CLASSROOM PEER EFFECTS AND STUDENT ACHIEVEMENT*

by

Mary A. Burke
Research Department
Federal Reserve Bank of Boston

Tim R. Sass
Department of Economics
Florida State University

This version: February 6, 2006

Preliminary Draft: Please do not cite without permission.

In this paper we analyze a unique micro-level panel data set encompassing all public school students in grades 3-10 in the state of Florida for each of the years 1999/2000-2003/2004. We are able to directly link each student and teacher to a specific classroom and can thus identify each member of a student's classroom peer group. The ability to track individual students through multiple classrooms over time and multiple classes for each teacher enables us to control for many sources of spurious peer effects such as fixed individual student characteristics and fixed teacher inputs, as well as to compare the strength of peer effects across different groupings of peers, across grade levels, and to compare the effects of fixed versus time-varying peer characteristics. We find mixed results on the importance of peers in the linear-in-means model, and resolve some of these apparent conflicts by considering non-linear specifications of peer effects. The results suggest that some grouping by ability may create Pareto improvements over uniformly mixed classrooms. In general we find that contemporaneous behavior wields stronger influence than peers' fixed characteristics.

---

*We wish to thank the staff of the Florida Department of Education's K-20 Education Data Warehouse for their assistance in obtaining and interpreting the data used in this study. The views expressed is this paper are solely our own and do not necessarily reflect the opinions of the Florida Department of Education. This research is part of a larger project assessing teacher quality being funded by grant R305M040121 from the U.S. Department of Education.

I. Introduction

The potential for peers to affect individual achievement is central to many education policy issues, including the likely effects of school choice programs, "tracking" policies, and racial and economic desegregation. Vouchers and other school choice programs may benefit some who obtain better peers by exiting the traditional public school system, but might also harm those left behind by diminishing the quality of their classmates (Epple and Romano (1998), Caucutt (2002)). Grouping students in classrooms by ability can have significant impacts on student achievement as well as on the competition between private and public schools, depending on the magnitude of peer influences (Epple, Newlon and Romano (2002)). Likewise, the impact of desegregation policies on achievement depends not only on potential spillovers from average ability, but on the ways in which peer influences might be mediated by race in the presence of race-specific behavioral norms (Angrist and Lang (2004), Cooley (2005), Fryer and Torelli 2005).

Recent empirical research has begun to shed some light on the role peers play in shaping individual educational outcomes. However, there exists considerable disagreement over the magnitude of peer effects construed broadly, as well as uncertainty about the relative importance of fixed peer characteristics, such as race, gender, and innate ability, versus variable (endogenous) factors such as peer effort levels and classroom comportment. Depending on the study, ability dispersion among peers may have positive effects on achievement, negative effects, or no significant impact at all. (See, for example, Glewwe 1997 and Vigdor and Nechyba 2004).

The conflicting evidence on peer effects in education is due in large part to the difficulty in measuring the relevant attributes of peers and separating the influence of these peer variables from other factors that can affect student achievement, including an individual's own ability and

the quality of their teachers and of the schools they attend. Given the complexity of the educational process and the particular challenges of peer effects estimation, the data requirements to reliably estimate peer effects are substantial. Indeed, the diversity of existing empirical evidence likely reflects the variety of ways in which researchers have approached a difficult estimation problem in the face of differing data limitations.

With a unique panel data set encompassing all public school students in grades 3-10 in the state of Florida, together with new analytical techniques, we have unprecedented resources with which to address the gaps and controversies concerning peer effects in the educational context. Unlike any previous study, we simultaneously control for the fixed inputs of students, teachers and schools in measuring peer influences on academic achievement at the classroom level. These controls sharply limit the scope for biases from "correlated effects" (Manski 1993), a well-known hazard in peer effects estimation. Our rich data set also allows us to compare the influence of classroom peers to the influence of grade-level peers, and estimate effects on both math and reading achievement separately for elementary school, middle school, and high school cohorts. With additional controls for variable teacher inputs and abundant information on time-varying student factors, we can disentangle the respective influences of endogenous factors (such as contemporaneous peer achievement and classroom disruption) and exogenous factors such as predetermined ability, race, and gender.

We find that peer effects operate differently for reading achievement than for math. Peer performance generally has stronger effects on individual achievement in math than in reading, particularly at the elementary school level. Further, measured peer influences vary considerably across grade levels and with the aggregation level of the peer group. Both mean exogenous peer characteristics as well as current average peer performance are found to have significant effects

on individual student achievement in a number of circumstances, though the results are not very robust. While higher mean peer achievement gains are generally associated with higher individual gains in the same period, increases in ability dispersion within a classroom tend to detract from student learning. Consistent with some other recent findings, we show that the linear-in-means model of peer effects, in which average peer-group factors are assumed to operate homogeneously across the peer group, may falsely indicate small or non-existent effects of peer achievement because peer effects on different sub-groups may cancel each other out.

II. Conceptual and Practical Issues in Peer Effects Estimation

Peer effects in general are hard to measure for several reasons.[1] First, in order to properly specify a model of peer effects it is necessary to identify potential peer influences and the mechanisms by which peers affect individual behavior. Second, one must determine the relevant peer group and accurately measure its salient characteristics. Third, endogenous selection of peers and correlated group effects must be taken into account in order to avoid spurious inferences. Fourth, one must avoid biases owing to the simultaneity of individual and peer outcomes. Fifth, in many specifications it is difficult to distinguish between the influence of endogenous peer outcomes and exogenous peer characteristics. In the education context one must also consider the estimation issues that derive from the cumulative nature of educational production, as discussed extensively in Todd and Wolpin (2003).

The hypothesized sources of peer influence on academic achievement are quite diverse. They include positive spillovers of knowledge, values, or motivation that are associated with student ability and family inputs, as well as negative effects of disruptive peer behavior (Lazear

[1] For a more thorough discussion of the generic hazards of identifying peer effects (in education as well as other contexts) see Manski (1993), Brock and Durlauf (2001) and Moffitt (2001).

(2001)). Peer pressure to reduce effort and conform to a group effort norm is another possibility, evidenced by pejorative labels such as "nerd" and "teacher's pet", although pro-achievement norms are possible as well (Akerlof and Kranton (2002)). Peer group characteristics such as gender, racial, or socioeconomic composition have been hypothesized to exert effects on individual achievement apart from their relationship to peer ability or achievement. Although these have been termed "exogenous" (or "contextual") peer effects because they work through predetermined peer characteristics, such traits may serve as proxies for race or gender-specific group effort norms or for other aspects of classroom disposition that produce achievement externalities (Hoxby 2000)).[2] Figlio (2004b) has pointed to racially-induced effects on teachers' expectations of students, which may spill over to children outside the racial group in the same classroom.[3] The socioeconomic status of peers may proxy for parental emphasis on education and thus peers' attitudes towards school, which may lead to contagion effects. Most prior studies have focused on spillovers from peer ability, as proxied by some achievement measure, and from peer demographic characteristics, rather than spillovers from other aspects of behavior such as classroom disruption. One exception is Figlio (2004a), who measures the impact of students who are suspended for misbehavior on their classmates' achievement.

The strength of peer influence may vary by individual ability and characteristics, as well as with the ability and demographic characteristics of the peers. For example, low-achieving students may benefit more from an increase in the average ability of peers than would the

[2] As is by now well known, it is difficult to distinguish between structural and reduced-form effects of race and gender composition. We discuss this and other identification issues below.

[3] However, racial composition may affect students of different races differently, and may exhibit non-linearities. For example Hoxby finds that percent black has a larger negative effect on achievement of black students than non-blacks, and finds that the effects of percent Hispanic reverse sign as the percent Hispanic increases, indicating "critical mass" effects. Cooley also finds peer effects to be stronger within-race than across. These factors, if not interacted with the primary race composition variables, may constitute another source of sensitivity in the estimation of race composition coefficients.

brightest students in a class. Similarly, girls might be more strongly influenced by other girls than by boys or black students may be affected more by their fellow black students than by their white classmates. In fact, two recent works (Cooley 2005, Hoxby and Weingarth 2005) indicate potentially fundamental flaws in the standard, "linear-in-means" model of peer effects, finding that peer influences operate heterogeneously depending on the individual as well as on the nature of the peer group.

The level of aggregation at which peer effects operate depends on the presumed transmission mechanism. If peer influences occur through knowledge spillovers and disruption effects, peer effects should be strongest at the classroom level, in which case class partitions could have important consequences for individual and aggregate achievement. In contrast, if a student's performance depends more on which "crowd" he hangs out with, experiences in the classroom may be less influential than the student's choice of social group, a choice which is less susceptible to policy manipulation (Akerlof and Kranton 2002). Unfortunately, the level at which peer effects have been analyzed has largely been driven by data availability, rather than ex-ante beliefs about the relevant peer group. The only previous study to consider both class-level and grade-level peer influences is Betts and Zau (2004).

Non-random assignment of students and teachers to schools and classrooms is a well-known obstacle to measuring peer effects in schools. Parents can choose a child's school through their choice of residential location, or in some cases, through various school choice programs, and within-school sorting can occur as well. In the primary grades parents may lobby to have their children assigned to a particular teacher and principals may assign students and teachers to classrooms to achieve objectives such as racial diversity, ability tracking, maximizing school-wide accountability measures, or fulfilling teacher preferences. In middle and high

school, students can typically exercise some choice over their courses and teachers. Non-random assignment leads to correlations within peer groups in unobserved student and family inputs affecting achievement, resulting in an upward bias on measured achievement spillovers. The standard method for dealing with unobserved student characteristics is to include student-level fixed effects. However, Vigdor and Nechyba (2004) rely on identifying classrooms where student assignments are "apparently" random. Cooley (2005) uses teacher fixed effects to control for the average student traits assigned to a given teacher, and includes a measure of time spent reading at home as a proxy for unmeasured individual reading ability.

Even if student characteristics are taken into account, unobserved teacher inputs can exert a common influence on achievement for all her students, introducing correlated effects bias on measured spillovers of contemporaneous achievement. Recent evidence suggests that the influence of teachers is considerable and yet is not closely related to traditional measures of teacher quality (e.g. Rivkin et al. (2005)). Effects of classroom demographic composition can also be biased if there is systematic matching of teachers to classrooms based on these compositional factors. Although controls for fixed school inputs are standard in the literature, controlling for teacher quality is particularly difficult since in many data sets teachers and students cannot be directly matched. As a result, some previous studies have had to ignore teacher quality altogether (Figlio (2004a)) or to study peer effects at the grade level only, relying on school-by-grade controls (Hoxby (2000), Hanushek, et al (2003)) to capture average teacher quality. Only a handful of peer effects studies have explicitly matched students to their specific classroom teacher. Betts and Zau (2004) and Vigdor and Nechyba (2004) employ observable teacher characteristics to control for teacher quality while Cooley (2005) utilizes teacher fixed

effects. No previous study has controlled simultaneously for student and teacher unobserved ability using both student and teacher fixed effects.

Including measures of peer behavior or "endogenous" peer effects can produce an additional source of bias, separate from the endogeneity of the peer group. Coined the "reflection problem" by Manski (1993), this stems from the mutuality of peer influence; any individual student may be both a producer and consumer of peer effects. The result is that measures of concurrent peer achievement or peer disruptive behavior are likely to be correlated with the unobserved factors determining individual achievement gains in the same period. Some authors have dealt with the problem by utilizing peers' prior test scores rather than contemporaneous ones (Betts and Zau (2004), Hanushek et al (2003), Vigdor and Nechyba (2004)) while others have employed instrumental variable methods (Figlio (2004a), Cooley (2005), Hoxby and Weingarth (2005)). In addition, it has been argued that biases stemming from the reflection problem are likely to be minor relative to other sources of bias such as mismeasurement of the relevant peer characteristics (Hanushek et al. (2003), Arcidiacono et al. (2005)).

The separation of endogenous peer influences, such as current effort levels or disruptiveness, from exogenous influences such as innate peer ability and classroom racial composition, cannot readily be achieved econometrically when the former constitute linear functions of the latter (Manski 1993). Some studies have side-stepped this issue by omitting endogenous peer variables (Arcidiacono, et al. 2005), capturing just the reduced form effects of fixed peer characteristics. These reduced form coefficients embed any direct effects of average peer characteristics, such as percent female, on individual achievement—for example, girls may be more well-behaved in the classroom and so promote a better learning environment for

everyone—together with "multiplier" effects that occur if there are spillovers from current peer achievement (that is, if girls raise mean achievement, and mean achievement further raises individual achievement). The direct and indirect effects cannot be unbundled, but the attribution matters for policy---for example, if the true effect is wholly indirect, schools should focus on harnessing achievement spillovers broadly rather than on distributing girls optimally across classrooms.

As emphasized by Cooley (2005), specifications that emphasize spillovers based on innate ability (for example with the use of lagged achievement measures, such as in Hanushek et al. and Vigdor and Nechyba) will miss any spillovers of shocks to individual effort and disruption levels, and may miss spontaneous aspects of classroom dynamics such as the emergence of group effort and behavioral norms. Such endogenous interactions, rather than passive externalities from fixed peer factors, constitute the central thesis of the theoretical literature on social interactions (see, for example, Becker and Murphy 2000, Bernheim 1994). However, inclusion of endogenous factors, such as current average peer test scores, leads to the reflection problem described above and necessitates the use of instrumental variable techniques to obtain unbiased estimates of model parameters.[4]

III. Previous Empirical Findings

The past five years have seen increasing academic attention to measuring peer effects in education, following the general surge in interest in social interactions and greater data availability. Hoxby (2000), Hanushek et al (2003), Betts and Zau (2004), Figlio (2004a), Vigdor

[4] In addition, current mean achievement (or mean disruptiveness, as described below) and individual achievement may be similarly influenced by an innovation to a common unobserved input, such as new computers in the classroom, producing spurious peer influence.

and Nechyba (2004), Cooley (2005), Hoxby and Weingarth (2005) and Zabel (2005), all exploit large individual-level longitudinal administrative databases to estimate the magnitude of peer effects in American public (primary or middle) schools.[5]

Hoxby (2000) uses data from Texas for grades 3 through 6, exploiting random variation in race and gender composition within and across cohorts as an exogenous source of variation in peer characteristics and ability. She finds significant positive spillovers of gains in mean peer achievement on individual gains on standardized math and reading tests, where interactions appear stronger within racial groups. The results indicate that the fraction of female peers exerts both a direct effect on individual achievement (perhaps via a beneficial effect on classroom atmosphere) as well as an indirect effect operating through peer achievement. However, the results are limited to grade-level interactions for grades 3 through 6.

Hanushek et al (2003) use the same Texas data, extending the analysis through to grade 8. Again, specific classroom peers are not observed, so interactions are measured at the grade level only, and only for math achievement. The results show significant positive effects of average peer achievement on individual achievement gains in mathematics, effects that do not appear to differ significantly across quartiles of the achievement distribution. The authors are careful to point out that their use of twice-lagged peer test scores may underestimate peer effects insofar as they mismeasure current peer achievement. However, they argue that the systematic components of peer achievement, as captured in the lagged scores, are likely to be more important than current innovations in peer achievement.

[5] The list is not exhaustive, but we consider these papers to be closest to ours in terms of the extent of the data and its focus on American primary and secondary education. Numerous additional studies treat European data, college-level data, or smaller samples from U.S. public schools.

Betts and Zau (2004) observe classroom level peers, in grades 2 through 5, for public school students in San Diego. They include fixed effects for individual students, the school, school-by-grade level, and an extensive list of teacher characteristics, but not teacher fixed effects. They find positive and statistically significant effects of mean peer achievement on individual test score gains in both reading and math. In contrast, grade-level peer achievement (holding class-level peer achievement constant) is not a significant determinant of individual achievement gains.[6] Betts and Zau also find some evidence of asymmetric peer effects; for a student whose long-run average test score initially equals the class average, a decrease in peer achievement is more harmful than an equivalent increase in peer quality.

Vigdor and Nechyba (2004) also observe classroom-level peer groups in an administrative panel data set from North Carolina, covering grades 3 through 5. They include teacher experience as a control for teacher quality and parent education to control for family inputs. Instead of using fixed effects for individual students, the authors compare estimation of peer effects on the full sample to estimation of effects on a restricted sample that meet the criteria for "apparent random assignment." Rather than using test score gains as their dependent variable they use test score levels and include an individual student's lagged test score as an explanatory variable. While this approach is more flexible than the standard gain model, the use of the lagged test score as a regressor in an ordinary-least-squres framework likely biases their estimates.[7] They find positive effects of mean (twice-lagged) peer test scores on individual achievement in math and reading tests as well as positive effects for dispersion in peer test scores

[6] When only grade-level peer achievement is included in the model, its marginal effect is positive and of greater magnitude than the class-level peer effect, suggesting that grade-level peer measures tend to overstate classroom peer effects.

[7] For a detailed discussion of this issue, as well as comparisons of estimates from the test-score-level and gain-score models, see Sass (2006).

for math using the full sample. However, these effects disappear in the restricted "random" sample and even become negative in the case of mean peer reading achievement.

Like Vigdor and Nechyba, Cooley (2005) estimates classroom-level peer effects using data on primary-school students in North Carolina. Unlike previous work, however, she employs a quantile regression approach to allow for differing peer effects across the distribution of student abilities. She controls for unobserved teacher quality via teacher fixed effects, but omits individual fixed effects because of the computational difficulty of estimating large numbers of fixed effects within the quantile regression framework. As an imperfect proxy for unobserved individual ability she includes self-reported measures of the proportion of the student's free time spent reading for pleasure. (She does not observe any equivalent proxy for math ability and therefore studies only reading test scores.) She exploits a policy change that raised the stakes of failing for low-achieving students in order to justify an exclusion restriction that aims to identify endogenous effort spillovers and to separate them from the effects of fixed peer characteristics. She finds that spillovers of contemporaneous effort are positive and significant, and much stronger than spillovers predicted by predetermined peer ability as measured with lagged test scores. She also finds that endogenous peer effects are stronger within race, and generally stronger for nonwhite students than white students. The quantile regression model reveals important non-linearities in peer effects, such that a simulated integration experiment results in some (but not a dramatic) narrowing of the race achievement gap. The results are provocative, but are likely to be sensitive to the quality of the reading-time variable as a control for unobserved ability given that inclusion of individual fixed effects has been shown to significantly alter estimates of peer effects in other studies (see, for example, Hanushek et al. 2003).

Angrist and Lang (2004) examine the effect of movements of students across schools resulting from racial integration policy in Boston. In agreement with one of Cooley's predictions, they find that achievement levels of white students were not negatively affected by the policy-induced influx of new, lower achieving non-white peers. However, they find some negative but transient effects of the integration policy on the achievement of nonwhite female students who attended the receiving school both before and after the influx of the new peers. The transience of the observed negative spillover agrees with the idea that moving imposes adjustment costs on movers that may have negative spillover effects on some non-moving students.

Figlio (2004a) focuses on the impact disruptive students have on their classmates. Using panel data from an unnamed "large Florida school district" for 1996/97-1999/00 Figlio estimates individual student mathematics achievement in sixth grade as a function of the fraction of students suspended for five days or more along with measures of the racial and gender composition of the class. Peer ability is captured by their third-grade test score. Figlio's model controls for unobserved student ability with student fixed effects but has no controls for teacher quality. To account for simultaneity between student outcomes and the disruptive behavior of peers Figlio utilizes a clever instrumental variable strategy. He demonstrates that boys with names more commonly given to girls are more prone to misbehavior and uses this as an instrument (hence the paper's title, "Boys Named Sue"). He finds statistically significant and quantitatively substantial negative effects of peer misbehavior on student achievement. Unfortunately, he doesn't report whether peer ability or the racial and gender composition of peers also have significant impacts on student achievement.

Hoxby and Weingarth (2005) study peer effects in Wake County, North Carolina, where a school reassignment plan produced exogenous changes in classroom composition and thereby created a set of natural experiments to study the impact of peers on student achievement. Their sample includes students in grades 3-8 for the years 1994/95-2002/03. Student achievement is measured by the sum of test scores in math and reading. While their focus is on a single county, they can track the past history of students who enter from other counties. Their paper is particularly noteworthy for its deviation from the typical linear-in-means peer specification. Not only do they allow the effect of average peer ability to vary with the (countywide) ability decile of the individual student, they also permit individual students at different ability levels to be impacted differently by the fraction of classmates at different deciles of the ability distribution. Student and school fixed effects control for unmeasured student and school characteristics and grade-by-school-by-year effects control for average teacher quality within a grade and school, but no measures of individual teacher quality are included. To isolate the effect of exogenous reassignments, they define a student's "simulated instrument cohort" to be the group of students who would be in his school/grade cohort if reassignments are allowed but potentially endogenous student movements (due to family relocations and the like) are disallowed. They compute means based on the simulated instrument cohort, and use the resulting variables as instruments for mean peer factors for the student's actual class.[8] They find that the linear-in-means model yields a positive effect of mean peer achievement on individual performance. Their least restrictive specification, with two-way interactions between own and peer ability, produces very noisy estimates but generally supports the “boutique model” (higher achievement when surrounded by peers of similar characteristics) and the focus model (peer homogeneity is

[8] This instrument is likely to flatten out variation in peer variables across classrooms within a grade, however, since values are based on the entire school-grade-year cohort minus a given individual student.

good, even if the student himself is not part of the homogeneous group) of peer effects. Exogenous peer characteristics (race, income, ethnicity) generally have very small effects in the non-linear peer effects model, but female classmates are more beneficial than boys.

Zabel (2005) estimates student math and reading achievement gains for a sample of 600 New York City elementary schools during the period 1995/96-1999/00. His data do not allow him to link teachers to classrooms so he has no controls for teacher quality. Nor does he exploit the panel nature of his data to include student fixed effects; rather he estimates separate models for fourth and fifth grades with school fixed effects. He uses the lagged reading test scores of peers as an instrument for contemporaneous math scores and likewise the lagged math test scores for contemporaneous peer reading achievement. To address the problem of student and teacher sorting among classrooms within a school he looks at schools with four or fewer classes per grade, which he argues afford less opportunity for sorting. He finds that mean peer achievement has a negative effect on individual achievement gains, while the effect of the standard deviation in peer achievement is weakly negative. Like Hoxby and Weingarth, he finds that the effects of exogenous peer characteristics are generally small and not consistently significant.

IV. Data and Sample Selection

In the present study we make use of a unique panel data set of school administrative records from Florida[9] that allows us to overcome many of the challenges to the estimation of peer effects. The data cover five school years, 1999-2000 through 2003-2004, and include all public-school students throughout Florida. Achievement test data are available for both math

[9] A more detailed description of the data is provided in Sass (2006).

and reading in each of grades 3-10.[10] Like databases in North Carolina and Texas, we can track individual students over time and thus control for unobserved student characteristics with student-specific fixed effects. However, unlike other statewide databases, we can precisely match both students and teachers to specific classrooms at all grade levels.[11] This classroom-level data, with consistent teacher identification over time, allows us to control for teacher characteristics via fixed effects. Another advantage of the data is that we can determine the specific classroom assignments of middle-school and high-school students, who typically rotate through classrooms during the day for different subjects. By observing different groups of students at the same grade level with the same teacher we avoid confounding peer variation with variation in the teacher, the grade level, the cohort, or the school.

Not only does our data directly link students and teachers to specific classrooms, it also provides information on the proportion of time spent in each class. This is potentially important for correctly determining the identity of teachers and peers at the elementary school level. While primary school students typically receive all of their academic instruction from a single teacher in a single "self-contained" classroom, this is far from universal. In Florida, five percent of elementary school students enrolled in self-contained classrooms are also enrolled in a separate

[10] The state of Florida currently administers two sets of reading and math tests to all 3rd through 10th graders in Florida. The "Sunshine State Standards" Florida Comprehensive Achievement Test (FCAT-SSS) is a criterion-based exam designed to test for the skills that students are expected to master at each grade level. The second test is the FCAT Norm-Referenced Test (FCAT-NRT), a version of the Stanford-9 achievement test used throughout the country. The scores on the Stanford-9 are scaled so that a one-point increase in the score at one place on the scale is equivalent to a one-point increase anywhere else on the scale. The Stanford-9 is a vertically scaled exam, thus scale scores typically increase with the grade level. We use FCAT-NRT scale scores in all of the analysis. The use of vertically scaled scores to evaluate student achievement is important since a one-unit change has the same meaning for low- and high-achieving students. Other types of measures, such as standard deviations from the mean score are potentially problematic; it is not clear that a 0.1 standard deviation increase in a test score, starting one standard deviation from the mean, is the same as a 0.1 standard deviation increase for someone with an initial score equal to the sample mean.

[11] Currently, the Texas data do not provide a way to link teachers and students to specific classrooms. For North Carolina, one can only (imperfectly) match specific teachers and students to classrooms at the elementary school level. Matching is done by identifying the person who administers each student the annual standardized test, which at the elementary school level is typically the classroom teacher.

math course, four percent in a separate reading course and four percent in a separate language arts course. In addition, nearly 13 percent of elementary students enrolled in self-contained elementary classes are also enrolled in some type of exceptional student education course, either special-education or gifted courses.[12]

We restrict our analysis of student achievement to students who receive instruction in the relevant subject area (math or reading/language arts) in only one classroom. At the elementary school level only students in "self-contained" classrooms are included. Elementary students spending less than one hour per day in the class (eg. special education students who spend nearly all of their day in a separate special-education classroom) are not considered members of the classroom peer group. At the middle and high-school level students who are enrolled in more than one course in the relevant subject area (mathematics or reading/language arts) are dropped, though all students enrolled in a course are included in the measurement of peer-group characteristics. To avoid atypical classroom settings and jointly taught classes we consider only courses in which 10-50 students are enrolled and there is only one "primary instructor" of record for the class. Finally, we eliminate charter schools from the analysis since they may have differing curricular emphases and student-peer and student-teacher interactions may differ in fundamental ways from traditional public schools.

In order to limit the computational time cost of estimation we restrict our analysis to randomly selected samples of 100 elementary, 100 middle and 100 high schools in Florida. Students who change schools are tracked, even if they move out of the initial sample of 100 schools. Thus the estimation samples actually include somewhat more than 100 schools in each

[12] Since previous studies lack data on students' complete course enrollments, they either ignore the fact that students may receive instruction outside their primary classroom or deal with the issue in an ad-hoc fashion. For example, Hanushek et al (2003) simply remove special-education students from their analysis of peer effects.

category. As we discuss below, the estimation techniques we employ do not require these sample restrictions. They are imposed simply to reduce computational time and we plan to increase the sample sizes in future work.

V. Econometric Model and Identification Strategies

A. General model of student achievement

In order to estimate the extent of peer effects we begin with a general specification of the "educational production function" that relates student achievement to vectors of time-varying student/family inputs (X), classroom-level inputs (C), school inputs (S) and time-invariant student/family characteristics (γ):

$$A_{it} - A_{it-1} = \Delta A_{it} = \boldsymbol{\alpha}_1 \mathbf{X}_{it} + \boldsymbol{\alpha}_2 \mathbf{C}_{ijmt} + \boldsymbol{\alpha}_3 \mathbf{S}_{mt} + \gamma_i + \varepsilon_{it} \quad (1)$$

The subscripts denote individuals (i), classrooms (j), schools (m) and time (t).

Equation (1) is a restricted, "value-added" form of the cumulative achievement function specified by Todd and Wolpin (2003) where the achievement level at time t depends on the individual's initial endowment (for example of innate ability or academic potential) and their entire history of individual, family and schooling inputs.[13] Although often not stated, there are a number of implicit assumptions underlying the education production function specified in (1). First, it is assumed that the cumulative achievement function does not vary with age, is additively separable, and linear. Family inputs are assumed constant over time, and the impact of parental inputs on achievement, along with the impact of the initial individual endowment on achievement, induce a (student-specific) constant increment in achievement in each period. This allows the combination of these time-invariant inputs to individual achievement gains to be

[13] It is important to note that while the dependent variable is the change in student achievement, equation (1) is a model of student achievement levels, not achievement growth. The lagged value of achievement on the left hand side serves to represent the cumulative effect of all prior schooling inputs on current achievement.

represented by the student-specific fixed component, γ_i. Third, the marginal impacts of all prior school inputs decay geometrically (at a rate greater than or equal to zero) with the time between the application of the input and the measurement of achievement at the same rate. Thus lagged achievement serves as a sufficient statistic for all prior schooling inputs. Fourth, to remove individual lagged score from the right-hand side of the gain equation, it is further assumed that the decay rate is actually zero—that is, that school inputs applied at any point in time have an immediate and permanent impact on cumulative achievement.[14] A thorough discussion of these assumptions and the derivation of the linear education production function model can be found in Todd and Wolpin (2003) and Sass (2006).

The vector of classroom inputs can be divided into four components: peer characteristics, $\mathbf{P}_{-ijmt}$ (where the subscript –i students other than individual i in the classroom), time-varying teacher characteristics (such as experience), $\mathbf{T}_{kt}$ (where k indexes teachers), time-invariant teacher characteristics (such as innate ability and pre-service education), δ_k. and non-teacher classroom-level inputs (such as books, computers, etc.), $\mathbf{Z}_j$. If we assume that, except for teacher quality, there is no variation in education inputs across classrooms within a school, the effect of $\mathbf{Z}$ becomes part of the school-level input vector, $\mathbf{S}_m$. If we further assume that school-level inputs are constant over the time span of analysis, they can be captured by a school fixed component, ϕ_m. The education production function can then be expressed as:

$$\Delta A_{it} = \boldsymbol{\beta}_1 \mathbf{X}_{it} + \boldsymbol{\beta}_2 \mathbf{P}_{-ijmt} + \boldsymbol{\beta}_3 \mathbf{T}_{kt} + \gamma_i + \delta_k + \phi_m + \nu_{it} \quad (2)$$

Where ν_{it} is a normally distributed, mean zero error.

[14] Thus, for example, the quality of a child's kindergarten must have the same impact on his cumulative achievement as of the end of the kindergarten year as it does on his achievement at age 18. While a strong assumption, this allows the impact of all prior schooling inputs to be captured by the lagged achievement score, A_{t-1}, on the left-hand side of the equation. Otherwise, equation (1) would contain a lagged dependent variable on the right hand side and thus could not be consistently estimated by ordinary least squares.

The vector of time-varying student characteristics, X, includes three measures of student mobility: the number of schools attended in the current year and indicators of "structural" and "non-structural" moves by students. Structural moves are defined as situations where a student moves from one school to another and at least 30 percent of his fellow students in the same grade at the initial school move to the same school. Thus this variable captures the effects of normal transitions from elementary to middle and middle to high school as well as the impact of significant school re-zonings. Correspondingly, a non-structural move occurs when a student attends a school different from the one attended at the end of the preceding school year but is not joined by at least 30 percent of her former schoolmates. This encompasses family relocations as well as movements between schools to attend "magnet" or other specialized programs.

Our vector of peer characteristics includes both "endogenous" peer variables (ie. student outcomes) as well as "exogenous" peer characteristics that are not a result of student actions. The two endogenous peer variables are student achievement gains and the number of discipline incidents. Exogenous peer characteristics include race and gender. The total number of peers is measured by the number of students in the classroom. The peer variables and associated estimation issues are discussed in detail below.

Time-varying teacher attributes are captured by a set of three dummy variables representing varying experience levels: 0 years of experience (ie. first-year teachers), 1 year of experience, and 2-4 years of experience. Teachers with five or more years of experience are the omitted category.

B. Computational Issues

Estimation of (2) is computationally challenging since it includes three levels of fixed effects: individual students (γ_i), teachers (δ_k) and schools (ϕ_m). Standard fixed effects methods

eliminate one effect by demeaning the data with respect to the variable of interest (eg. deviations from student means). Additional effects must then be explicitly modeled through the inclusion of dummy variable regressors. Given our data includes thousands of teachers and hundreds of schools, such standard methods are infeasible.

We combine two different approaches to solve the computational problem associated with estimating a three-level fixed effects model. First, we utilize the "spell fixed effects" method proposed by Andrews, Schank and Upward (2004) and combine the teacher and school fixed effects into a single effect, $\eta_{km} = \delta_k + \phi_m$. This combined effect represents each unique teacher/school combination or "spell." Not only does this facilitate computation, it allows the effects to be non-separable, allowing for the impact of teachers on student achievement to vary by school. The education production function thus becomes:

$$\Delta A_{it} = \boldsymbol{\beta}_1 \mathbf{X}_{it} + \boldsymbol{\beta}_2 \mathbf{P}_{-ijmt} + \boldsymbol{\beta}_3 \mathbf{T}_{kt} + \gamma_i + \eta_{km} + \nu_{it} \tag{3}$$

The second approach is an extension of the iterative fixed effects estimator recently proposed by Arcidiacono, et al (2005).[15] Taking deviations from the spell means, the achievement equation becomes:

$$(\Delta A_{it} - \overline{\Delta A}_{km}) = \boldsymbol{\beta}_1 (\mathbf{X}_{it} - \overline{\mathbf{X}}_{km}) + \boldsymbol{\beta}_2 (\mathbf{P}_{-ijmt} - \overline{\mathbf{P}}_{km}) + \boldsymbol{\beta}_3 (\mathbf{T}_{kt} - \overline{\mathbf{T}}_{km}) + (\gamma_i - \overline{\gamma}_{km}) + \nu_{it} \tag{4}$$

where the overbar and km subscript denote the mean of the relevant variable over all students and all time periods covered by teacher k at school m. Subtracting the de-meaned individual effect from both sides yields:

[15] Arcidiacono, et al derive their estimator in the context of a model with only fixed effects and no other covariates. However, it is straightforward to extend their approach to models with covariates. Details of the derivation are available upon request.

$$(\Delta A_{it} - \overline{\Delta A}_{km}) - (\gamma_i - \bar{\gamma}_{km}) = \boldsymbol{\beta}_1(\mathbf{X}_{it} - \overline{\mathbf{X}}_{km}) + \boldsymbol{\beta}_2(\mathbf{P}_{-ijmt} - \overline{\mathbf{P}}_{km}) + \boldsymbol{\beta}_3(\mathbf{T}_{kt} - \overline{\mathbf{T}}_{km}) + \nu_{it} \quad (5)$$

Equation (5) is estimated by ordinary least squares (OLS), using initial guesses for the individual fixed effects, γ_i and $\bar{\gamma}_{km}$. This produces estimates of $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and $\boldsymbol{\beta}_3$ which are then used to calculate predicted outcomes for each individual and to update the estimated individual effects by calculating an individual's residuals for each time period and computing their mean over all time periods. The model parameters are then re-estimated using the updated fixed effects, and then the fixed effects are further updated. The process is iterated until the coefficient estimates converge. Standard errors are obtained by bootstrapping.

C. Specification and Identification of Peer Effects

The inclusion of both student and teacher fixed effects as well as time-varying teacher characteristics should provide strong defense against selection bias due to the non-random assignment of teachers and students to classrooms.[16] However, as discussed above, unbiased estimation of peer effects also depends on the proper specification and measurement of peer inputs, $\mathbf{P}_{-ijmt}$, including the selection of the relevant peer group and inclusion of all relevant exogenous and endogenous peer variables. In addition, if endogenous peer variables are included in the model, one must account for the reflection problem associated with simultaneous determination of individual and peer outcomes.

Although our focus is on intra-classroom peer effects, we estimate models that measure peer characteristics at both the classroom and grade level. This allows us to see if there are

[16] While individual fixed effects will control for all time-invariant student/family inputs, any transient effect on the student input that spills over to her peers will not be controlled for. However, in our framework we do control for the effects of a change in the school and grade level.

significant spillovers beyond the classroom and enables us to determine if the grade-level measures used in some previous work may have produced biased estimates of classroom peer influences.

As noted by Vigdor and Nechyba and by Arcidiacono et al., test scores are a noisy measure of peer ability, reflecting both true (time-invariant) traits and an idiosyncratic error. While we estimate models that include mean peer achievement and the standard deviation in peer achievement as explanatory variables, we also estimate models that exclude peer achievement and instead utilize a direct measure of peer ability suggested by Arcidiacono, et al. They propose measuring peer ability as the mean of peers' fixed effects This is done by replacing peer characteristics with the estimated average fixed effect of peers in equation (4), producing:

$$(\Delta A_{it} - \overline{\Delta A}_{km}) = \boldsymbol{\beta}_1(\mathbf{X}_{it} - \overline{\mathbf{X}}_{km}) + \boldsymbol{\beta}_2(\bar{\gamma}_{-ijmt} - \bar{\gamma}_{km}) + \boldsymbol{\beta}_3(\mathbf{T}_{kt} - \overline{\mathbf{T}}_{km}) + (\gamma_i - \bar{\gamma}_{km}) + \nu_{it} \quad (6)$$

To estimate equation (6) by the iterative method one can subtract the individual fixed effect from both sides to obtain:

$$(\Delta A_{it} - \overline{\Delta A}_{km}) - \gamma_i = \boldsymbol{\beta}_1(\mathbf{X}_{it} - \overline{\mathbf{X}}_{km}) + \beta_2(\bar{\gamma}_{-ijmt}) + \boldsymbol{\beta}_3(\mathbf{T}_{kt} - \overline{\mathbf{T}}_{km}) + \beta_4(\bar{\gamma}_{km}) + \nu_{it} \quad (6)$$

where $\beta_4=(-1-\beta_2)$. At each iteration the individual fixed effects are re-estimated and used to obtain the mean peer fixed effect, $\bar{\gamma}_{-ijmt}$. Iteration continues until the coefficient estimates converge.

In addition to peer ability, compositional aspects of the peer group in terms of socioeconomic status, gender, and race may collectively affect the classroom atmosphere and

(justly or unjustly) teaching strategies, as discussed above. We include both the fraction of students who are female and the fraction who are black to account for these possible effects.[17]

Besides exogenous peer traits associated with (time-invariant) ability and demographic factors, we also incorporate (time-varying) peer behavior as a potential source of peer effects. In particular, we allow for both the academic performance of peers (measured by test score gains) as well disruptive behavior by peers to affect individual student achievement. We measure classroom disruption by the average number of disciplinary incidents per student in the class.[18] Lazear (2001) predicts that an unruly student hurts the achievement of his classmates by reducing the amount of class time devoted to learning. While many classroom disruptions (such as excessive questioning) will not result in disciplinary action, we take the incidence of such action within a peer group to be a proxy for the extent of classroom disruption. Teacher responses to an increase in disruptiveness could bias the estimated effect downward, however, and the specification works best if disciplinary action is positively correlated with the equilibrium level of disruption, which should be negatively related to individual achievement gains. Alternatively, disciplinary action may be a signal of a teacher's strictness or level of control rather than a signal of a negative classroom environment, in which case incidence of peer disciplinary action and individual achievement might be positively correlated. Given we control for time-invariant teacher inputs via fixed effects, the variation in disciplinary action should isolate peer-group-induced effects even if these operate partly through teacher actions.

---

[17]We also estimated some preliminary models that included the fraction of Hispanic students in the class, but ethnicity never appeared to have a significant peer effect. While we possess information on the proportion of students receiving free or reduced-price lunch, a frequently used proxy for family income, we did not include it in our vector of peer characteristics. Free lunch status is a very noisy measure of income and best and its use is particularly problematic at the high school level where participation is much lower than in elementary and middle school, perhaps due to social stigma.

[18] Disciplinary incidents can result from a variety of actions, including classroom disturbance, fighting, possession of weapons and drug use. A variety of actions may be taken, including corporal punishment, suspension or expulsion.

D. Capturing Endogenous Effects: Instrumental Variables Strategy

In order to deal with potential bias resulting from simultaneity of individual and peer actions, we employ an instrumental variable strategy, treating both peer achievement and peer disciplinary actions as endogenous. To form instruments for peer achievement we exploit variation in three aspects of the peer group that can reasonably be excluded from the individual achievement gain model: the proportion of classroom peers that moved to the school during the current year, the mean age of the current peer group (expressed in months), and the percentage of classroom peers with disabilities. Previous work (Bifulco and Ladd (2006), Sass (2006), among others) has shown that changing schools hurts individual student performance during the transition year. Following Argys and Rees (2004) and Angrist and Krueger (1991) we expect that the age of peers will affect their academic performance but have no direct impact on one's own achievement gains. Because we control for fixed student, school, and teacher inputs, as well as for the individual's own moving status, we avoid estimation biases that could occur if movers (or special education students, or relatively young students) tend to be assigned to certain types of teachers or types of students, or arrive more frequently at certain types of schools. (Special education status does not vary much over time in our sample, so we do not include it as a time-varying individual variable---we assume fixed individual effects embed special education status.) At the same time, we expect the proportion of peer "mover" students, mean peer age, and proportion of special education students---after controlling for fixed (and some variable) school, teacher, and individual inputs---to affect individual achievement gains only indirectly, insofar as they affect mean peer achievement. To account for exogenous factors that may predict peer disciplinary incidents we also include as instruments the prior-year number of disciplinary

incidents for the current peer group as well as the fraction of current peers with emotional disabilities.[19]

We control for the probability of moving and of experiencing a given proportion of movers with fixed individual, teacher-school spell, grade-level, and year fixed effects, so that the proportion of peer movers must be non-random only conditional on these factors. The teacher-school effects in particular should control for endogenous classroom assignment of new students within a school-grade level cohort. Assignment of movers to classrooms is not likely to be based on the current-year residuals in the individual achievement equation unless these are predictable *ex ante* and relevant to classroom assignment policies. The number of movers received at a given school in a given year (and so the proportion of movers experienced by a given student) could be endogenous to prior (or forecastable) changes, for good or ill, in a given school's resources that would affect all students at the school similarly, such as budget cuts or introduction of new computers or prior changes in the student body, as in reported instances of "white flight." While poorer school districts in general experience more moves on average, conditional on fixed school factors endogenous moves could predict either positive correlated effects on achievement (if an increase in movers is caused by an improvement at the school) or negative correlated effects (if the increase in movers signals an increase in transience in the neighborhood).

However, changes in the values of the instruments may induce changes in the (unobserved) time-varying components of teacher or other school inputs. For example, if teachers slow down to accommodate new ("mover") students, or if (as is known to occur) classes with a high number of special education students receive additional teacher resources (so that the

[19] This includes students labeled as "emotionally handicapped" or "emotionally disabled" as well as children diagnosed with autism.

overall student-teacher ratio falls) this may affect all students. However, the average effect of such adjustments is not obvious. These phenomena can be considered cases of indirect peer effects because the resource adjustments are caused by changes in the peer group. But if so, the instruments should then enter the individual equation as "contextual" or "exogenous" peer effects. For now, we maintain the exclusion assumptions, but if resource adjustment effects are significant they will be embedded in the effects of the instrumental variable for mean achievement. The effect of mean achievement should still represent genuine peer effects, but it may embed a combination of exogenous and endogenous effects.

VI. Results

We first consider results in the linear-in-means model, with instrumental variables for mean peer achievement gains as described above.[20] These models aim to capture any interactions or spillovers in contemporaneous behaviors that show up in current achievement gains and (in some specifications) the incidence of disciplinary action against students. We find that inclusion of teacher fixed effects alters the estimated coefficients in some cases.[21] (Refer to

[20] Refer to Table 1 for the results of the first stage regressions, which indicate significant explanatory power for the instruments, especially at the high school level. The fraction of peer "movers" has a significant impact on mean achievement in four of the six schooling level-by-subject equations. In three cases, the impact is significant and negative on achievement, as predicted, but for high school math the effect is strongly positive. The effects of mean age are negative (and significant) for math at all schooling levels, and positive and significant for reading at all levels. Fraction of peers designated as special education students has significant positive effects on mean achievement in each case except middle school math. If special education status generally predicts lower achievement at the individual level holding all other inputs constant, these results suggest that special education students receive more or better resources that compensate for any disadvantage they face, and that this may impose direct positive spillovers on peers. In future specifications we will experiment with including special education directly in the individual level equation rather than as an instrument. Note that in models with teacher fixed effects such effects also enter the first stage equation.

[21] School fixed effects are included in both models, but in the former case these are combined with the teacher effects to form the teacher-school spell effects. Teacher experience measures are included in all models.

Table 2 and Table 3 for the comparison.) For example, there are significant, positive spillovers of mean peer math achievement gains in elementary school under IV specifications both with and without teacher fixed effects. However, the coefficient point estimate is greater when teacher controls are included, suggesting a negative correlation between teacher quality and mean achievement gains in elementary level math. For middle-school math achievement, however, an insignificant coefficient becomes negative and significant when the teacher controls are included. In the case of reading achievement in middle school and high school, peer achievement spillovers are smaller when teacher-school effects are included than when they are not included. For high-school reading, the effect of mean achievement gains become insignificant with teacher fixed effects in the model. The differences in the results with and without fixed teacher effects suggest the possibility of either positive correlations between peer achievement gains and teacher quality (in the case of middle and high school reading) or negative correlations between these factors (in the case of elementary school math). Although such negative correlations are unexpected controlling for peer characteristics, these may reflect the ways that schools assign students to teachers. The results indicate that unobserved teacher inputs induce correlated effects on student achievement gains, and that models lacking such controls are prone to bias in either direction.

Unlike most previous studies of elementary school, we find no significant peer interactions on standardized reading scores at this level under any of the linear models. As we will see below in the analysis of the non-linear models, however, we cannot conclude that elementary school reading achievement spillovers are negligible in general. Also contrary to many previous findings, we obtain very few significant effects of the gender and racial composition variables. These results will prove consistent with the findings of the mean peer

fixed effect model discussed below. In some cases we observe either very large positive or very large negative effects of fraction of peers that are black, as in Tables 3 and 5. Table 3 in particular, insofar as the effect of fraction black peers becomes significant with the inclusion of teacher controls, suggests that teachers may be assigned to compensate for the negative impact of a high proportion of black students in a classroom, but these findings are not consistent across models and schooling levels. In addition, we have not yet estimated specifications in which effects of the race and gender composition variables are permitted to interact with individual race and gender. Based on earlier findings (such as Cooley 2005, Angrist and Lang 2004, Hoxby 2000) that the latter interactions are important, we are reluctant to make definitive pronouncements on race and gender effects based on our current findings.

Achievement spillovers for mathematics appear stronger at the grade level than at the classroom level in the IV models, as is seen by comparing estimated coefficients across Table 3 and Table 5. This could indicate either that effort and/or social learning in mathematics is more strongly influenced by grade-level peers outside of the classroom. However, as noted by Betts and Zau, the grade level model will yield biased estimates of the influence of grade level (but not classroom level) peers, since the model also embeds classroom-level influences. In addition, the grade-level instruments for peer achievement gains in math may be stronger proxies (than the classroom-level measures) for correlated innovations in unmeasured school, family, or student inputs.

In some cases the magnitude of the coefficients on the mean peer achievement gain is greater than one. Taken literally this implies unstable behavior of classroom mean achievement following any perturbation. Therefore we suspect that the IV coefficients on mean achievement are biased upward. A first guess might be that we are picking up correlations in unobserved,

time-varying inputs across students, or unobserved changes in teacher or school inputs. However, when we include the standard deviation of peer achievement gains in the model along with the mean (again using instrumental variables), as seen in Table 7, all spillovers from mean achievement gains become insignificant, and most point estimates decline in absolute magnitude. The only significant effect of the peer achievement dispersion measure occurs for high school math, and is positive (but less than one). The remaining point estimates are typically negative and insignificant. These findings suggest that the effect of peer achievement gains may vary across the individual achievement distribution. Dispersion itself does not in this model appear to have strong effects, but if effects are heterogeneous across students, the average effect of the mean will be weakest and least informative the greater the dispersion of ability in a given peer group. Accordingly we estimate a model in which peer effects are permitted to depend on a measure of individual student ability; specification and results are discussed below.

The effects of mean discipline incidents per peer are shown in Table 9. Most point estimates are negative, but a significant negative effect occurs only for high school math achievement gains. In the case of high school reading achievement, however, we observe a significant positive effect of peer discipline incidents. In most cases the inclusion of disciplinary incidents does not much alter the measured effects of mean peer achievement found in Table 3. However, in the case of middle school math the inclusion renders the negative peer effect of mean achievement gains insignificant (but still strongly negative), indicating a possible positive correlation between mean achievement gains and mean discipline incidents where the effect of the latter is negative. The noisiness of both coefficients in the model in which both are included suggests possible collinearity arising from the instrumental variables strategy. Additional

specifications will be required in order to get a clearer reading on the impact of peer disruptiveness.

Before discussing the non-linear specifications, we consider an alternative linear model that measures just spillovers from fixed peer characteristics, using an extension of the method developed by Arcidiacono et al., as described in detail above. The average peer fixed effect in this model represents the average fixed component of the test score gain among the peer group, consistent with the estimated individual fixed effects for each student in the peer group. This model aims to capture spillovers of peers' innate ability, or of any aspects of peer behavior that are consistent over time. However, insofar as the fixed individual effects predict effort and achievement gains in any given period, the mean peer fixed effects will also capture effort and achievement spillovers.

Looking at Tables 4 and 6, we see that the measured influence of mean peer fixed effects is significantly weaker than that of (instrumented) contemporaneous peer achievement gains. In the case of reading the mean peer fixed effects have a significant (small and positive) impact only in middle school at the classroom level. For math the effects are significant and positive (yet very small) for both elementary and high school at the classroom level, and significant (and yet smaller) at all levels of schooling measured at the broader grade level. In general these peer effects are at least an order of magnitude smaller than the effects of (instrumented) contemporaneous mean peer achievement gains.

These results agree with Cooley's findings that spillovers of contemporaneous effort appear significantly greater than spillovers from innate ability as measured by twice-lagged peer test scores. However, we feel it is too soon to claim that our results definitively confirm the greater importance of contemporaneous behavioral interactions among classroom and grade-

level peers, because in some cases the effects of mean peer achievement appear implausibly large, and in general are highly sensitive to the model specification. In addition, if the "constant-increment" specification of the individual fixed effects is incorrect (say, if the impact of innate ability on achievement follows a non-linear time trend), such fixed effects may be weak proxies for the impact of innate ability at a given point in time, both at the individual level and the mean peer level, and so the peer effects measured in this way could be biased downward.

The differences in the results between these two models indicate that the instrumental variables for mean achievement gains are not strongly correlated with the mean peer fixed effects. Again this could indicate that the component of achievement gains attributable to an individual's innate ability and parental inputs is not a constant value over time. Alternatively, the instruments may be isolating portions of current peer achievement gains that are orthogonal to the contribution of their fixed characteristics, such as adjustment costs of switching schools. However, with percent special education peers and mean peer age as added instruments, both being fixed characteristics, this seems less likely than the first explanation.

The magnitude and significance of the effects of mean achievement gains on individual gains changes rather dramatically when we allow the effects to depend on the individual student national percentile ranking based on his or her lagged test score. As seen in Table 10, we find large and highly significant effects of mean peer achievement gains, at all schooling levels and for both math and reading, among students in the lowest quintile. For the middle portion of the percentile ranking (from 21st to 80th) effects are still strong and positive for middle school and high school but not for elementary school. Where significant, the magnitudes are smaller than for the lowest ranked students. For top quintile students we observe both significant positive and significant negative effects, and on average the absolute magnitudes of the effects are smaller

than for the lower-ranked groups. Although again we hesitate to take coefficients on mean achievement that are greater than one at face value, the strong significance, as well as the predominance of positive spillovers (especially for reading), suggests that the non-linear specification is much more appropriate than the linear-in-means model. These results are consistent with the findings of Cooley (2005) as well as of Hoxby and Weingarth (2005), although Hanushek et al. (2003) did not find many significant differences in their measured peer effects across the achievement distribution. Considered in relative terms, it appears that lower-ability peers experience much stronger positive spillovers from increases in peer achievement levels. Even in cases where the highest-ranked students experience negative spillovers, as in elementary and high school math, these effects are more than outweighed by the positive effects at the lower levels. That is, the effect of shifting some high-performing peers to low-performing classrooms need not be zero-sum.

VII. Summary and Conclusions

This paper adds to a growing list of studies that use matched panel data in direct tests for peer effects among primary school students. As in the earlier studies, the panel data facilitate the identification of various peer influences by allowing a number of controls for endogenous variation in peer groups. Unlike many earlier studies, we are able to place students definitively within classroom groups with specific teachers, and we observe each teacher with more than one group of students. Our findings show that estimation of peer effects may be sensitive to the inclusion of teacher controls even if individual fixed effects are already included. We also find that peer effects are not "one-size-fits-all," as we observe differences in peer effects between mathematics and reading achievement (in general the effects for math are more robust), across

schooling levels, depending on whether we allow spillovers from contemporaneous peer outcomes or only from predetermined peer traits, and depending on the ability of the individual student. Therefore, caution should be exercised in generalizing from results of earlier studies using less comprehensive data. Achievement spillovers in reading do not appear significant until we permit interactions between mean achievement gains and individual ability. The non-linearities indicate biases in results under the linear-in-means specification, as well as the potential for Pareto improvements in outcomes from increasing the exposure of low-performing students to higher-achieving peers. However, we also find that greater ability dispersion among peers may impose some negative effects on achievement. Although in some cases the magnitude of achievement spillovers appears implausibly high, the results point to strong effects from contemporaneous behaviors that cannot be predicted on the basis of fixed peer factors. We also find some externalities on achievement from peer disruption, although they may be either positive or negative. We hope to resolve the questions raised by our findings as we continue to adjust the model specification and experiment with less restrictive models of cumulative achievement.

References

Akerlof, George A. and Rachel E. Kranton (2002). "Identity and Schooling: Some Lessons for the Economics of Education," 19 *Journal of Economic Literature* 1167-1201.

Andrews, Martyn, Thorsten Schank and Richard Upward (2004). "Practical Estimation Methods for Linked Employer-Employee Data," unpublished manuscript.

Angrist, Joshua D. and Alan B. Krueger (1991). "Does Compulsory Schooling Attendance Affect Schooling and Earnings?," 105 *Quarterly Journal of Economics*, 979-1014.

Angrist, Joshua D. and Kevin Lang (2004). "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program," 94 *American Economic Review* 1613-1634.

Arcidiacono, Peter, Gigi Foster, Natalie Goodpaster and Josh Kinsler (2005). "Estimating Spillovers in the Classroom with Panel Data," unpublished manuscript.

Argys, Laura M. and Daniel I. Rees (2004). "Searching for Peer Group Effects: A Test of the Contagion Hypothesis," unpublished manuscript.

Becker, Gary S. and Kevin M. Murphy (2000). *Social Economics: Market Behavior in a Social Environment*. Cambridge and London: Harvard University Press and Belknap Press.

Bernheim, B. Douglas (1994). "A Theory of Conformity," 102 *Journal of Political Economy* 841-877.

Betts, Julian R. and Andrew Zau (2004). "Peer Groups and Academic Achievement: Panel Evidence from Administrative Data," unpublished manuscript.

Bifulco, Robert and Helen F. Ladd (2006). "The Impacts of Charter Schools on Student Achievement: Evidence from North Carolina," 1 *Education Finance and Policy* xxx-xxx.

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

Caucutt, Elizabeth (2002). "Educational Vouchers When there are Peer Group Effects -- Size Matters," 43 *International Economic Review*, 195-222.

Cooley, Jane (2005). "Desegregation and the Achievement Gap: Do Diverse Peers Help?," unpublished manuscript.

Epple, Dennis, Elizabeth Newlon and Richard Romano (2002). "Ability Tracking, School Competition and the Distribution of Educational Benefits," 83 *Journal of Public Economics*, 1-48.

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

Figlio, David N. (2004a). "Boys Named Sue: Disruptive Children and Their Peers," unpublished manuscript.

Figlio, David N. (2004b). "Names, Expectations and Black Children's Achievement," unpublished manuscript.

Fryer, Roland G. and Paul Torelli (2005). "An Empirical Analysis of 'Acting White'," NBER working paper #11334.

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

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

Hoxby, Caroline M., (2000). "Peer Effects in the Classroom: Learning from Gender and Race Variation", National Bureau of Economic Research Working Paper 7867.

Hoxby, Caroline M. and Gretchen Weingarth (2005). "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects," unpublished manuscript.

Lazear, Edward. P. (2001). "Educational Production," 116 *Quarterly Journal of Economics,* 777-803.

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

Moffitt, R. (2001). "Policy Interventions, Low-Level Equilibria, and Social Interactions", in *Social Dynamics*, S. Durlauf and H. P. Young, eds., Cambridge Mass.: MIT Press, 45-82.

Rivkin, Steven G., Eric A. Hanushek and John F. Kain (2005). "Teachers, Schools and Academic Achievement," 73 *Econometrica*, 417-458.

Sass, Tim R. (2006). "Charter Schools and Student Achievement in Florida," 1 *Education Finance and Policy* xxx-xxx.

Todd, Petra E. and Kenneth I. Wolpin (2003). "On the Specification and Estimation of the Production Function for Cognitive Achievement," 113 *Economic Journal* F3-F33.

Vigdor, Jacob and Thomas Nechyba (2004). "Peer Effects in Elementary School: Learning from 'Apparent' Random Assignment," unpublished manuscript.

Zabel, Jeffrey E. (2005). "The Impact of Peer Effects on Student Outcomes in New York City Public Schools," unpublished manuscript.

Table 1
First-Stage Iterated IV Estimates of
Exogenous Individual and Classroom-Peer Characteristics on Mean
Classroom Peer Math and Reading Achievement Gains in Florida, 1999/2000-2003/2004
(100+ Elementary Schools, 100+ Middle Schools, 100+ High Schools)

	Math			Reading		
	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades 9-10)	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades9-10)
Fraction of Peers who are Female	-4.0359*** (9.34)	-0.2430 (1.03)	-2.0249*** (7.64)	3.4758*** (8.27)	0.8756*** (5.03)	-1.2591*** (4.81)
Fraction of Peers who are Black	4.6796*** (8.83)	-7.2043*** (25.81)	-0.2819 (0.89)	-0.1589 (0.31)	3.3408*** (14.97)	2.7235*** (7.99)
Fraction of Peers who Changed Schools	0.1832 (0.41)	-1.9975*** (8.67)	12.3887*** (55.49)	0.1263 (0.29)	-1.3933*** (8.21)	-6.3490*** (19.29)
Mean Age of Peers (in Months)	-0.2263*** (9.58)	-0.1431*** (9.96)	-0.3022*** (29.83)	0.1314*** (5.76)	0.1333*** (11.85)	0.1390*** (9.11)
Fraction of Peers who are Special Ed. Students	1.9486*** (4.16)	-3.2243*** (8.66)	2.7080*** (5.99)	7.0980*** (15.69)	1.5852*** (5.85)	2.6942*** (5.62)
Class Size	-0.0262** (2.11)	-0.0435*** (7.62)	-0.0437*** (7.09)	-0.1977*** (16.37)	-0.0113*** (2.77)	0.0309*** (4.78)
"Structural" Mover	0.1786 (0.36)	-0.1116 (0.79)	-1.8502*** (8.04)	0.4740 (0.99)	0.0250 (0.22)	0.0268 (0.12)
"Non-Structural" Mover	-0.2327** (2.03)	-0.0383 (0.36)	-1.3858*** (6.30)	0.0320 (0.29)	0.0722 (0.80)	0.3412 (1.59)
R^2	0.09	0.20	0.38	0.17	0.16	0.20
Number of Observations	38,715	67,716	62,950	39,023	105,708	49,635

Models also include indicators for teachers with no experience/1-year experience/2-4 years experience and year, grade level, and repeater-by-grade dummies. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test. Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores.

Table 2
Iterated IV Estimates of the Effects of Mean Classroom Peer Characteristics on Student Math and Reading Achievement in Florida, 1999/2000-2003/2004 (100+ Elementary Schools, 100+ Middle Schools, 100+ High Schools)

	Math			Reading		
	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades 9-10)	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades 9-10)
Mean Peer Achievement Gain	0.7664*	-0.5914	0.0177	-0.1447	1.7256**	0.5787***
	(1.69)	(1.19)	(0.35)	(0.24)	(2.55)	(3.69)
Fraction of Peers who are Female	-3.3972	-2.2945	-0.9565	3.4405	-0.3885	2.5735
	(0.75)	(1.26)	(0.53)	(0.84)	(0.23)	(1.19)
Fraction of Peers who are Black	2.4437	-0.6523	2.5904	5.8937	1.2527	2.0521
	(0.57)	(0.18)	(1.23)	(1.29)	(0.45)	(0.55)
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Fixed Effects	No	No	No	No	No	No
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Students	26,146	44,373	39,630	26,460	69,628	37,618
Number of Observations	43,278	73,439	69,776	43,578	117,297	57,668

Models include the following time varying student/class/teacher characteristics: number of schools attended in current year, "structural" move by student, "non-structural move" by student, indicator of a student repeating a grade, class size, indicators for teachers with no experience/1-year experience/2-4 years experience. All models also include year, grade level, and repeater-by-grade dummies. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test. Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores.

Table 3
Iterated IV Estimates of the Effects of Mean Classroom Peer Characteristics on Student Math and Reading Achievement in Florida, 1999/2000-2003/2004 (100+ Elementary Schools, 100+ Middle Schools, 100+ High Schools)

	Math			Reading		
	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades 9-10)	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades9-10)
Mean Peer Achievement Gain	1.1223*	-3.1672**	0.0348	-0.8779	1.1995**	0.0240
	(1.65)	(2.24)	(0.62)	(0.92)	(2.04)	(0.08)
Fraction of Peers who are Female	-3.7982	5.9484	-0.7134	-1.5283	0.4479	1.5096
	(0.44)	(1.61)	(0.35)	(0.22)	(0.24)	(0.73)
Fraction of Peers who are Black	6.4012	-19.2891**	2.9004	-4.3387	3.4363	3.6772
	(0.72)	(1.96)	(1.22)	(0.54)	(1.04)	(0.96)
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Students	26,146	44,373	39,630	26,460	69,628	37,618
Number of Observations	43,278	73,439	69,776	43,578	117,297	57,668

Models include the following time varying student/class/teacher characteristics: number of schools attended in current year, "structural" move by student, "non-structural move" by student, indicator of a student repeating a grade, class size, indicators for teachers with no experience/1-year experience/2-4 years experience. All models also include year, grade level, and repeater-by-grade dummies. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test. Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores.

Table 4
Iterated OLS Estimates of the Effects of Mean Classroom Peer Fixed Effects on Student Math and Reading Achievement in Florida, 1999/2000-2003/2004 (100+ Elementary Schools, 100+ Middle Schools, 100+ High Schools)

	Math			Reading		
	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades 9-10)	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades 9-10)
Mean Peer Fixed Effect	0.0205*	0.0078	0.0296**	0.0178	0.0152**	0.0128
	(1.93)	(0.87)	(2.04)	(1.50)	(2.21)	(1.47)
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Students	26,146	44,373	39,630	26,460	69,628	37,618
Number of Observations	43,278	73,439	69,776	43,578	117,297	57,668

Models include the following time varying student/class/teacher characteristics: number of schools attended in current year, "structural" move by student, "non-structural move" by student, indicator of a student repeating a grade, class size, indicators for teachers with no experience/1-year experience/2-4 years experience. All models also include year, grade level, and repeater-by-grade dummies. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test. Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores.

Table 5
Iterated IV Estimates of the Effects of Mean Grade-level Peer Characteristics on Student Math and Reading Achievement in Florida, 1999/2000-2003/2004 (100+ Elementary Schools, 100+ Middle Schools, 100+ High Schools)

	Math			Reading		
	Elementary (Grades 4-5)	Middle[a] (Grades 6-8)	High School (Grades 9-10)	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades9-10)
Mean Peer Achievement Gain	1.8696*** (3.67)	-0.6682 (1.00)	1.0585*** (7.16)	0.6655 (0.74)	-0.1469 (0.13)	0.0831 (0.10)
Fraction of Peers who are Female	7.9359 (0.40)	9.4703 (0.70)	-32.6184** (2.31)	12.6048 (0.81)	10.9490 (0.87)	7.3405 (0.35)
Fraction of Peers who are Black	-40.6621* (1.72)	39.5055* (1.87)	-2.3767 (0.21)	7.0328 (0.28)	-0.2566 (0.01)	-37.7522 (1.14)
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Students	26,146	20,755	39,630	26,460	69,628	37,618
Number of Observations	43,278	45,974	69,776	43,578	117,297	57.668

Models include the following time varying student/class/teacher characteristics: number of schools attended in current year, "structural" move by student, "non-structural move" by student, indicator of a student repeating a grade, class size, indicators for teachers with no experience/1-year experience/2-4 years experience. All models also include year, grade level, and repeater-by-grade dummies. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test. Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores. [a]Due to non-convergence of iterative-effects-model, this specification was estimated by standard fixed effects methods with explicit dummies for each teacher-school spell.

Table 6
Iterated OLS Estimates of the Effects of Mean Grade-level Peer Fixed Effects on Student Math and Reading Achievement in Florida, 1999/2000-2003/2004 (100+ Elementary Schools, 100+ Middle Schools, 100+ High Schools)

	Math			Reading		
	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades 9-10)	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades 9-10)
Mean Peer Fixed Effect	0.0058*	0.0071**	0.0138***	0.0067	0.0014	0.0052
	(1.91)	(2.35)	(2.94)	(1.37)	(0.58)	(1.21)
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Students	26,146	44,373	39,630	26,460	69,628	37,618
Number of Observations	43,278	73,439	69,776	43,578	117,297	57,668

Models include the following time varying student/class/teacher characteristics: number of schools attended in current year, "structural" move by student, "non-structural move" by student, indicator of a student repeating a grade, class size, indicators for teachers with no experience/1-year experience/2-4 years experience. All models also include year, grade level, and repeater-by-grade dummies. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test. Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores.

Table 7
Iterated IV Estimates of the Effects of Mean and Variance of Classroom Peer Characteristics on Student Math and Reading Achievement in Florida, 1999/2000-2003/2004 (100+ Elementary Schools, 100+ Middle Schools, 100+ High Schools)

	Math			Reading		
	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades 9-10)	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades9-10)
Mean Peer Achievement Gain	0.9064	-1.5511	-0.0275	-0.8003	1.1588	0.1420
	(0.66)	(0.42)	(0.45)	(0.81)	(1.14)	(0.06)
Standard Deviation in Peer Achievement Gain	-0.1330	-14.6418	0.6589***	-0.7556	-0.7872	4.9239
	(0.11)	(0.75)	(2.85)	(0.72)	(0.28)	(0.16)
Fraction of Peers who are Female	-5.8061	-15.9478	1.9116	-2.1357	-0.3801	11.5305
	(0.38)	(0.62)	(0.85)	(0.32)	(0.09)	(0.21)
Fraction of Peers who are Black	7.4327	-85.1418	6.2545**	-6.0765	1.3889	3.0828
	(0.59)	(0.73)	(2.47)	(0.74)	(0.11)	(0.17)
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Students	26,146	44,373	39,630	26,460	69,628	37,618
Number of Observations	43,278	73,439	69,776	43,578	117,297	57,668

Models include the following time varying student/class/teacher characteristics: number of schools attended in current year, "structural" move by student, "non-structural move" by student, indicator of a student repeating a grade, class size, indicators for teachers with no experience/1-year experience/2-4 years experience. All models also include year, grade level, and repeater-by-grade dummies. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test. Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores.

Table 8
Iterated IV Estimates of the Effects of Mean and Variance of Classroom Peer Fixed Effects on Student Math and Reading Achievement in Florida, 1999/2000-2003/2004 (100+ Elementary Schools, 100+ Middle Schools, 100+ High Schools)

	Math			Reading		
	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades 9-10)	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades 9-10)
Mean Peer Fixed Effect	0.0147	0.0285**	0.0574***	0.0160	0.0126	0.0333**
	(1.03)	(2.48)	(2.90)	(0.88)	(1.02)	(2.03)
Standard Deviation of Peer Fixed Effects	-0.0212**	-0.0071	-0.0354**	-0.0238***	-0.0203*	-0.0145
	(2.08)	(0.61)	(2.34)	(3.26)	(1.68)	(1.31)
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Students	26,146	44,373	39,630	26,460	69,628	57,668
Number of Observations	43,278	73,439	69,776	43,578	117,297	37,618

Models include the following time varying student/class/teacher characteristics: number of schools attended in current year, "structural" move by student, "non-structural move" by student, indicator of a student repeating a grade, class size, indicators for teachers with no experience/1-year experience/2-4 years experience. All models also include year, grade level, and repeater-by-grade dummies. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test. Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores.

Table 9
Iterated IV Estimates of the Effects of Mean Classroom Peer Characteristics on Student Math and Reading Achievement in Florida, 1999/2000-2003/2004 (100+ Elementary Schools, 100+ Middle Schools, 100+ High Schools)

	Math			Reading		
	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades 9-10)	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades9-10)
Mean Peer Achievement Gain	1.1983*	-3.3295	0.0224	-0.4519	1.1617**	0.0644
	(1.74)	(1.23)	(0.41)	(0.55)	(2.04)	(0.21)
Mean Discipline Incidents per Peer	-14.2506	-3.8485	-2.5485**	-0.5178	0.8810	2.0152*
	(1.43)	(0.90)	(2.33)	(0.05)	(0.91)	(1.71)
Fraction of Peers who are Female	-6.4839	3.1757	-1.9000	-1.1309	1.0516	2.6135
	(0.69)	(0.79)	(0.90)	(0.17)	(0.51)	(1.27)
Fraction of Peers who are Black	9.2224	-16.8124	4.5077*	-2.3834	2.5993	1.7084
	(0.98)	(1.17)	(1.81)	(0.33)	(0.75)	(0.40)
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Students	26,146	44,373	39,630	26,460	69,628	37,618
Number of Observations	43,278	73,439	69,776	43,578	117,297	57,668

Models include the following time varying student/class/teacher characteristics: number of schools attended in current year, "structural" move by student, "non-structural move" by student, indicator of a student repeating a grade, class size, indicators for teachers with no experience/1-year experience/2-4 years experience. All models also include year, grade level, and repeater-by-grade dummies. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test. Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores.

Table 10
Iterated IV Estimates of the Effects of Mean Classroom Peer Characteristics on Student Math and Reading Achievement in Florida by Lagged Student National Percentile Ranking, 1999/2000-2003/2004 (100+ Elementary Schools, 100+ Middle Schools, 100+ High Schools)

	Math			Reading		
	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades 9-10)	Elementary (Grades 4-5)	Middle (Grades 6-8)	High School (Grades9-10)
Mean Peer Achievement Gain × 1–20 Natl. Pct.$_{t-1}$	2.3199**	5.8335***	2.9330***	2.4410*	6.8134***	5.2057***
	(2.51)	(7.63)	(23.76)	(1.87)	(9.14)	(5.14)
Mean Peer Achievement Gain× 21–80 Natl. Pct. $_{t-1}$	0.3581	3.7587***	1.1449***	0.3404	4.7817***	1.3873*
	(0.41)	(4.79)	(14.28)	(0.26)	(6.29)	(1.89)
Mean Peer Achievement Gain × 81–99 Natl. Pct. $_{t-1}$	-1.8475**	1.1747	-0.3339***	-2.6368**	2.3469***	1.7140**
	(2.19)	(1.47)	(4.41)	(2.18)	(3.11)	(2.43)
Fraction of Peers who are Female	-16.9245	-10.6687***	-0.4435	-1.1554	0.2312	1.2998
	(1.55)	(3.55)	(0.20)	(0.20)	(0.11)	(0.52)
Fraction of Peers who are Black	6.6956	16.9899***	6.6261***	4.7021	-15.9300***	0.3469
	(0.66)	(2.96)	(2.78)	(0.41)	(4.28)	(0.09)
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Students	26,146	44,373	39,630	26,460	69,628	37,618
Number of Observations	43,278	73,439	69,776	43,578	117,297	57,668

Models include the following time varying student/class/teacher characteristics: number of schools attended in current year, "structural" move by student, "non-structural move" by student, indicator of a student repeating a grade, class size, indicators for teachers with no experience/1-year experience/2-4 years experience. All models also include year, grade level, and repeater-by-grade dummies. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test. Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores.