Economics of Education Review 54 (2016) 113–123

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

The role of testing noise in the estimation of achievement-based peer effects☆

Hongliang Zhang

Department of Economics, Hong Kong Baptist University, Hong Kong

ARTICLE INFO

Article history:
Received 26 October 2015
Revised 19 April 2016
Accepted 22 April 2016
Available online 30 May 2016

JEL Classification:
I20
I21
C13

Keywords:
Peer effects
Student achievement
Measurement error
Mean reversion

ABSTRACT

I demonstrate that in the value-added estimation of peer effects using lagged peer achievement, testing noise may generate another bias in addition to the well-known attenuation bias. Such a bias, which I refer to as the "reversion bias," may arise when some of a student's current peers happen to be his/her former peers whose performances in the baseline test were subject to the same common testing noise as the student's own. I propose a solution to overcome this problem by exploiting only the variation in the new peers' portion of the overall peer quality. Using real-world data, I illustrate the existence of this bias and demonstrate the proposed solution.

1. Introduction

Peer influences have been investigated for various outcomes in different settings,[1] among which the effects of classroom/school peers on a student's own achievement have received the greatest attention. While empirical studies abound, they provide mixed evidence of the existence, magnitude, and even sign of peer influences among students in schools (for surveys, see Epple & Romano, 2011; Sacerdote, 2011). The failure of previous studies to arrive at a consensus partly reflects the formidable identification challenges confronted in the estimation of peer effects (e.g., Angrist, 2014; Brock & Durlauf, 2001; Moffit, 2011). In particular, the reciprocal nature of peer interactions, known as the reflection problem (Manski, 1993), hinders differentiation between endogenous and contextual effects. To circumvent this problem, prior research has often resorted to estimating the reduced-form relationship between student achievement and predetermined measures of peer composition. One strand of research focuses on

☆ An earlier version of this paper was circulated under the title "Peer effects on student achievement: An instrumental variable approach using school transition data." I thank Joshua Angrist, David Autor, Abhijit Banerjee, Jiahua Che, Weili Ding, Esther Duflo, Amy Finkelstein, Chih-Sheng Hsieh, Tao Jin, Rongzhu Ke, Lars Lefgren, Steven Lehrer, Weifeng Li, Karen Polenske, Christopher Taber, Yu Zhu, and seminar participants at Fudan University, Hong Kong University, Lingnan University, MIT, Peking University, Zhejiang University, Royal Economic Society Conference, and CUHK Symposium on the Economics of Education for their valuable comments and suggestions. I am also grateful to the co-editor and two anonymous referees for helpful comments and feedback, and to the UK Department of Education for access to the National Pupil Database. I acknowledge financial support from the Hong Kong Research Grant Council General Research Fund (No. 458610). All remaining errors are my own.

E-mail address: zhang.hongliang@gmail.com

[1] Examples include college/military academy students' GPA (e.g., Foster, 2006; Han & Li, 2009; Lyle & David, 2007; Sacerdote, 2001; Stinebrickner & Stinebrickner, 2006), juveniles' behavior (e.g., Gaviria & Raphael, 2001; Glaeser, Sacerdote, & Scheinkman, 1996; Kremer & Levy, 2008), professionals' performances in tournaments (e.g., Guryan, Kroft, & Notowidigdo, 2009), worker productivity (Mas & Moretti, 2009), movie sales (Moretti & Enrico, 2011), etc.

http://dx.doi.org/10.1016/j.econedurev.2016.04.008
0272-7757/

examining the effects of contextual peer characteristics such as race, gender, immigration status, and family background (e.g., Angrist & Lang, 2004; Gould, Lavy, & Paserman, 2009; Hoxby, 2000; Lavy & Schlosser, 2011; McEwan, 2003). Another strand of research employs value-added models to estimate the effects of lagged peer achievement (e.g., Arcidiacono & Nicholson, 2005; Hanushek, Kain, Markman, & Rivkin, 2003; Lefgren, 2004; Vigdor & Nechyba, 2007). However, both strands largely ignore the measurement problem in the peer variables that may arise from missing data and/or measurement error. To the best of my knowledge, Ammermueller and Pischke (2009), Micklewright, Schnepf, and Silva (2012), and Sojourner (2013) are the only examples of studies that consider the measurement problem in the peer variables. While Ammermueller and Pischke (2009) and Micklewright et al. (2012) deal with both missing data and measurement error related to contextual peer characteristics in the first strand of research, in the second strand of research Sojourner (2013) considers only missing data related to lagged peer achievement.

In this paper, I extend the investigation of the role of measurement error to the estimation of achievement-based peer effects using lagged peer achievement. I point out that test scores of students from the same peer group (i.e., school or classroom) are subject to common testing noise arising from group-specific common influences having only transitory effects on test scores, e.g., a dog barking on the playground on the test day, a local flu pandemic, the coincidental overlap between the test and instruction contents, etc. While the existence of such common testing noise has been well documented in the school accountability literature (e.g., Betts & Danenberg, 2002; Kane & Staiger, 2002), in which conventional evaluation approaches are demonstrated to yield misleading assessments (Chay, McEwan, & Urquiola, 2005),[2] it has been underappreciated and largely ignored in the peer effects literature. To the best of my knowledge, this paper is the first to illustrate and highlight the relevance of common testing noise to the estimation of achievement-based peer effects. Specifically, I show that the conventional specifications using lagged peer achievement yield biased estimates of peer effects when (as is often the case) a student's current peer group consists of some former peers whose lagged performances were subject to the same transitory influences as experienced by the student him/herself. Given the existence of common testing noise in lagged performances, the continuing presence of a student's former peers in his/her current peer group leads to a spurious positive correlation between the student's own lagged achievement and mean peer lagged achievement, the workhorse variable in the linear-in-means model considered in this paper. That is, a higher mean peer lagged achievement implies more favorable testing noise in a student's own lagged achievement, which further indicates poorer achievement progress due to mean reversion, leading to a negative bias in the estimate of the coefficient on mean peer lagged achievement. Since this bias is the result of the mean-reversion property of test scores, I refer to it as the "reversion bias."[3]

Compared with the attenuation bias caused by missing data or classical measurement error considered in prior research, the reversion bias considered herein poses a more substantial challenge to the estimation of peer effects as it could even reverse the sign of the estimator. In addition to explicating the existence of the negative reversion bias, I propose a solution to overcome this problem by exploiting the variation in mean peer lagged achievement caused exclusively by new peers. Specifically, I partition mean peer lagged achievement into the old and the new peers' portions and use only the variation in the latter component (i.e., the product of the new peers' share and mean lagged achievement). I show that under some plausible mean independence conditions regarding the testing noise, mean ability of old peers, and unobserved determinants of learning, the estimated coefficient on the new peers' portion of mean peer lagged achievement is immune from the reversion bias.

To illustrate the existence of the reversion bias in the conventional estimators and demonstrate the proposed solution, I analyze achievement-based peer effects in England's secondary schools using the National Pupil Database (NPD) collected by the UK's Department of Education. The NPD contains students' test scores on both the Key Stage 2 (KS2) national exam taken at the end of primary school (sixth grade) and the Key Stage 3 (KS3) national exam taken in ninth grade at secondary school. As the same test is taken by all students of the same cohort, coincidental overlap between the test and instruction contents alone would lead to common testing noise in the scores of students from the same school. Using four cohorts of students in the NPD who finished ninth grade between 2005 and 2008, the standard school fixed-effect estimations show large, negative, and significant coefficients on mean peer lagged achievement, suggesting not only the existence of the reversion bias but also its dominance over the true peer effects (if the latter exist). The NPD data set also includes information on the primary school where students took their KS2 national exam, thus allowing a distinction to be made between old and new peers in secondary school. Performing my proposed estimation to overcome the reversion-bias problem yields modest, positive, and significant coefficients on the new peers' portion of mean lagged peer achievement, indicating that positive

[2] Chay et al. (2005) show that a difference-in-differences assessment of a school intervention program in Chile targeting low-performing schools overstates the program effect because of mean reversion in testing noise.

[3] Different from the reversion bias in the coefficient on mean peer lagged achievement considered here, Fruehwirth (2014) illustrates biases in the estimated coefficients on contextual peer characteristics conditional on lagged peer achievement. That is, when contextual peer characteristics and lagged peer achievement are both included in the estimation, the estimated coefficients on contextual peer characteristics are biased toward 0 or even take counterintuitive signs. The intuition is as follows: conditional on lagged peer achievement, more (less) favorable peer contextual characteristics partially capture a lower (higher) level of the unobservable peer quality that is not fully accounted for by lagged peer achievement, thus biasing the estimated coefficients on contextual peer characteristics toward 0 or even negative.

achievement-based peer effects do exist, although they are overshadowed by the severe negative reversion bias in the standard school fixed-effect estimates.

Although this paper is the first to demonstrate both econometrically and empirically the existence and severity of reversion bias, the peril of using mean peer lagged achievement in value-added models has been noted in some previous studies, in which the problem is bypassed with the use of alternative lagged achievement measures of peer quality that are not mechanically correlated with the student's own achievement progress. However, none of these studies has revealed and discussed formally the reversion-bias problem. To the best of my knowledge, Hanushek et al. (2003) may be the only authors who raise concern over the mechanical correlation due to common measurement error, albeit in a footnote only.[4] To circumvent this spurious correlation, these authors resort to using the twice-lagged mean peer achievement in estimating peer effects among third to sixth graders in Texas elementary schools. The same strategy is also employed by Vigdor and Nechyba (2007) in their analysis of peer effects in North Carolina public schools, though without any discussion of the reason for using the higher-order lagged mean peer achievement. Compared with their strategy, my proposed use of the new peers' portion of mean peer lagged achievement has two advantages. First, it needs test scores of only two periods instead of three periods. Second, it employs a more immediate and precise measure of contemporary peer quality and therefore is more likely to detect peer effects when they are present. In another paper using the same data source (i.e., the NPD) as the current paper (although covering different cohorts of students), Lavy, Silva, and Weinhardt (2012) also single out new peers from old peers and construct lagged achievement measures separately for the two groups. While separate measures of new and old peers' quality are always included simultaneously in their regressions, they only report the coefficients on new peers' quality and omit the coefficients on old peers' quality from all tables. Rather than the reversion-bias problem that I highlight in this paper, their justification for using only new peers' quality is to circumvent the reflection problem due to previous interactions among students belonging to the same former group. Moreover, their strategy of using the lagged achievement of new peers directly without accounting for the new peers' share results in an attenuation bias analogous to what Sojourner (2013) illustrates when applying the standard estimator to the subsample of observed-data peers in the presence of missing-data peers. In contrast, the solution proposed in this paper explicitly accounts for the new peers' share and is exempted from such an attenuation bias.

Despite the three aforementioned studies that circumvent the mechanical negative correlation between mean peer lagged achievement and the student's own achievement gain in estimating the value-added models of peer effects, the reversion-bias problem is never investigated formally in the literature, nor does it seem to be common knowledge in this field.[5] For example, in a study of peer effects among roommates at Williams College, Zimmerman (2003) finds no evidence that a student's college grade performance is influenced by his/her roommate's SAT scores from regressions of a student's college GPA on own SAT scores, roommate's SAT scores, and other individual control variables. However, some roommate pairs in his sample are mutually requested roommates, most likely among former high-school mates.[6] To the extent that the SAT scores of students from the same high school are subject to common transitory influences, the estimated roommate effects can be biased downward, which may explain why roommate's SAT scores have no effect.

The remainder of this paper is organized as follows. Section 2 derives the potential existence of the reversion bias in a value-added model with linear-in-means peer effects. Section 3 presents my proposed solution to the reversion-bias problem and discusses the conditions under which the proposed estimator can yield an unbiased estimate. Section 4 presents analysis of achievement-based peer effects in England's secondary schools to illustrate the existence of the reversion bias in the standard estimators and demonstrate the proposed solution. Section 5 concludes the paper.

2. Model setup and potential existence of reversion bias

2.1. Model setup

I begin with the simplest version of a value-added learning model with linear-in-means ability peer effects as follows:

$$\Delta a_{ics} = a^1_{ics} - a^0_{ics} = \lambda \bar{a}^0_{(-i)cs} + \mu_{ics}, \tag{1}$$

[4] Footnote 14 of Hanushek et al. (2003) says: "School specific non-random measurement error in the grade G-1 score may also be negatively correlated with grade G gains."

[5] Note that Hanushek et al. (2003), Vigdor and Nechyba (2007), and Lavy et al. (2012) all use students' scores from statewide or nationwide uniform end-of-grade tests. Due to the existence of an extra channel of coincidental overlap between the test and instruction contents, students' scores from uniform tests are to a greater extent subject to group-specific common transitory influences than scores from standardized individual tests (such as the SAT or ACT), although both are affected by other common classroom instructional practices having transitory effects on test scores, such as teaching to the test or teaching test-taking skills. As a result, when students' scores from uniform tests are used in estimating peer effects using the value-added specifications, the reversion bias is likely to be larger in magnitude and more noticeable, therefore alerting researchers to resort to alternative peer quality measures not mechanically correlated with the student's own achievement gain. In contrast, when students' scores from standardized individual tests are used, the reversion bias, may be smaller in magnitude (though still present) and thus more likely to be overlooked.

[6] Incoming students at Williams College were asked to fill out a housing preference form and indicate whether they would like to live with a particular roommate. In 2002, the only year for which the housing preference form was available, 5% of students requested for a particular roommate, of which 3.8% were mutual requests and were honored. As the housing preference forms were submitted before students met with other incoming freshmen, mutual requests should have come mostly from former schoolmates who had known each other before attending Williams College. The housing preference forms for the years (1990–2001) studied in Zimmerman (2003) were unfortunately destroyed, making it impossible to exclude the mutually requested roommate pairs from his sample.

where Δa_{ics} is the cognitive ability gain for student i of cohort c in school s, measured as the difference in cognitive ability between the exit (a^1_{ics}) and entry of school s (a^0_{ics}); $\bar{a}^0_{(-i)cs}$ is the mean cognitive ability of student i's peers predetermined at the entry of school s; and μ_{ics} is a stochastic term capturing the unobserved determinants of student i's learning progress. Two points are worth noting for the specification of the learning dynamics in Eq. (1). First, it assumes perfect persistence in learning, i.e., setting the parameter on a^0_{ics} to 1. I adopt this restriction for technical convenience as it simplifies the illustration of the reversion-bias problem; the main conclusions of the paper are not affected by this restriction and can be generalized to cases of imperfect persistence. Second, λ is a reduced-form parameter that captures the combination of contextual and endogenous effects. In this paper, I make no attempt to distinguish between these two types of peer effects and aim instead to separate their combined effects from correlated effects. To focus attention on the identification of λ, I shall assume for a moment that cognitive ability can be observed perfectly and defer discussion of its measurement problem to the next subsection.

Note that the stochastic term μ_{ics} in Eq. (1) contains unobserved school-cohort-level correlated effects on learning as students from the same school and cohort are subject to common influences not modeled directly here. These correlated effects can bias the estimate of λ if they are also correlated with $\bar{a}^0_{(-i)cs}$. Random assignment of students to groups, where a group refers to a cohort in a school in the framework considered here, can solve this problem because random assignment breaks the potential link between mean peer ability and group-level correlated effects. However, random assignment rarely exists outside experimental settings (e.g., Duflo, Dupas, & Kremer, 2011). In practice, parents select a school based on its quality and peer composition, and schools also have some discretion in choosing students. Hence, $\bar{a}^0_{(-i)cs}$ and μ_{ics} could be systematically correlated at the school level, causing the OLS estimator of λ to be biased. To extract causal peer influences, I adopt a commonly used strategy in the literature to employ school fixed effects and exploit only the cohort-to-cohort perturbation in peer composition within the same schools (see, e.g., Gould et al., 2009; Hanushek, Kain, & Rivkin, 2009; Hoxby, 2000).[7] The identification of the ability peer effects parameter λ in this school fixed-effect estimation requires that the expectation of the unobserved determinants be the same for all students in the same school regardless of their peers' mean cognitive ability, formally stated in the following mean independence condition (MIC):

$$E[\mu_{ics}|\bar{a}^0_{(-i)cs}, s] = E[\mu_{ics}|s] = \mu_s. \qquad \text{MIC-1}$$

[7] Prior literature also uses a variety of other empirical strategies to deal with the endogenous school selection and peer group formation, such as subgroup reassignment (e.g., Angrist & Lang, 2004; Hoxby & Weingarth, 2005), regression discontinuity (e.g., Abdulkadiroğlu, Angrist, & Pathak, 2014; Ding & Lehrer, 2007), and instrumental variables (Kang, 2007; Zabel, 2008).

Under MIC-1, the conditional expectation of Δa_{ics} given $\bar{a}^0_{(-i)cs}$ and s can be expressed as

$$E[\Delta a_{ics}|\bar{a}^0_{(-i)cs}, s] = \lambda\bar{a}^0_{(-i)cs} + E[\mu_{ics}|\bar{a}^0_{(-i)cs}, s] = \lambda\bar{a}^0_{(-i)cs} + \mu_s.$$

Therefore, λ is identified by a regression of Δa_{ics} on $\bar{a}^0_{(-i)cs}$ and school dummy ($\mathbf{1}_s$).

2.2. Potential existence of reversion bias

The analysis in Section 2.1 assumes perfect observation of cognitive ability, which in reality are latent values and cannot be observed directly. Instead, researchers observe only cognitive achievement, which measures cognitive ability with testing noise:

$$y^0_{ics} = a^0_{ics} + e^0_{ics}; \qquad (2a)$$

$$y^1_{ics} = a^1_{ics} + e^1_{ics}; \qquad (2b)$$

$$\bar{y}^0_{(-i)cs} = \bar{a}^0_{(-i)cs} + \bar{e}^0_{(-i)cs}. \qquad (2c)$$

In the proceeding equations, y^0_{ics} and y^1_{ics} denote the cognitive achievement of student i of cohort c at the entry and exit of school s, respectively; $\bar{y}^0_{(-i)cs}$ denotes the mean cognitive achievement of student i's peers at the entry of school s; and e^0_{ics}, e^1_{ics}, and $\bar{e}^0_{(-i)cs}$ denote the respective testing noise in y^0_{ics}, y^1_{ics}, and $\bar{y}^0_{(-i)cs}$. Here, I assume that each testing noise term (i.e., e^0_{ics}, e^1_{ics}, and $\bar{e}^0_{(-i)cs}$) is independent of all of the latent ability variables (i.e., a^0_{ics}, a^1_{ics}, and $\bar{a}^0_{(-i)cs}$) at both the student and school levels. Let e^*_{ics} denote any variable e^0_{ics}, e^1_{ics}, or $\bar{e}^0_{(-i)cs}$, and let a^*_{ics} denote any variable a^0_{ics}, a^1_{ics}, or $\bar{a}^0_{(-i)cs}$. That is, I assume both $e^*_{ics} \perp a^*_{ics}$ and $e^*_s \perp a^*_s$, where e^*_s and a^*_s denote the school's mean in e^*_{ics} and a^*_{ics}, respectively. Given these assumptions, each testing noise term and each latent ability variable are also uncorrelated conditional on s, i.e., $cov(e^*_{ics}, a^*_{ics}|s) = 0$.

Moreover, I assume that the expectations of the unobserved influences on both the learning component (μ_{ics}) and noise component (e^1_{ics}) of contemporary achievement are the same for all students in school s regardless of their peers' mean lagged achievement, i.e.,

$$E[\mu_{ics}|\bar{y}^0_{(-i)cs}, s] = E[\mu_{ics}|s] = \mu_s;$$
$$E[e^1_{ics}|\bar{y}^0_{(-i)cs}, s] = E[e^1_{ics}|s] = e^1_s. \qquad \text{MIC-2}$$

Given MIC-1 and the conditional uncorrelatedness between testing noise terms and latent ability variables, a sufficient condition for MIC-2 is that both μ_{ics} and e^1_{ics} are mean independent of $\bar{e}^0_{(-i)cs}$ conditional on s, i.e., $E[\mu_{ics}|\bar{e}^0_{(-i)cs}, s] = E[\mu_{ics}|s]$ and $E[e^1_{ics}|\bar{e}^0_{(-i)cs}, s] = E[e^1_{ics}|s]$. The former condition is the same in nature as MIC-1, and the latter is the same in nature as the assumption of conditional uncorrelatedness between testing noise terms and latent ability variables.[8] Re-expressing

[8] $E[e^1_{ics}|\bar{e}^0_{(-i)cs}, s] = E[e^1_{ics}|s]$ if e^1_{ics} and $\bar{e}^0_{(-i)cs}$ are independent at both the student and school levels, as assumed for other testing noise terms and latent ability variables.

Eqs. (2a)–(2c) in terms of a^0_{ics}, a^1_{ics}, and $\bar{a}^0_{(-i)cs}$, respectively, and substituting them into Eq. (1) yield

$$\Delta y_{ics} = y^1_{ics} - y^0_{ics} = \lambda \bar{y}^0_{(-i)cs} - \lambda \bar{e}^0_{(-i)cs} + \mu_{ics} + e^1_{ics} - e^0_{ics}. \tag{3}$$

Next, I consider a school fixed-effect regression of a student's achievement gain Δy_{ics} on mean peer lagged achievement $\bar{y}^0_{(-i)cs}$. Under MIC-2, the conditional expectation of Δy_{ics} given $(\bar{y}^0_{(-i)cs}, s)$ can be expressed as

$$\begin{aligned}
&E[\Delta y_{ics}|\bar{y}^0_{(-i)cs}, s]\\
&\quad= \lambda \bar{y}^0_{(-i)cs} - \lambda E[\bar{e}^0_{(-i)cs}|\bar{y}^0_{(-i)cs}, s] + E[\mu_{ics}|\bar{y}^0_{(-i)cs}, s]\\
&\qquad+ E[e^1_{ics}|\bar{y}^0_{(-i)cs}, s] - E[e^0_{ics}|\bar{y}^0_{(-i)cs}, s]\\
&\quad= \lambda \bar{y}^0_{(-i)cs} - \lambda \left[\frac{cov(\bar{e}^0_{(-i)cs}, \bar{y}^0_{(-i)cs}|s)}{var(\bar{y}^0_{(-i)cs}|s)} \bar{y}^0_{(-i)cs} + \alpha_s\right] + \mu_s + e^1_s\\
&\qquad- \left[\frac{cov(e^0_{ics}, \bar{y}^0_{(-i)cs}|s)}{var(\bar{y}^0_{(-i)cs}|s)} \bar{y}^0_{(-i)cs} + \beta_s\right]\\
&\quad= \left(\lambda - \frac{\sigma^2_{\bar{e}}}{\sigma^2_{\bar{a}} + \sigma^2_{\bar{e}}}\lambda - \frac{\tau}{\sigma^2_{\bar{a}} + \sigma^2_{\bar{e}}}\right)\bar{y}^0_{(-i)cs}\\
&\qquad+ (\mu_s + e^1_s - \lambda\alpha_s - \beta_s),
\end{aligned} \tag{4}$$

where $\sigma^2_{\bar{e}} = var(\bar{e}^0_{(-i)cs}|s)$, $\sigma^2_{\bar{a}} = var(\bar{a}^0_{(-i)cs}|s)$, and $\tau = cov(e^0_{ics}, \bar{e}^0_{(-i)cs}|s)$. In the second equality, the conditional expectations of $\bar{e}^0_{(-i)cs}$ and e^0_{ics} are expressed as the expected values of their respective regressions on $\bar{y}^0_{(-i)cs}$ and school dummy ($\mathbf{1}_s$).

Given Eq. (4), a school fixed-effect regression of Δy_{ics} on $\bar{y}^0_{(-i)cs}$ yields an estimator that contains two bias components compared with the ability peer effects parameter λ. The first bias component $(-\frac{\sigma^2_{\bar{e}}}{\sigma^2_{\bar{a}}+\sigma^2_{\bar{e}}}\lambda)$ is the classical "attenuation bias" as the result of measurement error $\bar{e}^0_{(-i)cs}$ contained in mean peer lagged achievement $\bar{y}^0_{(-i)cs}$. The second bias component $(-\frac{\tau}{\sigma^2_{\bar{a}}+\sigma^2_{\bar{e}}})$, which I refer to as the "reversion bias," arises if the testing noise in a student's own lagged achievement e^0_{ics} is correlated with mean peer lagged achievement $\bar{y}^0_{(-i)cs}$ through the latter's error component $\bar{e}^0_{(-i)cs}$. Note that such a correlation does not exist if e^0_{ics} is independently distributed across individuals, as considered in Ammermueller and Pischke (2009) and Micklewright et al. (2012) regarding measurement error in contextual individual characteristics. However, for students originating from the same former school, measurement error in the baseline test scores is *not* independently distributed as their performance was subject to school-cohort-level common transitory influences that lead to a positive intraschool correlation in testing noise. The presence of a student's former peers in his/her current peer group (if any) further carries over this correlation to between e^0_{ics} and $\bar{e}^0_{(-i)cs}$, resulting in a negative reversion bias, which could dominate the true peer effects and reverse the sign of the estimator to negative (i.e., when $\tau > \sigma^2_{\bar{a}}\lambda$).

It is worth noting the special case when $\tau = 0$, i.e., there are no former peers in a student's current peer group or there is no intraschool correlation among former peers in the baseline testing noise, in which the reversion-bias term vanishes and the school fixed-effect estimator in Eq. (4) becomes $\frac{\sigma^2_{\bar{a}}}{\sigma^2_{\bar{a}}+\sigma^2_{\bar{e}}}\lambda$. Although still subject to an attenuation bias compared with the ability peer effects parameter λ, this yields an unbiased estimate of peer effects operating through mean peer lagged achievement. In the next section, I consider the general case when $\tau \neq 0$ and develop an alternative identification strategy to overcome the reversion-bias problem in the generalized setting.

3. Identification

As the reversion bias is caused by the presence of former peers in a student's current peer group, an idea for bypassing this problem is to exploit the variation in the lagged achievement of new peers only, an approach used in Lavy et al. (2012). However, substituting the mean lagged achievement of all peers directly with that of new peers is analogous to applying the standard estimator to the subsample of observed-data peers in the presence of missing-data peers, which Sojourner (2013) illustrates to yield biased estimates. To circumvent the reversion bias caused by former peers, and yet without neglecting their presence, I adopt an alternative estimation strategy using the portion of mean peer lagged achievement caused exclusively by new peers. This strategy shares the same spirit as the so-called "*p-weight* estimation" Sojourner (2013) proposes to identify the linear-in-means peer effects with some missing-data peers, which exploits only the variation in the observed portion of mean peer lagged achievement (i.e., the product of the proportion and mean lagged achievement of observed-data peers). He demonstrates that under two mean independence conditions – i.e., the expectations of both the mean lagged achievement of missing-data peers and the unobserved influences are the same for all individuals in the same school, regardless of the proportion and mean lagged achievement of observed-data peers – this *p-weight* estimator can yield an unbiased estimate of peer effects despite the presence of missing-data peers.

Given the additive separability and linearity of the achievement production function, mean peer lagged achievement can be decomposed into the new and old peers' portions such that $\bar{y}^0_{(-i)cs} = p_{ics}\bar{y}^0_{(-i)cs,new} + (1 - p_{ics})\bar{y}^0_{(-i)cs,old}$, where p_{ics} denotes the proportion of new peers for student i and $\bar{y}^0_{(-i)cs,new}$ and $\bar{y}^0_{(-i)cs,old}$ denote the mean lagged achievement of student i's new and old peers, respectively. To circumvent the reversion-bias problem, the application of the *p-weight* estimation in the context considered in this paper exploits the variation in the former component $p_{ics}\bar{y}^0_{(-i)cs,new}$ only. Specifically, Eq. (3) can be written as follows:

$$\begin{aligned}
\Delta y_{ics} = y^1_{ics} - y^0_{ics} &= \lambda p_{ics}(\bar{y}^0_{(-i)cs,new} - \bar{e}^0_{(-i)cs,new})\\
&\quad+ \lambda(1 - p_{ics})\bar{a}^0_{(-i)cs,old} + \mu_{ics} + e^1_{ics} - e^0_{ics},
\end{aligned} \tag{5}$$

where $\overline{a}^0_{(-i)cs,old}$ denotes the (unobserved) mean cognitive ability of student i's old peers predetermined at the entry of school s.[9]

I next consider the identification conditions for this *p-weight* estimation:

$$\begin{aligned}
&(i)\ E[\overline{a}^0_{(-i)cs,old}|\overline{y}^0_{(-i)cs,new}, p_{ics}, s] = E[\overline{a}^0_{(-i)cs,old}|p_{ics}, s];\\
&(ii)\ E[\mu_{ics}|\overline{y}^0_{(-i)cs,new}, p_{ics}, s] = E[\mu_{ics}|p_{ics}, s];\\
&(iii)\ E[e^1_{ics}|\overline{y}^0_{(-i)cs,new}, p_{ics}, s] = E[e^1_{ics}|s] = e^1_s;\\
&(iv)\ E[e^0_{ics}|\overline{y}^0_{(-i)cs,new}, p_{ics}, s] = E[e^0_{ics}|s] = e^0_s;\\
&(v)\ E[\overline{e}^0_{(-i)cs,new}|\overline{y}^0_{(-i)cs,new}, p_{ics}, s]\\
&\quad = E[\overline{e}^0_{(-i)cs,new}|\overline{y}^0_{(-i)cs,new}, s]. \qquad \text{MIC-3}
\end{aligned}$$

Compared with the two aforementioned mean independence conditions Sojourner (2013) imposes for the mean lagged achievement of missing-data peers and the unobserved influences in dealing with missing data, assumptions (i) and (ii) are less restrictive, as they only require $\overline{a}^0_{(-i)cs,old}$ and μ_{ics} to be mean independent of $\overline{y}^0_{(-i)cs,new}$, but not necessarily of p_{ics}, conditional on s. Allowing μ_{ics} to be potentially correlated with p_{ics} is particularly important as some research has suggested that peer turnover can have an independent effect on students' achievement progress conditional on peer quality, although the sign of the effect of peer turnover remains ambiguous. Gibbons and Telhaj (2011) show that peer mobility has a detrimental effect on incumbent students in England's primary schools. However, using the Project STAR data, Luppino (2015) finds that high levels of classmate turnover generate positive externalities to students in center city schools but negative externalities to students in non-center city schools.[10] Assumptions (i) and (ii) adopted here are compatible with the heterogeneous findings of Luppino (2015) because not only do they allow the expectations of the mean cognitive ability of former peers ($\overline{a}^0_{(-i)cs,old}$) and the unobserved influences (μ_{ics}) to be correlated with p_{ics}, but the correlations may also vary across schools. Specifically, the conditional expectations of $\overline{a}^0_{(-i)cs,old}$ and μ_{ics} given p_{ics} and s can be expressed as their school-specific linear projections on p_{ics} as follows:

$$E[\overline{a}^0_{(-i)cs,old}|p_{ics}, s] = \pi_s + \gamma_s p_{ics}; \qquad (6a)$$

$$E[\mu_{ics}|p_{ics}, s] = \eta_s + \delta_s p_{ics}. \qquad (6b)$$

Assumptions (iii)–(v) in MIC-3 are additional assumptions related to the testing noise terms needed to address the challenges posed by measurement error in student achievement. Assumptions (iii) and (iv) require that, conditional on the school attended (s), testing noise in both the contemporary and lagged individual achievement be mean independent of the new peers' share (p_{ics}) and mean lagged achievement ($\overline{y}^0_{(-i)cs,new}$). In contrast, assumption (v) requires $\overline{e}^0_{(-i)cs,new}$ to be mean independent of p_{ics} only, as $\overline{e}^0_{(-i)cs,new}$ and $\overline{y}^0_{(-i)cs,new}$ are correlated by construction (with the former contained in the latter). Given the random nature of testing noise, all of these assumptions should hold in practice.

With MIC-3 and Eqs. (6a) and (6b), the conditional expectation of Δy_{ics} given ($\overline{y}^0_{(-i)cs,new}, p_{ics}, s$) can be expressed as

$$\begin{aligned}
&E[\Delta y_{ics}|\overline{y}^0_{(-i)cs,new}, p_{ics}, s]\\
&= \lambda p_{ics}\overline{y}^0_{(-i)cs,new} - \lambda p_{ics}E[\overline{e}^0_{(-i)cs,new}|\overline{y}^0_{(-i)cs,new}, p_{ics}, s]\\
&\quad +\lambda(1-p_{ics})E[\overline{a}^0_{(-i)cs,old}|\overline{y}^0_{(-i)cs,new}, p_{ics}, s]\\
&\quad +E[\mu_{ics}|\overline{y}^0_{(-i)cs,new}, p_{ics}, s]\\
&\quad +E[e^1_{ics}|\overline{y}^0_{(-i)cs,new}, p_{ics}, s] - E[e^0_{ics}|\overline{y}^0_{(-i)cs,new}, p_{ics}, s]\\
&= \lambda p_{ics}\overline{y}^0_{(-i)cs,new} - \lambda p_{ics}E[\overline{e}^0_{(-i)cs,new}|\overline{y}^0_{(-i)cs,new}, s]\\
&\quad +\lambda(1-p_{ics})(\pi_s+\gamma_s p_{ics}) + (\eta_s+\delta_s p_{ics}) + e^1_s - e^0_s\\
&= \lambda p_{ics}\overline{y}^0_{(-i)cs,new} - \lambda p_{ics}\\
&\quad \times\left[\frac{cov(\overline{e}^0_{(-i)cs,new}, \overline{y}^0_{(-i)cs,new}|s)}{var(\overline{y}^0_{(-i)cs,new}|s)}\overline{y}^0_{(-i)cs,new} + \kappa_s\right]\\
&\quad +(\lambda\pi_s + \eta_s + e^1_s - e^0_s)\\
&\quad +(\lambda\gamma_s + \delta_s - \lambda\pi_s)p_{ics} - \lambda\gamma_s p^2_{ics}\\
&= \frac{\sigma^2_{\overline{a}new}}{\sigma^2_{\overline{a},new}+\sigma^2_{\overline{e},new}}\lambda(p_{ics}\overline{y}^0_{(-i)cs,new}) + (\lambda\pi_s+\eta_s+e^1_s-e^0_s)\\
&\quad +(\lambda\gamma_s+\delta_s-\lambda\pi_s-\lambda\kappa_s)p_{ics} - \lambda\gamma_s p^2_{ics}, \qquad (7)
\end{aligned}$$

where $\sigma^2_{\overline{a}new} = var(\overline{a}^0_{(-i)cs,new}|s)$ and $\sigma^2_{\overline{e}new} = var(\overline{e}^0_{(-i)cs,new}|s)$. Similar to Eq. (4), $E[\overline{e}^0_{(-i)cs,new}|\overline{y}^0_{(-i)cs,new}, s]$ is expressed as the expected value of a regression of $\overline{e}^0_{(-i)cs,new}$ on $\overline{y}^0_{(-i)cs,new}$ and school dummy ($\mathbf{1}_s$).[11] The parameter $\frac{\sigma^2_{\overline{a}new}}{\sigma^2_{\overline{a},new}+\sigma^2_{\overline{e},new}}\lambda$ corresponds to the causal effect of a one-unit increase in mean peer lagged achievement attributable exclusively to the new peers' portion ($p_{ics}\overline{y}^0_{(-i)cs,new}$). Given Eq. (7), this *p-weight* estimation can be implemented by regressing Δy_{ics} on $p_{ics}\overline{y}^0_{(-i)cs,new}$, the set of school dummies ($\mathbf{1}_s$), their interactions with the new peers' share ($p_{ics}\cdot\mathbf{1}_s$), and their interactions with the new peers' share squared ($p^2_{ics}\cdot\mathbf{1}_s$). Compared with Sojourner (2013), I further include the interaction terms between school dummies and the new peers' share squared ($p^2_{ics}\cdot\mathbf{1}_s$) in the implementation of the *p-weight* estimation here because of the combined influence of two factors. First, the expectation of the mean cognitive ability of former peers ($\overline{a}^0_{(-i)cs,old}$) can be correlated with p_{ics} in any school-specific pattern. Second, the influence of

[9] The use of $\overline{a}^0_{(-i)cs,old}$ instead of $(\overline{y}^0_{(-i)cs,old} - \overline{e}^0_{(-i)cs,old})$ in Eq. (5) simplifies the assumptions laid out in MIC-3, in which only a mean independence condition for $\overline{a}^0_{(-i)cs,old}$ is imposed.

[10] Luppino (2015) explains the counterintuitive findings with a classroom learning model in which familiar classmates can be a resource in school environments with low-levels of behavioral problems (i.e., non-center city schools) but a distraction in school environments with high-levels of behavioral problems (i.e., center city schools).

[11] Here, the independence between the testing noise term and latent ability variable is assumed to also carry over to $\overline{e}^0_{(-i)cs,new}$ and $\overline{a}^0_{(-i)cs,new}$ such that $cov(\overline{e}^0_{(-i)cs,new}, \overline{y}^0_{(-i)cs,new}|s) = var(\overline{e}^0_{(-i)cs,new}|s) = \sigma^2_{\overline{e}_{new}}$ and $var(\overline{y}^0_{(-i)cs,new}|s) = var(\overline{a}^0_{(-i)cs,new}|s) + var(\overline{e}^0_{(-i)cs,new}|s) = \sigma^2_{\overline{a}_{new}} + \sigma^2_{\overline{e}_{new}}$.

$\bar{a}^0_{(-i)cs,old}$ on Δy_{ics} depends mechanically on the proportion of former peers in the current peer group $(1 - p_{ics})$.

4. Peer effects in England's secondary schools

In this section, I analyze achievement-based peer effects in England's secondary schools to illustrate the existence of the reversion bias in the standard estimations and demonstrate the proposed *p-weight* estimations. Section 4.1 describes the data. Section 4.2 shows the existence of the reversion bias in the standard school fixed-effect estimations. Section 4.3 presents the results of the proposed *p-weight* estimations.

4.1. Data description

The compulsory education in England is divided into two phases and four key stages. Primary school consists of KS1 (ages 5–7) and KS2 (ages 7–11). Secondary school consists of KS3 (ages 11–14) and KS4 (ages 14–16). The data come from the NPD and comprise four cohorts of pupils entering secondary school during the period 2002–2005 and taking their KS3 national exams during the period 2005–2008. The NPD dataset includes students' raw scores on both the KS2 and KS3 national exams and the schools where they took these exams, thus allowing the identification of and differentiation between new and old peers in secondary schools. As mentioned previously, Lavy et al. (2012) use the same database, although with different cohorts of students. Their identification strategy differs from mine in two important aspects. First, they use a student fixed-effect strategy exploiting the within-student variation in peers' lagged achievement across different subjects. Second, they substitute the lagged achievement measures of all peers' quality with those of new peers' quality directly without accounting for the new peers' share, which (as discussed previously) results in an attenuation bias proportional to the fraction of old peers. In contrast, the *p-weight* estimation used in the current paper is not subject to such an attenuation bias.[12]

Following the same spirit as Lavy et al. (2012), I impose a number of restrictions in selecting the analysis sample. First, I focus exclusively on comprehensive state schools,[13] which account for over 90% of secondary school enrollment in England. Second, I restrict the sample to schools with a cohort size between 25 and 150 during the entire study period. I exclude schools with a cohort size exceeding 150 because larger schools are more likely to implement ability grouping[14] and I measure peer composition at the school-grade level. Third, I further restrict the sample to schools with a maximum- to minimum-cohort-size ratio not exceeding 2 to exclude schools that had been exposed to large enrollment shocks in the study period, which may confound analysis. Finally, I focus only on students with valid test information for both the KS2 and KS3 national exams.[15] After imposing these restrictions, I am left with a sample of over 270,000 pupils from 671 comprehensive state secondary schools with complete information on the KS2 and KS3 exams as well as individual and peer characteristics. In the KS3 national exam, although all students are assigned to the same tier and take the same test for English, they are allocated into different tiers and take different tests for mathematics and science according to their learning levels.[16] Consequently, in England's secondary schools, ability grouping is more prevalent for mathematics and science than English (Kutnick et al., 2006). To further mitigate the ability grouping problem, I focus my analysis exclusively on students' English scores.[17]

Table 1 shows the descriptive statistics for the analysis sample. On average, students in the sample have a secondary school grade size of 118, and 87% of their schoolmates are new peers from a different primary school. As all of the students in the same cohort took the same English test in both the KS2 and KS3 national exams, I measure student achievement by the percentile scores, calculated according to the cohort-specific scores distribution. As the source of identification is the across-cohort variation within schools, Column 3 reports the standard deviations of the demeaned individual- and peer-level variables after subtracting their school-level means, which shows that a fair amount of cohort-to-cohort variation remains after accounting for school heterogeneity. Fig. 1 plots the histogram of the range of changes in schools' grade-level mean KS2 scores, the key peer variable of interest. During the study period, more than one-fifth of the schools in the sample, or 138 out of 671, experienced a change in the grade-level mean KS2 scores of more than 10 percentiles, which is approximately the difference in the school-level mean KS2 scores between the first (41.1) and second tertile (51.0) of England's comprehensive state secondary schools in the sample.

4.2. OLS and school fixed-effect estimation results

Table 2 reports the OLS and school fixed-effect regressions of a student's achievement gain (difference between the KS3 and KS2 scores) on the mean KS2 scores of his/her secondary school peers. Although I also estimate alternative specifications including additional

[12] In addition to estimating the linear-in-means model using mean peer lagged achievement, Lavy et al. (2012) also estimate nonlinear models of peer effects employing measures of peers' lagged achievement distribution. They find only significant negative effects of bad peers at the bottom of the achievement distribution, but little evidence that mean peer quality or good peers matter.

[13] The comprehensive state schools in the sample include community, foundation, voluntary-aided, and voluntary-controlled schools.

[14] Ability grouping, also known as tracking, is the practice of grouping students according to their initial achievement levels into different classes. Although no systematic data are available, anecdotal evidence (Kutnick et al., 2006) suggests that schools with larger cohort sizes are more likely to implement ability grouping in class assignment.

[15] I also include the percentage of a student's KS3 schoolmates with missing KS2 test scores as a measure of peer composition in specifications with peer-level control variables.

[16] For mathematics, the KS3 exam includes four different tiers: 3–5, 4–5, 6–7, and 6–8; for science, the KS3 exam includes two tiers: 3–6 and 5–7.

[17] The results, available upon request, are qualitatively similar for both mathematics and science.

Table 1
Descriptive statistics.

	Mean (1)	S.d. (2)	Within-school s.d. (3)
Panel A. Individual characteristics			
KS3 percentile scores	52.3	29.2	24.2
KS2 percentile scores	50.0	29.0	25.3
Female	0.488	0.500	0.439
Non-white British	0.207	0.405	0.405
English as an additional language (EAL)	0.110	0.313	0.259
Panel B. Peer/School-grade characteristics			
Proportion of new peers	0.865	0.153	0.111
Proportion of female peers	0.484	0.242	0.040
Proportion of non-white British peers	0.220	0.266	0.055
Proportion of peers with EAL	0.124	0.188	0.049
Proportion of peers with missing KS2 scores	0.075	0.058	0.031
Mean KS2 percentile scores of all peers	50.0	14.5	3.0
Mean KS2 percentile scores of new peers	49.8	14.7	3.5
Mean KS2 percentile scores of old peers[a]	49.6	18.9	13.1
School-grade size	117.5	21.7	9.3
Number of pupils		276,562	
Number of schools		671	

Notes: This table reports the means, standard deviations, and within-school standard deviations of the listed variables. The sample only includes pupils with valid KS2 and KS3 test scores as well as individual characteristics in state comprehensive secondary schools with a cohort size between 25 and 150 during the entire study period and a maximum- to minimum-cohort-size ratio not exceeding 2.

[a] A smaller sample ($n = 237,335$) is used to calculate the mean KS2 percentile scores of old peers as some pupils had no old peers in their secondary schools.

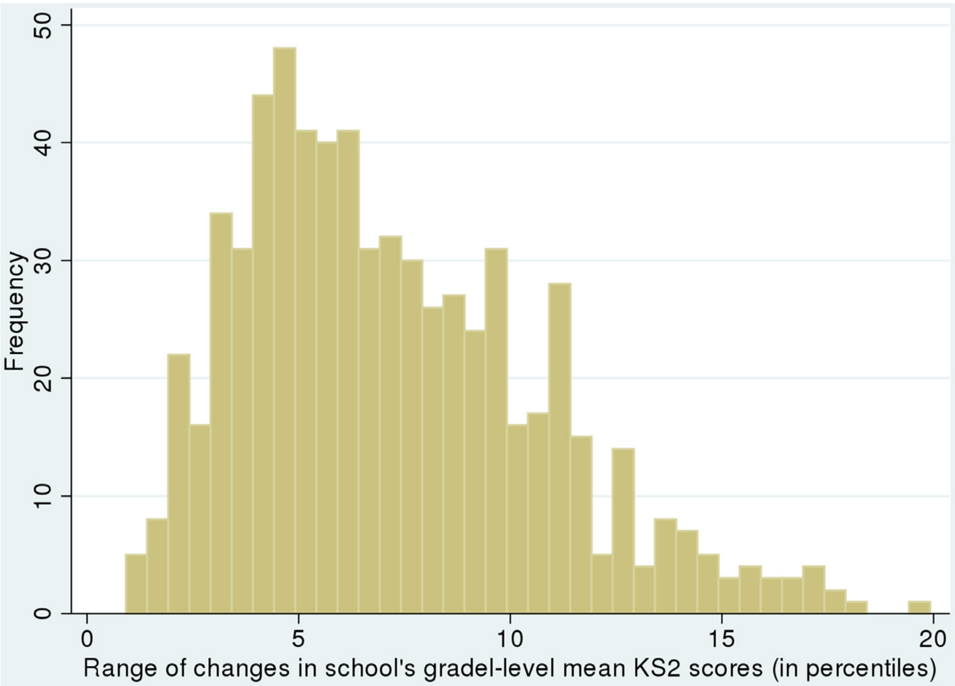

Fig. 1. Histogram of the range of changes in school's grade-level mean KS2 scores.

individual- and peer-level control variables,[18] their inclusion has little effect on the estimates of the coefficient on mean peer lagged achievement (hereafter *peer coefficient*) in all of the regressions. Therefore, I discuss only estimates of the peer coefficient with no control variables for the remainder of this section. The OLS estimate of the peer coefficient in Column 1 (0.085) shows a significant positive cross-sectional relationship between a student's achievement gain from the KS2 to KS3 national exam and the average KS2 scores of his/her secondary school peers.

[18] Individual-level control variables include dummy indicators for female, non-white British, and English as an additional language (EAL). Peer-level control variables include the proportion of female peers, the proportion of non-white British peers, the proportion of peers with EAL, and the proportion of peers with missing KS2 scores.

Table 2
OLS and school fixed-effect estimations.

	OLS			School fixed-effect		
	(1)	(2)	(3)	(4)	(5)	(6)
Mean KS2 percentile scores of all peers	0.085***	0.088***	0.089***	−0.310***	−0.309***	−0.320***
	(0.008)	(0.008)	(0.008)	(0.032)	(0.032)	(0.032)
Individual-level controls		X	X		X	X
Peer-level controls			X			X
Secondary school fixed effects				X	X	X
Number of observations		276,562			276,562	

Notes: The dependent variable is achievement gain in percentile scores of English between KS2 and KS3 national exams. Individual-level control variables include dummy indicators for female, non-white British, and English as an additional language (EAL). The peer-level control variables include the proportion of female peers, the proportion of non-white British peers, the proportion of peers with EAL, and the proportion of peers with missing KS2 scores. Robust standard errors clustered by secondary school interacted with KS3 exam year are reported in parentheses.
*** indicates statistical significance at the 1% level.

Table 3
School fixed-effect estimations using separate mean lagged achievement for new peers and old peers.

	(1)	(2)	(3)
Mean KS2 percentile scores of old peers	−0.223***	−0.222***	−0.223***
	(0.005)	(0.005)	(0.005)
Mean KS2 percentile scores of new peers	0.077***	0.076***	0.071***
	(0.027)	(0.027)	(0.027)
Individual-level controls		X	X
Peer-level controls			X
Number of observations		237,335	

Notes: The dependent variable is the achievement gain in percentile scores of English between the KS2 and KS3 national exams. All regressions include secondary school fixed effects. Individual-level control variables include dummy indicators for female, non-white British, and EAL. The peer-level control variables include the proportion of female peers, the proportion of non-white British peers, the proportion of peers with EAL, and the proportion of peers with missing KS2 scores. Robust standard errors clustered by secondary school interacted with KS3 exam year are reported in parentheses. A smaller sample is used in this table because some pupils had no old peers in their secondary schools.
*** indicates statistical significance at the 1% level.

However, the school fixed-effect estimate in Column 4 yields a markedly different result: $\widehat{\lambda}_{FE}$ (-0.310) is larger in magnitude than $\widehat{\lambda}_{OLS}$ but with the opposite sign. Although the literature has not yet reached a consensus regarding the existence or size of peer effects, it seems unlikely that mean peer lagged achievement could have an adverse effect of a magnitude of 0.3 as indicated by $\widehat{\lambda}_{FE}$. Hence, the surprisingly large negative $\widehat{\lambda}_{FE}$ suggests evidence of the existence of some negative bias.

To further illustrate that the negative sign of the estimated peer coefficient is indeed caused by the reversion bias due to the continuing presence of former peers, I replace mean peer lagged achievement with separate mean lagged achievements for old and new peers and rerun the school fixed-effect regressions in Table 3. The estimates of the coefficients on both mean measures are significant but have the opposite signs. While the negative coefficient on the mean lagged achievement of old peers (−0.223) is consistent with my hypothesis that the presence of old peers leads to a negative reversion bias, the positive coefficient on that of new peers (0.077) suggests that positive peer effects do exist, although they are sheltered by the severe negative reversion bias when new and old peers are pooled and estimated together.

4.3. *P-weight estimation results*

Following the approach presented in Section 3, I address the reversion-bias problem by performing a *p-weight* estimation, in which I regress a pupil's achievement gain between the KS2 and KS3 national exams on the product of the new peers' share (p_{ics}) and average KS2 scores ($\bar{y}^0_{(-i)cs,new}$), the set of secondary school dummies ($\mathbf{1}_s$), as well as their interactions with both the new peers' share ($p_{ics} \cdot \mathbf{1}_s$) and the new peers' share squared ($p^2_{ics} \cdot \mathbf{1}_s$). Table 4 reports the results of the *p-weight* estimations: the estimated peer coefficient is 0.131 in Column 1, significant at the 1% level and also robust to the inclusion of additional covariates in Columns 2 and 3. The remarkable difference between the positive $\widehat{\lambda}_p$ in Table 4 and the negative $\widehat{\lambda}_{FE}$ in Table 2 provides strong evidence that $\widehat{\lambda}_{FE}$ contains a negative reversion bias and emphasizes the need to correct this bias as it even reverses the sign of the estimate. Given the point estimate $\widehat{\lambda}_p$, a 10-percentile increase in mean peer lagged achievement – a magnitude of change corresponding to the difference in the school-level mean KS2 scores between the first and second tertiles of secondary schools in the sample – would increase a student's achievement gain over three years by 1.3 percentiles, a modest but still respectable improvement.

Table 4
School fixed-effect *p-weight* estimations.

	(1)	(2)	(3)
Proportion of new peers ∗ mean KS2	0.131***	0.130***	0.125***
Percentile scores of new peers	(0.035)	(0.035)	(0.035)
Individual controls		X	X
Peer-level controls			X
Number of observations		276,562	

Notes: The dependent variable is achievement gain in percentile scores of English between KS2 and KS3 national exams. The coefficient on *proportion of new peers ∗ mean KS2 percentile scores of new peers* corresponds to the p-weight estimates of the peer coefficient. All regressions control for secondary school fixed effects and their interactions with both the proportion of new peers and the proportion of new peers squared. Individual-level control variables include dummy indicators for female, non-white British, and EAL. The peer-level control variables include the proportion of female peers, the proportion of non-white British peers, the proportion of peers with EAL, and the proportion of peers with missing KS2 scores. Robust standard errors clustered by secondary school interacted with KS3 exam year are reported in parentheses.
*** indicates statistical significance at the 1% level.

5. Conclusion

Although the existence and consequences of common testing noise among students from the same classroom or school have been well acknowledged in the school accountability literature, the common testing noise problem attracts little attention in the peer effects literature and is largely ignored in prior research. In this paper, I investigate the role of common testing noise in the value-added estimation of peer effects using lagged peer achievement and find that the consequences may go beyond generating an attenuation bias. I show that a negative reversion bias could arise if a student's current peers include some of his/her former peers whose performances on the baseline test are subject to the same common transitory influences as experienced by the student him/herself. Using real-world data, I demonstrate that this concern is more than theoretical. In analysis of achievement-based peer effects in England's secondary schools, the standard school fixed-effect estimates of the coefficient on mean peer lagged achievement are negative and significant, suggesting not only the existence of the negative reversion bias but also its dominance over the true peer effects (if indeed any).

As the source of the reversion bias is the continuing presence of former peers in a student's current peer group, I propose an alternative estimation strategy to overcome this problem by exploiting only the variation in the portion of mean peer lagged achievement attributable exclusively to new peers. Performing the proposed estimation yields modest positive coefficient estimates, suggesting that positive peer effects do exist, although they are overwhelmed by the severe negative reversion bias in the standard estimates. While this paper focuses on the achievement context, its insight into the potential effect of correlated measurement error may also be important for understanding peer influences in other settings. For example, in workplaces or tournaments, the performances of former colleagues or players may be subject to correlated measurement error due to common environmental influences and these individuals may also encounter each other in future collaborations or competitions. The solution proposed in this paper also provides insight into overcoming the challenges imposed by measurement error in the identification of peer effects in such settings.

References

Abdulkadiroğlu, Angrist, A. J., & Pathak, P. (2014). The elite illusion: Achievement effects at boston and new york exam schools. *Econometrica, 82*(1), 137–196.
Ammermueller, A., & Pischke, J.-S. (2009). Peer effects in european primary schools: Evidence from the progress in international reading literacy study. *Journal of Labor Economics, 30*, 98–108.
Angrist, J. D. (2014). The perils of peer effects. *Labour Economics, 94*(5), 1613–1634.
Angrist, J. D., & Lang, K. (2004). Does school integration generate peer effects? evidence from boston's metco program. *American Economic Review, 94*(5), 1613–1634.
Arcidiacono, P., & Nicholson, S. (2005). Peer effects in medical school. *Journal of Public Economics, 89*(2–3), 327–350.
Betts, J. R., & Danenberg, A. (2002). School accountability in california: An early evaluation. *Brookings Papers on Education Policy 2002*, 123–197.
Brock, W. A., & Durlauf, S. N. (2001). Interactions-based Models. In J. J Heckman, & E. Leamer (Eds.), *Handbook of Econometrics: Vol. 5* (pp. 3297–3367). Amsterdam: Elsevier Science.
Chay, K. Y., McEwan, P. J., & Urquiola, M. (2005). The central role of noise in evaluating interventions that use test scores to rank schools. *American Economic Review, 95*(4), 1237–1258.
Ding, W., & Lehrer, S. F. (2007). Do peers affect student achievement in china's secondary schools? *Review of Economics and Statistics, 89*(2), 300–312.
Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. *American Economic Review, 101*(5), 1739–1774.
Epple, D., & Romano, R. E. (2011). Peer Effects in Education: A Survey of the Theory and Evidence. In J. Benhabib, A. Bisin, & M. O. Jackson (Eds.), *Handbook of social economics: Vol. 1B* (pp. 1053–1163). Amsterdam: Elsevier Science.
Foster, G. (2006). It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism. *Journal of Public Economics, 90*(8–9), 1455–1475.
Fruehwirth, C. J. (2014). Can achievement peer effect estimates inform policy? a view from inside the black box. *Review of Economics and Statistics, 96*(3), 514–523.
Gaviria, A., & Raphael, S. (2001). School-based peer effects and juvenile behavior. *Review of Economics and Statistics, 83*(2), 257–268.
Gibbons, S., & Telhaj, S. (2011). Pupil mobility and school disruption. *Journal of Public Economics, 95*(9–10), 1156–1167.
Glaeser, E. L., Sacerdote, B., & Scheinkman, J. e. A. (1996). Crime and social interactions. *Quarterly Journal of Economics, 111*(2), 507–548.
Gould, E. D., Lavy, V., & Paserman, M. D. (2009). Does immigration affect the long-term educational outcomes of natives? quasi-experimental evidence. *Economic Journal, 119*(540), 1143–1169.
Guryan, J., Kroft, K., & Notowidigdo, M. J. (2009). Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics, 1*(4), 34–68.
Han, L., & Li, T. (2009). The gender difference of peer influence in higher education. *Economics of Education Review, 28*(1), 129–134.

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

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2009). New evidence about brown v. board of education: The complex effects of school racial composition on achievement. *Journal of Labor Economics, 27*(3), 349–383.

Hoxby, C. M. (2000). *Peer effects in the classroom: Learning from gender and race variation*. NBER. Working Papers No. 7867.

Hoxby, C. M., & Weingarth, G. (2005). *Taking race out of the equation: School reassignment and the structure of peer effects*. Harvard University Working Paper.

Kane, T. J., & Staiger, D. O. (2002). The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives, 16*(4), 91–114.

Kang, C. (2007). Classroom peer effects and academic achievement: Quasi-randomization evidence from south korea. *Journal of Urban Economics, 61*(3), 458–495.

Kremer, M., & Levy, D. (2008). Peer effects and alcohol use among college students. *Journal of Economic Perspectives, 22*(3), 189–206.

Kutnick, P., Hodgekinson, S., Sebba, J., Humphreys, S., Galton, M., Steward, S., et al. (2006). Pupil grouping strategies and practices at key stage 2 and 3: Case studies of 24 schools in England. *Research Report 796*. United Kingdom: Department for Education and Skills.

Lavy, V., & Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics, 3*(2), 1–33.

Lavy, V., Silva, O., & Weinhardt, F. (2012). The good, the bad, and the average: Evidence on ability peer effects in schools. *Journal of Labor Economics, 30*(2), 367–414.

Lefgren, L. (2004). Educational peer effects and the chicago public schools. *Journal of Urban Economics, 56*(2), 169–191.

Luppino, M. (2015). Peer turnover and student achievement: Implications for classroom assignment policy. *Economics of Education Review, 46*, 98–111.

Lyle, D. S. (2007). Estimating and interpreting peer and role model effects from randomly assigned social groups at west point. *Review of Economics and Statistics, 89*(2), 289–299.

Manski, C. F. (1993). Identification RRROF endogenous social effects: The reflection problem. *Journal of Economic Studies, 14*(3), 115–136.

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

McEwan, P. J. (2003). Peer effects on student achievement: Evidence from chile. *Economics of Education Review, 22*, 131–141.

Micklewright, J., Schnepf, S. V., & Silva, P. N. (2012). Peer effects and measurement error: The impact of sampling variation in school survey data (evidence from PISA). *Economics of Education Review, 31*(2012), 1136–1142.

Moffit, R. A. (2011). Policy interventions, low-level equilibria, and social interactions. In S. N. Durlauf, & H. P. Young (Eds.), *Social dynamics*. , Cambridge, MA: MIT Press.

Moretti, E. (2011). Social learning and peer effects in consumption: Evidence from movie sales. *Review of Economic Studies, 78*(1), 356–393.

Sacerdote, B. (2001). Peer effects with random assignment: Results from dartmouth roommates. *Quarterly Journal of Economics, 116*(2), 681–704.

Sacerdote, B. (2011). Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education: Vol. 3* (pp. 249–277). New York: Elsevier.

Sojourner, A. (2013). Identification of peer effects with missing peer data: Evidence from project STAR. *Economic Journal, 123*(569), 574–605.

Stinebrickner, R., & Stinebrickner, T. R. (2006). What can be learned about peer effects using college roommates? evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics, 90*(8-9), 1435–1454.

Vigdor, J. L., & Nechyba, T. (2007). Peer effects in north carolina public schools. In L. Woessmann, & P. E Peterson (Eds.), *Schools and the equal opportunity problem* (pp. 73–102). Cambridge, MA: MIT Press.

Zabel, J. E. (2008). The impact of peer effects on student outcomes in new york city public schools. *Education Finance and Policy, 3*(2), 197–249.

Zimmerman, D. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Journal of Urban Economics, 85*(1), 9–23.