variable for new students in the class is insignificant. This implies that there is no positive impact of having a new student in the class with lower unauthorized absence than the class e.g. zero absenteeism. We have checked for nonlinearity in the absenteeism of the new student but do not find significant non-linear effects. Also, we find no significant impact of class size on the level of absenteeism in the class when we have controlled for individual fixed effects. However, our OLS

[17] For these calculation we have set the class size $n = 22$.

	OLS		Student fixed effect		Student-class fixed effects	
Avg. absenteeism	0.114***	0.088***	0.046***	0.024***	0.043***	0.022***
of new students	(4.65)	(3.96)	(4.64)	(2.77)	(4.41)	(2.63)
New students in class (0/1)	−0.004***	−0.002	−0.002**	−0.000	−0.002**	−0.000
	(−2.59)	(−1.22)	(−1.99)	(−0.22)	(−2.12)	(−0.37)
Class size		−0.004***		−0.000		−0.001
		(−3.53)		(−0.02)		(−0.95)
Class size squared		0.000***		0.000		0.000
		(2.13)		(0.22)		(1.27)
Class grade * year interactions	No	Yes	No	Yes	No	Yes
R^2	0.003	0.035	0.001	0.051	0.001	0.047
No. of observations	89,188	89,188	89,188	89,188	89,188	89,188

t-statistics clustered at class-year level are in parentheses. *: 0.10, **: 0.05, ***: 0.01

Table 2: Estimation results

estimate suggests a negative correlation between absenteeism and class size as we see in the descriptive analysis.

6.2 Robustness checks

In this section, we perform three robustness checks. We start by redefining our measure of absenteeism to also include authorized absence. The first column in Table 3 shows that the estimate of α_1 for the broader measure of absenteeism is 0.031 and significant. The estimate is slightly higher than in the baseline specification.

In the second column, we deviate from our baseline specification by including (primary) teacher-student-class fixed effects in the estimation. One concern would be if a class receives a problematic student and as a response allocates a good teacher to the class. To address this concern, we will control for (primary) teacher fixed effect, which means that we are only using the variation for classes that keep the same primary teacher. Our estimate of α_1 is almost unchanged; 0.020. We, therefore, conclude that our results are not sensitive inclusion of teacher fixed effects.

We also extend our sample to include students in classes that receive more than 3 new students at the same time. Here the concern is that the peer impact might be different if the class is receiving many new students at the same time. Another concern could be that classes that receive many new students are special in some unobserved ways. We check whether our results are sensitive to this. In column 3 we report the estimate of α_1. The estimate is only marginally lower and significant at the 5 percent level.

Finally, we report the estimates of a Poisson regression with student-class fixed effects in the fourth column of Table 3. As in the preferred specification in Table 2, we obtain a positive effect of lagged absenteeism which is significant at the 1 percent level. The main difference from the linear fixed effects model is that absenteeism is significantly declining in the class size.

6.3 Placebo tests

To further assess the robustness of our findings, we carry out two different placebo tests. Both placebo tests investigate whether our results are driven by spurious correlation due to unmeasured heterogeneous trends in absenteeism in the receiving class. Our results could be due to a correlation between unmeasured trends in the receiving class and absenteeism of the incoming students such that, e.g., students with high absenteeism are allocated to classes with increasing absenteeism. The first placebo test uses lagged information on absenteeism on the incoming student and lagged absenteeism for the receiving class. These should be uncorrelated because at time $t-1$ the new student has not yet entered the class. However, if new students with high absenteeism were allocated to classes with increasing absenteeism we would expect to see a positive correlation. The test is performed in the following regression model

$$Y_{it-1} = \gamma_0 + \gamma_1(\bar{Z}_{c(i),t-1} \times D_{c(i),t}) + \gamma_2 D_{c(i),t} + \gamma_3 S_{c(i),t-1} + \mu_i + \varepsilon_{it-1},$$

where we test the null: $H_0 : \gamma_1 = 0$.

The results from the first placebo test are reported in Table 4 in Panel A. We report the estimation results for the two different types of fixed effect estimations. From the table we see that there is no significant correlation between the absenteeism of incoming students the year before they enter the receiving class and

	Broader measure of absenteeism	Student-teacher fixed effects	Extended sample	Fixed effects Poisson regression
Avg. absenteeism of new students	0.031^{***}	0.020^{**}	0.019^{**}	0.002^{***}
	(3.22)	(2.44)	(2.13)	(2.68)
New students in class (0/1)	-0.001	-0.000	0.000	-0.002
	(-0.66)	(-0.18)	(-0.62)	(-0.15)
Class size	-0.001	-0.002^{*}	-0.002	-0.42^{**}
	(-0.58)	(-1.66)	(-1.36)	(2.15)
Class size squared	0.000	0.000^{*}	0.000	0.001^{***}
	(0.81)	(1.89)	(1.64)	(2.74)
Class grade * year interactions	Yes	Yes	Yes	Yes
R^2	0.033	0.039	0.047	
No. of observations	$89,188$	$87,089$	$91,143$	$69,394$

Student-class fixed effects are used in columns 1, 3, and 4. The lower number of observations in the fixed effects Poisson model is due to groups with only zeros or groups with only one observation. In columns 1-3, t-statistics clustered at class-year level are in parentheses. In column 4, t-statistics clustered at person-class level are in parentheses. *: 0.10, **: 0.05, ***: 0.01

Table 3: Robustness checks

	Panel A (dep var: lagged absence of stayers)		Panel B (dep. var: absence of leaving students)	
	Student fixed effects	Student-class fixed effects	Student fixed effect	Student-class fixed effects
Avg. absenteeism of new students	0.012	0.011	−0.009	−0.010
	(1.34)	(1.24)	(−0.49)	(−0.54)
New students in class (0/1)	−0.001	0.000	0.001	−0.000
	(−0.73)	(−0.45)	(−0.32)	(−0.23)
Class size	0.001	−0.000	0.002	0.006*
	(1.09)	(−0.15)	(0.70)	(1.70)
Class size squared	−0.000**	−0.000	0.000	0.000
	(−2.04)	(−0.73)	(−0.54)	(−1.43)
Class grade * year interactions	Yes	Yes	Yes	Yes
R^2	0.045	0.041	0.051	0.053
No. of observations	89,188	89,188	9,054	9,054

t-statistics clustered at class-year level are in parentheses. *: 0.10, **: 0.05, ***: 0.01

Table 4: Placebo tests

absenteeism of the students in the receiving class the year before receiving the incoming student. We find the largest effect in the model of 0.012, where we take student fixed effects but the effect is still far from significant. Hence, the test supports the assumption that our results are not due to unmeasured trends in absenteeism in the receiving class.

The second placebo test examines whether the absenteeism of students who left the receiving class is affected by incoming students' absenteeism (still measured the year before class change). There should be no correlation between the absenteeism of these students unless either students who leave the class are special and incoming students are sorted according to the leaving students. This could be the case if students with high absenteeism are sorted out of the class, e.g. because the class has a particular trend in absenteeism and that this trend for some reason also attracts students with a high level of previous absenteeism. The regression model for the second placebo test is given by

$$W_{it} = \tau_0 + \tau_1(\bar{Z}_{c(i),t-1} \times D_{c(i),t}) + \tau_2 D_{c(i),t} + \tau_3 S_{c(i),t-1} + \mu_i + \varepsilon_{it},$$

where W_{it} is the absence of the student leaving the class between year $t-1$ and t. The null hypothesis is $H_0 : \tau_1 = 0$.

The second placebo test is reported in Table 4, Panel B. The number of observations is much lower than in our baseline specification and robustness checks, because we only use observations for students leaving the class. We find no correlation between absenteeism of entering and leaving students. Thus, the test suggests that there is nothing special about those students who leave a class compared to those students who enter this class.

Taken together, our placebo tests indicate that our results are robust to the assumption of no correlation in the trend in absenteeism of the receiving class and the absenteeism of the incoming student.

7 Heterogeneous peer effects

In this section we extend our baseline specification to investigate whether peer effects are heterogeneous. Goldsmith-Pinkham and Imbens (2013) argue, that the homogenous linear-in-mean model is unlikely to hold in practice because peers are not equal and peer groups may be measured with errors as peer effects may primarily work through friendship networks, as suggested by Card and Giuliano (2013), and not entire school classes.

To begin with, we consider whether the peer effects are gender specific. These results are shown in Table 5. If we compare the peer effects for boys and girls (compare column 1 and column 3) we see that boys respond more strongly to the absenteeism of new students compared to girls. This is in line with the findings of Bertrand and Pan (2013) that boys respond more strongly to peer risky behavior (in our case, absenteeism) than girls. We also consider whether a student is more sensitive to classmates of the same gender. To do this, we break up the absence of the newcomers into absence of new girls and absence of new boys. Then, we can exam whether the absence of, for example, girls is more influenced by the absence of new girls compared to the absence of new boys (see results in column 2 and 4). Interestingly, we find that girls only respond to absenteeism of new girls in the class. This finding can be corroborated with the idea that peer effects in part work through friendship networks, such that girls more often form close friendships with other girls. This explanation is in line with the arguments in Fruehwirth (2012). This asymmetry is less pronounced for boys who are sensitive to peer absenteeism from both genders.

In order to see whether we can further separate out the different mechanisms driving peer effects in absenteeism, we investigate whether these are in part driven by joint leisure when being absent. To do this we exploit that we have information on the exact dates of absence. This allows us to investigate whether the students are unauthorized absent on the same days, which could indicate that they value joint leisure. However, we should emphasize that we will not be able to rule out other channels of peer effects.

The idea behind the following analysis is quite simple. We pick at random a boy and a girl within each class and then we examine whether the remaining students are more likely to be absent on days where the randomly selected boy is absent, the randomly selected girl is absent, or both are absent. Clearly, we cannot establish any causal direction from the randomly selected students to the remaining students in the class, but this exercise can shed light on whether joint leisure is an important factor in peer effects of absenteeism.

In table 6, we show the conditional probabilities separately for girls and boys. Only classes where both the randomly selected girl and boy have at least one day of absence have been selected. This reduces the sample of stayers used in the regressions by about 50 percent. In the first row of Table 6, we see that the average absenteeism for the students in the selected sample is 4.5 percent for girls and 4.9 percent for boys. If we compare this to a day where both the selected girl and boy are present, we see that the probability of being absent decreases to 3.9 and 4.3 percent, respectively. In contrast to this, when both the selected girl and boy are absent, the conditional probabilities of being absent increase to 16.5 percent for girls and 17.4 percent for boys. Obviously, other reasons than peer effects may explain this pattern, e.g. weather conditions, having at temporary substitute teacher, etc. However, if girls prefer to be absent together, then the probability of being absent will be higher when the selected girl is absent and the selected boy is present compared to the opposite case where the selected girl is present and the selected boy is absent. We find that this is the case, as the difference in conditional probabilities for girls is 1.5 percent whereas it is 1 percent for boys. Hence, similar to the regression results, we find a stronger asymmetry for girls than boys.[18] In sum, we interpret these differences

[18] Notice that for boys, the parameter estimate to the absenteeism of boys is larger than for girls although clearly not significantly different.

	Girls		Boys	
Avg. absenteeism of new students	0.014		0.032***	
	(1.48)		(3.10)	
Avg. absenteeism of new boys		−0.016		0.036*
		(−1.13)		(1.95)
Avg. absenteeism of new girls		0.044***		0.027**
		(3.13)		(2.17)
New students in class (0/1)	−0.000		−0.000	
	(−0.35)		(−0.36)	
New boys in class (0/1)		0.002		0.000
		(1.42)		(0.08)
New girls in class (0/1)		−0.002*		−0.000
		(−1.79)		(−0.53)
Class size	−0.002	−0.002	−0.000	−0.000
	(−1.39)	(−1.41)	(−0.21)	(−0.20)
Class size squared	0.000*	0.000*	0.000	0.000
	(1.67)	(1.69)	(0.51)	(0.50)
Class grade * year interactions	Yes	Yes	Yes	Yes
R^2	0.045	0.046	0.051	0.051
No. of observations	44,401	44,401	44,787	44,787

t-statistics clustered at class-year level are in parentheses. *: 0.10, **: 0.05, ***: 0.01.

Table 5: Gender effects in peer effects

Probability of being absent on a given day	Girls	Boys
Unconditional probability of being absent (selected sample)	0.045	0.049
Conditional on a randomly selected (among classmates)		
Girl and boy are absent*	0.165	0.174
Girl and boy are present	0.039	0.043
Girl is absent and boy is present	0.102	0.099
Girl is present and boy is absent	0.087	0.109
No. of observations	21,423	21,032

* Estimated using only observations where both the random selected boy and girl are absent (9,376 and 8,833).

Table 6: Peer effects and class enviroments

in conditional probabilities as support for the hypothesis of joint leisure.

Next, we examine whether the peer effects depend on the class environment. We only have a limited number of characteristics of the class environment such as grade and class size. In Table 7, we consider whether the peer effects are heterogeneous across classroom environments. First, we notice that peer effects are stronger in classes with class sizes below the median (22 students). This is in accordance with the Manski model, suggesting that the effect of a new student decreases with class size, because the influence of a new student is $1/n$ in the mean of classmates.

Finally, we split the sample according to the grades and we find larger peer effects for grades 1-5 compared to grades 6-9. The data do not allow us to examine why students in lower grades are more sensitive to the peers in the class, but as we noted in the section on the Danish School system, some subjects are taught outside the class from grade 7: see the section on the Danish school system. We also hypothesize that at least part of the larger effect in the low grades is due to the organization of leisure activities. For the vast majority of students, the set of peers potentially affecting the student is increasing in the class grade as leisure activities such as sports are organized outside schools. Furthermore, after-school centres at the schools are usually only available for children in the low grades. However, we cannot rule out that younger students are just more sensitive to peer effects.

To further investigate for heterogeneous peer effects we perform a quantile regression. We use Canay's

	Class size		Grade	
	Below median	Above median	1-5	6-9
Avg. absenteeism of new students	0.040^{***}	-0.001	0.038^{***}	0.013
	(3.20)	(-0.11)	(2.89)	(1.01)
New students in class (0/1)	-0.001	-0.001	-0.001	0.001
	(-0.80)	(-0.51)	(-1.02)	(0.54)
Class size	-0.005^{**}	0.011	0.002	-0.006^{***}
	(-2.24)	(1.07)	(1.05)	(-3.12)
Class size squared	0.000^{**}	-0.000	-0.000	0.000^{***}
	(2.36)	(-1.02)	(-0.90)	(3.68)
Class grade * year interactions	Yes	Yes	Yes	Yes
R^2	0.050	0.042	0.008	0.077
No. of observations	52,223	36,965	58,681	30,507

t-statistics clustered at the class-year level are in parentheses. *: 0.10, **: 0.05, ***: 0.01

Table 7: Peer effects and class enviroment

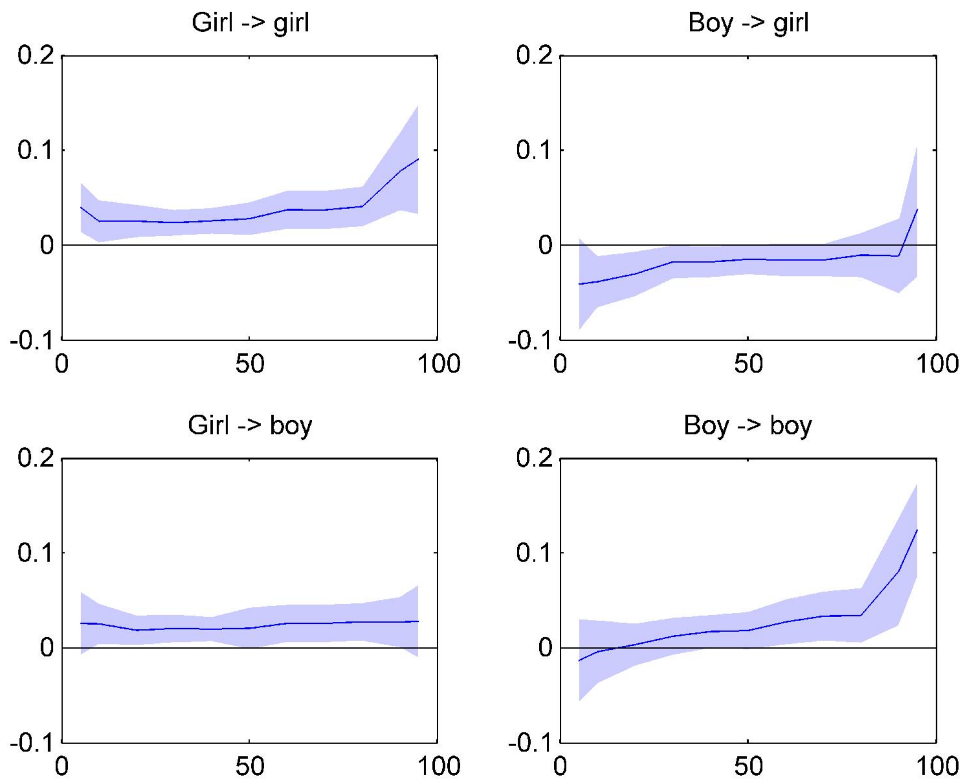

Figure 3: Quantile regressions

(2011) two-step procedure for panel data quantile regression. In the first step, student-class fixed effects are estimated from within estimation at the mean. In the second step, we estimate a pooled quantile regression with the dependent variable net of these estimated fixed effects. The results of the quantile regression are shown in figure 3.

The graphs show two distinct patterns. First, we find again the gender asymmetry pattern as documented in Tables 5 and 6. Students seem to be more sensitive to classmates of the same gender, although the absence of new girls in a class is borderline significant throughout the conditional distribution for boys. Second, we find that the peer effects tend to be increasing over the distribution within-gender.

In sum, we have provided evidence of significant peer effects that seem to be heterogeneous in several dimensions. We will try to summarize the evidence with respect to our potential mechanisms outlined in section two; these are; value of joint leisure, identity, disruptive students lowering the value of being in the class, and sanctions being harder to enforce when more students are absent. Gender specific peer effects lend support to the hypothesis that peer effects work through value of joint leisure as well as identity. On the other hand, there is no reason why there should be gender specific effects from disruptive students or from sanctions being more difficult to enforce with increased absenteeism.

Our analysis of the timing of absenteeism lends support to the hypothesis of value of joint leisure and may contradict the hypothesis of disruptive students as the presence of disruptive students may lower the value of being present in the class. Our finding that peer effects of absenteeism are larger in classes with fewer students also contradicts the hypothesis of disruptive students. It also contradicts the hypothesis that peer effects make sanctions harder to enforce. Overall we are not able to effectively rule out any of the mentioned mechanisms but we do find most support for the hypotheses that peer effects of absenteeism work through the value of joint leisure.

8 Concluding remarks

In this paper, we examine peer effects in absenteeism in compulsory schools. There exists a large literature on peer effects in cognitive outcomes, but our study is the first to document peer effects in absenteeism in compulsory schools. Research has shown that absenteeism lowers students' educational achievements. Therefore, investigating peer effects in absenteeism also sheds light on the channels through which peer effects in, e.g., test scores work in compulsory schools.

We derive an empirical strategy, which exploits that, with our data set, we can observe students' absenteeism when they move in between schools and classes in the metropolitan area of Copenhagen. The empirical strategy is to estimate the effect of incoming students' lagged absenteeism on the incumbent students' absenteeism. This strategy is robust to sorting based on student (or parent) type as we control for student-class fixed effects. Further, we show, using robustness checks and placebo tests, that challenges to our identification, such as non-random allocation of teachers to classes or differential heterogeneous absenteeism trends in classes, are not a concern.

The peer effects that we estimate are of similar size to those found in the literature on peer effects using other educational outcomes. We find heterogeneous effects and especially show that the peer effects work, in part, through same gender peers. The latter result is clearest for girls, who do not seem to respond to boys' absenteeism. We investigate this further using absenteeism measured on a daily level. Here, both genders are more likely to be absent when a student of same gender in the class is absent. We take this as evidence for value of joint leisure to be at least part of the mechanism that generates peer effects in absenteeism.

References

[1] Akerlof and Kranton (2002) Identity and Schooling: Some Lessons for the Economics of Education. *Journal of Economic Literature* vol 40 pp. 1167-1201.

[2] Angrist, Joshua (2014) The Perils of Peer Effects. *Labour Economics*. In Press.

[3] Bertrand, Marianne; Pan, Jessica (2013) The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior, *American Economic Journal: Applied Economics*, Volume 5(1), 32-64.

[4] Büttner, Bettina and Stephan L. Thomsen (2010) Are we Wasting our Time at School? Causal Evidence of the Impact of Shortening Secondary School Duration. Unpublished working paper.

[5] Burke, Mary A. and Tim R. Sass (2013) Classroom Peer Effects and Student Achievement. *Journal of Labor Economics*, 31,(1), pp. 51-82

[6] Canay, Ivan A. (2011) A Simple Approach to Quantile Regression for Panel Data. *The Econometrics Journal*, 14, pp. 368-386.

[7] Carrell, S.E. and Hoekstra, L.H., (2010) Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids. *American Economic Journal: Applied Economics*, Vol. 2, No. 1, pp. 211-228.

[8] Carlsson, Magnus, Gordon B. Dahl and Dan-Olof Rooth (2012) The Effect of Schooling on Cognitive Skills, NBER Working Paper 18484.

[9] Card, David and Laura Giuliano (2013) Peer Effects and Multiple Equilibria in the Risky Behavior of Friends. *The Review of Economics and Statistics*, Volume 95(4), 1130-49.

[10] Cicala, Steven, Roland G. Fryer and Jörg L. Spenkuch (2011) A Roy Model of Social Interactions. NBER Working paper, no. 16880.

[11] Dobkin, C., Gil, R. and Marion, J. (2010) Skipping Class in College and Exam Performance: Evidence from a Regression Discontinuity Classroom Experiment', *Economics of Education Review*, Vol. 29, No. 4, pp. 566–575.

[12] Durden, G. and Ellis, L. (1995) The Effects of Attendance on Student Learning in Principles of Economics, *American Economic Review*, Vol. 85, No. 2, pp. 343–346.

[13] Epple, Dennis and Richard E. Romano (2011) Peer Effects in Education: A Survey of the Theory and Evidence. In Jess Benhabib, Alberto Bisin and Matthew O. Jackson (eds.), *Handbook of Social Economics*, Volume 1B, Elsevier.

[14] Fitzpatrick, Maria D., David Grissmer and Sarah Hastedt. (2011) What a Difference a Day Makes: Estimating Daily Learning Gains during Kindergarten and First Grade Using a Natural Experiment. *Economics of Education Review* 30: 269–279.

[15] Fruehwirth, Jane Cooley (2012) Identifying Peer Achievement Spillovers: Implications for Desegregation and the Achievement Gap.*Quantitative Economics*, 4(1), pp. 85-124.

[16] Gibbons, Stephen and Telhaj, Shqiponja (2008) Peers and Achievement in England's Secondary Schools. Discussion Paper, No. 1, Spatial Economics Research Centre, London School of Economics.

[17] Goldsmith-Pinkham, Paul and Guido W. Imbens (2013) Social Networks and the Identification of Peer Effects. *Journal of Business & Economic Statistics.* 31(3): 253-64.

[18] Gottfried, Michael A. (2011) The Detrimental Effects of Missing School: Evidence from Urban Siblings *American Journal of Education*, Vol. 117, No. 2 (February 2011), pp. 147-182.

[19] Harris, Kamala, D. (2013) In School + On Track; Attorney General's 2013 Report on California's Elementary School Truancy & Absenteeism Crisis, Office of the Attorney General, California Department of Justice.

[20] Hoxby, Caroline (2000) *Peer Effects in the Classroom: Learning from Gender and Race Variation*, NBER Working Papers 7867, National Bureau of Economic Research, Inc.

[21] Jackson, David, Gary Marx and Alex Richards (2012). An Empty-desk Epidemic, Chicago Tribune, November 11.

[22] Kristoffersen, J.H., M.V. Krægpøth, H.S. Nielsen, and M. Simonsen (2014), Disruptive School Peers and Student Outcomes, working paper. School of Economics and Management, University of Aarhus

[23] Lavy, Victor, Olmo Silva and Weinhardt (2012) The Good, the Bad and the Average: Evidence on Ability Peer Effects in Schools. *Journal of Labor Economics*, 30(2): 367-414.

[24] Manski, Charles F. (1993) Identification of Endogenous Social Effects: The Reflection Problem, *Review of Economic Studies*, Vol. 60, No. 3, pp. 531-542.

[25] Pischke, Jörn-Steffen (2007) The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years, *Economic Journal*, 117(523): 1216-1242.

[26] Romer, D. (1993). Do Students Go to Class? Should They?, *Journal of Economic Perspectives*, Vol. 7, pp. 167–174.

[27] Sacerdote, Bruce (2011) Peer Effects In Education: How Might They Work, How Big are They and how Much Do We Know Thus Far? *Handbook of Economics of Education*. (Amsterdam: North Holland).

[28] Stanca, Luca. (2006) The Effects of Attendance on Academic Performance: Panel Data Evidence for Introductory Microeconomics. *Journal of Economic Education*. 27(3): 251-266.

[29] Stockard, J., and Mayberry, M. (1992) *Effective Educational Environments*. Newbury Park, CA: Corwin Press, Inc.

[30] Webbink, Dinand and Sander Gerritsen (2013) How Much do Children Learn in School? International Evidence from School Entry Rules. CPB Discussion Paper, 255.

[31] Wills, Gabrielle. 2014. The Effects of a Teacher Strike Activity on Student Learning in South Africa Primary Schools. Economic Research Southern Africa Working paper.

A Appendix

A.1 The outcome of the stayers

In this appendix, we derive the average class outcome for stayers. We use the model in equation (1). We start by taking the class average of equation (1). To derive the average class outcome in period 1 we use that

$\bar{Y}_{t(-i)} = \frac{1}{n-1}(n\bar{Y}_t - Y_{it})$,

$$\begin{aligned}
\bar{Y}_1 &= \beta_0 + \beta_1 \bar{Y}_{1(-i)} + \bar{\mu} + \bar{\varepsilon}_1 \Leftrightarrow \\
\bar{Y}_1 &= \beta_0 + \beta_1 \frac{1}{n-1}(n\bar{Y}_1 - \bar{Y}_1) + \bar{\mu} + \bar{\varepsilon}_1 \Leftrightarrow \\
\bar{Y}_1 &= \frac{\beta_0}{(1-\beta_1)} + \frac{\bar{\mu}}{(1-\beta_1)} + \frac{\bar{\varepsilon}_1}{(1-\beta_1)}
\end{aligned}$$

In period 2, a new student arrives and the class contains $n+1$ students. The new class average can be calculated in a similar way:

$$\frac{1}{n+1}\left(n\bar{Y}_2 + Z_2\right) = \frac{\beta_0}{(1-\beta_1)} + \frac{n\bar{\mu} + \mu_{n+1}}{(n+1)(1-\beta_1)} + \frac{n\bar{\varepsilon}_2 + \varepsilon_{n+1,2}}{(n+1)(1-\beta_1)}.$$

Note that the overall class average in period 2 is given by $\frac{1}{n+1}\left(n\bar{Y}_2 + Z_2\right)$. Using the Manski model, we also have an expression for the outcome of the new student

$$Z_2 = \beta_0 + \beta_1 \bar{Y}_2 + \mu_{n+1} + \varepsilon_{n+1,2}.$$

If we combine the two previous equations, we get

$$\frac{1}{n+1}\left(n\bar{Y}_2 + \beta_0 + \beta_1\bar{Y}_2 + \mu_{n+1} + \varepsilon_{n+1,2}\right) = \frac{\beta_0}{(1-\beta_1)} + \frac{n\bar{\mu} + \mu_{n+1}}{(n+1)(1-\beta_1)} + \frac{n\bar{\varepsilon}_2 + \varepsilon_{n+1,2}}{(n+1)(1-\beta_1)} \Leftrightarrow$$

$$\begin{aligned}
n\bar{Y}_2 + \beta_1\bar{Y}_2 &= \frac{\beta_0(n+1)}{(1-\beta_1)} + \frac{n\bar{\mu} + \mu_{n+1}}{(1-\beta_1)} + \frac{n\bar{\varepsilon}_2 + \varepsilon_{n+1,2}}{(1-\beta_1)} - \beta_0 - \mu_{n+1} - \varepsilon_{n+1,2} \Leftrightarrow \\
(n+\beta_1)\bar{Y}_2 &= \frac{\beta_0(n+\beta_1)}{(1-\beta_1)} + \frac{n\bar{\mu}}{(1-\beta_1)} + \frac{\beta_1\mu_{n+1}}{(1-\beta_1)} + \frac{n\bar{\varepsilon}_2 + \beta_1\varepsilon_{n+1,2}}{(1-\beta_1)} \Leftrightarrow \\
(n+\beta_1)\bar{Y}_2 &= \frac{\beta_0(n+\beta_1)}{(1-\beta_1)} + \frac{\bar{\mu}(n+\beta_1)}{(1-\beta_1)} + \frac{\beta_1(\mu_{n+1} - \bar{\mu})}{(1-\beta_1)} + \frac{n\bar{\varepsilon}_2 + \beta_1\varepsilon_{n+1,2}}{(1-\beta_1)} \Leftrightarrow \\
\bar{Y}_2 &= \frac{\beta_0}{(1-\beta_1)} + \frac{\bar{\mu}}{(1-\beta_1)} + \frac{\beta_1(\mu_{n+1} - \bar{\mu})}{(1-\beta_1)(n+\beta_1)} + \frac{n\bar{\varepsilon}_2 + \beta_1\varepsilon_{n+1,2}}{(1-\beta_1)(n+\beta_1)},
\end{aligned}$$

A.2 The probability limit of the fixed effect and OLS estimators

We derive the convergence in probability of the fixed effect estimator when the number of classes tends to infinity. To explicitly note that the students belong to different classes we introduce an index c to denote the class. We assume that we observe C classes. We also assume that $Z_{n+1,c,1}$ is a proxy for the true individual effect for the new student in class c: $\mu_{n+1,c}$.We formalize this by assuming that $Z_{n+1,c,1}$ is a measure of $\mu_{n+1,c}$ with a measurement error

$$Z_{n+1,c,1} = \mu_{n+1,c} + m_c,$$

where m_c is a classical measurement error with $E(m_c) = 0$ and $V(m_c) = \sigma_m^2$ and m is independent of the error terms and individual effects. To simplify the notation we make the following assumptions: i) we only have two time periods $T = 2$, ii) the number of students in all C classes is n in period 1 and $n + 1$ in period 2 where one new student has arrived. Our empirical model for the absence of stayers is given by

$$\begin{aligned} Y_{ic1} &= \alpha_0 + \mu_i + \upsilon_{ic1} \qquad i = 1, ..., n, c = 1, ..., C \\ Y_{ic2} &= \alpha_0 + \alpha_1 Z_{n+1,c,1} + \mu_i + \upsilon_{ic2} \qquad i = 1, ..., n, c = 1, ..., C, \end{aligned}$$

where $E(\mu_i) = 0$.[19] We estimate our parameter of interest α_1 by a fixed effect estimation, which is effectively the same as an OLS estimation on first differences:[20]

$$Y_{ic2} - Y_{ic1} = \alpha_1 Z_{n+1,c,1} + \Delta\upsilon_{ic2}.$$

The estimator is given by

$$\hat{\alpha}_1 = \frac{\sum_{c=1}^{C} \sum_{i=1}^{n} (Y_{ic2} - Y_{ic1}) Z_{n+1,c1}}{\sum_{c=1}^{C} \sum_{i=1}^{n} Z_{n+1,c1}^2} = \frac{\sum_{c=1}^{C} (\bar{Y}_{c2} - \bar{Y}_{c1}) Z_{n+1,c1}}{\sum_{c=1}^{C} Z_{n+1,c1}^2},$$

where $\bar{Y}_{ct}$ denotes the class average of the stayers. By using the expression for $\bar{Y}_{2c} - \bar{Y}_{1c}$ derived in equation (4) and the expression for $Z_{n+1,c,1}$ we get

$$\hat{\alpha}_1 = \frac{\frac{1}{C} \sum_{c=1}^{C} \left[\frac{\beta_1}{(1-\beta_1)(n+\beta_1)} (\mu_{n+1,c} - \bar{\mu}_c) + \frac{n\bar{\varepsilon}_{c2} + \beta_1 \varepsilon_{n+1,2}}{(1-\beta_1)(n+\beta_1)} - \frac{\bar{\varepsilon}_{c1}}{1-\beta_1} \right] \left[\mu_{n+1,c} + m_c \right]}{\frac{1}{C} \sum_{c=1}^{C} \left[\mu_{n+1,c} + m_c \right]^2}$$

We can then derive the convergence in probability of the numerator when C goes to infinity:

$$\begin{aligned} & \operatorname*{plim}_{C\to\infty} \frac{1}{C} \sum_{c=1}^{C} \left[\frac{\beta_1}{(1-\beta_1)(n+\beta_1)} (\mu_{n+1,c} - \bar{\mu}_c) + \frac{n\bar{\varepsilon}_{c2} + \beta_1 \varepsilon_{n+1,2}}{(1-\beta_1)(n+\beta_1)} - \frac{\bar{\varepsilon}_{c1}}{1-\beta_1} \right] \left[\mu_{n+1,c} + m_c \right] \\ = & \frac{\beta_1}{(1-\beta_1)(n+\beta_1)} \operatorname*{plim}_{C\to\infty} \frac{1}{C} \sum_{c=1}^{C} (\mu_{n+1,c} - \bar{\mu}_c) \left[\mu_{n+1,c} + m_c \right] \\ = & \frac{\beta_1}{(1-\beta_1)(n+\beta_1)} \operatorname*{plim}_{C\to\infty} \frac{1}{C} \sum_{c=1}^{C} (\mu_{n+1,c}^2 - \bar{\mu}_c \mu_{n+1,c}) \\ = & \frac{\beta_1}{(1-\beta_1)(n+\beta_1)} \left(E(\mu_{n+1,c}^2) - E(\bar{\mu}_c \mu_{n+1,c}) \right) \end{aligned}$$

The convergence in probability of the denominator is given by:

$$\operatorname*{plim}_{C\to\infty} \frac{1}{C} \sum_{c=1}^{C} \left[\mu_{n+1,c} + m_c \right]^2 = E(\mu_{n+1,c}^2) + \sigma_m^2.$$

[19] This is a normalization since we have an intercept in the model.

[20] In the case where $T = 2$, fixed effect estimation is the same as first difference.

The estimator $\hat{\alpha}_1$ will converge in probability to

$$\underset{C\to\infty}{\text{plim}}\hat{\alpha}_1 = \frac{\beta_1}{(1-\beta_1)(n+\beta_1)} \frac{\left(E(\mu_{n+1,c}^2) - E(\bar{\mu}_c \mu_{n+1,c})\right)}{E(\mu_{n+1,c}^2) + \sigma_m^2}.$$

We can then consider different cases. Note that if there are no peer effects, $\beta_1 = 0$, and the fixed effect estimator will converge in probability to zero. However, we can also have that the limit is zero even if peer effects exist. In the case where we have perfect sorting of new students into classes such that $\mu_{n+1,c} = \bar{\mu}_c$ we will have that $\text{plim}_{C\to\infty}\hat{\alpha}_1 = 0$. If we find that $\hat{\alpha}_1$ is positive, it indicates that $\beta_1 > 0$ and peer effects exist.

Furthermore if we have a positive correlation (positive sorting) or no correlation (no sorting) between $\bar{\mu}_c$ and $\mu_{n+1,c}$, we will have that

$$\underset{C\to\infty}{\text{plim}}\hat{\alpha}_1 < \frac{\beta_1}{(1-\beta_1)(n+\beta_1)}.$$

This implies that $\hat{\alpha}_1$ will be a downward biased estimate of $\frac{\beta_1}{(1-\beta_1)(n+\beta_1)}$. If the correlation is negative indicating negative sorting, then $\hat{\alpha}_1$ can be either a downward or an upward biased estimate of $\frac{\beta_1}{(1-\beta_1)(n+\beta_1)}$.

We can also derive the convergence in probability of the OLS estimator in a way similar to that of the FE estimator. To simplify the calculations we assume $\beta_0 = 0$ and therefore perform the estimation without a constant ($\alpha_0 = 0$). The OLS estimator can be derived by only using information from period 2. The OLS estimator of $\hat{\alpha}_{1,OLS}$ is given by

$$\hat{\alpha}_{1,OLS} = \frac{\sum_{c=1}^{C}\sum_{i=1}^{n} Y_{ic2} Z_{n+1,c1}}{\sum_{c=1}^{C}\sum_{i=1}^{n} Z_{n+1,c,i}{}^2} = \frac{\frac{1}{C}\sum_{c=1}^{C} \bar{Y}_{c2} Z_{n+1,c1}}{\frac{1}{C}\sum_{c=1}^{C} Z_{n+1,c,1}{}^2}$$

The convergence in probability is given by

$$\begin{aligned}
\underset{C\to\infty}{\text{plim}}\hat{\alpha}_{1,OLS} &= \underset{C\to\infty}{\text{plim}} \frac{\frac{1}{C}\sum_{c=1}^{C} \bar{Y}_{c2} Z_{n+1,c1}}{\frac{1}{C}\sum_{c=1}^{C} Z_{n+1,c,1}{}^2} \\
&= \frac{\beta_1}{(1-\beta_1)(n+\beta_1)} \frac{\left(E(\mu_{n+1,c}^2) - E(\bar{\mu}_c \mu_{n+1,c})\right)}{E(\mu_{n+1,c}^2) + \sigma_m^2} + \frac{E(\bar{\mu}_c \mu_{c,n+1})}{(1-\beta_1)(E(\mu_{n+1,c}^2) + \sigma_m^2)}
\end{aligned}$$

The asymptotic difference between the OLS and the FE estimator is determined by $\frac{E(\bar{\mu}_c \mu_{c,n+1})}{(1-\beta_1)(E(\mu_{n+1,c}^2)+\sigma_m^2)}$. This suggests that if the OLS estimator is larger than the FE estimator, the correlation between $\bar{\mu}_c$ and $\mu_{c,n+1}$is positive and it indicates positive sorting of new students into classes.