The Journal of Educational Research

ISSN: 0022-0671 (Print) 1940-0675 (Online) Journal homepage: http://www.tandfonline.com/loi/vjer20

Endogenous peer effects: Fact or fiction?

Ryan Yeung & Phuong Nguyen-Hoang

To cite this article: Ryan Yeung & Phuong Nguyen-Hoang (2016) Endogenous peer effects: Fact or fiction?, The Journal of Educational Research, 109:1, 37-49, DOI: 10.1080/00220671.2014.918528

To link to this article: http://dx.doi.org/10.1080/00220671.2014.918528

Published online: 28 Dec 2015.

THE JOURNAL OF EDUCATIONAL RESEARCH
2016, VOL. 109, NO. 1, 37–49
http://dx.doi.org/10.1080/00220671.2014.918528

Endogenous peer effects: Fact or fiction?

Ryan Yeung[a] and Phuong Nguyen-Hoang[b]

[a]Department of Public Administration, State University of New York, College at Brockport, Brockport, New York, USA; [b]School of Urban and Regional Planning & Public Policy Center, University of Iowa, Iowa City, Iowa, USA

ABSTRACT
The authors examine endogenous peer effects, which occur when a student's behavior or outcome is a function of the behavior or outcome of his or her peer group. Endogenous peer effects have important implications for educational policies such as busing, school choice and tracking. In this study, the authors quantitatively review the literature on endogenous peer effects through the use of meta-analytic methods. They find a significant and positive endogenous peer effect. It appears to be a genuine empirical effect but is dependent on the measure of educational outcomes, the peer group, publication status, and publication year.

ARTICLE HISTORY
Received 19 June 2013
Revised 23 March 2014
Accepted 21 April 2014

KEYWORDS
Analysis; metaregression; peer effects

American children spend approximately 6.5 hr a day, 180 days a year, in school (Silva, 2007). Much of this time is spent in the company of other children. This circumstance has not been lost on the research community. Prompted by the publication of the influential Coleman Report in 1966, a large body of research has examined the impact of academic peers on student outcomes. Harris (2010) defined a peer as "another student with whom the individual student comes in contact in school-related activities" (p. 1167). Peer effects in an educational setting occur when a student's performance or behavior is influenced by his/her interactions with other students or peers.[1] How do peers influence a student's performance?

Peers can influence a student's learning outcomes via two major effects: exogenous (also called compositional or contextual) and endogenous effects. These two effects come from the idea that a student's individual characteristics and behavior can facilitate or impede other students' learning. While exogenous effects occur when an individual student's educational achievement is influenced by, as the name suggests, exogenous peer characteristics (i.e., socioeconomic status, ethnicity, or gender), endogenous effects exist when the student's outcome is affected by peer behavior including outcomes (Manski, 1993). Exogenous and endogenous effects constitute peer effects or influences.

Research on the relationship between peers and student achievement is substantial. van Ewijk and Sleegers (2010a, 2010b) conducted meta-analytic reviews of studies that examined the role exogenous peer effects, in this case ethnicity and socioeconomic status, played in student achievement. This study, however, is the first meta-analysis (MA) to look at endogenous peer effects and student performance in a K–12 educational setting. Endogenous effects are important because only these effects, rather than exogenous ones, are a source of a social multiplier (Boucher, Bramoullé, Djebbari, & Fortin, 2014).[2] While the large majority of studies find significant endogenous peer effects, questions remain as to whether these significant estimates are a result of publication selection, and whether they vary across different research designs. We conducted a MA and a metaregression analysis (MRA) of empirical studies on endogenous peer effects in K–12 education to answer these questions. MRA corrects for publication selection bias in MA and helps uncover how different study characteristics affect estimates of endogenous peer effects. The empirical findings from the MA and MRA provide strong support for positive but small endogenous peer effects. The MRA results also indicate that some study characteristics might affect the estimate of endogenous peer effects.

Peer effects, if they exist, have important implications for education policy as educators attempt the social engineering of schooling. As with other common school inputs (e.g., teacher quality, class size, expenditures), many argue that peer composition is an important determinant of student achievement (Sacerdote, 2011). Endogenous peer effects are particularly important in understanding the effects of various educational reform initiatives that states across the United States have taken to boost academic achievement. Initiatives such as busing, school choice, and tracking (or ability grouping),[3] are based at least partly on the idea of peer effects. For a decade, the busing program in Wake County, North Carolina, was intended to promote diversity across schools so that poor-performing students could benefit academically from associating with high-performing students.

School choice can come in the form of education vouchers, private and charter schools, or inter-district school choice programs. A common concern is that school choice allows high-ability parents to flee to private or suburban schools. School choice programs have been accused of cream skimming (i.e., luring the best students from regular public schools; Altonji,

CONTACT Ryan Yeung ryeung@brockport.edu Department of Public Administration, The College at Brockport MetroCenter, 55 St. Paul Street, Rochester, NY 14604.

Huang, & Taber, 2010), creating schools with better performing students and schools with the remaining students who are of lesser performance and thus likely to be even worse off academically without their academically brighter peers around.

Tracking, or ability grouping, which has started to regain favor again recently (Yee, 2013), is believed to be beneficial to students academically because it allows teachers to tailor the pace and content of instruction better to students' needs. Also, ability grouping makes students more comfortable and engaged because they are surrounded by similar children, and high achievers flag when they are in classes with low performers (Westchester Institute for Human Services Research, 2002). The main argument for all of these education initiatives is that if low-achieving children benefit from heterogeneous classes, and high-achievers are not harmed, these initiatives would produce Pareto-superior outcomes. Put differently, if bad peers gain more from good peers than good peers are harmed by bad peers, such programs would create socially efficient outcomes.

The remainder of this paper is organized as follows. The next section presents background on endogenous peer effects including an economic model of endogenous peer effects. We briefly review the literature on endogenous peer effects in section three. In section four, we describe our process of selecting studies, discuss our measure of effect size for MA, and present an MRA model with relevant variables. Section five discusses the results and section six concludes.

Background on endogenous peer effects

Endogenous peer effects or influences were first formalized in a study by Erbring and Young (1979), and, later, in a seminal article by Manski (1993). According to Manski (1993), endogenous effects occur when, "the propensity of an individual to behave in some way varies with the behavior of the group" (p. 532). For example, a student's test score may depend on the average test score of his or her peers as in Equation 1.

Suppose Y represents an outcome or behavior like academic achievement (e.g., test score) of student i in group g.[4] A linear-in-means model to estimate peer effects is presented in Equation 1, which is adapted from Sacerdote (2011):

$$Y_{ig} = \sigma + \gamma X_{ig} + \alpha \overline{X}_{-ig} + \beta \overline{Y}_{-ig} + \varepsilon_{ig}, \qquad (1)$$

where σ is a constant term and ε is the error term. X_{ig} is a vector of the student's exogenous characteristics (e.g., free or reduced-priced lunch status [or parental income], limited English proficiency, parental education, gender, race, or ethnicity). $\overline{X}_{-ig}$ is a vector of his or her peers' average background (or exogenous) characteristics, and $\overline{Y}_{-i}$ represents his or her peers' average achievement. This model provides estimates of both exogenous (α) and endogenous (β) peer effects. In other words, in Equation 1, a student's achievement is a function of his or her free or reduced-priced lunch status (γ), the average free or reduced-priced lunch status of his or her group peers (α), and the average achievement of his or her group peers (β). The coefficient of endogenous peer effects, β, is of principal interest in this study.

Using an ordinary least squares estimation approach is likely to produce biased estimates of $\boldsymbol{\beta}$ from several sources of potential bias. The first source of bias results from the reflection problem (or simultaneity) as indicated in Manski (1993). The reflection problem occurs because a student's peer achievement, $\overline{\boldsymbol{Y}}$, affects his or her own achievement, $\boldsymbol{Y}$, and vice versa. While the reflection problem seems intractable on its face, scholars have proposed some empirical strategies to address it.

A popular approach is to use lagged peer achievement. The strength of this method depends on the strength of the relationship between current and past peer achievement. Lagged achievement is a good proxy for current achievement if there are no year-to-year shocks in current student achievement. This may not be the case. Studies have found that school test scores tend to show a significant component of random variation across years (Kane & Staiger, 2002). Also, lagged peer achievement does not capture the effect of current peer achievement on student achievement, thereby leading to downwardly biased results (Hanushek, Kain, Markman, & Rivkin, 2003). In addition, lagged average achievement is still likely to be endogenous due to selection into peer groups (Fruehwirth, 2013), or serial correlation with unobserved teacher, school and individual factors (Hanushek et al., 2003). Despite these potential problems, this lagged measure is, as indicated in Table 1, the most common method of addressing the reflection problem in the studies included in our MRA.

Another popular method to address this simultaneity issue is to use instrumental variables (IV) regression. The strength of the IV design is dependent on the validity of the IVs. A valid IV must satisfy two conditions. First, the endogenous variable, $\overline{Y}_{-ig}$, must be correlated strongly with the IV. Murray (2006) argued that weak IVs may have two problems: (a) two-stage least-squares (TSLS) estimates with weak IVs are biased, even in large samples; and (b) TSLS-estimated standard errors become too small, leading to overstated statistical significance in hypothesis testing. Second, IVs must not be correlated with the error term.

IVs that satisfy these two conditions are hard to come by. For instance, Zabel (2008) and Hanushek et al. (2003) instrumented current average peer test scores with lagged average peer test scores. Because lagged peer test scores were included directly into Equation 1 as adopted by several studies in Table 1, the validity of this IV (lagged peer scores) in these two studies was questionable given its potential failure to pass the second condition. Similarly, the average peer background characteristics that Gaviria and Raphael (2001) used as IVs for average peer behaviors (including dropping out while in Grade 11 or 12) are also questionable because they are usually estimated directly as exogenous peer effects in Equation 1.

Other sources of bias can be illustrated by dividing ε into several components (in addition to a random error capturing individual time-varying shocks) that may have effects on achievement: unobserved time-invariant individual (and family) factors (e.g., innate ability, self-motivation, parental involvement), and unobserved school, school-by-grade, and cohort-specific factors.[5] Of these, time-invariant omitted and mismeasured individual factors are one of the most difficult sources of bias to address (Hanushek et al., 2003). Student fixed effects can be used to address this bias. However, this student

Table 1. Studies selected for metaregression analysis.

Study	Endogenous peer effect measure	Reflection solution	Peer group	Time-series data?	Whole-population data?	Industrialized economy?	Published?
Asadullah & Chaudhury (2008)	MTS	IV	School	No	Yes	No	No
Atkinson, Burgess, Gregg, Propper, & Proud (2008)	MTS	L	Classroom	Yes	Yes	Yes	No
Atkinson et al. (2008)	RTS	L	Classroom	Yes	Yes	Yes	No
Betts & Zau (2004)	MTS	L	Classroom	Yes	Yes	Yes	No
Betts & Zau (2004)	MTS	L	Grade	Yes	Yes	Yes	No
Betts & Zau (2004)	RTS	L	Classroom	Yes	Yes	Yes	No
Betts & Zau (2004)	RTS	L	Grade	Yes	Yes	Yes	No
Boozer & Cacciola (2001)	CTS	IV	Classroom	No	Yes	Yes	No
Boucher et al. (2014)	MTS	IV	School	No	No	Yes	Yes
Boucher et al. (2014)	O	IV	School	No	No	Yes	Yes
Boucher et al. (2014)	O	IV	School	No	No	Yes	Yes
Boucher et al. (2014)	RTS	IV	School	No	No	Yes	Yes
Bradley & Taylor (2008)	CTS	L	School	Yes	Yes	Yes	No
Burke & Sass (2013)	MTS	RF	Classroom	Yes	Yes	Yes	Yes
Burke & Sass (2013)	MTS	RF	Grade	Yes	Yes	Yes	Yes
Burke & Sass (2013)	RTS	RF	Classroom	Yes	Yes	Yes	Yes
Burke & Sass (2013)	RTS	RF	Grade	Yes	Yes	Yes	Yes
Carman & Zhang (2012)	MTS	L	Classroom	Yes	Yes	No	Yes
Carman & Zhang (2012)	O	L	Classroom	Yes	Yes	No	Yes
Carman & Zhang (2012)	RTS	L	Classroom	Yes	Yes	No	Yes
Clark, Scafidi, & Swinton (2011)	O	IV	School	No	Yes	Yes	Yes
Clotfelter, Ladd, & Vigdor (2010)	CTS	RF	Classroom	Yes	Yes	Yes	Yes
Cook et al. (2007)	MTS	RF	Friends	No	Yes	Yes	Yes
Cook et al. (2007)	O	RF	Friends	No	Yes	Yes	Yes
Ding & Lehrer (2007)	CTS	L	School	Yes	Yes	No	Yes
Duflo et al. (2011)	CTS	L	Classroom	No	Yes	No	Yes
Duflo et al. (2011)	CTS	IV	Classroom	No	Yes	No	Yes
Duflo et al. (2011)	MTS	L	Classroom	No	Yes	No	Yes
Duflo et al. (2011)	MTS	IV	Classroom	No	Yes	No	Yes
Duflo et al. (2011)	RTS	L	Classroom	No	Yes	No	Yes
Duflo et al. (2011)	RTS	IV	Classroom	No	Yes	No	Yes
Dumay & Dupriez (2008)	RTS	L	School	Yes	Yes	Yes	Yes
Fortner (2010)	MTS	L	Classroom	Yes	Yes	Yes	No
Fortner (2010)	RTS	L	Classroom	Yes	Yes	Yes	No
Fruehwirth (2013)	RTS	IV	Classroom	Yes	Yes	Yes	Yes
Fruehwirth (2013)	RTS	L	Classroom	Yes	Yes	Yes	Yes
Freuhwirth (2013)	RTS	IV	Grade	Yes	Yes	Yes	Yes
Fruehwirth (2014)	RTS	RF	Grade	Yes	Yes	Yes	Yes
Fruehwirth (2014)	RTS	L	Grade	Yes	Yes	Yes	Yes
Gaviria & Raphael (2001)	O	IV	School	No	No	Yes	Yes
Gibbons & Telhaj (2008)	CTS	L	School	Yes	Yes	Yes	No
Gibbons & Telhaj (2012)	O	L	School	Yes	Yes	Yes	No
Gottfried (2010)	MTS	L	Classroom	Yes	Yes	Yes	No
Gottfried (2010)	MTS	L	Classroom	Yes	Yes	Yes	No
Gottfried (2010)	RTS	L	Classroom	Yes	Yes	Yes	No
Gottfried (2010)	RTS	L	Classroom	Yes	Yes	Yes	No
Graham (2008)	MTS	IV	Classroom	No	Yes	Yes	Yes
Graham (2008)	RTS	IV	Classroom	No	Yes	Yes	Yes
Halliday & Kwak (2012)	O	IV	Friends	No	No	Yes	Yes
Halliday & Kwak (2012)	O	IV	Grade	No	No	Yes	Yes
Hanushek et al. (2003)	MTS	L	Grade	Yes	Yes	Yes	Yes
Hoxby & Weingarth (2006)	CTS	L	Grade	Yes	Yes	Yes	No
Imberman et al. (2012)	CTS	IV	Grade	Yes	Yes	Yes	Yes
Jackson (2009)	CTS	IV	School	Yes	Yes	No	No
Kang (2007)	MTS	IV	Classroom	No	Yes	Yes	Yes
Kiss (2011)	MTS	L	Grade	Yes	Yes	Yes	No
Kramarz, Machin, & Ouazad (2010)	O	IV	Grade	Yes	Yes	No	No
Lai (2007)	CTS	RF	Classroom	Yes	Yes	No	No
Lavy et al. (2012)	CTS	L	Grade	Yes	Yes	No	Yes
Lefgren (2004)	MTS	IV	Classroom	Yes	Yes	Yes	Yes
Lefgren (2004)	RTS	IV	Classroom	Yes	Yes	Yes	Yes
Leiter (1983)	MTS	RF	Classroom	No	Yes	Yes	Yes
Leiter (1983)	RTS	RF	Classroom	No	Yes	Yes	Yes
Lin (2010)	O	IV	Friends	No	No	Yes	Yes
Link & Mulligan (1991)	MTS	L	Classroom	No	No	Yes	Yes
Link & Mulligan (1991)	RTS	L	Classroom	No	No	Yes	Yes
Liu, Patacchini, & Zenou (2013)	O	IV	Friends	No	No	Yes	No
Liu et al. (2013)	O	IV	Friends	No	No	Yes	No
Mora & Oreopoulos (2011)	O	RF	Classroom	No	Yes	Yes	Yes
Mora & Oreopoulos (2011)	O	RF	Friends	No	Yes	Yes	Yes
Mounts & Steinberg (1995)	O	RF	Friends	No	No	Yes	Yes

(*Continued*)

Table 1.(*Continued*)

Study	Endogenous peer effect measure	Reflection solution	Peer group	Time-series data?	Whole-population data?	Industrialized economy?	Published?
Nores (2006)	MTS	L	Classroom	Yes	No	Yes	No
Nores (2006)	MTS	L	School	Yes	No	Yes	No
Nores (2006)	RTS	L	Classroom	Yes	No	Yes	No
Nores (2006)	RTS	L	School	Yes	No	Yes	No
Patacchini, Rainone, & Zenous (2013)	O	L	Friends	No	No	Yes	No
Rangaprasad (2004)	CTS	IV	Classroom	Yes	No	Yes	No
Rangaprasad (2004)	CTS	IV	School	Yes	No	Yes	No
Ryabov (2011)	O	RF	School	No	No	Yes	Yes
Sojourner (2013)	CTS	L	Classroom	Yes	Yes	Yes	Yes
Sund (2009)	O	L	Classroom	Yes	Yes	Yes	Yes
Thomas & Webber (2001)	O	RF	School	No	Yes	Yes	Yes
Vigdor & Nechyba (2007)	MTS	L	Classroom	Yes	Yes	No	No
Vigdor & Nechyba (2007)	MTS	L	Grade	Yes	Yes	No	No
Vigdor & Nechyba (2007)	RTS	L	Classroom	Yes	Yes	No	No
Vigdor & Nechyba (2007)	RTS	L	Grade	Yes	Yes	No	No
Vigdor & Nechyba (2008)	MTS	L	Classroom	Yes	Yes	Yes	No
Vigdor & Nechyba (2008)	RTS	L	Classroom	Yes	Yes	Yes	No
Wang (2010)	MTS	L	Classroom	Yes	Yes	No	No
Zabel (2008)	MTS	IV	Classroom	Yes	Yes	Yes	Yes
Zabel (2008)	RTS	IV	Classroom	Yes	Yes	Yes	Yes
Zhang (2011)	MTS	IV	Grade	Yes	Yes	No	No
Zhang (2011)	MTS	IV	Grade	Yes	Yes	No	No
Zimmer (2003)	MTS	L	Classroom	Yes	No	Yes	Yes
Zimmer & Toma (2000)	MTS	RF	Classroom	Yes	No	Yes	Yes

Note, MTS = mathematics test score; RTS = reading test score; CTS = combined test score; O = other measures of peer achievement; IV = instrumental variables; L = lagged peer achievement; RF = reduced form.

fixed effects approach requires individual-level panel data. Potential bias from the remaining three components of ε can be addressed by using a set of different dummies (e.g., school-by-grade or cohort-by-grade-by-year dummies).

Literature review

The importance of peers to student outcomes has been identified as early as the seminal Coleman Report in 1966. Coleman et al. reported that "a pupil's achievement is strongly related to the educational backgrounds and aspirations of the other students in the school. Many studies have attempted to empirically examine peer effects, both exogenous and endogenous, in an educational setting. For example, an influential paper by Henderson, Mieszkowski, and Sauvageau (1978) found that average classroom IQ was positively associated with gains in French and mathematics test scores for a sample of Montreal students. This study did not identify endogenous peer effects per se because the IQ scores were different from the mathematics and French test scores used as the dependent variables. Given the interest of our study, we now focus on studies on endogenous peer effects since 1980.

Endogenous peer effect studies show variability in terms of whether they found evidence of endogenous peer effects, and if yes, how large the peer effects are. While a small number of studies did not find statistically significant endogenous peer effects (Cook, Deng, & Morgano, 2007; Lai, 2007; Zhang, 2011), the large majority of studies found significant effects of at least an average peer achievement measure on a student's own achievement. For instance, using school reassignment in Wake County, North Carolina, to identify the effect of peer achievement on student outcomes, Hoxby and Weingarth (2006) found that an increase of mean peer initial achievement by one point increased a student's own achievement by approximately 0.25 points. They also found that students in the extremes of the test score spectrum benefited from peers who had similar levels of achievement. Hanushek et al. (2003) using panel data from Texas found that a 0.1 standard deviation increase in lagged peer average mathematics test scores in the grade-cohort led to a roughly 0.02 standard deviation increase in a student's own mathematics score.

As in Hanushek et al. (2003), Fruehwirth (2013) found strong positive endogenous peer effects with data from North Carolina public elementary schools. Furthermore, Fruehwirth (2013) presented evidence that the endogenous effects from peers of the same race were larger than those from different races. Specifically, average reading scores of White and non-White students increased either by 0.22 and 0.07 standard deviations, respectively, for a one standard deviation increase in white peer achievement. However, the increase in reading scores for white and nonwhite students was 0.01 and 0.28 standard deviations, respectively, for a one standard deviation increase in non-White peer achievement. Relatively larger endogenous peer effects were found in Duflo, Dupas, and Kremer (2011), which reported using data from Kenya that a one standard deviation increase in average peer test score increased a student's own test score by 0.53 standard deviations.

Several studies estimated endogenous peer effects separately for several academic achievement measures. While many studies reported statistically significant endogenous peer effects for all academic achievement measures (Betts & Zau, 2004; Duflo et al., 2011; Zabel, 2008), some of these studies found that endogenous peer effects were significant only for one achievement measure. For instance, studies in Quebec, Canada, by Boucher et al. (2014) and in China by Carman and Zhang

(2012) found that only peer mathematics scores have significant effects on a student's mathematics scores.

As reported in Table 1, these studies have substantial variation in terms of measures of academic achievement, datasets (an American state vs. a developing country), and other study characteristics. Therefore, meta-analytic estimation methods are needed to provide evidence on whether there is a genuine endogenous peer effect and to understand how study characteristics affect estimated peer effects.

Methods

Selection of studies

The studies selected for inclusion in our MA and MRA are presented in **Table 1**. Fifty-three studies with 95 observations were selected for inclusion in our MA. To ensure interstudy comparability, we selected studies based on the following set of criteria:

1. The studies had to study academic outcomes. More specifically, the dependent variable had to be a measure of either academic achievement/performance (e.g., test scores, grade point average [GPA]) or academic attainment (e.g., dropout). These academic outputs are of particular interest to parental homebuyers who usually take into account school quality in their locational decisions.[6] Excluded were studies of peer effects on such outcomes as weight gain or obesity (Cohen-Cole & Fletcher, 2008; Trogdon, Nonnemaker, & Pais, 2008), smoking or drinking (Argys & Rees, 2008; Gaviria & Raphael, 2001), sexual activity (Duncan, Boisjoly, Kremer, Levy, & Eccles, 2005; Evans, Oates, & Schwab, 1992), labor market outcomes (Marmaros & Sacerdote, 2002), consumption of recreational activities (Bramoullé, Djebbari, & Fortin, 2009), or academic cheating (Carrell, Malmstrom, & West, 2008).
2. Because the focus of the present study is on endogenous peer effects, the selected studies had to specify an estimate of endogenous peer effects. Studies that explore only exogenous peer effects were excluded. Many of these studies were analyzed in van Ewijk and Sleegers (2010a, 2010b). Also, endogenous effects require peer achievement or attainment to be measured similarly to that for the dependent variable. For instance, if the dependent variable is student Grade 3 reading test scores, the endogenous peer effect variable could be current (or lagged) average peer-group third-grade reading test scores. As discussed in greater detail in the previous section, endogenous effects from peers are simultaneous in nature, meaning that a student's achievement is influenced by peer achievement and vice versa. We, therefore, did not include studies on the effect of academically related but nonendogenous peer behavior (e.g., peer school absences, as in Gottfried [2011], on student test scores).
3. The studies had to be on K−12 education. Excluded were studies on peer effects in higher education including Sacerdote (2001) and Zimmerman (2003).
4. Peer groups had to be within a school setting. Studies on the role of neighborhood peer groups on school outcomes (e.g., Ainsworth (2002), Mayer and Jencks (1989), were not included.
5. The studies had to be original academic research with regression analysis. We therefore exclude literature reviews (Wilkinson, Hattie, Parr, & Townsend, 2000), theoretical articles (Ryan, 2000), or policy briefs. We did include unpublished studies and doctoral dissertations if they met the other criteria. This strategy would reduce potential biases introduced by any nonrandom selection of studies (Stanley, 2001).
6. The studies had to be published after 1980 and written in English.

Selection of the studies for this study began with searches of the EconLit, Wilson Social Sciences Full Text, Education Information Resources Center (ERIC), Scopus, American Psychology Association (APA) PsycNET, Proquest Dissertations and Theses, and CSA Sociological Abstracts databases using the keywords *peer effects, peer influences, peer characteristics, school effects, school influences, school characteristics, classroom effects, classroom influences*, and *classroom characteristics*. The second stage involved identifying additional studies for inclusion using citations from the initial searches. This snowball search continued for another iteration before failing to produce additional new studies. As a final step, we searched for these key search terms in Google Scholar to uncover additional studies, particularly studies that had not been published yet.

These searches uncovered over 200 potential studies. Approximately 140 did not fit our criteria and were excluded from the analysis. Another seven did not have information that would allow us to calculate effect sizes. These studies were also excluded. Our final sample included 53 studies of endogenous peer effects. The authors of this study were separately responsible for coding of all of the studies; no research assistant was used. The kappa statistic, a measure of the degree of agreement reliability of ratings, was equal to 0.92, suggesting a high degree of agreement between the coders (Sim & Wright, 2005). We examined the discrepancies carefully, and mutually agreed on the final dataset.

In some cases, we included more than one estimate per study. We included multiple estimates when the estimates differed according to the moderator (or metaindependent) variables (to be discussed subsequently). These additional estimates were necessary to help us uncover how differences in study design affected the estimate of endogenous peer effects. For instance, a type of moderator variable is the measure of peer academic achievement. Academic achievement can be measured in terms of mathematics and reading test scores. Estimates for both of these measures reported by a study, as in Burke and Sass (2013), were included in our metaregression.

We did not include all the estimates from a study as studies with multiple estimates often present benchmark results that are usually biased, or subsample estimates (e.g., peer effects estimated separately for third-, fourth-, and fifth-grade students). Including these estimates would unduly influence MAs (Havránek, 2010; Krueger, 2003). There is no standard rule regarding whether to include a single estimate or multiple estimates in a metaregression (Melo, Graham, & Noland, 2009).

Following the practice advised in Stanley (2001), and adopted in Nelson (2006), Havránek (2010), and Viscusi and Aldy (2003), we only coded the best estimate for that moderator as argued or preferred by authors, as in the case of Imberman, Kugler, and Sacerdote (2012), and Zabel (2008). In cases of no preference expressed, we chose the best fit or most robust result discussed in the abstract or conclusion as in the cases of Hanushek et al. (2003) and Lavy, Silva, & Weinhardt (2012). As discussed above, we often coded more than one best estimate to increase variation in the dataset of our moderator variables. None of the studies used logistic regression and there was no need to transform different measures of effect.

Meta-analysis

This study relies on MA to estimate and test endogenous peer effects on student achievement. MA, according to Hunt (1997), is how science takes stock. MA relies on the calculation of effect sizes, i.e., measures of effect that can be compared between and within studies. Each of these estimates also has an associated variance. Together the effect size and its variance can be used to compute a weighted mean of effect sizes or overall effect size (Borenstein, Hedges, Higgins, & Rothstein, 2009).

While effect sizes are the typical dependent variable in any meta-analytic study, they can be measured in various ways.[7] We used partial correlations as our measure of effect size. Partial correlations measure the strength of the effect of endogenous peer achievement on student outcomes, controlling for all other covariates. As in Greene (2011, p. 37), the partial correlation coefficient, r, can be calculated using Equation 2:

$$r = \frac{t}{\sqrt{t^2 - df}} \tag{2}$$

where t represents the t statistic of the regression coefficient of peer achievement (i.e., β in Equation 1), and df represents degrees of freedom, which is equal to [sample size − (number of variables + 1)]. $\sqrt{t^2 - df}$ represents the standard error of the partial correlation. Hence, dividing t by $\sqrt{t^2 - df}$ standardizes the t, so that it can be compared across different studies with different units.

As studies rarely report the degree of freedom, we, therefore, replaced it with sample size. Because the calculation of r is not affected considerably by imprecise values of df, the values of r obtained with sample size in Equation 3 are almost identical to those with df.[8] According to Table 2, the average partial correlation across selected estimates is 0.03, which suggests a positive and significant peer effect. There is, however, considerable variation in this variable ranging from −0.095 to 0.3.

As an indicator of *statistical* effects, the partial correlation coefficient does not represent the marginal size of an effect (Stanley & Doucouliagos, 2012), we chose partial correlations as the dependent variable for two major reasons. First, it is a unitless measure, enabling us to compare estimates from studies with different measurement units of quantitative impact.[9] Let us take, for example, measures of peer academic achievement used to estimate endogenous peer effects. As shown in column 2 of Table 1 and discussed later in the paper, studies utilize different measures of academic achievement. These measures may be test scores (mathematics, reading, or combined scores), GPA, or rates of educational attainment. Even among studies with the same achievement measure (say, mathematics scores), these scores collected in different American states and countries that have different testing and grading standards are incomparable and thus produce incomparable estimates. Second, partial correlations have been used as effect size measures in studies including Alptekin and Levine (2012), Askarov and Doucouliagos (2013), Doucouliagos and Paldam (2011, 2013), Efendic, Pugh, and Adnett (2011), Haile and Pugh (2013), and Mekasha and Tarp (2013).

Table 2. Descriptive statistics.

Variable	*M*	*SD*	Min.	Max.
Dependent variable				
Partial correlation (*r*)	0.030	0.052	−0.095	0.300
Independent variables				
Standard error of the partial correlation (S)	0.012	0.012	0.001	0.058
Educational outcomes				
Mathematics test score	0.337	0.475	0	1
Reading test score	0.295	0.458	0	1
Combined test score	0.158	0.367	0	1
Other	0.211	0.410	0	1
Reflection problem solutions				
Instrumental variables	0.337	0.475	0	1
Lagged peer achievement	0.484	0.502	0	1
Reduced form	0.179	0.385	0	1
Peer groups				
School	0.189	0.394	0	1
Grade	0.189	0.394	0	1
Classroom	0.526	0.502	0	1
Friends	0.095	0.294	0	1
Data types				
Time-series or panel data? (1 = yes, 0 = no)	0.642	0.482	0	1
Whole-population data? (1 = yes, 0 = no)	0.758	0.431	0	1
Industrialized economies? (1 = yes, 0 = no)	0.768	0.424	0	1
Published? (1 = yes, 0 = no)	0.579	0.496	0	1
Publication year	2007.9	5.81	1983	2014

Note. For descriptive purposes, we present independent variables either as dummy variables or uncentered variables.

We also calculate an overall effect size by computing a weighted mean (M) where the weight assigned to each estimate (w) is the inverse of that estimate's variance.[10] More specifically, the numerator for M is equal to the sum of the product of our measure of effect size r for each estimate i with its associated weight w, whereas the denominator is the sum of all the weights. In notation terms, M is represented by Equation 3:

$$M = \frac{\sum_{i=1}^{k} W_i r_i}{\sum_{i=1}^{k} W_i}, \tag{3}$$

where k is the number of studies. We estimated a random-effects MA. The fixed-effects MA requires the strong assumption that the genuine effect is the same in all studies. This is especially the case in this study with its broad range of dependent variables and other differences in study design. In addition, a test of heterogeneity indicated that the findings are significantly heterogeneous ($Q = 1{,}108$ on 94 degrees of freedom, $p = .00$), providing additional support for the random effects model.[11]

Metaregression model

To examine the roles played by study design and other moderator variables in endogenous peer effects, we also estimated a MRA. Metaregression, as the name would suggest, is a form of MA. MRA is the most common meta-analytic technique in economics (Longhi, Nijkamp, & Poot, 2010), and has been used in fields as diverse as psychology (Tak et al., 2011), health (Lawlor & Hopker, 2001), education (van Ewijk & Sleegers, 2010a, 2010b; Van der Sluis, Van Praag, & Vijverberg, 2008), public finance (Ballal & Rubenstein, 2009; Yeung, 2009), and economics (Card & Krueger, 1995), to quantitatively review and summarize the literature on a certain topic. MRA is particularly appropriate for studies utilizing regression methods (Stanley, 2001), as the majority of studies on endogenous peer effects does. Indeed, it can be described as a multiple regression of multiple regressions (Yeung, 2009).

A metaregression model was formally presented in a seminal paper by Stanley and Jarrell (1989), and later further developed and adopted by various researchers. Following the MRA literature such as Stanley and Doucouliagos (2012), and Askarov and Doucouliagos (2013), our MRA regression model is represented by Equation 4. Specifically, study j's estimate of endogenous peer effects, r_j, is a function of the true effect size (r^*), moderator variables indicating various study characteristics (Z), and the standard error of $r_j(S)$:

$$r_j = r^* + \sum_{k=1}^{K} \delta_k Z_j + \beta_0 S_j + \varepsilon_j, \tag{4}$$

where ε is an error term. More intuitively, Equation 4 represents factors that influence estimated endogenous peer effects. Specifically, r^* and r_j are positively correlated. Larger estimated effects, r_j, are more likely to be obtained as a result of the larger true effect, r^*. Other moderator variables that will be discussed in greater detail later include how endogenous peer achievement and peer groups are measured or defined, and measures of research or data quality. As in our MA, our MRA is a random-effects estimator.[12]

The standard error of the effect size, S, is included in Equation 4 to control for publication selection or bias, if any. Publication selection refers to the practice of publishing research papers based on the statistical significance of their findings. Researchers and peer-reviewed journal editors favor statistically significant results. Therefore, authors of small-sample studies have temptations to carry out specification and data searches until they find statistically significant results (Stanley, Doucouliagos, & Jarrell, 2008). These small-sample studies tend to produce large and significant effect-size results due to the large standard errors associated with small-sample data. In contrast, it is easier for authors of large-sample studies to find statistically significant estimates of smaller effect size. To put it simply, effect size and standard errors are likely to be strongly correlated. Including S in Equation 4, hence, serves as a test for publication selection and a means of controlling for it. Various studies have used this method to address the problem of selection bias including Doucouliagos and Paldam (2006), Stanley (2008), and Stanley, Doucouliagos, and Jarrell (2008).

Attention needs to be paid to the interpretation of the coefficients in the estimating Equation 4, especially the coefficient β_0, and the intercept, r^*. The conventional t test of β_0 provides a test for publication selection bias, and its estimate indicates the direction of this bias, if any. Failure to reject the null hypothesis (i.e., $\beta_0 = 0$) suggests that publication selection is not likely. The intercept term is also of special interest. If it is significantly different from 0, it suggests that there is an authentic endogenous peer effect beyond (or corrected for) publication selection. Finally, the coefficients, δ_k, of the moderator variables (Z_k) are also of interest and indicate how endogenous peer effects vary with different study designs and characteristics. The following subsection discusses in greater detail which moderator variables were included.

We estimated Equation 4 with weighted least squares. As suggested by Stanley and Doucouliagos (2012), we weighted our regression with precision squared ($1/S^2$) to address the problem of heteroskedasticity.

Moderator variables

Moderator variables, Z_k, are study characteristics that can affect the estimate of an endogenous peer effect. Moderator variables are effect coded and specify different measures of endogenous peer effects, identification strategies to address the reflection problem, definitions of a peer group, data types, in particular whether a study used a time-series dataset, and whether its measure of peer group was derived through sampling, whether a study was set in a developing or transitional economy, and whether it is published.[13] Table 1 documents how selected individual studies differed according to these moderator variables and Table 2 provides their summary descriptive statistics.

Educational outcomes

Studies used different measures of academic achievement for the dependent variable and thus the endogenous peer effect variable. They include mathematics test scores, reading test scores, combined (multiple subjects) test scores, GPA, and educational attainment (usually measured by dropout rates). The most and least common types of achievement peer effect, according to Table 2, were mathematics and combined test scores (about 34% and 16%, respectively, of the total 95 observations). In our metaregression estimations, combined test scores are effect-coded as the omitted category. As reviewed earlier, these effect codes are specified in Equation 4 because endogenous peer effects may vary significantly across different peer achievement measures.

Reflection problem solutions

Studies adopted two major empirical solutions to the reflection problem: lagged peer achievement (46 observations, approximately 48%) or instrumental variables (32 observations). Other studies (with 17 observations) follow neither of these identification strategies. These studies include current peer achievement measures but do not treat them as endogenous. We call this strategy the reduced form.[14] While a few studies in this category did not acknowledge the problem of simultaneity (Cook et al., 2007; Ryabov, 2011), some studies in this category noticed potential bias from this problem but argued against treating the current peer achievement variable as endogenous.

For instance, by using nonreciprocating friends' achievement, Mora and Oreopoulos (2011) argued that simultaneity was nonexistent because peer effects flow in one direction from nonreciprocating friend to friend. Burke and Sass (2013) assumed that peers influence each other's test scores only through their fixed ability or aptitude. They argue that as a result, simultaneity is no longer a problem with their use of peer fixed effects. These estimation methods rely on different assumptions, and may therefore lead to very different estimates and are worth controlling for.

Peer groups

The choice of peer group also may play a role in the estimation of endogenous peer effects. The majority of estimates in the sample were interested in classroom peer effects (50 observations), followed by school peer effects (18 observations), grade peer effects (18 observations, the omitted category in our MRA) and finally interactions among friends (nine observations). Halliday and Kwak (2012) found that definitions of peer groups impact estimations of peer effects. Their study argues that school-grade cohorts as peer groups may lead to underestimates of peer effects relative to friends.

Data types

Different types of data may produce differential estimates of endogenous peer effects, all else being equal. We have two effect-coded variables in this category. The first variable was equal to 1 if a study employed time-series data while the omitted group included studies with cross-sectional data. As discussed previously, time-series or panel data allow a researcher to control for various types of fixed effects (including student fixed effects) that are able to control for time-invariant omitted factors correlated with endogenous peer effects.

The second variable was coded 1 for studies employing data for the whole population of the student body of interest. This variable was equal to −1 for sample survey data. Datasets in Burke and Sass (2013) and Betts and Zau (2004), for example, consisted of all students in the state of Florida and in San Diego Unified School District, respectively, whereas only a portion of students were sampled in the National Education Longitudinal Study as in Gaviria and Raphael (2001), or the National Longitudinal Study of Adolescent Health Survey as in Halliday and Kwak (2012) and Ryabov (2011). This variable was specified based on the finding by Micklewright, Schnepf, and Silva (2012) that the estimated effects from a sample of peers in a survey dataset for England, namely the Program for International Student Assessment, are biased downward by about one third relative to complete administrative data with all students.

Industrialized economy

We included an effect-coded variable representing if a study explored endogenous peer effects in a industrialized economy (equal to 1 if yes, and −1 otherwise). We controlled for this variable to see if the relationship between peers and student outcomes differed between industrialized countries and developing or transition countries. Over one fifth of the estimates (22 observations) came from studies on developing or transition economies.

Publication status

We included an effect-coded variable indicating whether a study was published in a peer-reviewed journal or not (= −1 for unpublished studies). Research published in a peer-reviewed journal is usually considered to be of higher quality as it must go through a review process. In addition, this variable helped control for publication selection (Égert & Halpern, 2006; Zelmer, 2003), given the tendency for journals to publish significant results. It is important to note that this variable is not sufficient to completely control for publication selection. Publication selection bias still needed to be tested formally with β_0 in Equation 4 because this bias exists even among unpublished papers (Stanley & Doucouliagos, 2012).

Publication year

Finally, topics become hot at various times. This may mean that articles that study this topic may be more likely to be published in some years and not in others. In addition, there may be temporal changes in the way peer effects are modeled and analyzed in the literature. For these reasons, we control for publication year. The average publication date in the sample was approximately 2008.

We centered continuous moderator variables (publication year and standard error, *S*) using their mean values for ease of interpretation. Centered publication year and *S*, together with the remaining effect-coded variables, imply that the constant term in our MRA equals the grand mean of partial correlations at the mean standard error and in the mean publication year.

Results

Tests of publication bias

We begin by presenting tests for publication bias. The mean overall effect should reflect the true population effect in which we are interested. If however, the sample of estimates we use to conduct our MA is biased, the results of our MA will also be biased. As a result, it is important to examine the level of publication bias in our data.

Our first test of publication bias is the funnel plot, which we present in **Figure 1**. This chart plots the partial correlations against the inverse of their standard errors (or precision). In the absence of a publication bias, estimates should be

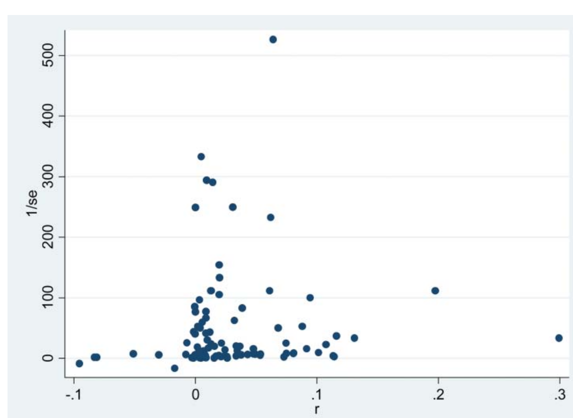

Figure 1. Funnel plot.

distributed symmetrically around the mean effect size, since the sampling error is random (Borenstein et al., 2009). The interpretation of the funnel plot is entirely subjective, and in our case, it is not clear if the data exhibits publication bias. Most of the smaller studies (which have larger standard errors) are compressed at the bottom. There may be a clustering of data around 0, but it is not clear.

As the results of the funnel plot are both subjective and unclear, we also conduct a trim-and-fill analysis of publication bias. According to Borenstein et al. (2009), this analysis uses an iterative procedure to remove the most extreme small studies from the positive side of the funnel plot, re-computing the effect size at each iteration until the funnel plot is symmetric about the (new) effect size. In theory, this will yield an unbiased estimate of the effect size (Duval & Tweedie, 2000a, 2000b).[15] Our trim-and-fill analysis resulted in no studies needing to be filled to create a symmetric funnel plot. This is strong evidence that publication bias is not driving our results.

Meta-analysis results

As reported in **Table 2**, the unweighted average partial correlation is 0.03. The weighted average partial correlation with M derived using Equation 2 was slightly smaller and equal to 0.02. Both values are consistent with the funnel plot suggesting a clustering of estimates near 0. We used the Z value of the weighted mean to test the null hypothesis that the mean effect is zero. The Z-value we obtained was 80.2 ($p = .00$). Given the lack of evidence of publication bias, and the statistically significant results of our MA, we can conclude at this stage of analysis that there is a genuine empirical endogenous effect of peers on student outcomes. Better peers result in better outcomes. Doucouliagos (2011) provided some guidelines for interpreting partial correlations: “A partial correlation that is less than ±0.07 can be regarded as small …, *even if it is statistically significant* (p. 10, emphasis in the original). Using this guideline, the true empirical endogenous peer effect can be considered small.

Metaregression results

Table 3 presents the results of our MRA with partial correlations as the dependent variable with unweighted (column 1) and weighted (column 2) effect coding strategies.[16] The results are similar in both columns. Weighted effect coding is more appropriate because sample sizes differ across moderator groups and our sample consists of most, if not all, the studies in the population. We therefore focus our discussion on column 2.

One of the key variables in our MRA is the standard error of the partial correlation (S). This variable, a measure of publication selection, is not statistically different from zero, suggesting that there is no publication selection bias in the endogenous peer effects literature. This finding is consistent with the funnel plot and trim-and-fill analysis of the data presented in the previous section. As discussed previously, the constant term has special meaning in our MRA with the inclusion of the standard error. It captures the true or genuine empirical effect. **Table 3** shows that the constant is highly significant at the .01 level. This finding suggests that there is a genuine empirical endogenous peer effect on student outcomes, even controlling for differences in study quality and characteristics. Also, the regression results presented in column 2 of **Table 3** are set up so that the constant is equal to the weighted grand mean of the dependent variable at the mean standard error in the mean year. The constant is estimated at 0.03, which is slightly higher than the overall or grand mean we found in our MA (both weighted and unweighted) but is still considered small in magnitude by Doucouliagos (2011).

We are also interested in whether the genuine empirical effect varies with different study designs. The answer appears to be yes. According to **Table 3**, five of the moderator variables are significant. First, the educational outcome of reading test score makes it less likely to find a positive endogenous peer effect while “other” makes it more likely to find a positive

Table 3. Random-effects metaregression results.

Variable	Unweighted effect-coded moderator variables (1)	Weighted effect-coded moderator variables (2)
Constant	0.036**	0.030**
	(3.89)	(5.90)
Standard error of the partial correlation	−0.99	−0.97
	(1.62)	(1.62)
Educational outcomes (combined test score is the omitted category)		
Mathematics test score	−0.0075	−0.0049
	(0.86)	(0.65)
Reading test score	−0.018†	−0.016†
	(1.97)	(1.85)
Other	0.023†	0.026†
	(1.84)	(1.95)
Reflection solution (reduced form is the omitted category)		
Lagged peer variable	0.014	0.010
	(1.66)	(1.63)
Instrumental variable	−0.0039	−0.0075
	(0.45)	(0.90)
Peer group (grade is the omitted category)		
Class	−0.010	−0.0028
	(1.02)	(0.49)
School	−0.0063	0.0010
	(0.59)	(0.09)
Friends	0.041*	0.048*
	(2.30)	(2.22)
Data type		
Time−series or panel data? (1 = yes, −1 = no)	−0.0031	−0.0022
	(0.39)	(0.39)
Whole−population data? (1 = yes, −1 = no)	0.0085	0.0041
	(1.16)	(1.16)
Industrialized economy? (1 = yes, −1 = no)	−0.0097	−0.0045
	(1.32)	(1.32)
Published? (1 = yes, −1 = no)	0.011†	0.0095†
	(1.80)	(1.80)
Publication year	−0.0031**	−0.0031**
	(2.75)	(2.75)
Adjusted R^2	12.22%	12.22%

Note. There are 95 observations. The dependent variable is the partial correlation coefficient. The absolute values of t statistics are in parentheses. While the standard error of partial correlations and publication years are centered, all of the other moderator variables are effect-coded with unweighting (column 1) or weighting (column 2). The regression is weighted by the inverse of the variance of the correlation.
† $p < .10$.
* $p < .05$.
** $p < .01$.

endogenous peer effect. While peer groups at the class and school levels do not have a significant moderating effect on endogenous peer effects, we find that studies with friends as peers are more likely to find a positive endogenous peer effect. We also find that articles published in peer-reviewed journals are more likely to find a positive endogenous peer effect. Finally, we find that each additional publication year is associated with a 0.003 decrease in the partial correlation. This result, significant at the .01 level, suggests more recent publications are less likely to find a positive endogenous peer effect. Somewhat surprisingly, we do not find the solution to the reflection problem has any significant statistical impact on the partial correlation of endogenous peer effects.

Conclusion

The growth in large administrative datasets that are available for research use has resulted in a sizable rise in the research on endogenous peer effects in educational settings. Endogenous peer effects occur when a peer's behavior or outcome influences a student's behavior or outcome. This research has developed along multiple lines. Scholars have utilized various strategies to resolve the difficult reflection problem, and have examined the effects of various peer groups. This is the first study, to our knowledge, to synthesize the research on these endogenous peer effects in a quantitative and analytical way. In this paper, we use meta-analytic procedures to summarize the results of 53 studies from 1980 to 2014 on the subject of endogenous peer effects in education.

Do children actually perform better surrounded by better peers? Our finding appears to be an unequivocal yes. We find a highly significant endogenous peer effect. This positive and significant association is, however, small in magnitude. In addition, although there is evidence for positive endogenous peer effects, an unanswered question in the field is whether the existence of endogenous peer effects that have been found in the literature is the result of publication selection bias. Various tests we conducted in this study suggest that our result is a genuine empirical effect. Students are influenced by their peers. As their peers improve, they also improve. When peers regress, they also regress.

Through the use of MRA, we also find that endogenous peer effects are study and setting dependent. We find that the type of educational outcome, the choice of peer group, and the publication status of a study have moderating effects on endogenous peer effects. Additionally, peer effects appear to have a temporal component as more recent studies have been less likely to find positive effects.

Our study has several potential policy implications. The answer to the question as to whether children actually perform better surrounded by better peers serves as ammunition for both proponents and opponents of social policies such as tracking, busing, and school choice, which are based, at least partially, on a theory of peer composition effects. The two sides of these policy debates often pick and choose empirical findings (yes or no) with regard to this question from a whole host of studies to support their positions. Our study helps provide a definitive yes to this question. We also find that positive endogenous peer effects are more likely at the lowest level of peer group composition, when the peer group consists of friends. Unfortunately, because schools and school administrators cannot assign friends, they cannot take advantage of this peer effect mechanism. Additionally, it is important to note that endogenous peer effects, if any, are small in nature. Proponents of policies based upon endogenous peer effects should not expect a panacea for their education system's ills.

This study was not without limitations. Ideally, we would like to have used elasticities as our measure of effect. However, elasticities that can be used to estimate the mean effect size are neither readily available nor possible to calculate. Almost all of the selected studies do not use a log-log functional form in which the coefficient is the elasticity. Rather, they use a level-level functional form studies and do not report the means of dependent and independent variables needed to calculate elasticities. In addition, we did not examine how the composition of social groups not in a school context (e.g., neighborhoods and families, affects student performance); this is an area ripe for future meta-analyses.

Notes

1. Depending on the context, peer effects may be called in different terms, such as peer influences (Steinberg, Fletcher, & Darling, 1994), neighborhood effects (Kling, Liebman, & Katz, 2007), or herd behavior (Banerjee, 1992).
2. A social multiplier can be a desirable aspect for a social policy or intervention. It occurs in the presence of positive spillovers or strategic complementaries (Glaeser, Scheinkman, & Sacerdote, 2003), thereby creating aggregate social effects greater than individual effects (Becker & Murphy, 2000).
3. Technically, tracking and ability grouping are different in terms of scale and permanence (Slavin, 1987). Ability grouping is smaller and more informal.
4. Group *g* can be defined at the grade, class, or school level or as friends that students interact with. Also, researchers can include multiple group levels in their estimation depending on their data structure. For instance, Hoxby and Weingarth (2006) included both grade and class groups.
5. If a peer group is defined at the school level, an example of unobserved common factors within a peer group correlated with peer achievement is average teacher and principal quality. These effects are what Manski (1993) called correlated effects.
6. Given this parental demand, houses in neighborhoods associated with better-quality schools are, all else being equal, higher-priced than those in worse-quality school neighborhoods. See Nguyen-Hoang and Yinger (2011) for a critical review of studies on the capitalization of school quality into house prices.
7. For instance, meta-analytic studies in medicine can use a wide range of effect sizes, namely Cohen's *d*, the odds ratio, Glass's *g*, log-odds, and log-risk (Stanley & Doucouliagos, 2012).
8. $r = \frac{t}{\sqrt{t^2 - df}}$s the formula for th. e partial correlation. We can further break *df* into *n-k-1* where *n* is the of observations and *k* is the number of independent variables in the model. This gives us $r = \frac{t}{\sqrt{t^2 - (n-k-1)}}$ As *n* approaches infinity, *n-k* converges to *n*. Even at large *n*, the difference between *n* and *n-k-1* is minimal. For instance, suppose a *t* statistic is 2.1, then $r = .0035$ for *df* of 363,901. The same *r* (after rounding) can be obtained for different values of *df* ranging between 349,927 and 370,505. The practically insignificant difference between degrees of freedom and sample size in the calculation of partial correlations is already noted the literature, such as in Stanley and Doucouliagos (2012).
9. Another unitless measure of effect size is the *t*-statistic. However, this measure, if used, is still to be interpreted in terms of partial correlations (Efendic, Pugh, & Adnett, 2011). Also, the use of *t* statistic as an

effect size in MRA has given way to the use of partial correlations (Haile & Pugh, 2013).

10. See page 73 of Borenstein et al. (2009) for a detailed exposition of how *w* is calculated.
11. We used the Stata user-written program METATRIM by Steichen (2010) for this test. Also, see Borenstein et al. (2009) for further discussion on this heterogeneity test.
12. We used the METAREG Stata command written by Harbord and Steichen (2009) for this purpose.
13. We could use dummy coding (0,1) for these moderator variables. However, the dummy coding could subject the constant term that captures the genuine peer effect to potential extrapolation bias. This is the case when there does not exist a study that possesses all of the omitted (coded 0) characteristics. The interpretation of effect coding (−1,1) is, however, different from that of dummy coding. With effect coding, we compare how a group is doing relative to the grand mean, regardless of which group is omitted. We thank a reviewer for this idea.
14. The term *reduced form* may be different in different peer effect contexts. For instance, reduced form in Ammermueller and Pischke (2009) means an estimation equation in which exogenous and endogenous effects are combined, rather than separated.
15. We used the Stata user-written program METATRIM written by Steichen (2010) for this purpose.
16. Weighting is done for omitted categories. Specifically, an omitted category is weighted by the product of −1 and a ratio of the sample size of the group coded 1 and the sample size of the omitted category. Weighted effect-coded variables have a mean of 0. See page 123 of Aguinis (2004) for detailed exposition on weighted effect coding procedures. However, for ease of exposition, we indicate omitted groups as coded −1, regardless of whether they are (un)weighted.

Acknowledgments

The authors would like to thank T. D. Stanley, Bob Bifulco, and members of the Association for Education Finance and Policy for their helpful thoughts. Special thanks go to an anonymous reviewer and a consulting editor.

References

Aguinis, H. (2004). *Regression analysis for categorical moderators*. New York, NY: Guilford Press.

Ainsworth, J. W. (2002). Why does it take a village? The mediation of neighborhood effects on educational achievement. *Social Forces*, *81*, 117−152.

Alptekin, A., & Levine, P. (2012). Military expenditure and economic growth: A meta-analysis. *European Journal of Political Economy*, *28*, 636−650.

Altonji, J. G., Huang, C.-I., & Taber, C. R. (2010). *Estimating the cream skimming effect of school choice. NBER Working Paper Series* (Working paper). Cambridge, MA: National Bureau of Economic Research.

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*, *27*, 315−348.

Argys, L. M., & Rees, D. I. (2008). Searching for peer group effects: A test of the contagion hypothesis. *Review of Economics and Statistics*, *90*, 442−458.

Asadullah, M. N., & Chaudhury, N. (2008). *Social interactions and student achievement in a developing country: An instrumental variables approach* (Working paper). Washington, DC: World Bank.

Askarov, Z., & Doucouliagos, H. (2013). Does aid improve democracy and governance? A meta-regression analysis. *Public Choice*, *157*, 601−628.

Atkinson, A., Burgess, S., Gregg, P., Propper, C., & Proud, S. (2008). *The impact of classroom peer groups on pupil GCSE results* (CMPO Working Paper Series). Bristol, UK: Centre for Market and Public Organisation.

Ballal, S., & Rubenstein, R. (2009). The effect of tax and expenditure limitations on public education resources: A metaregression analysis. *Public Finance Review*, *37*, 665−685.

Banerjee, A. V. (1992). A Simple Model of Herd Behavior. *Quarterly Journal of Economics*, *107*, 797−817.

Becker, G. S., & Murphy, K. M. (2000). *Social economics: Market behavior in a social environment*. Cambridge, MA: Belknap Press.

Betts, J. R., & Zau, A. (2004). *Peer groups and academic achievement: Panel evidence from administrative data* (Working paper). Retrieved from http://www.iza.org/iza/en/webcontent/events/transatlantic/papers_2004/betts.pdf

Boozer, M. A., & Cacciola, S. E. (2001). *Inside the “black box” of Project STAR: Estimation of peer effects using experimental data* (Working paper). New Haven, CT: Economic Growth Center.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *An introduction to meta-analysis*. New York, NY: Wiley.

Boucher, V., Bramoullé, Y., Djebbari, H., & Fortin, B. (2014). Do peers affect student achievement? Evidence from Canada Using Group Size Variation. *Journal of Applied Econometrics*, *29*, 91−109.

Bradley, S., & Taylor, J. (2008). *Do peers matter? Estimation of peer effects from pupil mobility between schools* (Working paper). Lancaster, UK: Department of Economics, Lancaster University.

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

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

Card, D., & Krueger, A. B. (1995). Time-series minimum-wage studies: a meta-analysis. *American Economic Review*, *85*, 238−243.

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

Carrell, S. E., Malmstrom, F. V., & West, J. E. (2008). Peer effects in academic cheating. *Journal of Human Resources*, *43*, 173−207.

Clark, C., Scafidi, B., & Swinton, J. R. (2011). Do peers influence achievement in high school economics? Evidence from Georgia’s economics end of course test. *Journal of Economic Education*, *42*, 3−18.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *Journal of Human Resources*, *45*, 655−681.

Cohen-Cole, E., & Fletcher, J. M. (2008). Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic. *Journal of Health Economics*, *27*, 1382−1387.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, F., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of Educational Opportunity*. Washington, DC: U.S. Government Printing Office.

Cook, T. D., Deng, Y., & Morgano, E. (2007). Friendship influences during early adolescence: The special role of friends’ grade point average. *Journal of Research on Adolescence*, *17*, 325−356.

Ding, W., & Lehrer, S. F. (2007). Do peers affect student achievement in China’s secondary schools? *Review of Economics and Statistics*, *89*, 300−312.

Doucouliagos, H. (2011). How large is large? Preliminary and relative guidelines for interpreting partial correlations in economics (Working paper). Victoria, Australia: School of Accounting, Economics and Finance, Deakin University. Retrieved from http://www.deakin.edu.au/buslaw/aef/workingpapers/papers/2011_5.pdf

Doucouliagos, H., & Paldam, M. (2006). Aid effectiveness on accumulation: A meta study. *Kyklos*, *59*, 227−254.

Doucouliagos, H., & Paldam, M. (2011). The ineffectiveness of development aid on growth: An update. *European Journal of Political Economy*, *27*, 399−404.

Doucouliagos, H., & Paldam, M. (2013). Explaining development aid allocation by growth. *Journal of Entrepreneurship and Public Policy*, *2*, 21−41.

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*, 1739−1774.

Dumay, X., & Dupriez, V. (2008). Does the school composition effect matter? Evidence from Belgian data. *British Journal of Educational Studies*, *56*, 440−477.

Duncan, G. J., Boisjoly, J., Kremer, M., Levy, D. M., & Eccles, J. (2005). Peer effects in drug use and sex among college students. *Journal of Abnormal Child Psychology*, *33*, 375−385.

Duval, S., & Tweedie, R. (2000a). A nonparametric "Trim and Fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95* (449), 89−98.

Duval, S., & Tweedie, R. (2000b). Trim and Fill: A simple funnel-plot−based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455−463.

Efendic, A., Pugh, G., & Adnett, N. (2011). Institutions and economic performance: A meta-regression analysis. *European Journal of Political Economy*, *27*, 586−599.

Égert, B., & Halpern, L. (2006). Equilibrium exchange rates in Central and Eastern Europe: A meta-regression analysis. *Journal of Banking & Finance*, *30*, 1359−1374.

Erbring, L., & Young, A. A. (1979). Individuals and social structure contextual effects as endogenous feedback. *Sociological Methods & Research*, *7*, 396−430.

Evans, W. N., Oates, W. E., & Schwab, R. M. (1992). Measuring peer group effects: A study of teenage behavior. *Journal of Political Economy*, *100*, 966−991.

Fortner, C. K. (2010). *Within classroom peer effects and tracking: Assessing the impact of classroom peer variation and ability tracking with disaggregated high* school (Doctoral dissertation). Georgia State University, Atlanta.

Fruehwirth, J. C. (2013). Identifying peer achievement spillovers: Implications for desegregation and the achievement gap. *Quantitative Economics*, *4*, 85−124.

Fruehwirth, J. C. (2014). Can achievement peer effect estimates inform policy? A view from inside the black box. *Review of Economics and Statistics*, *96*, 514−523.

Gaviria, A., & Raphael, S. (2001). School-based peer effects and juvenile behavior. *Review of Economics and Statistics*, *83*, 257−268.

Glaeser, E. L., Scheinkman, J. A., & Sacerdote, B. I. (2003). The social multiplier. *Journal of the European Economic Association*, *1*, 345−353.

Gibbons, S., & Telhaj, S. (2008). *Peers and achievement in England's secondary schools* (SERC discussion paper). Retrieved from http://eprints.lse.ac.uk/30854/1/sercdp0001.pdf

Gibbons, S., & Telhaj, S. (2012). *Peer effects: Evidence from secondary school transition in England* (IZA discussion paper). Retrieved from http://ftp.iza.org/dp6455.pdf

Gottfried, M. A. (2010). *On educational outcomes in an urban society*. (Doctoral dissertation). University of Pennsylvania, Philadelphia.

Gottfried, M. A. (2011). Absent peers in elementary years: The negative classroom effects of unexcused absences on standardized testing outcomes. *Teachers College Record*, *113*, 1597−1632.

Graham, B. S. (2008). Identifying social interactions through conditional variance restrictions. *Econometrica*, *76*, 643−660.

Greene, W. H. (2011). *Econometric analysis* (7th ed.). Englewood Cliffs, NJ: Prentice Hall.

Haile, M. G., & Pugh, G. (2013). Does exchange rate volatility discourage international trade? A meta-regression analysis. *The Journal of International Trade & Economic Development*, *22*, 321−350.

Halliday, T. J., & Kwak, S. (2012). What is a peer? The role of network definitions in estimation of endogenous peer effects. *Applied Economics*, *44*, 289−302.

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.

Harbord, R. M., & Steichen, T. J. (2009). METAREG. Boston, MA: Boston College.

Harris, D. N. (2010). How do school peers influence student educational outcomes? Theory and evidence from economics and other social sciences. *Teachers College Record*, *112*, 1163−1197.

Havránek, T. (2010). Rose effect and the Euro: Is the magic gone? *Review of World Economics*, *146*, 241−261.

Henderson, V., Mieszkowski, P., & Sauvageau, Y. (1978). Peer group effects and educational production functions. *Journal of Public Economics*, *10*, 97−106.

Hoxby, C. M., & Weingarth, G. (2006). *Taking race out of the equation: School reassignment and the structure of peer effects* (Working paper). Retrieved from http://www.hks.harvard.edu/inequality/Seminar/Papers/Hoxby06.pdf

Hunt, M. M. (1997). *How science takes stock: The story of meta-analysis*. Thousand Oaks, CA: Sage.

Imberman, S. A., Kugler, A. D., & Sacerdote, B. I. (2012). Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. *American Economic Review*, *102*, 2048−2082.

Jackson, C. K. (2009). *Peer quality or input quality?: Clean evidence from Trinidad and Tobago* (Working paper). Retrieved from http://www.econ.yale.edu/seminars/labor/lap09/jackson-090911.pdf

Kane, T. J., & Staiger, D. O. (2002). Volatility in school test scores: Implications for test-based accountability systems. *Brookings Papers on Education Policy*, 235−283.

Kang, C. (2007). Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea. *Journal of Urban Economics*, *61*, 458−495.

Kiss, D. (2011). *The impact of peer ability and heterogeneity on student achievement: Evidence from a natural experiment* (Working paper). Retrieved from http://econstor.eu/bitstream/10419/44398/1/64516979X.pdf

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, *75*, 83−119.

Kramarz, F., Machin, S., & Ouazad, A. (2010). *Using compulsory mobility to identify heterogeneous school quality and peer effects* (Working paper). London, UK: Centre for Economic Performance.

Krueger, A. B. (2003). Economic considerations and class size. *Economic Journal*, *113*, 34−63.

Lai, F. (2007). *How do classroom peers affect student outcomes?: Evidence from a natural experiment in Beijing's middle schools*. New York, NY: New York University.

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*, 367−414.

Lawlor, D. A., & Hopker, S. W. (2001). The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. *British Medical Journal*, *322* (7289), 1−8.

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

Leiter, J. (1983). Classroom composition and achievement gains. *Sociology of Education*, *56*, 126−132.

Lin, X. (2010). Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics*, *28*, 825−860.

Link, C. R., & Mulligan, J. G. (1991). Classmates' effects on black student achievement in public school classrooms. *Economics of Education Review*, *10*, 297−310.

Liu, X., Patacchini, E., & Zenou, Y. (2013). *Peer effects: Social multiplier or social norms?* (Working paper). Retrieved from http://people.su.se/˜yvze0888/Aggregate_vs_Average_18_02_2013.pdf

Longhi, S., Nijkamp, P., & Poot, J. (2010). Meta-analyses of labour-market impacts of immigration: key conclusions and policy implications. *Environment and Planning C: Government and Policy*, *28*, 819−833.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, *60*, 531−542.

Marmaros, D., & Sacerdote, B. (2002). Peer and social networks in job search. *European Economic Review*, *46* (4−5), 870−879.

Mayer, S. E., & Jencks, C. (1989). Growing up in poor neighborhoods: How much does it matter? *Science*, *243* (4897), 1441−1445.

Mekasha, T. J., & Tarp, F. (2013). Aid and growth: What meta-analysis reveals. *Journal of Development Studies*, *49*, 564−583.

Melo, P. C., Graham, D. J., & Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics*, *39*, 332−342.

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*, 1136−1142.

Mora, T., & Oreopoulos, P. (2011). Peer effects on high school aspirations: Evidence from a sample of close and not-so-close friends. *Economics of Education Review*, *30*, 575−581.

Mounts, N. S., & Steinberg, L. (1995). An ecological analysis of peer influence on adolescent grade point average and drug use. *Developmental Psychology*, *31*, 915−922.

Murray, M. P. (2006). Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives*, *20*, 111−132.

Nguyen-Hoang, P., & Yinger, J. (2011). The capitalization of school quality into house values: A review. *Journal of Housing Economics*, *20*, 30−48.

Nelson, J. P. (2006). Cigarette advertising regulation: A meta-analysis. *International Review of Law and Economics*, *26*, 195−226.

Nores, M. (2006). Gaps in early childhood attainment: Learning the A, B, Cs: Analysis of the 1998 cohort (Doctoral dissertation). Columbia University, New York, NY.

Patacchini, E., Rainone, E., & Zenous, Y. (2013). *Strong versus weak ties in education* (Working paper). Retrieved from http://people.su.se/˜yvze0888/Patacchini_Rainone_Zenou_07_April_2013.pdf

Rangaprasad, V. (2004). *Essays on the determinants of school quality and student achievement* (Doctoral dissertation). Southern Methodist University, Dallas, TX.

Ryabov, I. (2011). Adolescent academic outcomes in school context: Network effects reexamined. *Journal of Adolescence*, *34*, 915−927.

Ryan, A. M. (2000). Peer groups as a context for the socialization of adolescents' motivation, engagement, and achievement in school. *Educational Psychologist*, *35*, 101−111.

Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics*, *116*, 681−704.

Silva, E. (2007). *On the clock: Rethinking the ways schools use time*. Washington, DC: Education Sector.

Sim, J., & Wright, C. C. (2005). The Kappa Statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy*, *85*, 257−268.

Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, *57*, 293−336.

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

Stanley, T. D. (2001). Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of Economic Perspectives*, *15*, 131−150.

Stanley, T. D. (2008). Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics and Statistics*, *70*, 103−127.

Stanley, T. D., & Doucouliagos, H. (2012). *Meta-regression analysis in economics and business*. London, England: Routledge.

Stanley, T. D., Doucouliagos, C., & Jarrell, S. B. (2008). Meta-regression analysis as the socio-economics of economics research. *Journal of Socio-Economics*, *37*, 276−292.

Stanley, T. D., & Jarrell, S. B. (1989). Meta-regression analysis: A quantitative method of literature surveys. *Journal of Economic Surveys*, *3*, 161−170.

Steichen, T. (2010). METATRIM. Boston, MA: Boston College Department of Economics.

Steinberg, L., Fletcher, A., & Darling, N. (1994). Parental monitoring and peer influences on adolescent substance use. *Pediatrics*, *93*, 1060−1064.

Sund, K. (2009). Estimating peer effects in Swedish high school using school, teacher, and student fixed effects. *Economics of Education Review*, *28*, 329−336.

Tak, L. M., Cleare, A. J., Ormel, J., Manoharan, A., Kok, I. C., Wessely, S., & Rosmalen, J. G. M. (2011). Meta-analysis and meta-regression of hypothalamic-pituitary-adrenal axis activity in functional somatic disorders. *Biological Psychology*, *87*, 183−194.

Thomas, W., & Webber, D. J. (2001). "Because my friends are": The impact of peer groups on the intention to stay on at sixteen. *Research in Post-Compulsory Education*, *6*, 339−354.

Trogdon, J. G., Nonnemaker, J., & Pais, J. (2008). Peer effects in adolescent overweight. *Journal of Health Economics*, *27*, 1388−1399.

Van der Sluis, J., Van Praag, M., & Vijverberg, W. (2008). Education and entrepreneurship selection and performance: A review of the empirical literature. *Journal of Economic Surveys*, *22*, 795−841.

van Ewijk, R., & Sleegers, P. (2010a). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, *5*, 134−150.

van Ewijk, R., & Sleegers, P. (2010b). Peer ethnicity and achievement: a meta-analysis into the compositional effect. *School Effectiveness and School Improvement*, *21*, 237−265.

Vigdor, J., & Nechyba, T. (2007). Peer effects in North Carolina public schools. In L. Woessmann & P. E. Peterson (Eds.), *Schools and the equal opportunity problem* (pp. 73−101). Cambridge, MA: MIT Press.

Vigdor, J., & Nechyba, T. (2008). *Causal inference when assignment may have been random: Peer effects in North Carolina elementary schools* (Working paper). Durham, NC: Duke University. Retrieved from https://www.aeaweb.org/assa/2009/retrieve.php?pdfid=369

Viscusi, W. K., & Aldy, J. E. (2003). The value of a statistical life: A critical review of market estimates throughout the world. *Journal of Risk and Uncertainty*, *27*, 5−76.

Wang, L. C. (2010). *Peer effects in the classroom: Evidence from a natural experiment in Malaysia* (Working paper). San Diego, CA: Department of Economics, University of California. Retrieved from http://economics.missouri.edu/seminars/files/2010/100129.pdf

Westchester Institute for Human Services Research. (2002). *Ability grouping*. White Plains, NY: Westchester Institute for Human Services Research.

Wilkinson, I. A. G., Hattie, J. A., Parr, J. M., & Townsend, M. A. R. (2000). *Influence of peer effects on learning outcomes: A review of the literature*. Auckland UniServices Limited. Retrieved from http://www.eric.ed.gov/PDFS/ED478708.pdf

Yee, V. (2013, June 9). Grouping students by ability regains favor in classroom. *New York Times*. Retrieved from http://www.nytimes.com/2013/06/10/education/grouping-students-by-ability-regains-favor-with-educators.html?pagewanted=all&_r=0

Yeung, R. (2009). The effects of fiscal decentralization on the size of government: A Meta-analysis. *Public Budgeting & Finance*, *29* (4), 1−23.

Zabel, J. E. (2008). The impact of peer effects on student outcomes in New York City public schools. *Education Finance and Policy*, *3*, 197−249.

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics*, *6*, 299−310.

Zhang, H. (2011). *Peer effects on student achievement: An instrumental variable approach using school transition data* (Working paper). Hong Kong: Department of Economics, The Chinese University of Hong Kong. Retrieved from http://ihome.cuhk.edu.hk/b121320/assetts/files/Zhang_PeerEffects_May2011.pdf

Zimmer, R. (2003). A new twist in the educational tracking debate. *Economics of Education Review*, *22*, 307−315.

Zimmer, R. W., & Toma, E. F. (2000). Peer effects in private and public schools across countries. *Journal of Policy Analysis and Management*, *19*, 75−92.

Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics*, *85*, 9−23.