HHS Public Access

Author manuscript

Soc Sci Res. Author manuscript; available in PMC 2016 July 01.

Published in final edited form as:

Soc Sci Res. 2015 July ; 52: 494–507. doi:10.1016/j.ssresearch.2015.03.004.

Social Interactions and College Enrollment: A Combined School Fixed Effects/Instrumental Variables Approach

Jason M. Fletcher◆

Associate Professor, La Follette School of Public Affairs, Department of Sociology, University of Wisconsin-Madison, 1225 Observatory Drive, Madison, WI 53706

Jason M. Fletcher: jmfletcher@wisc.edu

Abstract

This paper provides some of the first evidence of peer effects in college enrollment decisions. There are several empirical challenges in assessing the influences of peers in this context, including the endogeneity of high school, shared group-level unobservables, and identifying policy-relevant parameters of social interactions models. This paper addresses these issues by using an instrumental variables/fixed effects approach that compares students in the same school but different grade-levels who are thus exposed to different sets of classmates. In particular, plausibly exogenous variation in peers' parents' college expectations are used as an instrument for peers' college choices. Preferred specifications indicate that increasing a student's exposure to college-going peers by ten percentage points is predicted to raise the student's probability of enrolling in college by 4 percentage points. This effect is roughly half the magnitude of growing up in a household with married parents (vs. an unmarried household).

Keywords

Demand for Schooling; Human Capital; Peer Effects; College Enrollment

Introduction

There is a large and growing literature in economics and other social sciences that seeks to quantify the importance of peer influences on individual choices. Education and health outcomes have been of particular interest. One reason for this interest is that finding

◆The author thanks Steven Lehrer for helpful comments and for research support by core grants to the Center for Demography and Ecology at the University of Wisconsin-Madison (P2C HD047873). This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

evidence of social interactions suggests that policies that influence an individual's decision are predicted to "spill over" and influence peers' decisions. This spillover effect suggests that traditional cost-effectiveness analyses of programs that influence important choices (e.g. smoking, college enrollment) do not capture the full benefit of the programs and that these programs will likely be under-utilized as a result. In addition, finding evidence of social interactions also indicates that the composition of peer groups matters for individual and aggregate outcomes. This fact has implications for policies that shift the composition of groups, such as school choice policies, busing, and voucher programs.

Unfortunately, although the policy significance of finding social interactions in important decisions such as college enrollment is large, the empirical challenges with credibly assessing the presence of social interactions are equally large. Three principal difficulties include (1) the endogeneity of the peer group, (2) separately identifying different types of social influences, and (3) the presence of group-level unobservables that could lead to spurious estimates of social influences. The first two problems have been the focus of much current research.[1] However, one concern with much current research in the areas of peer effects and social interactions is the problem of group-level unobservables (Durlauf 2004). The presence of group-level unobservables could allow the researcher to conclude that social interactions exist when in fact there are 'third factors' that simultaneously affect the outcomes of all group members, leading to false conclusions of the importance of social interactions in the individual's decision-making process (Manski (1993) uses the term 'correlated effects').

Some of the most convincing research on peer influences in education outcomes has come from administrative data that is able to use school and/or school-by-grade fixed effects.[2] In many cases, this requires administrative datasets because of both the need for multiple cohorts and the need for large sample sizes so that enough variation is present in the data to identify the importance of social influences. Unfortunately, using administrative datasets limits both the individual-level observables available (e.g. family background variables) as well as the outcomes that can be examined.

In this paper, I combine the benefits of using rich longitudinal data that contains multiple cohorts of individuals in the same school and large enough samples to permit school-level fixed effects in the analysis to control for group-level unobservables. I combine this methodology with an instrumental variables approach to identify the parameters of interest to examine the importance of social influences on individual college enrollment decisions. The preferred instrumental variable strategy uses plausibly exogenous across-grade, within-school variation in the exposure to peers whose parents strongly encourage college enrollment. This encouragement by peers' parents is assumed to affect an individual's

[1]Endogeneity of the peer group has been confronted by using instrumental variables (Evans et al. 1992), random assignment of peers (Sacerdote 2001), school level fixed effects (McEwan 2003, Lavy and Schlosser 2007), or comparing groups based on mobility (Gaviria and Raphael 2001). The separate identification of different types of social influences has come from the use of instrumental variables (Gaviria and Raphael 2001) as well as the use of lagged peer measures (Hanushek et al. 2003, Clark and Loheac 2007).

[2]In particular, group level unobservables have been controlled in some research that has access to multiple cohorts (Hoxby 2000, Arcidiacono and Nicholson 2005, Hanushek et al. 2003, Lundborg 2006, Lavy and Schlosser 2007), where, for example, school or school/grade fixed effects are controlled and idiosyncratic variations of peer characteristics and/or peer outcomes are used to assess the importance of social influences.

college enrollment related choices only through his peers' college enrollment related choices, controlling for other factors such as peer family income, maternal education, race, etc. This assumption is consistent with the intuition that peers' parent investments that occur at home are distinct from other peer variables, such as income, and thus only affect an individual's college enrollment decision through the impacts on his/her peers (i.e. are "excludable" from the individual's decision equation). This paper provides evidence that variation in this peer family characteristic is quasi-random across cohorts within schools and also conducts falsification tests that replace peer college choices with the college choices of individuals in adjacent cohorts in the same school (Lavy and Schlosser 2007, Bifulco et al. 2011).

Using the combined instrumental variables/fixed effects approach indicates that college enrollment decisions are influenced by peers' choices. In particular, increasing a student's exposure to college-going peers by ten percentage points is predicted to raise the student's probability of enrolling in college by 4 percentage points. This effect is similar in magnitude to half the estimated benefit of growing up in a household with married parents (vs. unmarried) or having a mother with one year of additional schooling (vs. the sample average).

Background

For several decades in economics and other social sciences, researchers have made a substantial effort to incorporate social influences in modeling and estimating the decisions of individuals. Many point to the Coleman Report (1966) as a seminal work, as it drew attention to the family and social determinants of youth outcomes. Dozens of researchers have followed this shift of emphasis toward examining peer influences on education outcomes (e.g. Hoxby 2000, Hanushek et al. 2003, Sacerdote 2001, Foster 2006, Cooley 2007, Ding and Lehrer 2007, Lavy and Schlosser 2007, Ammermueller and Pischke 2007, Zabel 2008, Bifulco et al. 2011).

Before outlining relevant research that examines the importance of social influences on education outcomes, it is necessary to define relevant concepts in the social interactions literature. Manski (1993, 2000) distinguishes among the following types of social effects: endogenous effects, contextual effects, and correlated effects. *Endogenous effects* occur when the propensity of an individual to behave in some way varies with the behavior of the reference group. *Contextual effects* (also called exogenous effects in the sociology literature) occur when the propensity of an individual to behave in some way varies with the exogenous characteristics of the reference group. *Correlated effects* occur when individuals in the same group tend to behave similarly because they have similar individual characteristics or face similar institutional environments. Consider the case of college enrollment. Endogenous effects can occur if an individual is more likely to enroll in college if his classmates enroll—that is, if their decisions are interdependent. Contextual effects can occur if an individual is more likely to enroll in college if he or she is surrounded by classmates with highly educated parents. Correlated effects, which are not social in nature, can occur if individuals in the same school choose to enroll in college because they are geographically close to a college.

Distinguishing among these categories of effects (i.e. endogenous, contextual, and correlated effects) is important for several reasons, especially because of their different implications for potential policies. Policies that take advantage of endogenous effects are likely to produce a *social multiplier*. For example, if college enrollment is subject to endogenous effects, a policy that increases the propensity to enroll in college of an individual or a group of individuals within a school will affect other individuals who were not directly targeted by the policy—the effect of the policy is multiplied through social interactions. On the other hand, contextual changes may not imply the same multiplier effect responses to an exogenous shock. For example, adding higher income students to a low-income school would benefit the students in the receiving school, but the students at the sending school would be worse off. Thus, the gains to the former school would offset the losses from the latter school so that there would be no predicted aggregate multiplier effect.[3] Finally, correlated effects (i.e. group level unobservables) are not necessarily social in nature, and failure to adequately control for these effects can lead to spurious conclusions about the importance of social influences on individual choices.

Recently, there has been renewed interest in estimating social interactions and/or peer effects in the context of education outcomes[4]. For example, peer decisions and/or peer characteristics have been shown to be important in predicting elementary school achievement (Hanushek et al. 2003, Hoxby 2000, Ammermueller and Pischke 2007, Lavy and Schlosser 2007, Cooley 2007, Graham 2008), middle school achievement (McEwan 2003, Lavy and Schlosser 2007), high school achievement (Ding and Lehrer 2007), and achievement during college (Sacerdote 2001, Zimmerman 2003, Fletcher and Tienda 2008, Carrell et al. 2008, 2009, Arcidiacono et al. 2012). However, there is little credible published work examining potential endogenous peer effects related to attending college. For example, Sokatch (2007) examines correlations between classmate college expectations and own college-going behaviors though has no way to control for omitted group level characteristics (i.e. school fixed effects). Bifulco et al. (2011) examines the influence of peer characteristics (e.g. peer maternal education) on own college-going rates but does not attempt to directly estimate the influence of peer college behaviors on own college behaviors. The authors find evidence that students in school grades with higher proportions of highly educated mothers are more likely to attend college. The current paper controls for peers' maternal education and examines the complementary question of whether peers' behavioral choices related to college enrollment affect their grademates' college-going.

Unfortunately, credibly estimating peer influences on individual decisions is fraught with difficulty, and most previous work has been unable to address critical challenges in identifying a primary parameter of interest—*endogenous social effects.* Determining the importance of this effect is policy relevant because a positive endogenous social effect suggests that (1) the composition of peer groups is important determinants of adolescent college enrollment decisions and (2) interventions that increase the college enrollment propensities of any one student will increase the college enrollment propensities of her peers

[3]Hoxby (2000) and others have pointed out that contextual effects that are non-linear could imply distributional consequences of changing the composition of schools.
[4]See Epple and Romano (2011) for a recent review.

as well. This paper seeks to determine the answer to whether there are endogenous social effects in college enrollment decisions by using a combined instrumental variables/fixed effects strategy to confront the numerous empirical difficulties in estimating models of social interactions in college enrollment.

Data and Instrument Description

The data in this study come from the restricted version of the National Longitudinal Study of Adolescent Health (Add Health). Add Health is a school-based, longitudinal study of the health-related behaviors of adolescents and their outcomes in young adulthood. Beginning with an in-school questionnaire administered to a nationally representative sample of students in grades 7 through 12 in 1994–95, the study follows up with a series of in-home interviews of students approximately one year and then six years later. Other sources of data include questionnaires for parents, siblings, fellow students, and school administrators. By design, the Add Health survey included a sample stratified by region, urbanicity, school type, ethnic mix, and size. Preexisting databases (e.g. census data) have been linked with the individuals in the sample and provide information about neighborhoods and communities.[5]

Of the over 14,000 individuals who completed the Wave 1 and Wave 3 surveys[6], approximately 13,500 have valid school identification codes and reported their educational attainment, and nearly 11,000 also reported a parent's expectations about college attendance during Wave 1.[7] I include individual, family, school, and neighborhood level variables in the empirical models. For individual and family-level variables, I include age, gender, race, grade level, an ability proxy (Peabody Picture Vocabulary Test score)[8], marital status of parents, mother's educational attainment, family income, rural status, and the number of siblings.

For a measure of parental encouragement to enroll in college, I construct an indicator that is equal to one if the parent[9] reports that she would be "very disappointed" if the child did not graduate from college and the child reports that his parent would be very disappointed if he did not graduate from college.[10] This measure is meant to capture investments made directly to (and received by) the child from the parents that arguably would not directly affect the child's grademates once other school level (fixed effects) and peer-level characteristics are

[5]See Udry (2003) for full description of the Add Health data set. Some $7^{th}/8^{th}$ graders were sampled from "feeder" middle schools for the high school and other $7^{th}/8^{th}$ graders were enrolled in schools that had larger grade spans (i.e. 7–12 or K-12).

[6]The full Wave I sample includes 20,000 students. In unreported results, very few variables predict attrition from the sample by Wave III, including male, test scores, and growing up in a with married parents. Importantly, college encouragement and peer college encouragement do not predict attrition.

[7]In order to keep nearly 3,000 students, I impute family income and maternal education and create a dummy variable for missing parental data. The results are qualitatively unchanged with other methods of imputing the data, but the peer effect estimates increase if I drop individuals with missing data.

[8]The Add Health Picture Vocabulary Test (AHPVT) is a computerized, abridged version of the Peabody Picture Vocabulary Test-Revised (PPVT-R). The AHPVT is a test of hearing vocabulary, designed for persons aged 2 1/2 to 40 years old who can see and hear reasonably well and who understand standard English to some degree. The test scores are standardized by age. Some psychologists interpret PVT scores as a measure of verbal IQ. Information on the test is provided online at http://www.cpc.unc.edu/projects/addhealth/files/w3cdbk/w3doc.zip.

[9]This report is nearly always provided by the mother.

[10]Specifically, the child is asked, "On a scale of 1 to 5, where 1 is low and 5 is high, how disappointed would she be if you did not graduate from college?". I code a "5" as very disappointed. The parent is asked, "How disappointed would you be if your child did not graduate from college?" The parent's choices include "very disappointed", "somewhat disappointed" and "not disappointed."

controlled, including the school-grade racial distribution, average family income and maternal education, and proportion in a household with married parents.

This measure of parental encouragement, aggregated to the grade level, is the instrumental variable used below for predicting the proportion of grademates who attend college. The intuition behind selecting this instrument is the desire for a variable that captures family processes and investments in education that do not directly affect other grademates[11]. While family income and education, and their deviation from school-level averages, may be observable to grademate parents, the more proximate measure of investment used as the instrument is arguably less so, especially its deviation from school-level averages[12]. Indeed, I show evidence below that peer parental encouragement is conditionally uncorrelated with a large set of individual-level observable variables (race, academic performance, income), suggesting that, within schools, this measure is quasi-random. In addition, other grademate-level variables that I argue are more easily observable, such as race, education, and academic performance, are not found to be uncorrelated with individual level characteristics, which is consistent with the relative usefulness of the chosen instrument. As with any instrumental variable strategy, there are two key conditions that must hold for the results to be valid. First, a testable condition is that the instrument is statistically related to the proportion of grademates who attend college following high school. Second, it must be true that the only effect of grademates' parental encouragement on own college enrollment operates through grademates' college enrollment behaviors and not through other channels (i.e. the instrument is uncorrelated with the error term in the main equation predicting own college enrollment).

This second criterion is an untestable, maintained assumption. There are plausible situations where it could be violated. For example, parent encouragement may help determine school resources directed at college attendance. These resources could affect all students' likelihood of attending college. The assumption in this paper is that these channels are adequately captured by school fixed effects and grade level characteristics. That is, while parent encourage might determine *school* resources (captured by school fixed effects), encouragement might not determine *school-grade* resources. The evidence presented below is highly suggestive of the instrument's validity, although to the extent that the second condition is not met, the results may be inflated.

Unweighted summary statistics are presented in Table 1. Fifty-six percent of students report some college enrollment by Wave 3[13], when they are on average 22 years old. Nearly 40% of the students were strongly encouraged to go to college by their parents, based on the measure described above. The PVT score has been standardized on the sample of Wave 3 respondents, so that the mean value of 0.08 indicates that the analysis sample is slightly

[11]See also similar strategies in the cases of estimating peer influence on adolescent smoking (Fletcher 2010) and drinking (Fletcher in 2012) decisions.

[12]That is, although families may sort into schools with high levels of parental encouragement, the assumption in the research design is that across-grade, within school variation in parental encouragement of peers is not related to own unobservable determinants of college enrollment.

[13]This measure is taken from the question: What is the highest grade or year of regular school you have completed? Individuals reporting more than 12 years of schooling are in the category of "college" and those with 12 or fewer years are in the category of "no college".

positively selected on underlying ability.[14] The range of the variable capturing the percent of college enrollment at the school-grade level ranges from 0 to 1, which means that children in some school-grades have zero peers who attend college and others have peers who all go to college.

Empirical Methods

The primary empirical specification for this paper is the linear-in-means model of social interactions (Case and Katz 1991, Manski 1993):

$$Y_{igs}=X_{igs}B+\overline{X}_{-igs}\delta+W_s\theta+\alpha\overline{Y}_{-igs}+\varepsilon_{igs} \quad (1)$$

In this equation, Y_{igs} is the college enrollment choice of individual i in grade g in school s; individual and family characteristics are contained in a vector X; peer characteristics are measured as grade-level averages of the X vector excluding the individual, labeled $\bar{X}_{-igs}$. School-level factors are measured in the vector, W_s, which is eliminated in school-level fixed effects specifications. Finally, $\overline{Y}_{-igs}$ the grade-level average outcome excluding the individual (i.e. the proportion of individuals in the same grade and school who enroll in college).

Following the theoretical model in Brock and Durlauf (2001), the implicit assumption in the empirical application is that students are making choices based on their *expectations* of their peer groups' actions, rather than their actual actions. To close the model, an assumption that follows is that, in equilibrium, these expectations are in fact equal to the peer group actions, on average. An advantage of this interpretation (and consistent with the timing of the variables captured in the analysis) is a reduction in the likelihood that peer outcomes (college enrollment) affect peer parental expectations and encouragement (the instrument). A disadvantage of this approach is the possible failure to account for intermediate behavioral choices of the peers that affect both peer college enrollment and the (updating) of peer parental encouragement strategies[15]. Dynamic empirical models of peer interactions and peer effects face these issues, generically, even in the presence of randomization to peer groups (Bayer, Ross, and Topa 2009); these advantages and disadvantages of the approach should be used when considering the empirical results.

The main coefficient of interest is the endogenous effect, α, which indicates the extent to which individuals are influenced by their peers' choices to enroll in college. If α is estimated to be positive, interventions that change the college enrollment behavior of individuals (or subsets of individuals) within a reference group would be predicted to spillover on non-treated individuals in the same reference group.[16]

[14]Summary statistics for the dropped sample are shown in the appendix. Briefly, the dropped sample come from families who are somewhat more disadvantaged in comparison to the analysis sample, such as maternal education, family income, and other measures.

[15]A specific example is peers' enrollment (and achievement) in college preparatory courses, which could affect the course decisions of an individual. This process could happen either through the updating of expectations about peer college enrollment decisions in the future and also through learning (or mimicking) about the value of college preparation courses during high school on achievement or other important outcomes. These, potentially important, feedback effects during high school are not modeled in the current paper and should be the subject of future research.

As noted above, α in equation (1) is not identified without the use of instrumental variables or other methods (Manski 1993, Brock and Durlauf 2001). All results use robust standard errors, where ε_{igs} are allowed to be arbitrarily correlated across individuals in the same school. In order to accommodate school fixed effects and because the outcome is not high or low prevalence, I use linear probability models for the binary outcome instead of logit or probit models.

Results

In Table 2, I present standard OLS regression results that examine the individual-level, family-level, and peer-level determinants of college enrollment decisions. Results in Column 1 show that males are less likely to enroll in college and minority students are more likely to enroll (controlling for income, ability, and other characteristics). Importantly, the measure of college encouragement by parents is positively related to college enrollment. Maternal education, family income, and living in a household with married parents also increase the probability of enrolling. The ability measure (PVT Score) increases the likelihood of enrolling in college—a one standard deviation increase in this score increases the likelihood of college enrollment by over 10 percentage points. At the peer level, racial composition, maternal education, proportion Hispanic, and living in a household with married parents are strongly related to individual decisions to enroll in college.

Columns 2 and 3 stratify the sample by gender. Being in a household with married parents is positively associated with college enrollment for females but less so for males. The college encouragement measure is a strong predictor of college enrollment for both genders.

Table 3 stratifies the results from Table 2 by race. Maternal education and married parents have stronger relationships with college enrollment for white and blacks than Hispanics. The relationship between peer-level maternal education and enrollment is greatest for black and Hispanic students and small for whites.

Instrumental Variable/Fixed Effects Results

Before presenting the results using instrumental variables, I show that there is adequate variation within schools in the peer measures and that the preferred instrumental variable is plausibly exogenous. In Table 4, I show the R-squared from a regression that predicts a selection of peer level variables with a complete set of school dummies and grade dummies. While the racial composition variables are nearly completely determined by the school and grade dummies, the peer college variables have reasonable variation.

In Table 5A, I show that the instrumental variable, peers' parental college encouragement, is conditionally uncorrelated with ten important individual characteristics, including race, test scores, maternal education, family income, and parent's marital status (Lavy and Schlosser 2007 and Bifulco et al. 2011 also use this strategy)[17]. In order to provide suggestive evidence that school-grade unobservables are not driving the results, Table 5A also shows

[16]Previous research has used variation in adjacent cohorts' peer characteristics within a school to examine the effects of peer characteristics (but not outcomes) on individual outcomes, where the coefficient of interest is δ rather than α (e.g. Hoxby 2000, Hanushek et al. 2003, Lavy and Schlosser 2007).

that students' reports that their teacher cares or mother cares[18] about them are uncorrelated with the peer measure. This test is in the spirit of examining usually unobservable variables (which would be in the error term) in addition to the more common characteristics, such as family income. These results add confidence that the preferred instrument, peer's parental college encouragement, is quasi-random within schools and suggest that parents are not responding in a measurable way to this peer characteristic by removing their children from a particular school-grade. Overall, this evidence is suggestive, but not conclusive, that the instrument may also be uncorrelated with unobservable factors related to college enrollment, following the logic of Altonji et al. (2005), which is a maintained untestable assumption for the analysis to be valid.

An additional issue with the specific instrument used in this paper is whether it is different than other family characteristics that could also be used as instruments. What privileges parental encouragement over other variables, such as parental education, as valid instruments? In order to further examine this issue of choosing one instrument over other instruments, Table 5b shows selective associations between other contextual factors, such as test scores, racial composition, and parental education, with individual characteristics. The evidence in this table is consistent with the claim that these other contextual factors are not good instruments because of their correlation with own observables. This is suggestive of the assumption made in this paper that parental encouragement is a qualitatively different from other family characteristics.

In Table 6, I begin an examination of whether there is evidence of endogenous social effects for college enrollment decisions—Do the enrollment decisions of individuals affect the enrollment decisions of their peers? In Column 1 in Table 6 the results of 2SLS regression analysis are shown. The results begin by following much of the literature on social interactions (Gaviria and Raphael 2001, Powell et al. 2005), by assuming that contextual (peer-level characteristics) effects are non-existent ($\delta = 0$ from Eq. 1) and are thus available to use as instruments.[19] The results of this specification suggest large endogenous effects in college enrollment—increasing the exposure to college-going peers by 10 percentage points is predicted to increase an individual's propensity to enroll in college by nearly 5 percentage points. The magnitude of the relationship is nearly as large as the size of the association between growing up in a household with married parents and enrolling in college. While the F-statistic in this specification is quite large (nearly 45)[20], as shown elsewhere (e.g. Fletcher in 2012), using contextual variables as instruments in social interactions models is often questionable. In fact, the over-identification test strongly rejects the validity of the instruments. The results are also subject to the criticism that there are likely school-level unobservables that simultaneously affect the college decisions of all students in the school

[17]Seemingly unrelated regression (SUR) analysis, which allows the error terms across outcomes to be correlated, is done in order to accommodate the potential for correlated errors across the outcomes in the tests.
[18]Both "teacher cares" and "mother cares" are five-point scales reported by the student respondent.
[19]Gaviria and Raphael (2001) defend this assumption by arguing that students are more affected by neighborhood contextual effects than school-level contextual effects.
[20]Appendix Table 2 A presents the first stage results for Table 6. The reader will also note the large differences by grade level in enrolling in college. These coefficients reflect both cohort differences as well as an artifact of the sample design, where some students were initially surveyed in grade 12 and others in grade 7. Of course, individuals who survive to grade 12 are more likely to enroll in college than those who survive to grade 7, which conflates the effects of differential dropout opportunities/risk with birth cohort effects.

(e.g. high school quality and distance to nearest college). These omitted variables likely bias upward the endogenous social effect estimates in this specification.

In Column 2, I add school-level fixed effects to the specification. Now, individuals in the same school but in different grades who face different sets of peers are compared. The use of fixed effects may also suggest that contextual effects may now be validly excluded from the main equation. That is, once time-invariant school-level factors are controlled, the idiosyncratic differences in peer characteristics across grades within schools are more credibly excluded from the individual's propensity to enroll in college (Lundborg 2006). These peer characteristics are assumed to determine peer enrollment outcomes but not individual enrollment decisions after controlling for school-level fixed effects. However, the results show this may not be a reasonable assumption in the present case. The F-statistic is dramatically reduced to 4.7 because of the reduced variation in the instruments (see % black in Table 4, for example). Additionally, the over-identification test still rejects the validity of the instruments at the <5% level.

Next, in Column 3, I examine the reduced form effects of contextual variables on the likelihood an individual attends college following high school. This investigation cannot separate exogenous and endogenous social effects but can provide reduced form evidence of a relationship between the instrument, other contextual factors, and the outcome, which could suggest some form of peer effect. The results suggest that a move from 0% to 100% of classmates who receive parental encouragement increases own college-going by 8 percentage points.

Finally, in Column 4, I combine the school-level fixed effects with the preferred instrument —the proportion of peers whose parents encourage them to go to college. The endogenous peer effect coefficient is now 0.457, which implies that a 10% increase in peer college enrollment increases the likelihood of own college attendance by 4.6 percentage points; of some interest is the reduction of the effect of the proportion of classmates with married parents, compared with Column 3, suggesting a possible mechanism linking this contextual effect with the outcome[21]. The modest F-statistic of 6.1 suggests that this coefficient may be slightly upwardly biased, though (Stock and Staiger 1997)[22].

Additional Issues

One issue not yet addressed in the empirical work is that the data do not contain complete rosters of grademates within the schools.[23] This missing data creates measurement error in the peer variables. Luckily, the data were gathered randomly within grade, in which case Ammermueller and Pischke (2009) provide a relatively simple calculation to adjusted peer

[21]I thank an anonymous reviewer for suggesting this point.

[22]In unreported results, I attempted to add a second instrument, "parent's discouragement", which indicates students who report that their parents would not be upset if they did not graduate from college (and parents agree); 4% of the sample have discouraging parents. This second instrument was weaker than the current instrument (as were other, similar, instruments I created). The results from Table 6, Column 3 using two instruments were: 0.414 point estimate (p-value<0.13), first stage F-statistic of 3.3 and the failure to reject the validity of the instruments using an over-identification test (p-value<0.38). This is additional suggestive evidence of the validity of the specification, but not conclusive due to the weakness of the instruments.

[23]The median school/grade size was 220 students (range 20–700), while the median surveyed school/grade size was 43 (range 20–575). The mean proportion surveyed was 37.5%, with median 24%.

effect coefficients.[24] The authors show, under these conditions, the peer effect estimate will be biased by a factor of $(N_{sample} - 1)/(N_{actual} - 1)$, which is approximately 0.4 in this sample, suggesting actual peer effects could be as much as 2.5 times larger than those in the results tables (see also Aizer 2009). In addition to this result, I also estimated the preferred specification for only those schools that had "saturated" samples of students, where the mean proportion of those sampled against those on the roster was nearly 90%. The peer effect estimate for this sample was 0.75 (compared to 0.45 for the full sample), which is consistent with the importance of measurement error in attenuating results.

In order to provide additional confidence that grade-level peers matter rather than other peers in the same school as well as provide evidence that some types of unobservable school factors are not driving the results, I follow Lavy and Schlosser (2007) by estimating two additional placebo specifications in Table 7. In these results, the actual grade-level peer measures are replaced by the peer measures of the adjacent cohort, so that 10^{th} graders are combined with either 9^{th} grade peer measures or 11^{th} grade peer measures from their school. Unfortunately, this shrinks the sample because 12^{th} graders do not have a upper adjacent grade to use in this specification and 9^{th} graders do no have a lower adjacent grade, as shown in columns 2 and 3, which repeats columns 1 results on the smaller samples. Then in columns 4 and 5, the actual peer measures are replaced by the peer measures for the adjacent grade in the same school. The results suggest that the consistent peer effect is found by using the grade-level peers and not adjacent peer measures. However, these placebo exercises are only potentially falsifying a small domain of hypotheses. For example, one might be worried that there are within-school variations not picked up by the school-fixed effects (e.g. shocks affecting grades 9 in some schools that are correlated with the instrument) or other similar stories. This exercise shows evidence against these narrow hypotheses. It also is a (somewhat weak) test of whether actual grademates matter vs. adjacent gradements. That is, are we just capturing noise or other omitted factors when we measure grademates as the peer group? The results are consistent with the main results from the paper.

Conclusions

In this paper, I use a social interactions framework to detect whether individual college enrollment decisions are influenced by peer college enrollment decisions, providing some of the first evidence on this research question. I address several large empirical challenges in addressing this question, including the endogeneity of school (and thus peers) through residential location choices, 'third factors' such as school-level unobservables that influence individual and peers' choices simultaneously, and the difficulty of the identification of parameters in empirical models of social interactions (Manski 1993, Brock and Durlauf 2001). In particular, I use an instrumental variables/fixed effects methodology that compares students in different grade levels within the same high school who face a different set of grademates and grademates' decisions. Preferred specifications suggest that increasing the proportion of peers who enroll in college by 10% will increase the likelihood an individual enrolls by approximately 4 percentage points.. A limitation of these results is that several

[24]See also Sojourner (2009) who describes a more general approach without random missingness.

estimates are only statistically different from zero at a 10% level (rather than a more traditional 5% level).

While this paper is able to address many of the relevant econometric issues in estimating the importance of social influences on college enrollment decisions, there are several limitations. First, while I follow much of the current literature in assuming that the relevant reference group for adolescents are grademates, it could be the case that this reference group is too broad, too narrow, or of limited importance in influencing the decisions for typical adolescents.[25] To the extent that the school-grade measure of peers is too broad, the results presented here may be conservative. On the other hand, as school-grade level composition is a potentially more important policy lever than other alternatives, such as friendship networks, there are other practical reasons to want to understand the impacts of broader measures of composition. A second limitation is that there could be school-grade specific correlated effects that are not eliminated through using school fixed effects, so that group-level factors (correlated effects) could still bias the results. However, evidence discussed above suggests that variation in teachers within schools is not driving the results.

With these limitations in mind, this paper is still likely the most compelling current evidence that there are social interactions present in adolescent college enrollment decisions. These findings suggest that policies that influence one person's decision to enroll in college will also affect the decisions of the person's peers through social interactions. Further, these results suggest that traditional cost-effectiveness analysis of college enrollment intervention programs that do not take into account the spillover effect on untreated individuals likely understate the total social benefits of the programs and suggest that too few programs will be undertaken.

References

Aizer, Anna. Brown University Working Paper. 2009. Peer Effects, Institutions and Human Capital Accumulation: The Externalities of ADD.

Altonji, Joseph G.; Elder, Todd E.; Tabor, Christopher R. Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. Journal of Political Economy. 2005; 113:151–184.

Ammermueller, Andreas; Pischke, Jorn-Steffen. Working Paper 2007. 2007. Peer Effects in European Primary Schools: Evidence from PIRLS.

Ammermueller A, Pischke J. Peer effects in European primary schools: Evidence from the Progress in International Reading Literacy Study. Journal of Labor Economics. 2009; 27(3):315–348.

Arcidiacono, Peter; Nicholson, Sean. Peer Effects in Medical School. Journal of Public Economics. 2005; 89:327–350.

Arcidiacono, Peter; Foster, Gigi; Goodpaster, Natalie; Kinsler, Josh. Estimating Spillovers using Panel Data, with an Application to the Classroom. with Quantitative Economics. Nov; 2012 3(3):421–470.

Bifulco, Robert; Fletcher, Jason; Ross, Stephen. The Effect of Classmate Characteristics on Individual Outcomes: Evidence from the Add Health. American Economic Journal: Economic Policy. 2011

Brock W, Durlauf S. Discrete Choice with Social Interactions. Review of Economic Studies. 2001:68.

Carrell, Scott; Fullerton, Richard; West, James. Does Your Cohort Matter? Measuring Peer Effects in College Achievement. Journal of Labor Economics. 2009; 27(3):439–464.

[25]However, Weinberg (2008) shows that 74% of friendship nominations occur within the same grade level using this dataset.

Carrell, Scott; Malmstrom, Frederick; West, James. Peer Effects in Academic Cheating. Journal of Human Resources. 2008; XLIII(1):173–207.

Case, Anne; Katz, Lawrence. NBER Working Paper 3705. 1991. The company you keep: the effects of family and neighborhood on disadvantaged youth.

Clark, Andrew; Loheac, Youenn. 'It Wasn't Me, It Was Them!' Social Influence in Risky Behavior by Adolescents. Journal of Health Economics. 2007

Coleman, James, et al. Equality of Educational Opportunity. Washington, DC: United States Government Printing Office; 1966.

Cooley, Jane. University of Wisconsin-Madison Working Paper. 2007. Desegregation and the Achievement Gap: Do Diverse Peers Help?.

Ding, Weili; Lehrer, Steven. Do Peers Affect Student Achievement in China's Secondary Schools? Review of Economics and Statistics. 2007

Epple, Dennis; Romano, Richard. Peer effects in education: A survey of the theory and evidence. Handbook of Social Economics. 2011; 1(11):1053–1163.

Fletcher, Jason M. Social Interactions and Smoking Decisions: Evidence Using Multiple Cohorts, Instrumental Variables, and School Fixed Effects. Health Economics. 2010; 19(4):466–484. [PubMed: 19382102]

Fletcher, Jason M. Peer influences on adolescent alcohol consumption: evidence using an instrumental variables/fixed effect approach. Journal of Population Economics. 2012

Fletcher, Jason M. Similarity in peer college preferences: New evidence from texas. Social Science Research. 2012b; 41(2):321–330. [PubMed: 23017754]

Fletcher, Jason M.; Tienda, Marta. High School Classmates and College Success. Sociology of Education. 2009; 82(4):287–314. [PubMed: 23172979]

Foster, Gigi. It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism. Journal of Public Economics. 2006; 90(8–9):1455–1475.

Gaviria, Alejandro; Raphael, Steven. School-Based Peer Effects and Juvenile Behavior. Review of Economics and Statistics. 2001; 83(2)

Graham, Bryan. Identifying Social Interactions Through Conditional Variance Restrictions. Econometrica. 2008; 76(3):643–660.

Hanushek EA, Kain JF, Markman JM, Rivkin SG. Does peer ability affect student achievement? Journal of Applied Econometrics. 2003; 18(5):527–544.

Hoxby, Caroline. The Effects of School Choice on Curriculum and Atmosphere. In: Mayer, Susan; Peterson, Paul, editors. Earning and Learning: How Schools Matter. Brookings Institute Press; 1999.

Hoxby, Caroline. NBER Working Paper 7867. 2000. Peer Effects in the Classroom: Learning from Gender and Race Variation.

Lavy, Victor; Schlosser, Analia. NBER Working Paper 13292. 2007. Mechanisms and Impacts of Gender Peer Effects at School.

Lundborg, Peter. Having the Wrong Friends? Peer Effects in Adolescent Substance Use. Journal of Health Economics. 2006; 25:214–233. [PubMed: 15964090]

Manski, Charles. Identification of Endogenous Social Effects: The Reflection Problem. Review of Economic Studies. 1993:60.

Manski, Charles. Identification Problems in the Social Sciences. Harvard University Press; Cambridge, MA: 1995.

Manski, Charles. Economic Analysis of Social Interactions. Journal of Economic Perspectives. 2000; 14:3.

Powell, Lisa; Taurus, John; Ross, Hana. The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior. Journal of Health Economics. 2005; 24(5):950–968. [PubMed: 15990184]

Sacerdote, Bruce. Peer Effects with Random Assignments: Results from Dartmouth Roommates. Quarterly Journal of Economics. 2001; 116:2.

Sokatch, Andrew. Peer influences on the college-going decisions of low socioeconomic status urban youth. Education and Urban Society. 2006; 39(1):128–146.

Sojourner, A. Working Paper. 2009. Inference on Peer Effects with Missing Peer Data: Evidence from Project STAR.

Stock, James; Staiger, Douglas. Instrumental Variables Regression with Weak Instruments. Econometrica. 1997; 65(3):577–586.

Udry, JR. The National Longitudinal Study of Adolescent Health (Add Health), Waves I and II, 1994–1996; Wave III, 2001–2002 [machine-readable data file and documentation]. Carolina Population Center, University of North Carolina at Chapel Hill; Chapel Hill, NC: 2003.

Weinberg, Bruce. Ohio State University Working Paper. 2008. Social Interactions and Endogenous Associations.

Zabel, Jeffrey E. The Impact of Peer Effects on Student Outcomes in New York City Public Schools. Education Finance and Policy. 2008; 3(2):197–249.

Zimmerman D. Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. Review of Economics and Statistics. 2003; 85(1)

Appendix

Table 1A

Summary Statistics of Individuals with Missing College Encouragement Information

Variable	Wave	Obs	Mean	Std.	Dev.	Min
Individual Level Variables						
College Enrollment	3	2358	0.49	0.50	0	1
College Encouragement	1					
Male	All	2360	0.47	0.50	0	1
Age	3	2360	22.59	1.73	18	27
Black	All	2360	0.24	0.43	0	1
Hispanic	All	2360	0.20	0.40	0	1
Other Race	All	2360	0.17	0.38	0	1
Number of Siblings	1	2360	1.79	1.62	0	12
Has Older Siblings?	1	2357	0.53	0.50	0	1
PVT Score	1	2204	−0.16	1.01	−6	2
Maternal Education Level^	1	2360	12.95	2.13	0	17
Family Income^	1	2360	43.27	19.98	0	600
Married Household^	1	2360	0.62	0.29	0	1
Parent Missing Data Dummy	1	2360	0.96	0.20	0	1
Grade Level = 7	1	2360	0.08	0.27	0	1
Grade Level = 8	1	2360	0.09	0.29	0	1
Grade Level = 9	1	2360	0.13	0.34	0	1
Grade Level = 10	1	2360	0.20	0.40	0	1
Grade Level = 11	1	2360	0.23	0.42	0	1
Grade Level = 12	1	2360	0.26	0.44	0	1
School Level Variables						
% College Enrollment	3	2358	0.55	0.19	0	1
% College Encouragement	1					
% Black	1	2359	0.23	0.26	0	1
% Hispanic	1	2360	0.20	0.22	0	1
% Other Race	1	2356	0.12	0.13	0	1

Variable	Wave	Obs	Mean	Std.	Dev.	Min
% Married Household	1	2360	0.69	0.11	0	1
Mean Family Income^	1	2360	44.88	13.82	20	143
Mean Maternal Education^	1	2360	13.12	0.81	11	16
Rural Status	1	2360	0.22	0.41	0	1
Number Sampled in Grade	1	2360	1.41	1.91	0 c	

Table 2A

Main and First Stage Results for Table 6

Outcome	College Enrollment	% College Enrollment	College Enrollment	% College Enrollment	College Enrollment	% College Enrollment
Sample	Full		Full		Full	
Method	2SLS	First Stage	2SLS	First Stage	2SLS	First Stage
Instruments	All Contextual		All Contextual		% College Encouragement	
Fixed Effects	None		School		School	
% College	0.482*** (0.042)		0.544*** (0.142)		0.457* (0.274)	
College Encouragement	0.111*** (0.009)	0.011** (0.005)	0.105*** (0.009)	−0.002 (0.001)	0.105*** (0.009)	−0.000 (0.002)
Male	−0.074*** (0.010)	−0.004 (0.003)	−0.074*** (0.010)	0.002 (0.001)	−0.074*** (0.009)	0.002 (0.001)
Age	−0.075*** (0.007)	−0.004 (0.003)	−0.077*** (0.007)	0.002** (0.001)	−0.076*** (0.007)	0.002** (0.001)
Black	0.044*** (0.012)	0.008 (0.006)	0.050*** (0.013)	−0.000 (0.003)	0.049*** (0.013)	−0.001 (0.002)
Hispanic	0.038** (0.016)	0.016** (0.007)	0.015 (0.022)	0.001 (0.003)	0.014 (0.021)	0.001 (0.003)
Other Race	0.060*** (0.023)	−0.012 (0.009)	0.082*** (0.019)	0.000 (0.003)	0.080*** (0.019)	−0.001 (0.003)
Number of Siblings	0.002 (0.004)	−0.001 (0.002)	0.002 (0.004)	0.001 (0.001)	0.002 (0.004)	0.001 (0.001)
PVT Score	0.096*** (0.008)	0.012* (0.006)	0.099*** (0.007)	−0.001 (0.001)	0.099*** (0.007)	−0.001 (0.001)
Maternal Education	0.032*** (0.003)	0.001** (0.001)	0.032*** (0.003)	0.000 (0.000)	0.032*** (0.003)	0.000 (0.000)
Family Income	0.001*** (0.000)	0.000** (0.000)	0.001*** (0.000)	0.000** (0.000)	0.001*** (0.000)	0.000** (0.000)
Married Household	0.080*** (0.011)	0.009*** (0.003)	0.076*** (0.011)	0.002 (0.002)	0.077*** (0.011)	0.003 (0.002)
% Black		0.081 (0.053)		−0.146 (0.145)	−0.066 (0.109)	−0.124 (0.134)
% Hispanic		0.259*** (0.051)		−0.018 (0.148)	0.056 (0.130)	−0.033 (0.145)
% Other Race		−0.048 (0.151)		0.079 (0.117)	−0.152 (0.121)	0.084 (0.112)
% Married Households		0.402*** (0.088)		0.187** (0.082)	0.140 (0.092)	0.214*** (0.081)
Mean Family Income		0.003*** (0.001)		0.002** (0.001)	0.000 (0.001)	0.002** (0.001)

Outcome	College Enrollment	% College Enrollment	College Enrollment	% College Enrollment	College Enrollment	% College Enrollment
Mean Maternal Education		0.115*** (0.016)		0.044*** (0.016)	−0.027 (0.017)	0.039** (0.016)
Rural Status		−0.006 (0.010)		−0.002 (0.003)	−0.007 (0.014)	−0.001 (0.002)
Grade Level = 8	0.095*** (0.013)	0.043** (0.018)	0.094*** (0.014)	0.045** (0.017)	0.096*** (0.018)	0.047*** (0.017)
Grade Level = 9	0.174*** (0.019)	0.069*** (0.023)	0.175*** (0.021)	0.078*** (0.027)	0.176*** (0.027)	0.074*** (0.026)
Grade Level = 10	0.277*** (0.026)	0.126*** (0.024)	0.277*** (0.030)	0.126*** (0.028)	0.278*** (0.042)	0.123*** (0.027)
Grade Level = 11	0.353*** (0.030)	0.148*** (0.023)	0.352*** (0.037)	0.148*** (0.027)	0.358*** (0.050)	0.146*** (0.026)
Grade Level = 12	0.450*** (0.038)	0.198*** (0.026)	0.451*** (0.047)	0.202*** (0.028)	0.458*** (0.065)	0.199*** (0.027)
% College Encouragement						0.172** (0.070)
Observations	10638	10638	10638	10638	10638	10638
R-squared	0.235	0.611	0.135	0.333	0.138	0.348
Number of Schools			121	121	121	121
F-stat	44.948		4.780		6.087	
J-Stat P-value	0.0000681		0.042			

Robust standard errors clustered at the school-grade level,
*** 1%,
** 5%,
* 10%.

Additional Controls: Grade dummies, Constant, Dummy for Missing Parent Information, Number of Siblings, Rural Status, Number of students sampled in school

Note: aggregate variables (e.g. % black) at the school-grade level

Highlights

- Little evidence has been published that examines the importance of peer influences on college enrollment behaviors.
- This paper uses within-school, across-cohort variation in child-specific college investments by parents as an instrument for exposure to college-going peers.
- The results suggest that a 10 percentage point increase in college-going peers increases the likelihood of own college going by approximately 4 percentage points.

Table 1

Summary Statistics: Add Health Data N=10,638

Variable	Wave	Mean	Std Dev	Min	Max
Individual Level Variables					
College Enrollment	3	0.56	0.50	0	1
College Encouragement	1	0.37	0.48	0	1
Male	All	0.47	0.50	0	1
Age	3	21.84	1.71	18	27
Black	All	0.22	0.41	0	1
Hispanic	All	0.16	0.36	0	1
Other Race	All	0.08	0.27	0	1
Number of Siblings	1	1.59	1.41	0	13
Has Older Siblings?	1	0.49	0.50	0	1
PVT Score (standardized)	1	0.08	0.96	−6	3
Maternal Education Level^ (years)	1	13.25	2.29	0	17
Family Income^ ($1,000s)	1	46.11	42.95	0	990
Household with Married Parents^	1	0.74	0.44	0	1
Mother Cares	1	4.86	.47	1	5
Teacher Cares	1	3.52	.98	1	5
Parent Missing Data Dummy	1	0.19	0.40	0	1
Grade Level = 7	1	0.15	0.36	0	1
Grade Level = 8	1	0.14	0.35	0	1
Grade Level = 9	1	0.18	0.39	0	1
Grade Level = 10	1	0.20	0.40	0	1
Grade Level = 11	1	0.19	0.39	0	1
Grade Level = 12	1	0.14	0.34	0	1
School/Grade Level Variables					
% College Enrollment	3	0.55	0.20	0	1
% College Encouragement	1	0.36	0.14	0	1
% Black	1	0.20	0.25	0	1
% Hispanic	1	0.16	0.21	0	1

Variable	Wave	Mean	Std Dev	Min	Max
% Other Race	1	0.10	0.12	0	1
% Household with Married Parents^	1	0.70	0.12	0	1
Mean Family Income^ ($1,000s)	1	44.96	13.95	18	144
Mean Maternal Education^	1	13.18	0.81	10	16
Rural Status	1	0.27	0.44	0	1

^ imputed variable

Soc Sci Res. Author manuscript; available in PMC 2016 July 01.

Table 2

Individual, Family, and Peer-Level Determinants of College Enrollment Full Sample and by Gender

Outcome	College Enrollment	College Enrollment	College Enrollment
Sample	Full	Males	Females
Method	OLS	OLS	OLS
Fixed Effects	None	None	None
College Encouragement	0.114***	0.134***	0.093***
	(0.010)	(0.014)	(0.013)
Male	−0.076***		
	(0.009)		
Age	−0.079***	−0.067***	−0.091***
	(0.007)	(0.009)	(0.010)
Black	0.052***	0.040*	0.060***
	(0.014)	(0.024)	(0.018)
Hispanic	0.020	0.026	0.014
	(0.020)	(0.024)	(0.028)
Other Race	0.070***	0.103***	0.039
	(0.019)	(0.028)	(0.030)
Older Siblings	0.007	0.006	0.006
	(0.009)	(0.014)	(0.012)
PVT Score (std)	0.104***	0.100***	0.108***
	(0.007)	(0.011)	(0.007)
Maternal Education	0.033***	0.030***	0.035***
	(0.003)	(0.004)	(0.004)
Family Income	0.001***	0.001***	0.000***
	(0.000)	(0.000)	(0.000)
Married Household	0.083***	0.063***	0.095***
	(0.011)	(0.018)	(0.016)
% Black	0.097**	0.089	0.100**
	(0.044)	(0.062)	(0.049)
% Hispanic	0.245***	0.227***	0.258***
	(0.044)	(0.048)	(0.058)
% Other Race	−0.168	−0.212**	−0.128
	(0.106)	(0.106)	(0.125)
% Married Households	0.404***	0.348***	0.450***
	(0.083)	(0.098)	(0.104)
Mean Family Income	0.002**	0.002*	0.001

Outcome	**College Enrollment**	**College Enrollment**	**College Enrollment**
	(0.001)	(0.001)	(0.001)
Mean Maternal Education	0.041***	0.043**	0.040**
	(0.015)	(0.018)	(0.019)
Number Sampled in Grade	−0.014**	−0.013*	−0.015**
	(0.006)	(0.007)	(0.007)
Observations	10638	5024	5614
R-squared	0.227	0.222	0.230

Robust standard errors clustered at the school-grade level,

*** 1%,

** 5%,

* 10%.

Additional Controls: Grade dummies, Constant, Dummy for Missing Parent Information,

Number of Siblings, Rural Status

Note: aggregate variables (e.g. % black) at the school-grade level

Table 3

Individual, Family, and Peer-Level Determinants of College Enrollment Results by Race

Outcome	College Enrollment	College Enrollment	College Enrollment
Sample	White	Black	Hispanic
Method	OLS	OLS	OLS
Fixed Effects	None	None	None
College Encouragement	0.120***	0.061***	0.108***
	(0.014)	(0.016)	(0.023)
Male	−0.077***	−0.084***	−0.075***
	(0.013)	(0.019)	(0.021)
Age	−0.083***	−0.091***	−0.079***
	(0.010)	(0.013)	(0.021)
Older Siblings	0.014	0.016	−0.006
	(0.011)	(0.020)	(0.021)
PVT Score	0.114***	0.085***	0.104***
	(0.009)	(0.011)	(0.020)
Maternal Education	0.040***	0.037***	0.019***
	(0.003)	(0.005)	(0.007)
Family Income	0.001***	0.000	0.001
	(0.000)	(0.000)	(0.001)
Married Household	0.086***	0.072***	0.047
	(0.015)	(0.021)	(0.032)
% Black	0.008	0.102	0.190*
	(0.062)	(0.071)	(0.105)
% Hispanic	0.219***	0.144	0.318***
	(0.068)	(0.114)	(0.080)
% Other Race	−0.169	0.334*	0.012
	(0.135)	(0.182)	(0.134)
% Married Households	0.352***	0.540***	0.271
	(0.094)	(0.138)	(0.214)
Mean Family Income	0.002*	−0.000	0.002
	(0.001)	(0.001)	(0.002)
Mean Maternal Education	0.023	0.076***	0.049
	(0.018)	(0.025)	(0.032)
Number Sampled in Grade	−0.016**	−0.040***	−0.016
	(0.007)	(0.010)	(0.010)
Constant	0.746**	0.338	0.618

Outcome	College Enrollment	College Enrollment	College Enrollment
	(0.318)	(0.379)	(0.612)
Observations	5909	2344	1649
R-squared	0.235	0.259	0.165

Robust standard errors clustered at the school-grade level,

*** 1%,

** 5%,

* 10%.

Additional Controls: Grade dummies, Constant, Dummy for Missing Parent Information,

Number of Siblings, Rural Status

Note: aggregate variables (e.g. % black) at the school-grade level

Table 4

Proportion of Cohort Variable Variation Explained by School and Grade Fixed Effects

Variation in Classmate-Level Variables	R-Squared
% College	0.85
% College Encouragement	0.78
% Black	0.98
% Hispanic	0.99
% Other Race	0.96
% Married	0.80
Family Income	0.87
Mean Mom Education	0.90

Note: The table displays the adjusted R-squared from a regression predicting the grade-level variables

Each row represents a separate regression

Controls: school and grade fixed effects

Table 5A

Balancing Tests of Exposure to Peers' Parental Encouragement

Outcome	SUR Test	Age	Black	Hispanic	PVT Score	Maternal Education	Family Income	Teacher Cares	Mother Cares	Married	Number of Siblings
% College Encouragement		0.059 (0.100)	0.035 (0.048)	−0.048 (0.038)	0.068 (0.136)	−0.383 (0.340)	6.484 (7.129)	−0.045 (0.216)	−0.001 (0.076)	−0.086 (0.069)	0.043 (0.185)
Observations		10638	10638	10638	10638	10638	10638	10588	10637	10638	10638
F-Test/Chi-2	8.47	0.349	0.524	1.633	0.248	1.267	0.827	0.043	0.000	1.533	0.054
P-Value	0.78	0.556	0.471	0.204	0.619	0.263	0.365	0.836	0.991	0.218	0.816

Each Column represents a separate regression that examines the association between the cohort variable and individual level characteristics. School and grade dummies are controlled but not presented. The first column presents the joint test from a seemingly unrelated regression specification.

Table 5B

Selected Balancing Tests of Exposure to Peers' Characteristics

Outcome	SUR Test	Mom Cares	PVT Score	Outcome	SUR Test	Black	Number of Sibs	Outcome	SUR Test	Other Race	Married
% Black		−0.172* (0.093)	−0.717 (0.711)	Average Parent Education		−0.024 (0.020)	0.081 (0.057)	Average PVT Score		0.004* (0.003)	−0.004* (0.002)
Observations		10,637	10,638	Observations		10,638	10,638	Observations		10,638	10,638
F-Test/Chi-2	17.33	3.421	1.015	F-Test/Chi-2	12.11	1.502	2.008	F-Test/Chi-2	16.49	3.136	3.837
P-Value	0.044	0.067	0.316	P-Value	0.21	0.223	0.159	P-Value	0.057	0.079	0.052

Each column represents a separate regression that examines the association between the cohort variable and individual level characteristics. School and grade dummies are controlled but not presented. The first column in each set presents the joint test from a seemingly unrelated regression specification.

Table 6

Individual, Family, and Peer-Level Determinants of College Enrollment Two-Stage Least Squares and Reduced Form Results with and without School-Level Fixed Effects

Outcome	**College Enrollment**	**College Enrollment**	**College Enrollment**	**College Enrollment**
Sample	Full	Full	Full	Full
Method	2SLS	2SLS	OLS	2SLS
Instruments	All Contextual	All Contextual	None	% College Encouragement
Fixed Effects	None	School	School	School
% College	0.482***	0.544***		0.457*
	(0.042)	(0.142)		(0.274)
% College Encouragement			0.081*	
			(0.042)	
College Encouragement	0.111***	0.105***	0.105***	0.105***
	(0.009)	(0.009)	(0.009)	(0.009)
Male	−0.074***	−0.074***	−0.073***	−0.074***
	(0.010)	(0.010)	(0.009)	(0.009)
Age	−0.075***	−0.077***	−0.075***	−0.076***
	(0.007)	(0.007)	(0.007)	(0.007)
Black	0.044***	0.050***	0.049***	0.049***
	(0.012)	(0.013)	(0.013)	(0.013)
Hispanic	0.038**	0.015	0.014	0.014
	(0.016)	(0.022)	(0.021)	(0.021)
Other Race	0.060***	0.082***	0.080***	0.080***
	(0.023)	(0.019)	(0.019)	(0.019)
PVT Score	0.096***	0.099***	0.098***	0.099***
	(0.008)	(0.007)	(0.007)	(0.007)
Maternal Education	0.032***	0.032***	0.032***	0.032***
	(0.003)	(0.003)	(0.003)	(0.003)
Family Income	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Married Household	0.080***	0.076***	0.078***	0.077***
	(0.011)	(0.011)	(0.011)	(0.011)
% Black			−0.123	−0.066
			(0.146)	(0.109)
% Hispanic			0.074	0.056
			(0.189)	(0.130)
% Other Race			−0.084	−0.152
			(0.172)	(0.121)

Outcome	College Enrollment	College Enrollment	College Enrollment	College Enrollment
% Married Households			0.240***	0.140
			(0.089)	(0.092)
Mean Family Income			0.001	0.000
			(0.001)	(0.001)
Mean Maternal Education			−0.010	−0.027
			(0.017)	(0.017)
Observations	10638	10638	10,638	10638
R-squared	0.235	0.135	0.149	0.138
Number of Schools		121	121	121
F-stat	44.948	4.780		6.087
J-Stat P-value	0.0000681	0.042		

Robust standard errors clustered at the school-grade level,

*** 1%,

** 5%,

* 10%.

Additional Controls: Grade dummies, Constant, Dummy for Missing Parent Information, Number of Siblings, Number of Older Siblings, Rural Status, Number of students sampled in school.

Note: aggregate variables (e.g. % black) at the school-grade level

Table 7

Preferred Results and Placebo Tests Two-Stage Least Squares Results with School-Level Fixed Effects

Outcome	**College Enrollment**	**College Enrollment**	**College Enrollment**	**College Enrollment**	**College Enrollment**
Sample	Full	Grade Above	Grade Below	Grade Above	Grade Below
Method	2SLS	2SLS	2SLS	2SLS	2SLS
Instruments	% College Enrollment	% College Enrollment	% College Enrollment	% College Enrollment	% College Encouragement
Peer Definition	Same Grade	Same Grade	Same Grade	Grade Above	Grade Below
Fixed Effects	School	School	School	School	School
% College	0.457*	0.455	0.445	0.244	−0.206
	(0.274)	(0.392)	(0.402)	(0.737)	(0.335)
Observations	10638	7119	7971	7119	7971
Number of Schools	121	121	121	121	121
R-squared	0.138	0.123	0.133	0.133	0.146
F-stat	6.087	3.812	2.705	2.129	2.863

Robust standard errors clustered at the school-grade level,

*** 1%,

** 5%,

* 10%.

Additional Controls: Grade dummies, Constant, Dummy for Missing Parent Information, Number of Siblings, Rural Status, Number of students sampled in school

Note: aggregate variables (e.g. % black) at the school-grade level

Grade above: replaces the actual peer exposure measure with the peer measure from the grade above in the same school (11^{th} grade measure is used for 10^{th} graders). Thus 12^{th} graders are dropped

Grade below: replaces the actual peer exposure measure with the peer measure from the previous grade in the same school (9^{th} grade measure is used for 10^{th} graders). Thus 9^{th} graders or 7^{th} graders are dropped (depending on school grade span)