Queen's Economics Department Working Paper No. 1047

Do Peers Affect Student Achievement in China's Secondary Schools?

Weili Ding
Queen's University

Steven F. Lehrer
Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

3-2006

DO PEERS AFFECT STUDENT ACHIEVEMENT IN CHINA'S SECONDARY SCHOOLS?

Weili Ding and Steven F. Lehrer*

March 1, 2006

Abstract

Peer effects have figured prominently in debates on school vouchers, desegregation, ability tracking and anti-poverty programs. Compelling evidence of their existence remains scarce for plaguing endogeneity issues such as selection bias and the reflection problem. This paper is among the first to firmly establish the link between peer performance and student achievement, using a unique dataset from China. We find strong evidence that peer effects exist and operate in a positive and nonlinear manner; reducing the variation of peer performance increases achievement; and our semi-parametric estimates clarify the tradeoffs facing policymakers in exploiting positive peers effects to increase future achievement.

* Both authors are affiliated with the School of Policy Studies and Department of Economics, Queen's University.

We are grateful to Dennis Epple and Thomas Rawski for detailed comments, suggestions and helpful discussions which have greatly improved this paper. We also wish to thank Martin Gaynor, two anonymous referees, seminar participants at the Hong Kong University of Science and Technology, Santa Clara University, University of Pittsburgh, the 2002 North American Summer Meeting of the Econometric Society and the 2003 Canadian Economics Association Meetings for helpful comments and suggestions. Ding wishes to thank the University of Pittsburgh China Council for research support. Lehrer wishes to thank SSHRC for research support. We are responsible for all errors.

JEL Codes: I2, Z13 and P36.

1 Introduction

A revival of interest has developed in recent years among social scientists to ascertain whether non-market (or social) interactions affect a variety of individual outcomes. These interactions demonstrate that interdependence among individuals may go well beyond their responses to private economic incentives such as price and income. The impact of these interactions upon outcomes of interest have been given among other names "social capital," "contagion," "neighborhood effects" and for the focus of the present paper "peer group effects." For example, each student influences her classmates not only through knowledge spillovers and how teachers respond to her, but also in how she affects classroom standards. A less disciplined student is more likely to disrupt her classmates, forcing the teacher to devote more time in class to disciplining rather than transmitting knowledge. Therefore a student's performance in school may be influenced by the characteristics and behavior of her peers. If these peer group effects are substantial, government policy may exploit them by optimally grouping students in different classrooms to achieve desired socioeconomic outcomes.

The rekindled interest in non-market interactions is motivated in part by recent theoretical developments in the economics literature which demonstrate that the composition of peer groups affect the optimal organization of schools, neighborhoods and other arenas where agents interact. Recent theoretical analyses of school choice showing that the introduction of vouchers could increase competition and improve the efficiency of public

schools rely on the existence of positive peer effects (e.g. Epple and Romano (1998), Caucutt (2001), and Nechyba (2000)). Further, a number of recent macroeconomic growth models (e. g. Benabou (1996) and Kremer (1993)) depend on certain form of peer effects for their optimal organization of neighborhoods and jobs of different skill levels respectively. Peer group effects have also played a prominent role in policy debates concerning ability tracking, school desegregation, affirmative action and anti-poverty programs in both rural areas and urban ghettos.

While substantial theoretical developments have been made recently on the effects of peers (and other social interactions), there does not exist much compelling empirical evidence on whether and how peer group characteristics affect a variety of student outcomes.[1] Empirical analyses on the effects of peer groups have been plagued by various conceptual and data problems.[2] Many commonly used datasets suffer from significant attrition bias as they fail to track those students who switched in between and out of schools, which results in choice based samples.[3] Moreover, such datasets often lack information on important factors such as teacher quality and student innate ability. The analyses are further complicated by various endogeneity issues such as the reflection problem and selection bias. A reflection problem arises when student and peer achievement are determined simultaneously, which inherently convolutes the measure of peers' influence. Selection bias leads to a correlated unobservables problem when families select into neighborhoods and schools for their children based in part on some unobserved group characteristics they

favor. Thus an observed positive association between one's own outcome and that of the peers may not be causal but due to some unknown factors that affect both peer quality and one's own educational performance. This issue plagues studies using data from the United States and many other countries where parents have much freedom in selecting schools.

Researchers have attempted to overcome the selection bias in one of three ways. The most popular method is the use of instrumental variables (see Evans, Oates and Schwab (1992) and Gaviria and Raphael (2001)) to mitigate the correlation between unobservables and peer variables. But the statistical properties and economic validity of these instruments are of debate.[4] Recently, several studies have exploited credible exogenous variations to identify the effects of peers on student outcomes. Sacerdote (2001) has provided some of the strongest evidence that peer effects exist at the college level by exploiting the fact that roommates are conditionally randomly assigned at Dartmouth college.[5] Hoxby (2000) introduces two empirical strategies based on the assumption that students are randomly assigned across classrooms based on gender and race to recover estimates of peer effects free of selection bias. Alternatively, researchers have acquired or collected rich datasets (see Hanushek, Kain, Markman and Rivkin (2003)) that provide strong controls to capture the selection process. We follow the last strategy and use a unique dataset from a county in China's Jiangsu Province.

The primary goal of this study is to examine whether peer groups affect students'

academic achievement in China's secondary schools.[6] The secondary school system in China differs markedly from that of the United States. Its features and institutional structure yield significant benefits for research aimed at identifying the impact of peer group characteristics on academic achievement. In particular, i) we are able to separate the effect of teacher quality from peer effects, ii) attrition bias is not a concern in our dataset, and iii) students are assigned into schools based primarily on observable test scores. Therefore, the majority of cross sectional variation in a students' peer group is NOT generated by selection on unobservables in our setting. A second aim of the study is to investigate in detail how peer groups affect student achievement and discuss the policy implications.

This paper is organized as follows. In the next section, we describe China's secondary education system and the dataset that we have collected. In section 3 we introduce the model that guides our estimation. We demonstrate that the institutional structure of China's secondary school system in combination with our econometric methodology overcomes most of the problems associated with peer effects estimation. Our empirical results are presented and discussed in section 4. There is strong evidence that peer effects exist in China's secondary schools. We find that students benefit from having higher achieving schoolmates and from having less variation in the quality of peers in their schools. Peer effects operate in a heterogeneous manner: students who score in the top quantiles on the college entrance examination benefit nearly twice as much from their

peers as students who score in the lowest quantiles from similar peers. Our semiparametric estimates unveil the difficulties facing policy makers when they choose educational policies that assigns students of different achievement levels across schools (streaming or mixing) to yield higher future achievement. Finally, we demonstrate that linear peer effects are not identified in a value added model using a fixed effects estimator with a one way error component where the peer variable is defined at the same level as the error component.[7] We introduce an alternative strategy that permits identification of linear peer effects with unobserved fixed error component. A concluding section summarizes our findings and discusses directions for future research.

2 China's Secondary Education System

Our research focuses on secondary schools in one county in China's Jiangsu Province.[8] The generous cooperation of local officials allowed us to collect a unique dataset that follows nearly 1,300 students' academic histories from the completion of junior high school in 1995 through the admission to colleges in 1998. The county has a population of approximately one million in the sample period. It is an affluent county typical of the booming coastal regions of China: the per capita incomes for both urban and rural households were over three times the national averages for each sector in 1997.[9] This data set is composed of individual scores on high school and college entrance exams in all six subject areas, matched teacher data from administrative records for over 1,000 teachers, annual local

government investment (spending) in each school as well as other sources of school input.

In China, entrance examination scores are by far the most important determinant of admission to both senior middle schools (10^{th} grade to 12^{th} grade)and colleges.[10] Within each county, there is a clear ranking of high schools and there is keen competition to attend higher ranked schools. It is common knowledge among the population whether a school is a national model school, a provincial keynote school or a school that focuses on teaching students trade skills. Students compete for positions in the higher ranked schools by writing a municipal level high school entrance examination at the completion of junior middle school (7^{th} grade to 9^{th} grade).[11] This three-day examination covers material in six subject areas.[12] Since 1995, the entrance examination have underwent continuous re-design to reduce the effectiveness of drilling and teaching towards the test.[13] The subject areas are Chinese, English, mathematics, physics, chemistry, and politics. Scores on this exam determine high school admission. Exceptions fall under the category of showing exceptional ability in a subject area, fine arts, music or athletics.

In their final year of senior middle school (12^{th} grade), college bound students list their preferences for various majors at colleges and universities, which is typically followed by a three-day nationwide college entrance examination that encompasses material in six subject areas.[14] There are two versions of the college entrance exams. One is for students wishing to major in the arts and is composed of questions in Chinese, English, mathematics, geology, history, and politics. The other, for science majors, covers material

in Chinese, English, mathematics, physics, chemistry and biology. Both exams were scaled out of 750 in 1998. The scores on these exams together with the preference lists, determine which major of a university or college the student can attend.

In China, the competition for positions in colleges or universities is very keen. In 1999, only 10.7% of the population between 18-25 were either in the process of completing or had obtained a tertiary degree.[15] Since there are many more applicants than there are positions available in colleges and universities this may result in some strategic choices made by the students when filling out their preference lists.

We have collected the cumulative scores on the municipal high school entrance examination for all junior middle students graduating in 1995. For eleven of the sixteen senior middle schools in the county we know of the junior middle school each incoming student was from and have indicators if the student was admitted based on exceptional skills in art, music or athletics (in these situations admission test scores play a smaller role). The records also indicate that some students were admitted prior to the entrance exam based on their strong academic records in junior high schools.[16] These records further provide us with several demographic variables about each student and their family. Finally, we have collected the scores on the 1998 National College Entrance Examination for the same cohort of students. The scores obtained on these exams are almost the sole determinant of admission to college. For each of these students, we are aware of the particular major of a college or university which granted admission. This information is matched with the

incoming test scores so that we are able to follow nearly 1300 students from completion of junior high school through admission to a tertiary education institute.

We have collected information on school inputs from the administrative records of ten of the county's sixteen secondary schools from 1995 to 1998. These records provide us not only with individual level information on each teacher's demographic variables, education and teaching history, but most importantly contain information on a measure of teacher quality. In China, the local education bureau assesses each teacher using both objective and subjective performance measures of factors including classroom instruction, work ethics and student performance. Based on these assessments, teachers are ranked within the education system and can be promoted from intern (newly hired) to third-class, second-class, first-class and finally superior-class teacher. These rankings together with years of teaching experience uniquely determine components of teachers' salaries.[17] This information is matched with the student data so that we obtain accurate measures of school inputs.[18]

Table 1 presents summary statistics on the characteristics of students and school inputs available to students at schools of different rankings in 1998. Notice that schools that are nationally or provincially ranked have a higher percentage of instructors who are in the superior class. Interestingly, average teaching experience is lower in these ranked schools than in the unranked schools. This occurs since ranked schools can attract younger teachers with higher degrees. Teachers in these ranked schools are also more likely to

teach the subjects in which they hold their degrees. It appears that teachers in the national model school teach fewer sections and courses. Test scores for the incoming classes are higher and variation in these scores lower in national and provincially ranked schools versus the unranked schools. In particular, there is little variation in performance at the national model school. Students at the national model school are most likely to obtain tertiary education and at the Bachelor's level. In most schools the majority of students are accepted into college certificate programs. Due to incomplete data on college enrollment in school 7 it appears that school 2 and school 4, which are not ranked, place a greater percentage of students in tertiary institutions than this provincially ranked school.

The local government invests substantially more funds in national and provincial model schools than in unranked schools. This is not surprising since local Chinese governments usually associate economic growth with the education level of its population. Annual reports even advertise the number of superior teachers in the local school system to attract outside investment. Furthermore, nationally and provincially ranked schools tend to receive more external donations than unranked schools, which exacerbates the inequities across the schools.[19] On our visit, we observed that ranked schools tend to have more modern facilities. Overall, this data assists greatly in estimating education production functions since we have i) valid outcome measures that accurately proxy achievement due to the comprehensiveness of the mandatory exams in different subject matters and strong incentives for students to perform to their best on the exam, ii) no attrition bias as not

a single student left the sample, iii) reduced omitted variable bias problems as we have actual measures of teacher quality and iv) complete information on the format of the school assignment algorithm.

3 Empirical Model

In this section, we provide a simple model that guides our estimation and describe how our dataset aids in overcoming many problems associated with peer effects estimation that have plagued empirical researches. We view the education process at period T in two stages. In stage 1, a student is matched to a school. Denote S_{iT} to be the school matched to child i at time T. To distinguish our empirical strategy from those commonly used in the literature we consider two matching processes. A school *selection* process takes place if the education system does not permit schools to exclude students based on their qualities (achievement, abilities, etc.). In this case, altruistic parents select the optimal school j^* for child i ($S_{iT} = j^*$) which provides the highest indirect utility for their household V_{ij}^*,

$$V_{ij} \equiv V_{ij}(X_i, C_j, P_j, Q_j, A_{iT-1}, I_i), \text{ for each } j \text{ available to child } i \tag{1}$$

where X_i are observable person-specific and family characteristics of the child i; C_j is the cost of attending school j, which include the cost if living in a good school district; P_j captures the characteristics of peers in school j, assuming parents have perfect foresight

at stage one (or peer characteristics do not vary much over time); Q_j is school-specific characteristics; A_{iT-1}indexes child i's measured achievement at the stage of decision making; and I_i is child i's innate abilities. The availability of schools to a child is described by the school admission rules in the local areas where parents can commute to work daily.

When schools admit students based on their qualities, a school *assignment* process takes place. School j has the jth highest minimum quality requirement, τ_{jT}, for admission if we sort schools based on some quality function $Q_{iT} \equiv q(I_i, A_{iT-1})$ that maps each child's abilities and past achievement into a quality score. Each child i is assigned as follows:

$$S_{iT} = j \text{ if } \tau_{jT} \leq Q_{iT} < \tau_{j-1T}, \tag{2}$$

In the second stage, conditional on the selection or assignment of school j at the first stage, the complete history of inputs $[(X_{iT}...X_{i0}), (P_{-ij_T}...P_{-ij_0}), (Q_{j_T}...Q_{j_o})]$, and independent random shocks $(\epsilon_{iT}...\epsilon_{i0})$, the child gains knowledge as measured by a score on an achievement test:

$$A_{ijT} = f(X_{iT}...X_{i0}, P_{-ij_T}...P_{-ij_0}, Q_{j_T}...Q_{j_o}, I_{i,}\epsilon_{iT}...\epsilon_{i0}) \tag{3}$$

Linearizing the above achievement relationship (equation 3) yields

$$\begin{aligned} A_{ijT} &= \beta_{0T}+\beta_{1T}X_{iT}+\beta_{2T}P_{-ij_T^*}+\beta_{3T}Q_{j_T^*}+\beta_{4T}I_i \\ &\quad +(\sum_{t=0}^{T-1}\alpha_{0t}+\alpha_{1t}X_{it}+\alpha_{2t}P_{-ij_t}+\alpha_{3t}Q_{j_t}+\alpha_{4t}I_i+\delta_{it}) + \epsilon_{iT} \end{aligned} \tag{4}$$

where $\delta_{it} = \alpha_{5t}\varepsilon_{it}$ for some coefficient α_{5t}. The components of equation 4 may include higher order and interaction terms.

Following, Boardman and Murnane (1979) we re-express the achievement function as

$$A_{ijT} = \beta_0 + \beta_1 X_{iT} + \beta_2 P_{-ijT} + \beta_3 Q_{jT} + \beta_4 I_i + \beta_5 A_{ijT-1} + \epsilon_{iT} \tag{5}$$

where the vector X contains individual characteristics (gender interacted with the residential status of each student's family where males from urban areas are treated as the comparison group), the vector I contains indicator variables for individual skills that capture portions of innate abilities and P_{-ijT} is a vector of variables that captures the prior achievement of all other students in the same cohort within that school. The inclusion of A_{ijt-1} in the regression is to pick up a variety of confounding influences including the prior, and often unrecorded as well as unobserved history of parental, school and community effects. The implied restrictive assumption for this type of model is that A_{ijt-1} is a *sufficient statistic* of all the previous influences. This assumes that A_{ijt-1} is a state variable that follows a Markov process.[20] Note, in our baseline specification we include the mean, the mean squared and the variance of the peers' high school entrance examinations scores.

3.1 Estimation

The general empirical challenge facing researchers in estimating problem equation 5 is endogeneity of the peer variable which can arise and be formulated in more than one manner. In the economics of education literature, the most discussed form of endogeneity is the endogenous group membership problem. When public school admission decisions are

mostly based on residential location, the conditional expectation of the error term when school j is chosen , $E(\epsilon_{ijT}|S_{iT} = j \ if \ V_{ijT} >= V_{ikT} \ for \ all \ k \neq j)$, does not typically equal zero.[21] Since V_{ijT} can include many such factors to decision making that are similar within a community and not typically contained in a dataset, concerns regarding biased coefficient estimates become serious.

In contrast to the typical US experience of a school selection process, a school assignment process based primarily on characteristics observable in our data matches students to senior middle schools in China. Students can attend a given school from anywhere within the county and dormitories are provided at very low cost to accommodate students from neighboring communities or rural areas. For this reason there is no need for people to form communities within this county based on characteristics that henceforth influence school composition.[22] This result is originally pointed out by Barnow, Goldberger and Cain (1981) : "Unbiasedness is attainable when the variables that determined the assignment rule are known, quantified and included in the [estimating] equation." Let S^C_{iT} to indicate the school assigned to child i at time T in China. In this setting if we consider the effects of school assignment on student achievement (equation 5), the conditional expectation is given as:

$$E[A_{ijT}|S^C_{iT} = j] = \beta_0 + \beta_1 X_{iT} + \beta_2 P_{-ijT} + \beta_3 Q_{jT} + \beta_4 I_i + \beta_5 A_{iT-1} + E(\epsilon_{ijT}|S^C_{iT} = j) \tag{6}$$

$$= \beta_0 + \beta_1 X_{iT} + \beta_2 P_{-ijT} + \beta_3 Q_{jT} + \beta_4 I_i + \beta_5 A_{iT-1} + E(\epsilon_{ijT}|\tau_{jT} \leq q(I_i, A_{iT-1}) < \tau_{j-1T}) \tag{7}$$

Since all the variables to the right of the conditional operator are known, and all the schools follow the same deterministic assignment rule, $E(*) = 0$, a simple regression would obtain unbiased parameter estimates.

To determine admission each senior middle school is provided with each student applicant's high school entrance examination and information on exceptional skills from the local education bureau. The files are sent in a sequential manner ensuring that the higher ranked schools have the privilege of recruiting the best students.[23] Administrators at each senior high school convert this information to a single index and grant admission to students whose index is above a cutoff score. To verify whether students were truly assigned to type of schools based on the observables in our data we used an ordered probit regression of the rank of the school (unranked, affiliated with the ranked schools, provincial school and national school) on the incoming test scores and indicator variables for early admission or talent in subject areas. We then used our estimates to construct the predicted probability that each student was assigned to each school type. Assuming that each student attended the school type to which she had the highest predicted probability we compared these predicted assignments to the actual assignment. In total we find that 96.67% of the matches are accurate,[24] leaving little room for any unobserved factor to potentially bias the estimates of equation 5.[25]

Simultaneity problem in peer effect estimation was first coined the "reflection problem" in Manski (1993), as it is similar to the problem of interpreting the almost simultaneous

movements of a person and her reflection in a mirror. The problem occurs if the peer variable measures classmates' achievement at time T, which is obtained at the same time as one's own achievement. In our education production function, all the regressors are known (predetermined) at the time of regression, which in theory avoids simultaneity problems. Moreover, our peer variable is constructed using academic scores measured immediately prior to any interactions among classmates, which captures the background of one's peers. We argue that a student's current performance is not a function of the current outcomes of his or her peers, but rather depends on peer quality used to form the student body. Manski notes that if the transmission of peer effects really follows this temporal pattern, the identification problem is alleviated.

4 Empirical Results

4.1 Do Peer Effects Exist?

Ordinary least squares estimates of our baseline equation 5 are reported in Table 2. The first column presents the results for the full sample. With the exception of the tuition variable, each regressor enters in a statistically significant manner.[26] The results suggest that a 1% increase in peer quality will increase one's college entrance examination score by 0.088%. In other words, the effect of an increase in peer's incoming test scores is equivalent to about one twelfth (0.078%) of the effect of a corresponding increase in

one's own incoming test score. The sign on the peer squared term suggests a concave relationship in peer effect. Furthermore, the results suggest that individuals respond negatively to an increase in the variation of their peers' performance.

Not surprisingly, winning an academic award or being awarded early admission is positively related to future student test performance, while skills in music and athletics are correlated with poorer performance. Urban females score lower than rural females. This may be explained by the greater accessibility to high school education for urban females than their rural counterparts as the negative impact of being a female may not affect a top student as much as it affects an average student.

The school quality variables included are mainly teacher quality variables. The estimated coefficients in both regressions indicate that students benefit most from higher ranked teachers, which suggests the importance of teacher inputs.[27] The natural collinearity between the school quality variables may explain the implausibly large magnitude of the associated coefficients. It is worth noting that even with the removal of the school quality variables a concave relationship in peers exists. The results suggest that a 1% increase in the quality of peers will increase one's score on the college entrance examination by 0.41%. This estimate provides an upper bound for the effect of peers since school quality is positively correlated with peer quality. The finding that second class teachers benefit students slightly more than superior teachers is driven by the provincial school 7, which has many young teachers ranked at the second class. If that school is removed

from the analysis the results correspond to our priors with the ordering of the effect of each type of teachers consistent with the definition.

We examined how good the measures of teacher qualities are at explaining the differences in between schools other than the peers. We find that having a higher percentage of superior and first class teachers accounts for $35 - 50\%$ of the variation in the school fixed effect. We explored the relationship with extrinsic measures of teacher quality such as highest degree attained and teaching experience and find that they could only explain $5 - 10\%$ of the variation. This result is consistent with other studies examining school quality which suggest that a good measure of teacher quality is essential to any such studies.[28] In fact replacing the teacher quality variables with these extrinsic measures suggested that a 1% increase in the quality of peers will lead to an implausible 0.80% increase in one's score on the college entrance examination. This suggests that the extrinsic measures may be negatively correlated with actual teaching quality.

The second column of Table 2 reports the results from exploiting the natural regression discontinuities created by the entrance exam cutoff rules for entry into different ranked secondary schools.[29] This addresses the concern that our ability, past achievement and demographic variables may not capture all the individual differences across students that affects future academic outcomes. We compare the outcomes for students who scored similarly (within 20 points or 2.9%) and tightly around the cutoff scores of the provincial model schools on the high school entrance examination. Within a tight range of 20 points

on a 685-point exam, the academic abilities and achievements of the students are hardly discernibly different, but they were enrolled in different schools with different peers due to keen competition described in section 2. The use of this sample is similar to a natural experiment in that most of the difference between college entrance examination scores of this group of students who attend different schools must be attributed to differences in the behavior and composition of the peer groups and differences in school inputs.

In this subsample students scored similarly and within a twenty point interval on the high school entrance examination.[30] Those who attended provincial schools in this sample entered with an average score of 579.31 (standard deviation =7.92) which was less than one point higher than students in the lower ranked schools (578.46 with standard deviation= 7.92). Yet these students had access to higher quality peers (and different school resources[31]) and scored an average of 20 points (approximately 3% higher) greater on the college entrance examination[32]. Mann Whitney tests confirm that this difference in performance on the college entrance examination is statistically significant (z = -3.036, Prob > |z| = 0.0024).

The OLS estimates for this subsample are carried out on 65% fewer students and two fewer schools and as a result peer quality has significantly less variation within and between schools. The peer variance continues to be negatively and significantly related to higher scores on the college entrance examination, while the signs on the mean peer variable and the quadratic term have changed from those obtained in the full sample.

The marginal peer effect remains positive and is significantly larger for this subsample (0.518), with a convex relationship on peer effects. Since the tuition variable is a perfect predictor of attending provincial schools it is not included in the analysis.

4.2 An Examination of Heterogeneous Responses

The concavity in the full sample and the convexity in the subsample suggests that heterogeneity in students' response to their peers may exist. To get a complete picture of the heterogeneous responses, we employ a partial linear model where the mean peer variable is entered as a nonparametric function to be estimated instead of a simple quadratic function. The nonparametric estimate would not only explain the difference between the shapes in the full sample versus the subsample but would also lend evidence to policy discussions on whether to "stream" or "mix" students of different abilities. Streaming is to place students of similar abilities into the same group while mixing is to randomly assign students into groups based on their abilities.[33]

A partial linear model allows us to enter the mean peer effect variable in a nonparametric fashion. To estimate a partial linear model we modify equation 5 as follows

$$A_{ijT} = \beta_0 + \beta_1 X_{iT} + h(P_{-ijT}) + \beta_{2v} VP_{jT} + \beta_3 Q_{jT} + \beta_4 I_i + \beta_5 A_{ijT-1} + \epsilon_{iT} \qquad (8)$$

where h is a nonparametric function of the mean peer variable to be estimated and VP_{jT} is the variance of the peer variable. As before, we include the peer variance as well as the remaining educational inputs as linear regressors. If a concave relationship is

estimated this would suggest that to increase total academic achievement, mixing students of different abilities is preferred.[34] Similarly a convex relationship provides support for streaming students. Entering the peer variable in a nonparametric manner will improve parameter estimates of the individual $\beta's$ if the peer variable is stochastic or if peers affect achievement in a manner that can not be captured by only two terms. Since improved parameter estimates are obtained this has a secondary benefit of improving the estimation of the shape of the peer effect function: $h(P_{-ijT})$.[35]

The nonparametric estimates of peer effects are presented in Figure 1. The shape is statistically different from a horizontal line which further verifies the existence of peer effects in China's secondary schools. The graph of the conditional achievement on average peer quality starts as concave and alternates between convexity and concavity as peer quality increases. Moreover, neither a concave or convex figure could lie within a 90% confidence interval of the estimated curve. The results explain the different shapes we obtained using the full sample and subsample near the discontinuity. Further, they demonstrate the dilemma educational policy makers may face in determining whether streaming or mixing students will obtain higher future average achievement. Students at the low end of the conditional achievement benefit substantially from mixing with better peers whereas students at the high end gain from streaming, while students in the middle may prefer to be mixed with certain peers but streamed from other peers. The decision on streaming or mixing may depend not only on the particular group of students involved

but also on policy makers' preferences in trading off achievements of different groups of students, which is beyond the scope of positive economic analysis.[36]

The above findings have focused on how peers affect average performance. The impact of observed factors may vary over the distribution of unobserved factors (i.e. ability) that affect achievement.[37] To examine this hypothesis we reestimate equation 5 via quantile regression using quantiles ranging from 0.05 to 0.95 at each 0.05. Quantile regression provides a more flexible approach to characterizing the effects of observed covariates such as peer groups on different percentiles of the conditional achievement distribution providing a richer pattern on how peer groups affect student achievement. Implicitly we are allowing peer groups and ability to be two separate factors in the generation of achievement (equation 4) to interact in unknown ways . If ability and peer groups are substitutes we would expect the marginal returns on peers to decrease with ability. If ability and peer groups are complements then marginal returns to peers would be higher for the more able.

Figure 2 presents the results corresponding to our baseline specification for regressors such as the student's own incoming test score and peer effect variables. The solid line in each graph represents the ordinary least squares estimate of the mean effect. Each triangle represents one of the 19 quantile regression coefficient estimates. The effect of incoming test scores presents an inverted U graph where the largest gains are for those students in the middle quantiles.

While the mean peer variable demonstrates that the strongest gains accrue to those in the lowest quantile, the quadratic peer variable shows that these individuals also achieve the smallest benefits to achievement. Without common support of the peer variables across different quantiles, we cannot calculate comparable marginal effects as a common set of peer values are needed. If we assign the average of all peers as the common peer, a graph of the marginal effect across quantiles demonstrates the largest gains from having higher quality peers appear in the middle quantiles. Students in the lowest quantiles gain the least-only half of what the students gain in the higher quantiles. Finally, all individuals benefit from less variation in their classmates' performance, although those at the higher quantiles are not as negatively affected by increases in this variation.

It is clear that the heterogeneity of peer group effects operates in at least two dimensions. First, the quantile regression results demonstrate that students of different unobserved abilities respond to the same peer differently. and we are able to reject the hypothesis of homogeneity in the returns to peer groups. Second, estimates from the partial liner model demonstrate that same student will respond in a heterogeneous manner to different compositions of their peer groups.

4.3 Are the results robust?

To address the concerns (despite the advantages of the data and the controls we employ) that there may still exist some unobserved or omitted differences across schools that drive

the results in Tables 2, we consider fixed effect estimation. This can be accommodated in our framework by relaxing the assumption that the school quality variables are measured perfectly (Q^t_{jT}) and allow for the possibility that they are measured with error (Q^m_{jT}) as given by the relationship below:

$$Q^m_{jT} = Q^t_{jT} + u_{jT} \tag{9}$$

If we were to allow the random noise term (u_{jT}) to be correlated with some exogenous regressors in equation 5 and employ a fixed effects estimator (which are regularly used in empirical studies of peer effects) we are unable to uniquely identify the mean peer effect coefficient.[38] The coefficients of nonlinear peer variables can still be identified. To demonstrate this consider substituting equation 9 into equation 5,

$$A_{ijT} = \beta_0 + \beta_1 X_{iT} + \beta_2 P_{-ijT} + \beta_3 (Q^t_{jT} + u_{jT}) + \beta_4 I_i + \beta_5 A_{ijT-1} + \in_{iT}$$

Rewriting this equation at the mean level for each school j, we get

$$\begin{aligned}
\overline{A_{ijT}} &= \beta_{0j} + \beta_1 \overline{X_{iT}} + \beta_2 \overline{P_{-ijT}} + \beta_3 Q_{jT} + \beta_4 \overline{I_i} + \beta_5 \overline{A_{ijT-1}} + \overline{\in_{iT}} \\
where\ \beta_{0j} &= \beta_0 + \beta_3 u_{jT}, P_{-ijT} = \frac{\sum A_{-ijT}}{N_j - 1} \text{and} \\
\overline{P_{-ijT}} &= \overline{A_{ijT-1}} \text{ if } P_{-ijT} \text{ is the mean peer effect.}
\end{aligned}$$

The fixed effect regression is carried out on the following equation

$$\begin{aligned} A_{ijT}-\overline{A_{ijT}} &= \beta_1(X_{iT}-\overline{X_{iT}}) + \beta_2(P_{-ijT}-\overline{P_{-ijT}}) + \beta_4(I_i-\overline{I_i}) \\ &\quad +\beta_5(A_{ijT-1}-\overline{A_{ijT-1}})+(\epsilon_{iT} - \overline{\in_{iT}}) \\ &= \beta_1(X_{iT}-\overline{X_{iT}}) + \beta_2\frac{-(A_{ijT-1} - \overline{A_{ijT-1}})}{N_j - 1}+\beta_4(I_i-\overline{I_i}) \\ &\quad \beta_5(A_{ijT-1}-\overline{A_{ijT-1}})+(\epsilon_{iT} - \overline{\in_{iT}}) \end{aligned} \tag{10}$$

Notice that β_2 cannot be identified separately from β_5 due to perfect correlation between the transformed mean peer variable and one's past achievement as $(N_j - 1)$ does not vary for each individual. The mean peer effect can not be identified in this setting. Further, if one were to use current test scores to define the peer variables as considered in several studies (recall this also introduces a reflection problem), the transformed peer effect variable would perfectly explain the transformed dependent variable.

The coefficients on the nonlinear peer variables can still be "identified" in these fixed effect regressions as such perfect co-movement does not apply to the quadratic or the variance terms.[39] However only qualitative interpretation of these coefficients is meaningful as the marginal effect cannot be calculated. Moreover the results are still sensitive to any variation in school sizes and may suffer from lack of variation within schools.

A strategy that helps us to identify whether mean peer effects exist when the random noise term, u_{jT}, is correlated with some exogenous regressors in equation 5 is as follows. We randomly assign with replacement each student a classmate from the same school.

We then recalculate the peer effect variable for each student omitting oneself as before but **also** the randomly assigned classmate from that school. This serves to break the perfect correlation between the mean peer and earlier achievement regressors to achieve identification, although multicollinearity arises instead which renders mean peer estimate imprecise. Since this newly constructed peer variable can be viewed as if it were measured with error, the estimates are biased towards zero in absolute value and statistical rejection.

Column 1 of Table 3 presents summary results from one hundred fixed effects regressions with different classmates randomly removed with replacement. The average of the estimated peer effect coefficients, while large in scale due to the limited variation in the peer variables within a school, are exactly the same in terms of signs and statistical significance as those in OLS and random effects. This renders it unlikely that something unobserved or omitted is driving the results, at least qualitatively. We find that the variance of the peer quality is negatively related to student performance at the 1% level in each of the 100 individual regressions. We also find that in 72 of the 100 individual regressions both the mean peer and the quadratic peer are statistically significantly and in a concave manner. Moreover, the three peer variables are always jointly significant in all the regressions. Since the design of our identification strategy may bias these coefficients towards zero in absolute value and statistical rejection, the findings strongly reinforce the existence of peer effects in secondary schools. Further, they demonstrate conclusively that the negative and statistically significant effects of higher order moments of the peer

variable on student achievement are not driven by omitted school characteristics.

Column 2 of Table 3 presents fixed effects regression results for the subsample near the entrance cutoff discontinuity.[40] Since we have only 406 students in seven schools consistent estimates can not be obtained due to the incidental parameters problem.[41] To remedy the situation we run school type fixed effects (provincial vs. lower ranked schools). In the first column we omit teacher quality variables which vary across schools within school types. The pattern across both columns is consistent with the OLS results. Peer effects enter significantly in a convex manner and the peer variance remains negatively and significantly related to higher scores on the college entrance examination.

5 Conclusions

In this paper, we employ a unique and rich dataset from a county in China's Jiangsu province to investigate whether peer group effects exist. The secondary school system in China differs markedly from that of the United States and its features and institutional structure yield significant benefits to identify the impacts of peer group characteristics on student achievement. In particular, i) we are able to separate the effect of teacher quality from peer effects ii) attrition bias is of little concern iii) we have valid outcome measures that are accurate and comprehensive and iv) students are assigned into schools based primarily on observables. Therefore, the majority of cross sectional variation in students' peers is not generated by self-selection in our setting. Our methodology overcomes the

reflection problem and deals with the possibility that the school quality variables may be measured with error.

We find strong and robust evidence that peer effects exist. We find that peer groups contribute significantly and positively to student performance measured by test scores. Our evidence is strongest in showing that peer groups operate in a nonlinear manner and all individuals respond negatively to an increase in the variation of peer quality. The marginal effect of a one percent increase in the quality of peers on student achievement is equivalent to between $8-15\%$ of a one percent increase in one's own earlier achievement.

We find that peer effects operate in a heterogenous manner. High ability students benefit more from having higher achieving schoolmates and from having less variation in peer quality than students of lower ability. Further, our semiparametric estimates clarify the trade-offs facing educational policy makers in constructing a policy that exploits the existence of peer effects to increase future average test scores.

Finally, we demonstrated that the magnitude of our results are sensitive to the school quality variables included in the estimating equation. The plausible economic significance of the peer effects relies on the use of teacher quality variables that are strongly correlated with actual teaching performance.

While we argue that our data and methodology overcome the majority of statistical problems associated with peer effects estimation, we are aware of the limitations of this study. First, there is substantially less variation in student ability in China than that

in the US public high schools. Second, the county is only representative of the affluent coastal regions of China.

To summarize, this study provides the most compelling evidence to date on the existence of peer effects at the secondary school level. The institutional structure of China's secondary education system mitigates traditional biases related to selection of schools and attrition assisting in identifying the role of school inputs including peer groups. All students appear to benefit from having higher achieving schoolmates in a non-linear manner and from having less variation in the quality of students within their schools. Heterogeneity arises from students of different abilities responding to same peers differently and same students responding to different peer qualities differently.

Notes

[1]The literature on peer group effects in education dates back to the publication of the Coleman Report (Coleman et al. 1966). Numerous studies have since been conducted in both the economics, sociology and education literatures which find either weak or non-existent peer effects. Recently, Hoxby (2000) and Sacerdote (2001) found positive peer effects employing sources of exogenous variation in peer group composition which are outlined later in this section.

[2]See Manski (1993), Manski (2000) and Moffitt (2001) for comprehensive discussion.

[3]This is especially problematic for studies using metropolitan US datasets, where approximately 20% of the students are believed to switch in between and out of schools annually. For example, Hanushek, Kain and Rivkin (2004) report that in Texas approximately 50% of children switch schools at least once between grades four and seven even after excluding changes due to the transition from elementary to middle school.

[4]Heckman (1997) considers the economic interpretation of instrumental variable estimators in general. The statistical concerns extend beyond the weak instrument problem. For example, Rivkin (2001) argues that the use of metropolitan wide aggregate data as instruments in the Evans, Oates and Schwab (1992) study exacerbates the specification error in peer effect estimates.

[5]This methodology has also been used in Zimmerman (2003) and Kremer and Levy (2003) who use data from different selective higher education institutions where roomates

and dormmates are conditionally randomly assigned.

[6]To the best of our knowledge this is the first such study using micro data from China and the developing world. See Hanushek (1995) or Kremer (1995) for recent surveys of school efficacy in developing countries.

[7]If measures of peers are taken at the school (classroom) level, the coefficient on mean peer effects would not be identified if school (classroom) indicators are included in a fixed effects regression with two periods of data per student using a levels or linear growth specification of the education production function.

[8]Per our agreement with the local government we do not identify the county by name.

[9]Source: County Statistical Yearbook on Education, 1995-2000.

[10]Current education reforms in China aim at the ultimate elimination of high school entrance examination as educators argue that it overburdens students.

[11]A municipality in China differs markedly from one in North America. There are several counties and one city within a municipality, which bears similarity to a Metropolitan Statistical Area (MSA) in the US.

[12]These six subject areas account for a total of 640 points. The remaining 45 points are scores based on physical fitness. Rosen (1982) explains that the introduction of admission examinations resulted in students spending an inordinate amount of time learning facts they became less politically and physically active. Beginning in 1981, the criteria used for admission changed to include athletic ability since many argued that a student needs

a well-balanced education intellectually, morally and physically.

[13]Discussions with teachers and local education bureau officials gave us a strong impression that educators are increasingly aware of these changes and are feeling the pressure from such changes.

[14]There are some areas in China where students do not compete in this manner. Beginning in July 2002, college entrance examination will be restricted to four subject areas instead of six in many parts of China to reduce the burden imposed on students.

[15]Further, this figure comes after the government introduced a massive program boosting the number of admissions in colleges by mandating colleges to accept more students. Source: *Zhongguo Tongji Nianjian* (China Statistical Yearbook) 2000.

[16]These students still complete the high school entrance examination and are encouraged to perform to the best of their abilities. This occurs frequently as the inclusion of these students will raise the average test score for the corresponding junior middle schools. It is clear that these students do not face the same incentives as other test takers.

[17]See Ding and Lehrer (2002) for a detailed explanation and analysis of China's teacher compensation system.

[18]Class sizes average 52 - 56 students in the data. There is very little variation in class size within and across schools. We are able to match students in four of the nine schools to their teachers and school investment data. The local education bureau claimed that since teacher rankings were based in part on student performance, students were quite

randomly assigned across classrooms for fairness purposes. On our return trip to China this spring we collected classroom data. We verified the claimed randomization using simple regressions and tests for differences between classes.

[19]Accurate information on the exact size of these external donations was not recorded by either the school board or individual schools.

[20]This assumption requires that the effect of observed and unobserved factors in the production process to be adjusting over time at the same rate as no past inputs and shocks are left unrepresented by A_{ijt-1}. Due to the limitation of almost all education data sets, this assumption is adopted for practical reasons and has not been tested to the best of our knowledge.

[21]In this case when a school selection process matches children to school, parents who care about child's education are willing to pay higher housing prices to move to a good school district (see Black (1999)) and are more likely to engage in actions to improve their child's performance.

[22]Social network (*Guanxi)* in China plays a limited role in determining who gets access to the better senior middle schools. School assignment based on public information (test scores) reflects in part the strong common preference for fair assignment process, which is monitored by each family whose child is concerned. In counties such as the one studied in this paper where extreme importance is attached to education, the scope with which network can affect the assignment process is minimized. Moreover, any back-door admis-

sions are generally rationalized as if the student had exceptional ability in a particular subject area, fine arts, music or athletics, for which we have indicators to control.

[23]Students are restricted to apply to one school at each quality level, a factor that we exploit in one of our empirical strategies described below. The philosophy behind national and provincial model schools is that they are necessary to effectively train top-level manpower needed for China's development. In certain parts of China, such as Shenzhen, Xiamen, Shanghai and Changchun such schools have drawn criticism for being elitist and have been abolished conceptually.

[24]We received unofficial formulas that confirmed that assignment was based solely on these characteristics and they were treated in a linear fashion with weights. The coefficient estimates are available from the authors by request. The concordance rates between predicted and actual assignment by school type were 97.55%, 89.74%, 99.19% and 93.7% respectively.

[25]Potential candidates in our setting for such unobservables may exist due to parents' preference to be geographically close to their child or the selection process by the student prior to the high school entrance examination. To simplify the high school admission process, students are required to choose one of the three provincial model schools as their most preferred school before their High School Entrance Examination. Although all the provincial schools are of the same rank they may have different cutoff scores for admission due to the number of applications. Thus a student with score 588 may attend a lower

ranked school if his preferred provincial school has the cutoff score at 590, even though his score exceeded the cutoff of another provincial school (say 585). The choices of students among the provincial schools reflect their locational preferences and strategic guessing (i.e. which school may have a lower cutoff score), which do not correlate with academic performances.

[26]The bottom 25% of the incoming students in terms of admission scores are required to pay higher tuition fees than other incoming students in each school. The local education bureau claimed that in this county where education is highly valued, almost no one refused to pay and went to either a lower ranked school or dropped out of school. Although there is no selection in tuition payment, those students may face different incentives as their parents may expect more in the results. See Ding and Lehrer (2002) for a discussion of survey evidence which demonstrates that families in urban areas of China rarely refuse to pay higher tuition fees.

[27]The excluded group is the sum of percentage of third class and new teachers. Nearly all new teachers are promoted to third class within their first year and as such we assume that teaching quality would not differ between these ranks.

[28]This point was originally made in Murnane (1975) and stressed in the influential survey by Eric Hanushek (1986).

[29]Regression discontinuity estimators exploit discontinuities in the treatment assignment mechanism that may develop because of rules governing access as in this study (or

as in Angrist and Lavy (1999) study of the effects of class size on test scores) or geographic boundaries (as in Black (1999) who estimated parents' willingness to pay for higher quality schools). The approach only identifies treatment effects locally at the point where the probability of receiving treatment changes discontinuously. Even though the rule in our case is deterministic we are not provided with the threshold score for each provincial school. Thus, we can not use it as instrumental variable for the probability of receiving the treatment (attending a provincial school).

[30]The scores included are those between 568 to 588.

[31]Mann Whitney tests for this sample demonstrate that there are significant differences in peer ability between provincial and non- provincial schools (z =-20.165, Prob > |z| = 0.00). For the school resources presented in table 1 significant differences between provincial and non provincial school exist only for years of teaching experience (z = 1.640, Prob > |z| = 0.10) and teachers who instructed the subject in which they were awarded their degree (z = -1.94, Prob > |z| = 0.05).

[32]Provincial school students scored 453.48 (s. d.=54.90) versus a score of 431.04 (s.d.= 51.75) for students who attended lower ranked schools in this subsample.

[33]See Hallam (2002) and the references cited in Glewwe (1997) for a recent survey of the ability grouping literature in education and economics respectively.

[34]Glewwe (1997) introduced a non-linear least squares (NLLS) estimator to determine whether streaming is preferred to mixing. This estimation takes account the distribution

of peer group characteristics and draws on a Box-Cox flexible functional form of the mean peer variable. An advantage of this approach is ease of interpretation of the results as the sign of one parameter estimate is enough to determine whether streaming is preferred to mixing. The downside to this approach is that the functional form assumed is still restrictive. This approach assumes that policymakers have a utilitarian welfare function and that individuals only gain utility from the level of achievement they obtain. NLLS estimates with our data (available from the authors on request) suggest that mixing is preferred to streaming. Yet neither a concave or convex function could fit the nonparametric function estimated in Figure 1 even with an addition of a 90% confidence interval.

[35]Methods to estimate this equation are introduced in Robinson (1988). We follow the "double differencing" approach suggested in Robinson (1988) using a local linear regression estimator in place of the traditional kernel regression estimator. As shown in Fan (1992) the bias of the local linear regression estimator does not depend on the design density of the data. Second, the local linear regression estimator avoids the boundary bias problem associated with kernel regression estimators. We experimented with four different bandwidths and there were no major differences in the shape of the Figure.

[36]This requires that a social planner knows the exact form of the social welfare function defined for each possible achievement distribution for the population (i.e. it is strongly separable with respect to each individual achievement value). Whether such a utilitarian social welfare function is the right one for society is a question that is left to the reader

and beyond the current scope of this paper.

[37]In all specifications we assume separability between student characteristics, school quality variables and peer group effects. Quantile regression allows students with different unobserved factors to respond differently to observed factors. If the coefficient estimates differ across quantiles heterogeneous effects are present. Note that at each of the 19 quantiles each of the three peer quality variables were statistically significant.

[38]We have also considered the case where this random noise term (u_{jT}) is normally distributed with zero mean and uncorrelated with all the other exogenous regressors in equation 5 and incorporate a random effects estimator to achieve consistent estimates. The patterns of estimates were similar in magnitude, sign and significance to those obtained with OLS for both the full and subsample. The results (available from the authors by request) suggest that a 1% increase in the quality of peers will increase one's score on the college entrance examination by 0.154%, which is equivalent to more than one seventh (0.151%) the effect of a corresponding increase in one's own earlier achievement score for the full sample and 0.354 for the subsample.

[39]In theory when the full rank condition is violated, identification is not possible. In practice, canned software packages typically achieve identification by dropping one of the perfectly correlated variables or a single observation.

[40]Fixed effects regressions are identified with this subsample since the peer variable is defined over the full sample rather than the subsample in the regression which prevents

perfect correlation. However, multicollinearity may render the estimate on the mean peer variable imprecise.

[41] Asymptotics require either the number of students within a school or the number of schools to go to infinity. There exists substantial Monte Carlo evidence that fixed effects estimates perform poorly when neither dimension is greater than 50. In our subsample the number of students within a school averages 58 with 3 schools having fewer than 40 students.

REFERENCES

Angrist, Joshua D., and Victor Lavy, "Using Maimonides' Rule to Estimate the Effect of Class Size on Student Achievement," *Quarterly Journal of Economics 114:2 (1999), 533-575.*

Barnow, Burt S., Glen C. Cain and Arthur Goldberger, "Selection on Observables," in E. Stromsdorfer and G. Farkas (eds) *Evaluation Studies Review Annual* 5 (1981), 43-59.

Benabou, Roland , "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review* 86:3 (1996), 584-609.

Black, Sandra, "Do Better Schools Matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics* 114:2 (1999), 577-600.

Boardman, Anthony E., and Richard J. Murnane, "Using Panel Data to Improve

Estimates of the Determinants of Educational Attainment," *Sociology of Education* 52:2 (1979), 113-121.

Brock, William A., and Steven N. Durlauf, "Discrete Choice with Social Interactions," *Review of Economic Studies* 68:2 (2001), 235 - 260.

Caucutt Elizabeth, "Peer Group Effects in Applied General Equilibrium," *Economic Theory* 17:1 (2001), 25-51.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York, *Equality of Educational Opportunity* (Washington D.C.: U.S. Government Printing Office, 1966).

Ding, Weili, and Steven F. Lehrer, "Incentives and the Quality of Teachers and Schools," forthcoming in Emily Hannum and Albert Park (eds.), *Education and Reform in China* (New York: Routledge Press, 2002).

Epple, Dennis and Richard Romano, "Competition Between Private and Public Schools, Vouchers and Peer Group Effects," *American Economic Review* 88:1 (1998), 33-62.

Evans, William, Wallace Oates, and Robert Schwab, "Measuring Peer Group Effects: A Study of Teenage Behavior," *Journal of Political Economy* 100:5 (1992), 84-117.

Fan, Jianqing, "Design Adaptive Nonparametric Regression," *Journal of the American Statistical Association* 87:420 (1992), pp. 998–1004.

Gaviria, Alejandro, and Steven Raphael, "School-Based Peer Effects and Juvenile Behavior," *Review of Economics and Statistics* 83:2 (2001), 257-268.

Glewwe, Paul, "Estimating the Impact of Peer-Group Effects on Socioeconomic Outcomes: Does the Distribution of Peer Group Characteristics Matter?" *Economics of Education Review* 16:1 (1997), 39-43.

Hallam, Susan, *Ability Grouping in Schools: a Literature Review* (London: University of London, Institute of Education, 2002).

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin, "Disruption versus Tiebout Improvement: The Costs and Benefits of Switching Schools," *Journal of Public Economics* 88:9 (2004), 1721-1746.

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin, "Does the Ability of Peers Affect Student Achievement?," *Journal of Applied Econometrics* 18:5 (2003), 527-544.

Hanushek, Eric A., "Interpreting Recent Research on Schooling in Developing Countries," *World Bank Research Observer* 10:2 (1995), 227-246.

Hanushek, Eric A., "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature* 24:3 (1986), 1141-1177.

Heckman, James J., "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations," *Journal of Human Resources* 32:3 (1997), 441-462.

Hoxby, Caroline M., "Peer Effects in the Classroom: Learning from Gender and Race Variation," *NBER working paper,* no. 7867 (2000).

Kremer, Michael and Dan M. Levy, "Peer Effects from Alcohol Use Among College Students," *NBER working paper* no. 9876 (2003).

Kremer, Michael, "Research on Schooling: What We Know and What We Don't: A Comment on Hanushek," *The World Bank Research Observer* 10:2 (1995), 247-254.

Kremer, Michael, "The O-Ring Theory of Economic Development," *Quarterly Journal of Economics* 108:3 (1993), 551-576.

Manski, Charles, "Economic Analysis of Social Interactions," *Journal of Economic Perspectives* 14:3 (2000), 115-136.

Manski, Charles, "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies* 60:3 (1993), 531-542.

Moffitt, Robert, "Policy Interventions, low-level equilibria, and social interactions", in Steven Durlauf and Peyton Young (eds.), *Social Dynamics*, (Cambridge: MIT Press, 2001).

Murnane, Richard J., *Impact of School Resources on the Learning of Inner City Children* (Cambridge: Ballinger, 1975).

Nechyba, Thomas J., "Mobility, Targeting, and Private-School Vouchers," *American Economic Review* 90:1 (2000), 130-146.

Rivkin, Steven G., "Tiebout Sorting, Aggregation, and the Estimation of Peer Group Effects," *Economics of Education Review* 20:3 (2001), 201-209.

Robinson, Peter M., "Root-N-Consistent Semiparametric Regression", *Econometrica*

56:4 (1988), 931-954.

Rosen, Stanley, "Obstacles to Educational Reform in China," *Modern China* 8:1 (1982), 3-40.

Sacerdote, Bruce, "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics* 116:2 (2001), 681-704.

Zimmerman, David J., "Peer Effects in Higher Education: Evidence From a Natural Experiment," *Review of Economics and Statistics* 85:1 (2003), 9-23.

Table 1: Summary Statistics on Teachers, School Investment and Students

School	1	2	3	4	5	6	7	8	9
Ranking	Regular	Regular	New Provincial School	Regular	National School	Affiliated with national	Provincial School	Regular	Provincial School
% of Superior teachers	13.04%	10.27%	9.61%	11.36%	26.61%	9.38%	6.09%	10.00%	21.59%
% of First class teachers	24.22%	32.19%	35.90%	39.39%	22.94%	41.67%	30.43%	40.00%	35.23%
% of Second class teachers	28.57%	33.56%	28.21%	28.03%	25.69%	37.5%	50.43%	22.00%	27.27%
% of Third class teachers	21.12%	7.53%	13.46%	6.06%	9.17%	4.17%	3.48%	12.00%	2.27%
Teacher Salary (*yuan*)	11590.4 (2431.9)	9900.76 (1594.32)	9715.87 (1675.56)	9137.45 (2202.9)	10148.4 (1900.4)	9890.43 (1508.45)	9178.63 (1985.00)	9927.72 (1730.36)	10334.29 (1793.48)
Teaching Experience	11.13 (10.83)	11.22 (10.52)	10.41 (9.98)	13.05 (10.78)	9.35 (9.65)	10.52 (8.13)	7.32 (7.21)	11.68 (10.56)	10.45 (9.06)
% of teachers with university	22.36%	40.41%	25.64%	25.00%	75.23%	30.21%	66.96%	39.00%	63.64%
% of teachers with college	60.87%	53.42%	71.79%	50.00%	21.10%	68.75%	31.30%	56.00%	29.54%
% of teachers teaching their specialty	74.67%	73.29%	80.77%	61.36%	95.41%	70.83%	93.91%	77.00%	88.64%
% of teachers aching only one subject	91.30%	87.67%	N. A.	97.72%	100.0%	91.67%	100.0%	100.0%	N. A.
Average teaching load	10.02 (2.79)	10.66 (2.51)	N. A.	10.35 (2.55)	9.32 (3.00)	10.62 (2.64)	9.95 (2.52)	9.49 (2.38)	N. A.
Number of classes taught weekly	2.90 (2.09)	2.90 (1.96)	N. A.	2.97 (2.48)	2.62 (1.65)	3.01 (2.25)	3.01 (1.83)	N. A.	N. A.
% of Female Teachers	29.19%	30.82%	39.74%	28.03%	33.95%	38.54%	46.09%	23.00%	27.27%
Number of Teachers	150	128	156	128	106	84	115	100	88
1996 per Pupil Investment	3838.23	5557.09	3620.32	4670.65	8180.22	2785.70	8539.57	3431.36	4522.07
1997 per Pupil Investment	4769.49	9029.15	6228.72	3419.32	11657.8	3425.76	16861.1	4705.12	6187.96
1998 per Pupil Investment	6340.30	10671.1	7212.99	8743.96	12309.4	4216.22	14055.3	5498.33	5126.18
Average Score on HSEE for Entering Class	566.71 (26.16)	565.47 (24.09)	576.81 (24.96)	551.48 (35.01)	638.52 (6.54)	464.88 (78.13)	606.85 (32.05)	554.17 (34.06)	615.76 (14.60)
% of students in higher education	31.28%	50.66%	52.66%	44.40%	94.95%	39.33%	40.19%	35.50%	66.24%

Note: Standard Deviations in Parentheses

Table 2: Ordinary Least Squares Estimates of the Achievement Equation

Variable	Full Sample	Subsample near Cutoff Score
Score on High School	1.130	0.840
Entrance Examination (A_{ijT-1})	(0.165)	(0.347)
Rural Male	−1.054	−1.419
	(0.456)	(0.894)
Rural Female	−2.295	−2.076
	(0.428)	(0.886)
Urban Female	−3.658	−2.873
	(0.543)	(1.017)
Tuition	1.290	Not included
	(0.984)	in specification
Peer Mean	18.304	−50.325
	(1.981)	(13.454)
$\frac{\text{Peer Squared}}{100}$	−10.713	30.128
	(1.250)	(8.219)
Peer Variance	−0.455	−0.278
	(0.056)	(0.072)
Early Admission	4.798	4.879
	(0.061)	(0.906)
Academic award	6.917	7.646
	(0.142)	(2.293)
Music skills	−7.511	−7.910
	(0.980)	(2.504)
Athletic Ability	−7.006	−6.557
	(0.366)	(2.313)
% of Superior and First	Not included	8.010
Teachers	in specification	(27.631)
% of Superior	377.311	Not included
Teachers	(33.556)	in specification
% of 1st Class	354.127	Not included
Teachers	(31.017)	in specification
% of 2nd Class	472.067	467.611
Teachers	(42.893)	(113.915)
Constant	−1183.087	−1148.745
	(121.577)	(280.893)
R squared	0.682	0.677
Observations	1241	1241

Note: Corrected standard errors at the school level in parentheses for OLS estimates.

Table 3: Fixed Effects Estimates of the Achievement Equation

	Peer Measures Omit a Classmate Chosen at Random	Subsample near Admission Cutoff Score	Subsample with School Level Teacher Quality
Entrance Exam Score (A_{ijT-1})	0.929	1.512	0.836
	(0.025)	(0.543)	(0.362)
Rural Male	−0.937	−1.299	−1.683
	(0.023)	(1.111)	(0.879)
Rural Female	−2.124	−1.958	−2.145
	(0.024)	(0.953)	(0.875)
Urban Female	−3.547	−3.743	−4.070
	(0.031)	(1.244)	(1.014)
Tuition	0.637	Not included	Not included
	(0.025)	in specification	in specification
Peer Mean	332.593	−43.968	−2600.228
	(152.535)	(11.818)	(742.808)
$\frac{\text{Peer Squared}}{100}$	−201.889	26.418	1536.389
	(93.516)	(6.948)	(438.725)
Peer Variance	−18.561	−0.198	−4.479
	(0.188)	(0.065)	(1.241)
Academic award	6.861	Not included	Not included
	(0.069)	in specification	in specification
Early Admission	4.714	Not included	Not included
	(0.029)	in specification	in specification
Music skills	−5.942	−7.215	−6.787
	(0.082)	(0.577)	(2.606)
Athletic Ability	−7.896	−6.651	−6.636
	(0.268)	(1.011)	(1.955)
% of Superior and 1st Class Teachers	N.A.	Not included in specification	31.170 (9.694)
Constant	−13139.850	1767.465	−1543.588
	(6195.781)	(538.534)	(1121.098)
Level of Fixed Effect	School	School Type	School Type
R squared	N.A.	0.233	0.254
Observations	1241	402	402

Note: The mean and standard deviation in parentheses of the coefficient estimates from 100 individual regressions are presented in column 1. Column 2 and 3 present fixed effects coefficient estimates of the achievement equation with standard errors in parentheses.

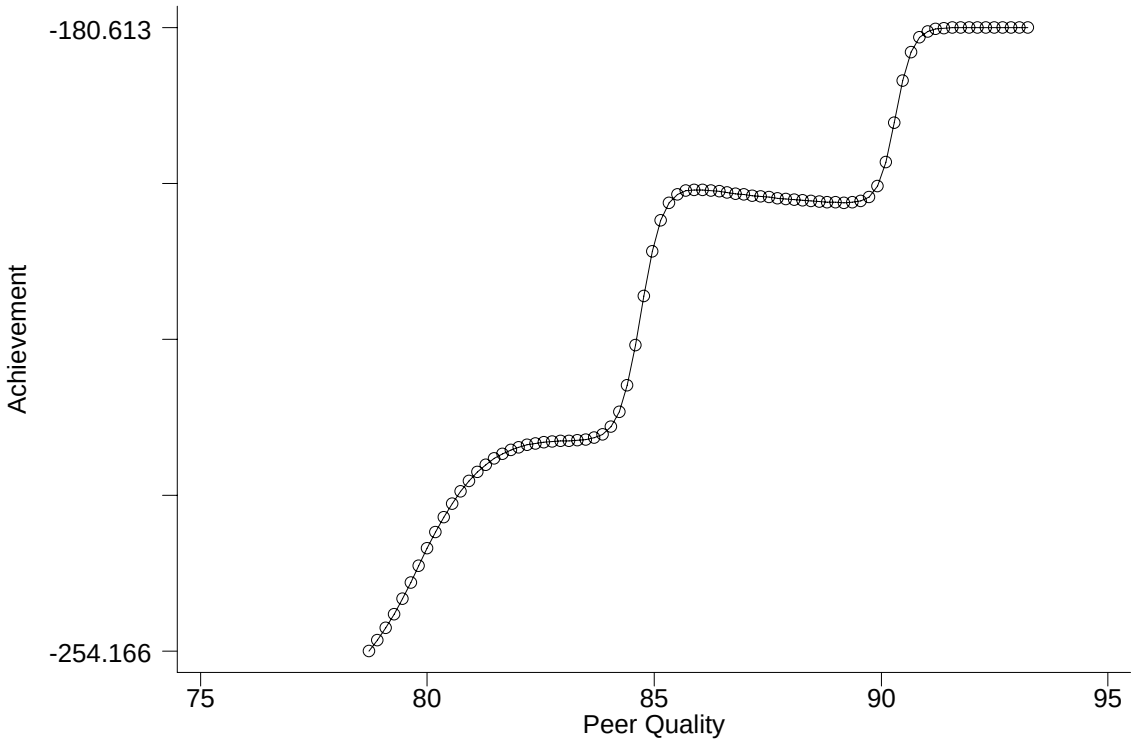

Figure 1: Non Parametric Estimate of Peer Quality

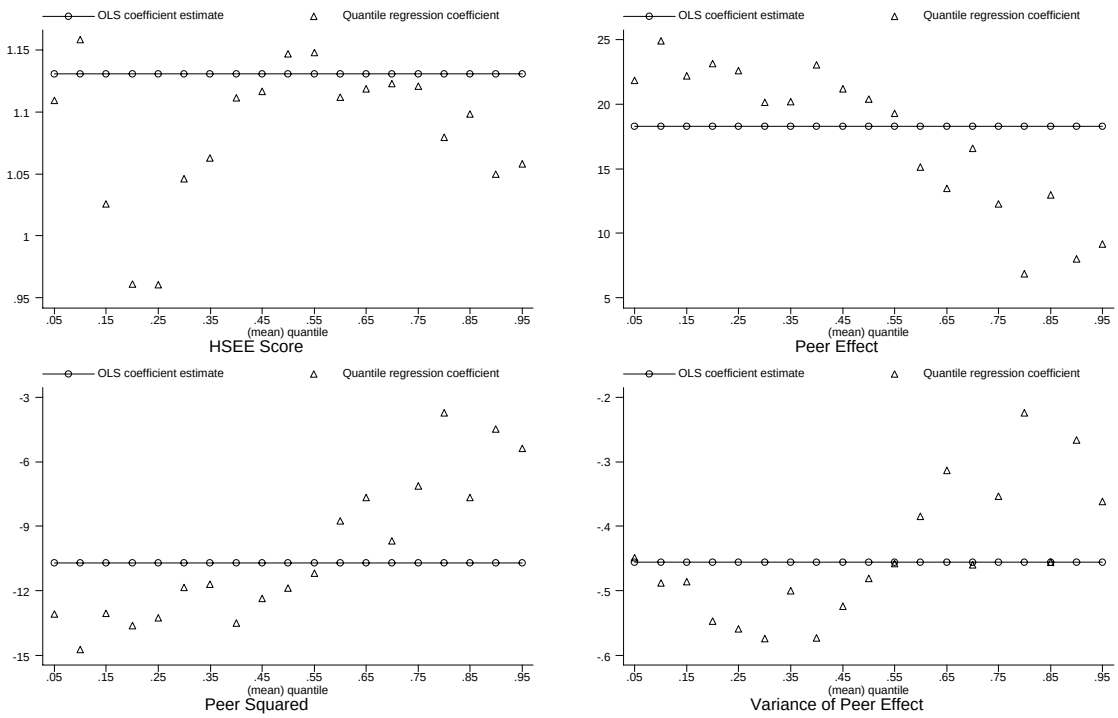

Figure 2: Quantile Regression Results for CSEE Score Baseline Specification