# Disentangling School and Peer Effects: A Non-linear Approach with School Lotteries

Mariana Laverde * Minseon Park† Suk Joon Son‡

March 2026

PRELIMINARY AND INCOMPLETE. PLEASE DO NOT CIRCULATE

**Abstract**

We develop a unified framework to jointly estimate and separately identify school effects and peer effects on student achievement. Identification exploits random tie-breaker lotteries in centralized school choice systems, which generate exogenous variation in both school assignments and peer composition. To flexibly model peer effects while preserving statistical power, we summarize the full distribution of peers' baseline achievement using functional principal component analysis (FPCA). Endogeneity in school and peer assignments is addressed by adapting the recentering approach of Borusyak and Hull (2023), which isolates lottery-driven variation in realized assignments. Applying the framework to New York City middle schools, we find that peer effects are predominantly negative and strongly heterogeneous by baseline achievement. A simulation exercise for new entrants illustrates that peer and school effects can operate in opposing directions, implying meaningful trade-offs in the value of school assignments.

**Keywords:** Peer Effects, School Value-Added, Functional Principal Component Analysis, Educational Production Function

*Boston College, laverdem@bc.edu.
†University of Michigan, minseonp@umich.edu.
‡University of Tokyo, sjson@e.u-tokyo.ac.jp.

# 1 Introduction

A central question in the economics of education is how different inputs translate into student learning. Schools and peers are among the most studied determinants of academic achievement, yet their independent contributions remain understudied. Measuring their relative importance is consequential because they imply different margins for intervention. If peer composition is a primary driver of achievement, reforms that reshape classroom composition—such as tracking—may be effective in inducing changes in student achievement. If instead school-level inputs, such as teacher quality or institutional practices, play the dominant role, policies that improve school resources or management will yield higher returns.

In this paper, we propose a framework that jointly estimates peer and school effects by exploiting randomness generated by algorithmic school assignments. We show that widely used assignment mechanisms generate two distinct types of exogenous variation—randomness in school assignments and randomness in peer composition within schools—both of which are required for separate identification of peer and school effects. Furthermore, our unified model allows for flexible peer effects where student outcomes depend on the full distribution of peers' test scores, rather than only their mean or a few summary statistics. By relaxing the standard linear–in–means assumption, the framework accommodates richer peer composition effects and does not impose that aggregate peer effects average to zero.

Estimating our proposed model raises two main challenges. First, to operationalize our flexible model of peer effects—where student outcomes may depend on the full distribution of peer test scores—we require a parsimonious representation of peer composition. Second, identifying causal effects requires exogenous variation in both school assignments and peer composition. Under a flexible specification of peer effects, this requirement is more demanding, as identification must rely on variation across the entire distribution of peer achievement rather than a limited set of moments.

We address the first challenge by applying Functional Principal Component Analysis (FPCA) (Chang, Durlauf, Lee, and Park, 2023) to summarize the high dimensionality of peer test score distributions. FPCA represents each distribution as a weighted combination of a small number of orthogonal component functions that capture the dominant

dimensions of variation in peer composition. The corresponding weights—individual-specific loadings on these components—provide a low-dimensional summary of each student's peer environment. This representation allows us to flexibly model rich peer effects while maintaining tractability in the regression analysis.

Regarding the second challenge, we show that algorithmic school assignments produce distinct types of exogenous variation—in school assignments and in peer composition within schools—that together allow separate identification of school and peer effects. The former arises when students face uncertainty over their assigned school, while the latter arises even for students with deterministic assignments, as changes in other students' assignments alter their realized peer composition.[1] We show that identification does not rely on the nonlinearity of peer effects or their summarization through FPCA.

Even when such exogenous variation is available, estimation remains challenging due to the presence of high-dimensional confounders—namely, student preferences and priorities. To efficiently control for these confounders, we use expectation-based controls as in Borusyak and Hull (2023). Intuitively, preferences and priorities jointly determine each student's assignment probabilities across schools, as well as their expected peer distributions. Conditional on these expectations, realized school assignments and co-assigned peer compositions vary only through lottery tie-breakers, providing the exogenous variation needed for identification.

Deviations between assignments and enrollments may be systematically related to students' potential outcomes, and thus may bias the estimates. To address this concern, we estimate our model under an assumption similar to that in Angrist, Hull, Pathak, and Walters (2024). Specifically, we assume that enrollment decisions and the resulting peer compositions are independent of students' potential outcomes, conditional on student demographics, assignment probabilities at all schools, and expected peer distributions. We assess the plausibility of this assumption empirically.

We apply our framework to data from New York City's centralized middle school choice system. We find that exposure to higher-achieving peers is associated with lower

[1] In our setting, on average, 9.8% of students (over 21,000 students) face uncertainty in school assignments across two sets of lottery draws. Among students without assignment uncertainty, 7.7% (over 13,500 students) experience changes in their peer distribution exceeding 5% in total variation distance (TVD). TVD equals one-half of the $L^1$ distance between two discrete distributions and can be interpreted as the minimum fraction of probability mass that must be relocated to transform one distribution into the other. We compute TVD using 20 percentile-based bins of the peer score distribution.

test score outcomes relative to exposure to lower-achieving peers in the full sample. This negative relationship displays modest departures from linearity but is well summarized by a downward-sloping peer effect function. Disaggregating by baseline achievement reveals sharp differences. For lower-achieving students, peer effects are close to zero on average, suggesting limited scope for gains or losses through peer composition. In contrast, higher-achieving students experience pronounced and consistently negative peer effects: replacing 10% of same-scoring peers with peers scoring one standard deviation higher reduces test scores by approximately 0.2 standard deviations. These findings are consistent with rank or relative-position effects documented in prior work (Carneiro, Cruz-Aguayo, Salvati, and Schady, 2025; Carrell, Sacerdote, and West, 2013; Murphy and Weinhardt, 2020).

Turning to school effects, estimates of school effectiveness change once peer composition is accounted for. In the full sample, a one standard deviation increase in school effectiveness raises student test scores by 0.1 standard deviations, when peer effects are accounted for. When peer effects are not included, the effect drops to about 0.08 standard deviations, driven by downward shifts in the right tail of the value-added distribution. This pattern reflects the fact that high value-added schools tend to enroll higher-achieving peers, and peer effects are negative at the top of the achievement distribution. When peer composition is omitted, these negative peer effects are absorbed into estimated school effects, compressing differences across schools. Accounting explicitly for peer composition separates these forces, leading to a wider dispersion of estimated school effectiveness.

Finally, because estimates of the peer effect function alone do not reveal how much peer-induced gains actually differ across schools and how these gains compare to those generated by school value-added, we complement our analysis with a simulation exercise. The exercise evaluates, for a new entrant at different points of the baseline achievement distribution, the gains associated with assignment to each school through both school effects and peer composition, holding incumbent assignments fixed.

The simulation highlights that peer effects offer limited scope for gains among lower-achieving students, whose predicted outcomes are driven primarily by school value-added. For higher-achieving students, by contrast, peer effects vary in magnitude similarly to school effects and operate in opposing directions, generating meaningful trade-offs across schools. The remainder of the introduction outlines how our framework accommodates rich peer composition effects and how lottery-based variation enables separate identifica-

tion of school and peer effects.

**Related Literature** Our framework and findings contribute to several strands of the literature on peer effects, school effectiveness, and assignment mechanisms.

First, we contribute to the peer effects literature by moving beyond linear–in–means models that summarize peer composition using a single moment and instead depend on rich features of the peer achievement distribution. A small number of studies allow for nonlinearities by discretizing peer characteristics into pre-specified summaries of the achievement distribution—such as measures of dispersion or the share of peers above or below fixed percentile thresholds (Hoxby and Weingarth Salyer, 2006; Gibbons and Telhaj, 2016; Lavy, Silva, and Weinhardt, 2012)— but in practice are constrained to a limited number of such summaries to preserve statistical power. In contrast, we use FPCA to represent peer achievement distributions using a finite number of components that capture the dominant modes of variation present in the data, while retaining statistical power. Substantively, our findings add evidence of the presence of rank effects, consistent with recent experimental results (Carneiro, Cruz-Aguayo, Salvati, and Schady, 2025).

Second, we relate to a growing literature that exploits random variation in centralized school choice algorithms to estimate school effectiveness (Abdulkadiroğlu, Angrist, and Pathak, 2014; Abdulkadiroğlu, Angrist, Narita, and Pathak, 2017, 2022; Angrist, Hull, Pathak, and Walters, 2024). While this literature focuses primarily on identifying school value added, our work clarifies the conditions under which lottery–induced variation can be used to recover *both* school and peer effects. Our empirical strategy builds on the recentering approach of Borusyak and Hull (2023) and adapts it to settings with endogenous peer composition, showing how conditionally random assignments generated by known assignment rules can be leveraged for this purpose, even when ideal compliance conditions are not fully satisfied.

More broadly, our identification strategy relates to a long tradition of using quasi-random peer assignments to study peer effects, including Boozer and Cacciola (2001) and Duflo, Dupas, and Kremer (2011) in primary and secondary education, and Sacerdote (2001), Zimmerman (2003), Carrell, Fullerton, and West (2009), and Lyle (2009) in higher education.

Finally, our work contributes to a small but growing literature that jointly identifies

and estimates school and peer effects. Only a limited number of recent papers, including Allende (2019) and Woodcock, Friesen, and Meilman Lomaz Cohn (2025), take steps in this direction while keeping an assumption of linear-in-mean peer effects. To our knowledge, we are the first to jointly identify peer and school effects under a nonlinear model of peer effects.

# 2 Model and Methods

This section introduces our model of test score production and the empirical strategy used to estimate school and peer effects. We first present a production function in which outcomes depend on a school effect and a flexible peer effect that is a functional of the distribution of peers' baseline achievement. We then describe how we summarize high-dimensional peer distributions using Functional Principal Component Analysis (FPCA), yielding a low-dimensional set of peer-composition measures that can be incorporated into regression analysis.

## 2.1 Test Score Production Function

We model student test scores as the sum of a school effect, a flexible peer effect that depends on the distribution of peers' achievement, and an idiosyncratic component capturing observed and unobserved ability.

Specifically, the production function for student $i$'s test score is given by

$$y_i = \alpha^g + \int \beta^g(r)\, f_i(r)\, dr + \mu^g_{j(i)} + \varepsilon_i, \tag{1}$$

where $g \equiv g(i) \in \mathcal{G}$ indexes the demographic group to which student $i$ belongs, and $\mathcal{G}$ partitions students into disjoint and exhaustive groups. All model parameters are allowed to vary with $g$, allowing for heterogeneity in peer and school effects across student subgroups. This framework nests a homogeneous-effects model as a special case when $\mathcal{G}$ contains a single group encompassing all students. In our empirical application, we primarily emphasize heterogeneity by students' baseline achievement.

The variable $r = \tau - \tau_i$ denotes the difference between a peer's baseline test score $\tau$ and student $i$'s own baseline test score $\tau_i$. The function $f_i(r)$ is the probability density function of these relative test scores among student $i$'s co-enrolled peers, excluding $i$ herself. The function $\beta^g(r)$ captures the marginal effect of peers with relative achievement $r$ on student $i$'s outcome, and the integral aggregates these effects over the entire distribution of relative peer achievement.

The term $j(i)$ denotes the school in which student $i$ enrolls, and $\mu^g_{j(i)}$ captures the school effect for students in demographic group $g$. This component summarizes the effects of school-level inputs including teachers, leadership, and other institutional characteristics. The term $\alpha^g$ is a group-specific intercept and the residual term $\varepsilon_i$ captures idiosyncratic determinants of achievement, reflecting student ability and other observed and unobserved factors.

Building on the definitions of school effects and residual outcomes above, we conceptually distinguish between school and peer effects as follows. School effects reflect characteristics of a school that are invariant to peer composition in the short run. In contrast, peer effects capture both direct effects of peers and school's instantaneous response to changes in peer composition. For example, teachers may adjust curriculum or instructional practices in response to the distribution of academic preparedness among enrolled students, the impact of these cohort-specific responses would be summarized by peer effects, rather than school effects.

The peer effect function $\beta^g(\cdot)$ admits a clear counterfactual interpretation. Holding the school and individual components fixed, consider reallocating mass within the peer achievement distribution. For any two points $r$ and $r'$ in the support of $f_i(\cdot)$ and any mass $m$, the expression $(\beta^g(r') - \beta^g(r))\, m$ represents the change in $y_i$ induced by removing a mass $m$ of peers with relative achievement approximately equal to $r$ and reallocating that same mass to peers with relative achievement $r'$.[2] The level of $\beta^g(\cdot)$ is normalized so that $\int \beta^g(r)\, dr = 0$, reflecting the fact that the level of $\beta^g(\cdot)$ is not separately identified from the intercept $\alpha^g$: any constant shift in $\beta^g(\cdot)$ are absorbed by the intercept $\alpha^g$, since $f_i(\cdot)$ integrates to one.

Modeling peer effects as a function of relative rather than absolute baseline achievement implies that peer effects depend on distance in achievement, rather than on a stu-

[2]See Appendix A.4 for a formal proof.

dent's absolute position in the achievement distribution. For example, a high-achieving peer need not have the same effect on low- and high-achieving students. At the same time, two pairs of peers who differ by the same amount in baseline achievement are assumed to exert the same peer effect on one another, regardless of their absolute location in the distribution. In addition, we estimate a version of the model in which $\mathcal{G}$ groups students by baseline achievement level $\tau_i$. Since $r = \tau - \tau_i$ differs from $\tau$ only by the approximately constant shift $\tau_i$ within each group, the two specifications are essentially equivalent within each group.

Our model of peer effects nests many commonly studied peer effect specifications. The widely used linear–in–means model, which restricts peer effects to depend only on average peer achievement is a special case of our specification

$$y_i = \alpha + b \underbrace{\int r \, f_i(r) \, dr}_{=E(r)} + \mu_{j(i)} + \varepsilon_i, \tag{2}$$

where $b$ is a constant. While standard linear–in–means model uses mean absolute score $E[\tau]$ rather than mean relative score $E[r] = E[\tau - \tau_i]$, the two models are equivalent if we include $\tau_i$ as a control.[3] Under this restriction, the linear–in–means model is a special case of our model where $\beta^g(r) = br$.

More generally, our framework accommodates peer effects that depend on higher-order moments of the peer achievement distribution. Any linear combination of moments of relative test scores, which can be written as $\sum_{l=1}^{k} b_l \int \tau^l f(\tau) \, d\tau$, is representable within our model by setting $\beta^g(\tau) = \sum_{l=1}^{k} b_l \tau^l$.

While our framework is flexible to incorporates many prior peer effect models as special cases, it does not directly capture peer-composition effects that are centered around a group mean rather than the student's own score. For example, the within-school variance of achievement—a moment centered around the school mean—cannot be represented by $\int \beta(r) f_i(r) \, dr$. While we can include it explicitly or introduce an additional peer-effect term based on scores normalized by the group mean (e.g., $\tau - \bar{\tau}_{j(i)}$), we do not pursue this for parsimony, as the second moment of relative test score distribution—which we already capture with our baseline specification—is likely highly correlated with within-

[3] Let $f_i^{\tau}(\cdot)$ denote the density of peers' *absolute* baseline scores $\tau$, and let $f_i^{r}(\cdot)$ denote the density of *relative* scores $r = \tau - \tau_i$. By a change of variables, we have $b \int \tau f_i^{\tau}(\tau) \, d\tau = b \int (r + \tau_i) \, f_i^{r}(r) \, dr = b \int r f_i^{r}(r) \, dr + b\tau_i$.

school variance.

## 2.2 Summarizing Peer Distributions with FPCA

Our model allows student outcomes to depend on the entire distribution of peers' baseline test scores, $f_i$. While this flexibility is central to our approach, it poses a practical challenge: each peer distribution is a high-dimensional object that cannot be directly included as a regressor in standard regression analyses. To address this challenge, we approximate each peer distribution using a small number of summary functions, allowing us to represent $f_i$ with a finite-dimensional vector that can be incorporated directly into the outcome equation.

We construct these summaries using functional principal component analysis (FPCA), following Chang, Durlauf, Lee, and Park (2023). FPCA identifies a set of orthogonal basis functions that capture the dominant patterns of variation across peer score distributions in the data. Each peer density function is represented as a weighted combination of the leading principal components (functions), yielding a low-dimensional representation that preserves the most informative features of peer composition while remaining tractable for regression analysis. The resulting weights (loadings) summarize student-specific peer compositions and serve as the regressors in our empirical analysis.

Intuitively, FPCA can be viewed as applying principal component analysis (PCA) to a finely discretized representation of each peer distribution. Specifically, we evaluate each function $f_i(\cdot)$ on a dense grid of relative test scores $(r_1, \ldots, r_T)$ and treat the resulting vector $(f_i(r_1), \ldots, f_i(r_T))$ as the object to which PCA is applied. PCA then extracts orthogonal components that capture the main sources of variation across these discretized distributions. By incorporating individual loadings on the components in a regression and inverting back the coefficients on these loadings into the functional space, we estimate the function $\beta(\cdot)$ that represents the impact of the full distribution of peers $f_i(\cdot)$ –not its finite approximation– on test scores.

This approach has several advantages. First, by construction, the leading functional principal components provide the $M$-dimensional linear approximation of the peer distributions that minimizes average squared approximation error, ensuring that the finite-dimensional representation loses as little information as possible on average. Second,

FPCA yields loadings (which enter our regressions as regressors) that are mutually uncorrelated. This implies improved statistical efficiency and no omitted-variable bias due to FPCA approximation error (Chang, Durlauf, Lee, and Park, 2023).[4]

In the remainder of this section, we describe the construction of these components, the estimation of student-specific loadings, their inversion to recover the functional coefficient $\beta(\cdot)$, and the criteria used to select the number of components retained.

We now show how FPCA yields a finite-dimensional representation of the peer distribution that can be incorporated directly into equation 1. Let $\langle \beta, f_i \rangle := \int \beta(r) f_i(r)\, dr$ denote the peer effects term in the test score production function in equation 1. FPCA yields a sequence of orthonormal basis functions $\{v_m\}_{m=1}^{\infty}$ spanning the relevant function space. In practice, we approximate each peer distribution using the leading $M$ components.

We therefore write

$$f_i(r) \approx \sum_{m=1}^{M} \langle v_m, f_i \rangle v_m(r).$$

The coefficient $\langle v_m, f_i \rangle$ represents the "signed length" of the projection of $f_i$ onto the direction defined by $v_m$—that is, it measures how much of $f_i$ can be explained by the $m$-th component, with positive (negative) values indicating alignment with (opposition to) the pattern captured by $v_m$. These projection coefficients $\langle v_m, f_i \rangle$—called the FPCA loadings under standard terminology—enter our regressions as regressors. Under this approximation, the peer effects term can be written as

$$\langle \beta, f_i \rangle \approx \sum_{m=1}^{M} \langle v_m, \beta \rangle \langle v_m, f_i \rangle,$$

which yields the finite-dimensional regression equation

$$y_i \approx \alpha + \pi(\beta)' \pi(f_i) + \mu_{j(i)} + \varepsilon_i, \tag{3}$$

where $\pi(f_i) := (\langle v_1, f_i \rangle, \ldots, \langle v_M, f_i \rangle)'$ and $\pi(\beta)$ is defined analogously. Estimating this regression yields an estimate $\widehat{\pi(\beta)}$ of the coefficients associated with the FPCA loadings.

[4] See Appendix A for details.

An estimate of the functional peer effect $\beta(\cdot)$ is then obtained by

$$\hat{\beta}(\cdot) := \sum_{m=1}^{M} \widehat{\pi(\beta)}_m v_m(\cdot).$$

As shown in Chang, Durlauf, Lee, and Park (2023), $\hat{\beta}(\cdot)$ consistently estimates the projection of $\beta(\cdot)$ onto the space spanned by the first $M$ principal components under standard regularity conditions.

In our application, we select $M = 3$, which captures more than 92% of the variation in peer test score distributions. Monte Carlo simulations reported in Appendix C.2 confirms that this choice also provides an accurate approximation of $\beta(\cdot)$. See Appendix A for more details on the implementation of the FPCA.

One alternative approach for capturing non-linear peer effects is to classify peer achievement into discrete bins (Carrell, Sacerdote, and West, 2013; Hoxby and Weingarth Salyer, 2006). In Section 4.3, we discuss peer effects estimated following this approach. This provides a useful comparison to our FPCA approach: the broad qualitative patterns are similar. The binning approach also highlights how conclusions may depend on how achievement bins are defined.

# 3 Identification

Identifying causal school and peer effects requires addressing potential endogeneity in both school assignments and peer composition. If unobserved student ability is correlated with the schools students attend or the peers they encounter, naïve comparisons may incorrectly attribute differences in outcomes to schools or peers rather than to selection. When it comes to identifying causal school effects, recent studies have shown one can use random assignment embedded in school choice lotteries (Abdulkadiroğlu, Angrist, Narita, and Pathak, 2017, 2022; Angrist, Hull, Pathak, and Walters, 2024).

Our identification builds on and extends these ideas, noting that the lotteries used in the school assignment process provide distinct sources of variation for school assignment and for peer composition. Roughly speaking, lotteries generate uncertainty not only in an individual student's school assignment but also in peer composition, even among those

whose assignment is deterministic. The two types of variation are not collinear, allowing peer and school effects to be separately identified. Notably, several features of our empirical implementation—such as allowing for nonlinear peer effects or imposing time-invariant school effects across cohorts—are not required for identification. While these features enhance precision, we can identify both school and peer effects based solely on the structure of the assignment mechanism and the variation it generates.

In the rest of this section, we first discuss how we reduce the dimension of confounders by adapting the recentering approach of Borusyak and Hull (2023). We then present our main identification argument, which proceeds in two steps. In the first step, we leverage uncertainty in a student's school assignment to identify the composite school-and-peer effect. In the second step, variation in peer composition identifies the peer effect within the composite, with the remainder being the school effect. We also discuss the asymptotic rate conditions under which both effects can be consistently estimated. Lastly, we then turn to practical issues including undersubscribed schools and imperfect compliance with assignments, and the identifying assumptions that address them.

**Reducing the Dimension of Confounders** In our setting, both school assignments and peer composition are endogenous. The endogenous component of school assignment can be controlled for by conditioning on students' preferences and priorities, which are high-dimensional objects that the literature has shown can be summarized using assignment probabilities derived from students' school application list (Rosenbaum and Rubin, 1983; Abdulkadiroğlu, Angrist, Narita, and Pathak, 2017). A similar challenge arises for peer composition, whose endogenous component is high-dimensional and jointly determined with school assignments through the assignment mechanism.

We resolve this challenge by adapting the recentering approach developed by Borusyak and Hull (2023), which applies to continuous and high-dimensional treatments. Their framework is designed for "formula treatments"—treatments that are deterministic functions of multiple sources of variation according to a known rule. Our setting fits this structure naturally: both school assignments and peer compositions are deterministic functions of students' preferences, school priorities, and random lottery draws. The key insight of the recentering approach is to isolate purely exogenous lottery-driven variation by recentering assigned treatments around their expected values. These expectations are taken over the exogenous variation—the lottery draws in our case—and therefore serve as con-

trol functions that absorb the endogenous component of assignments.

More specifically, let $s_i^a \in \{0,1\}^J$ denote the vector of binary school-assignment indicators for student $i$, with $J$ being the number of schools. Let $f_i^a$ denote the corresponding co-assigned peer distribution. Let $s_i$ and $f_i$ represent the *enrolled* school and co-enrolled peer distribution. The DA assignments can be written as known functions of all students' preferences and priorities $\theta = (\theta_{i'})_{i'}$ and the vector of random lottery draws $\ell = (\ell_{i'})_{i'}$:

$$\begin{pmatrix} s_i^a \\ f_i^a \end{pmatrix} = \begin{pmatrix} g_i(\theta, \ell) \\ h_i(\theta, \ell) \end{pmatrix} \tag{4}$$

This known functional form is crucial because it allows us to simulate counterfactual assignments by drawing alternative lottery realizations $\ell'$ while holding the observed $\theta$ fixed. Doing so enables us to compute the expected assignment vector (assignment probabilities) and the expected peer distribution, $E[s_i^a \mid \theta] =: p_i$ and $E[f_i^a \mid \theta]$. In practice, we approximate these expectations by simulating the DA algorithm across 1,000 independent lottery draws and computing sample averages of the resulting assignments and peer distributions. Intuitively, these expectations average out all lottery-induced randomness and therefore capture the purely endogenous component of assignments.

The recentered assignments $s_i^a - E[s_i^a \mid \theta]$ and recentered assigned peer distributions $f_i^a - E[f_i^a \mid \theta]$—which reflect purely exogenous lottery-induced variation—can then serve as instruments for the endogenous enrolled school and peer composition $(s_i, f_i)$. Provided that these recentered instruments are relevant, Borusyak and Hull (2023) show that the resulting instrumental-variables estimator is consistent for the structural parameters under standard regularity conditions.

**Identifying Peer and School Effects Separately** To clarify the sources of identification, we consider a more flexible model than what we use for empirical analysis:

$$y_{ij} = \mu_j^{\tau_i, g(i)} + \int \beta^{\tau_i, g(i)}(\tau) f_j(\tau) d\tau + \epsilon_i.$$

This model is more flexible compared to our empirical analysis in Equation 1 in several respects. First, note that instead of $\beta^{g(i)}(r)$ with $r = \tau - \tau_i$ representing the difference between a peer's baseline test score $\tau$ and the student $i$'s own test score $\tau_i$, we have $\beta^{\tau_i, g(i)}(\tau)$. The key

difference is that $\beta^{\tau_i,g(i)}(\tau)$ is a flexible function of own and peer's scores $(\tau_i, \tau)$, whereas $\beta^{g(i)}(r)$ restricts that the effect only works through the difference in scores $r = \tau - \tau_i$. This makes $\beta^{g(i)}(r)$ a special case of $\beta^{\tau_i,g(i)}(\tau)$. With a slight abuse of notation, $f_j(\tau)$ now denotes the (absolute) peer test score distribution within school $j$. To make the identification even more (ostensibly) difficult, we do not use the leave-one-out distribution for the peer distribution. Also, school effects $\mu_j^{\tau_i,g(i)}(i)$ are allowed to vary across $(\tau_i, g(i))$ rather than just $g(i)$; this means that school effects can depend arbitrarily on the student's own score $\tau_i$ even within the same group $g(i)$.

Note that we only observe $y_{ij}$ for $i$'s enrolled school $j(i)$:

$$y_{ij(i)} = \sum_j s_{ij} \underbrace{\left(\mu_j^{\tau_i,g(i)} + \int \beta^{\tau_i,g(i)}(\tau) f_j(\tau) d\tau\right)}_{\delta_j^{\tau_i,g(i)}} + \epsilon_i.$$

Condition on individuals with $\tau_i = \tau', g(i) = g$. For all schools $j$, consider $s_{ij}^a - E_l[s_{ij}^a|\theta]$ as an re-centered instrument for $s_{ij}$. Following the arguments of Borusyak and Hull (2023), the instrument reflects the lottery-induced variation only and therefore is exogenous. Assume that rank condition is satisfied, so that there is nondegenerate variation in the instrument for all $j$. Then, the usual IV argument of instrumenting $s_{ij}$ with $s_{ij}^a - E_l[s_{ij}^a|\theta]$ allows us to identify $\delta_j^{\tau',g}$ for all $(\tau', g)$. Aggregate school effect for group $(\tau', g)$, $\delta_j^{\tau',g}$, corresponds to the school VA in the conventional sense as used in the literature, which reflects both peer and school effects in our model.

Having identified the aggregate school effect $\delta_j^{\tau',g}$, we turn to separating the two terms—peer and school effects—within it. Fix $(\tau', g)$ as before. Consider now a school-level equation:

$$\delta_j^{\tau',g} = \int \beta^{\tau',g}(\tau) f_j(\tau) d\tau + \mu_j^{\tau',g}$$

The first component captures the part of aggregate school effect that systematically varies with peer composition, and the second absorbs the part that can not be explained by peer composition. Consider instrumenting $f_j(\tau)$ with $f_j^a(\tau) - E_l[f_j^a(\tau)|\theta]$ for all $\tau$. Following the arguments of Borusyak and Hull (2023), the re-centered instrument reflects the lottery-induced variation only and is therefore uncorrelated with the school effect $\mu_j^{\tau',g}$. Assuming that the rank condition holds—that there is variation for each percentile—, the terms in this equation can be identified by the usual IV argument. Intuitively, we are com-

paring schools with similar expected peer distributions but different realized peer distributions due to the lottery; say that one school has more lottery-induced mass of peers at percentile $x$ and less mass at percentile $y$ than the other school. If the aggregate school effect is higher for the former school, then we infer that $\beta^{\tau',g}(x) - \beta^{\tau',g}(y) > 0$.

Note that the identification is achieved even without leave-one-out distributions, nonlinearity of peer effects, availability of multiple cohorts data (and assuming time-constant school effects), or assuming constant school effects within groups $g$. Meanwhile, we do have these additional sources of identification, which only adds to the identification strength.

Taking stock, student-level and school-level lottery-induced variation allow us to disentangle peer and school effects. In particular, the two effects being "subsumed" by the school fixed effects in the student-level IV regression can be resolved by the second-stage IV regression at the school level. While two rank conditions are assumed—one for each level—they are not about the separability of the two effects, but rather about the power of the instruments to provide variation for each "element" *within* each of the two effects ($j$ for school effects and $\tau$ for peer effects).

While both rank conditions can hold simultaneously, they impose competing demands on the number of schools $J$ relative to the number of students $N$. The first step needs $J$ small relative to $N$ (i.e. $J/N \to 0$): fewer schools mean more students per school, sharpening each school's composite effect estimate. The second step, by contrast, needs $J$ large: more schools provide more school-level observations, and each school is smaller (with school size $N/J$), so its realized peer distribution is more dispersed around the population mean (Angrist, 2014). These two forces compound: under simplifying assumptions,[5] the total peer-composition variation across schools scales as $J^2/N$, so consistency of peer effect estimates requires $J/\sqrt{N} \to \infty$. Combined with the first-step requirement $J/N \to 0$, both effects can be consistently estimated only in the sandwich region where $J$ grows faster than $\sqrt{N}$ but slower than $N$.

**Practical Issues and Identifying Assumptions** One practical issue of identifying our production function is that many schools in centralized assignment systems are undersubscribed, meaning they admit all applicants and are subject to little or no lottery-based

[5]To illustrate, the rates can be derived under equal school sizes $q_n = N_n/J_n$ along a sequence $n$, with each school seat assigned to a student drawn i.i.d. from a superpopulation of students.

variation; in such a case, $s^a_{ij} - E[s^a_{ij}|\theta]$ is nearly constant for such $j$. This is a well-recognized limitation in the literature (Angrist, Hull, Pathak, and Walters, 2024). Therefore, we use the following assumption as in the literature:

*Assumption* 1 (Risk Controlled Conditional Independence Assumption (RC-CIA)).

$$\varepsilon_i \perp\!\!\!\perp s_i \mid (x_i, p_i)$$

This assumption states that, conditional on observed characteristics and *propensity scores*, a student's unobserved ability is independent of their school enrollment. Propensity scores are admissions probability for each school, which can be calculated given that school assignment procedure is known. Put differently, RC-CIA imposes that observable characteristics $x_i$ drive noncompliance with the assignment. Our $x_i$ includes students' 5th grade standardized test scores (lagged test scores), gender, ethnicity, Free-or-Reduced lunch eligibility, special education status, and English Learner status. Note that, under perfect compliance, this assumption is satisfied automatically (Angrist, Hull, Pathak, and Walters, 2024), and this is what makes RC-CIA different from conditional independence assumptions in other value-added models that do not have thecondition on (known) propensity scores (Rosenbaum and Rubin, 1983; Chetty, Friedman, and Rockoff, 2014). In that sense, RC-CIA is a statement about compliance with the assignment.[6] Angrist, Hull, Pathak, and Walters (2024) validates the assumption by comparing the results under RC-CIA with IV estimates based on randomized school offers in the context same as ours, NYC middle school assignment.

In a similar spirit to RC-CIA and Borusyak and Hull (2023), we also consider the following assumption:

*Assumption* 2 (Augmented Risk Controlled Conditional Independence Assumption (aRC-CIA)).

$$\varepsilon_i \perp\!\!\!\perp (s_i, f_i) \mid (x_i, p_i, E[f^a_i|\theta])$$

Relative to RC-CIA, the main content of the assumption is that a student's co-enrolled peer distribution is independent of their unobserved ability, conditional on the observed characteristics, propensity scores, expected co-assigned peer distribution, and enrolled

[6] Angrist, Hull, Pathak, and Walters (2024) shows that RC-CIA is implied by $\varepsilon_i \perp\!\!\!\perp s_i \mid (x_i, p_i, s^a_i)$.

school.[7] There would be a violation of aRC-CIA if there is coordinated compliance decision on unobserved ability not explained by $x_i, p_i, s_i, E[f_i^a|\theta]$. Like the RC-CIA, this assumption is essentially a compliance condition: under perfect compliance, the school and peer effects are uncorrelated with students' unobserved ability. The aRC-CIA assumption motivates the following OLS specification with:

$$y_i = \sum_{m=1}^{M} b_m c_m^{ij} + \sum_j \mu_j s_{ij} + \sum_{m=1}^{M} \rho_m E[z_m^{ij}] + \sum_j \gamma_j p_{ij} + X_i'\alpha + \epsilon_i \tag{5}$$

where $c_m^{ij}$ are FPC coordinates of realized peer distribution, $E[z_m^{ij}]$ are expected coordinates of the co-assigned peer distribution, $s_{ij}$ are school indicators, and $p_{ij}$ are propensity scores. This represents our main specification for the empirical analyses. As discussed in section 2.2, given the estimates $(\hat{b}_m)_m$ of $(b_m)_m$, we can derive an estimate of the functional parameter $\beta(\cdot)$ by $\hat{\beta}(\cdot) := \sum_{m=1}^{M} \hat{b}_m v_m(\cdot)$.

Alternatively, one can run the RC-CIA assumption motivates the following 2SLS procedure that instruments realized distribution of co-enrolled peer with the distribution ofco-assigned peer distribution, controlling for expected co-assigned peer distribution (Borusyak and Hull, 2023):

$$\text{Second stage: } y_i = \sum_{m=1}^{M} b_m c_m^{ij} + \sum_j \mu_j s_{ij} + \sum_{m=1}^{M} \rho_m E[z_m^{ij}] + \sum_j \gamma_j p_{ij} + X_i'\alpha + \epsilon_i \tag{6}$$

$$\text{First stage: } c_m^{ij} = \delta_m z_m^{ij} + \text{controls} + e_{im} \tag{7}$$

where $z_m^{ij}$ are the FPCA coordinates of the assigned peer distribution. Given the relatively limited exogenous variation compared to OLS, estimating this equation separately for different demographic groups is challenging. Accordingly, we employ 2SLS as a robustness specification in our analyses.

[7] More specifically, aRC-CIA can be decomposed into the following two assumptions:

$$\varepsilon_i \perp\!\!\!\perp s_i \mid (x_i, p_i, E[f_i^a|\theta])$$
$$\varepsilon_i \perp\!\!\!\perp f_i \mid (x_i, p_i, s_i, E[f_i^a|\theta])$$

# 4 Application to NYC Middle School

## 4.1 Data

We use student-level data from the New York City Department of Education (DOE) as our primary data source. We use three middle school application cohorts—2014–15, 2015–16, and 2016–17, which consists of 216,525 students in total. We observe students' test scores on the New York State Math and English Language Arts (ELA) exams, as well as their submitted rank-ordered choice lists, final school assignments, and admissions priorities at each school program. In addition, we observe detailed demographic information, including race/ethnicity, English Language Learner status, and eligibility for subsidized lunch. 201,602 students have non-missing baseline characteristics, which we treat as population in Table 1.

Panel A of Table 1 reports summary statistics for student characteristics. After data cleaning discussed in the appendix B.3 in detail, the analysis sample includes 146,972 students. The majority of the dropped observations, 49,982 observations out of 54,630, have either missing assigned or enrolled schools. The mean baseline test score, which is the average of Math and ELA scores, is 0.03, indicating that students in our analysis sample are slightly higher achieving than the overall population of New York City students who took the standardized exams. Fifty percent of students in the sample are male, while 21%, 41%, and 18% are Black, Hispanic, and Asian, respectively.

We estimate peer effects at the school-by-cohort level. There are 509 unique middle schools overall and 507 schools in the analysis sample, yielding 1,473 school-by-cohort observations. Panel B of Table 1 reports summary statistics at the school-by-cohort level. The average sixth-grade cohort contains 103 students, with a standard deviation of 89, reflecting substantial variation in cohort size across schools. Comparing student-level and school-level average test scores suggests that larger schools tend to enroll higher-achieving students. Although we estimate the effect of school peers, the relevant random variation in admissions occurs at the program level, which is the unit of middle school assignment. There are 807 unique programs in the analysis sample, and 58% of them are Unscreened, which use random lotteries to decide which applicants to admit when oversubscribed. Screened programs use non-random tie-breakers such as attendance and standardized test

scores, while Audition programs additionally rely on program-specific tests or auditions. Even students who enroll in Screened or Audition programs may face random variation in peer composition because our peer measure is defined at the school rather than the program level. Once students enter a middle school through different programs, they share common spaces, extracurricular activities, and classes, and NYC administrative data track students only at the school level after school assignment. Given this institutional background, it is natural to interpret peer effects as reflecting not only direct interactions among peers but also indirect effects operating through short-run resources, pedagogy, and school management, as discussed in Section 2.1.

## 4.2 Identifying Variation and Identification Assumptions

In this section, we show that we have sufficient variation to obtain peer effects even after controlling for a rich set of characteristics in our main specification 5, and present evidence that supports our identifying assumptions in Section 3.

**Source of variation** In section 3, we established that lottery-induced variation provides separate identifying variation for peer and school effects, *provided* the two rank conditions at the student and school level, respectively.

Here, we investigate the feasibility of this identification approach with finite data. We ask three questions: whether there exists sufficient variation to (1) separate the effects of peers and schools, (2) identify the effects of *all* schools, and (3) identify the effects of the *whole* distribution of the peers.

Before proceeding to empirical results, we first examine why DA lotteries may provide more identifying variation than initially apparent. As emphasized by Abdulkadiroğlu, Angrist, Narita, and Pathak (2017), even schools with strict priority orderings (i.e., no within-priority ties and thus no direct use of lottery tie-breakers) can experience effectively random fluctuation in enrolled composition through rejection chains. Applicants who are randomly rejected at oversubscribed, lottery-using schools are redirected to lower-ranked options, propagating randomness to schools that otherwise would not use tie-breakers. Furthermore, as we discussed above, our focus on peer effects provides supplementary sources of variation: we leverage randomness not only in students' own assignments but

also in the assignments of their peers.

Now we turn to empirical evidence regarding the three questions. The first question asks whether there exists sufficient separate variation for both peer and school effects. This would be the case if there are sufficiently many students who, due to the lotteries, (1) switch schools, and (2) among non-switchers, experience nonnegligible changes in their peer distribution.

Table 2 summarizes the lottery-induced variation across our three-year sample (N=221,225 students). Across two sets of lottery draws, 9.16% of students ($\approx$ 20,265 students) switch their assigned schools. Moreover, even among non-switchers—students assigned to the same school under both lottery draws—6–30% experience peer distribution changes exceeding the 0.05 total variation distance (TVD[8]) threshold, depending on the measurement approach (kernel density estimation with different bandwidths or histogram-based binning).

Turning to the third question about whether the lottery provides enough variation to vary all components representing distributions, Panel B decomposes variation by examining individual FPC coordinates. For all students, lottery-induced changes (SD($\Delta$)) represent 15–30% of the total cross-sectional variation (SD(raw)), with higher-order FPCs showing proportionally larger lottery effects. Among non-switchers, per each FPC, 1–1.6% of students (approximately 1,600–2,500 students) exhibit FPC coordinate changes large enough to generate TVD-equivalent shifts[9] exceeding 0.05 across all four components. For switchers, the effects are much larger: 38–64% of students cross the 0.05 TVD threshold, and the standard deviation of lottery-induced changes is 55–96% of the overall cross-sectional dispersion. Taken together, the lottery provides identifying variation in both school switches and peer composition (for each principal component) within schools.

The second question concerns whether there exists sufficient variation to identify the effects of *all* individual schools. We address this challenge using established solutions discussed in Section 3.

[8]TVD equals one half of the $L^1$ distance between two discrete distributions and can be interpreted as the minimum fraction of mass that must be relocated to transform one distribution into the other.

[9]Since peer distributions are decomposed as $f_i(r) = \mu(r) + \sum_{m=1}^{m} c_{im} v_m(r)$, the difference between two lottery draws is $f_i^{(1)}(r) - f_i^{(2)}(r) = \sum_{m=1}^{m} \Delta c_{im} v_m(r)$. By the triangle inequality, TVD satisfies $\text{TVD}(f_i^{(1)}, f_i^{(2)}) \leq \frac{1}{2} \sum_{m=1}^{m} |\Delta c_{im}| \cdot \|v_m\|_{L^1}$, where $\|v_m\|_{L^1} = \int |v_m(r)| dr$ measures how much probability mass basis function $v_m$ can shift. Thus $|\Delta c_{im}| \cdot \|v_m\|_{L^1}$ approximates component $m$'s contribution to distributional change, making it directly comparable to TVD thresholds.

**Testing the Identifying Assumption** Our identification relies on Assumption 2 (aRC-CIA), which requires that, conditional on observed characteristics, propensity scores, and the expected co-assigned peer distribution, realized peer composition is independent of students' unobserved ability.

The primary threat to this assumption is coordinated non-compliance; although lottery assignment generates random variation in co-assigned peers, students may not ultimately enroll in their assigned schools. If students with systematically higher (or lower) unobserved ability are more likely to deviate from their assignment through appeals, transfers, or strategic enrollment decisions, then realized peer composition may remain correlated with unobserved ability.

To assess the plausibility of aRC-CIA, we conduct a placebo test by running equation 5 where students' 4th grade test scores is the outcome. While unobserved ability is not observed, 4th grade achievement should provide a strong proxy. Under the identifying assumption, conditional on expected co-assigned peer distribution, propensity scores, observed characteristics, and enrolled schools, realized co-enrolled peer distributions should not predict 4th grade test scores. For the ease of interpretation, we present results where we use peers' absolute baseline test scores as independent variable. We present results with relative test scores—which aligns with our main specification better—in the appendix Figure C4, D1, and D3, and the results remain the same.

Figure 1 illustrates the results of this placebo exercise. The "Unconditional" plots describe the estimated peer-effect function from a regression of 4th grade test scores on realized peer distributions without conditioning on covariates. The resulting curve is strongly increasing in peer achievement, suggesting that students with higher underlying ability—proxied by higher 4th grade test scores—tend to enroll in middle schools with peers with high 5th grade test scores.

The "Conditional" panels repeat the exercise after controlling for expected co-assigned peer distribution, propensity scores, observed characteristics, and school enrollment. Once these controls are included, the estimated peer-effect function becomes essentially flat and statistically indistinguishable from zero across the support of peer achievement. The disappearance of the slope suggests that the residual variation in realized peer composition—our identifying variation—is no longer correlated with predetermined achievement, supporting our identifying assumption.

Appendix Tables D1, D2 and D3 report the corresponding placebo regressions under alternative specifications. Appendix Table D1 presents results for the linear-in-means model, analogous to conventional balance tables in that framework. Appendix Table D2 and D3 extends the analysis to our FPCA specification. The coefficients underlying Figure 1 correspond to columns (3) and (5) of Tables D1 and D2, respectively.

Across both approaches, the conclusions are consistent. For co-assigned peer characteristics (either the mean or the full distribution), controlling for their expected values is sufficient to eliminate correlations with baseline achievement, supporting Assumption 1 (RC-CIA). In contrast, for co-enrolled peer characteristics—the object of interest in our analysis—controlling only for expected co-assigned peers is insufficient due to endogenous compliance. Once we additionally control for propensity scores, observed characteristics, and school enrollment, co-enrolled peer distributions are no longer correlated with students' 4th grade test scores, supporting Assumption 2 (aRC-CIA).

### 4.3 Peer Effects Estimates

Applying our methodology to data from New York City, we find that peer effects in the full sample are negative, while also exhibiting departures from linearity. Figure 2 presents our main estimates under the augmented conditional independence assumption (aRC-CIA). The blue curve shows the nonparametric estimate of the peer effect function $\beta(\tau - \tau_i)$, which measures the marginal effect on student $i$ of peers whose baseline test scores differ from $i$'s own baseline score by $\tau - \tau_i$. For comparison, the red line reports the corresponding estimate from a linear-in-means specification, which differs only in that it summarizes peer composition using the mean of the peer distribution rather than a finite-dimensional representation of the full distribution. The shaded gray area depicts the empirical distribution of $\tau - \tau_i$ in the sample, illustrating where the variation in relative peer achievement is located.

The downward-sloping shape of both curves indicates a negative peer effect: holding own achievement fixed, exposure to higher-achieving peers is associated with lower outcomes relative to exposure to lower-achieving peers. At the same time, Figure 2 reveals modest departures from linearity in the peer effect function. In particular, the marginal effect of peers who are slightly higher achieving than the student—corresponding to relative test scores just above zero—is relatively flat, suggesting that small increases in peer

achievement in this range have similar impacts. By contrast, peers who are slightly lower achieving than the student—those with relative test scores just below zero—exert more positive effects relative to peers closer to zero. Quantitatively, replacing 10% of same-scoring peers with peers scoring one standard deviation higher reduces test scores by approximately 0.4% of a standard deviation, while replacing those same peers with peers scoring one standard deviation lower increases test scores by about 2% of a standard deviation.

The full-sample estimates mask substantial heterogeneity by baseline achievement. Panel (a) of Figure 3 focuses on lower-achieving students—those with baseline test scores below the sample mean—and shows little systematic sensitivity to peer composition. While the shape of the blue line suggest mildly negative peer effects over parts of the support, these effects are imprecisely estimated and close to zero on average. In contrast, Panel (b) shows that higher-achieving students—those with baseline test scores above the mean—exhibit pronounced and consistently negative peer effects across much of the support, indicating a substantially stronger response to peer composition. For high performing students, replacing 10% of same-scoring peers with peers scoring one standard deviation higher reduces test scores by approximately 20% of a standard deviation, a sizable effect. This pattern of negative peer effects is consistent with the literature on rank or relative-position effects, which emphasizes that students may be discouraged by exposure to higher-achieving peers (Carneiro, Cruz-Aguayo, Salvati, and Schady, 2025). Appendix C.4 shows heterogeneity by gender, income level, and more granular groups of baseline achievement. Appendix C.5 shows our estimates as a function of absolute peer test scores $\tau$ rather than relative peer scores $r = \tau - \tau_i$.

As an alternative way to summarize peer distributions, we also estimate a class of discrete specifications that classify peer achievement into bins. Specifically, we regress outcomes on the shares of peers in each bin and their interactions with students' own bin. As in our main analysis, we control for the corresponding expected peer shares computed from assignment simulations in order to isolate the causal component of peer composition. We further include school fixed effects, propensity scores, baseline student observables, and special education indicators.

Appendix Table D4 first reports a specification closely mimicking Carrell, Sacerdote, and West (2013), in which peer composition is summarized by the shares of co-enrolled peers in the bottom and top quartiles of the baseline score distribution, and these effects

are allowed to vary by students' own achievement tercile. The pooled coefficients (column (3)) are imprecise, although suggesting overall negative peer effects. Once we allow for heterogeneity by own tercile in column (6), however, a clearer pattern emerges. In particular, the coefficient on the share of bottom-quartile peers is large and positive for high-achieving students, at 0.269 (s.e. 0.078), while other coefficients are close to zero. This negative peer effects pronounced particularly among high-achieving students is consistent with the patterns uncovered by our FPCA estimates.

At the same time, Appendix Table D5 shows that the results depend on how achievement bins are defined. We report three natural variants of the same basic approach: terciles for both peers and students ($T3 \times T3$), quartiles for both peers and students ($Q4 \times Q4$), and peer quintiles interacted with own terciles ($Q5 \times T3$). Both under $T3 \times T3$ and $Q4 \times Q4$, the positive effect of the lowest achieving peers on high-achieving students arises. Under Q5 $\times$T3, though, the main trade-off for the highest achieving peers arises between peers in the third and fourth quintile, with the effects of the lowest achieving peers in quintiles 1 and 2 being muted. These are all reasonable variants of the same underlying discrete specification, yet they emphasize different margins of the peer distribution and can suggest different narratives. This sensitivity to researcher-imposed cutoffs is precisely where FPCA has an advantage: it uses the full peer achievement distribution and avoids requiring a researcher to decide which bins are substantively meaningful.

## 4.4 School Effect Estimates

To understand how accounting for peer composition affects estimates of school effectiveness, we compare school value-added estimates obtained from our joint peer-and-school model to those from a specification that abstracts from peer effects. The latter follows the empirical approach commonly used in the recent school value-added literature, including Angrist, Hull, Pathak, and Walters (2024), in which student outcomes are regressed on school indicators and controls, including assignment propensity scores, but peer composition is not modeled explicitly. In particular, this specification omits both the flexible peer effect term and controls for the expected peer distribution derived from the assignment mechanism.

For each specification, we recover school fixed effects as estimates of school value added. Because school-level estimates are noisy—particularly for smaller schools—we apply Bayesian

shrinkage following Walters (2024). This procedure combines the estimated school effects with their sampling variances to obtain posterior mean estimates that optimally shrink noisy school-level estimates toward the overall mean.

Figure 4 plots the shrunk school value-added estimates from the joint model against those from the school-only specification. The figure shows that while the two sets of estimates are highly correlated, the dispersion of school effects increases once peer composition is accounted for, going from having a standard deviation of 0.08 in the school-only model to 0.1 in the school and peer model. This difference is driven primarily by schools in the right tail of the value-added distribution: schools that appear highly effective under the joint model tend to have lower estimated value added in the school-only specification.

This pattern is consistent with the peer effect estimates discussed above. High value-added schools tend to enroll higher-achieving peers, and peer effects are negative at the top of the achievement distribution. When peer composition is omitted, these negative peer effects are absorbed into estimated school effects, compressing the upper tail of the distribution. Accounting explicitly for peer composition separates these components, reallocating variation from peers back to schools and resulting in a wider dispersion of estimated school effectiveness.

## 4.5 Varying School Assignments

Having separately estimated peer and school effects, we now use these estimates to characterize the value of alternative school assignments. In this section, we study how peer composition and school quality jointly contribute to student achievement under counterfactual assignments of students to schools.

We begin by considering the assignment of a new entrant into the system. Holding the school assignments of incumbent students fixed, we simulate the peer and school values that this entrant would experience at each school in the city. Using the estimates from the heterogeneous peer-effect model (Figure 3), Figure 5 plots peer values and shrunk school effects as functions of school selectivity—measured by the mean baseline test score of enrolled students—for entrants at different points of the achievement distribution.

Panels A and B show that low-achieving entrants (at the 10th and 30th percentiles) experience limited gains from peer composition across the range of school selectivity. Peer

effects are small and display little systematic relationship with selectivity. By contrast, school effects vary substantially, with larger gains concentrated in more selective schools. As a result, for lower-achieving students, enrollment in more selective schools is predicted to yield higher test-score gains primarily through school quality, with peer effects playing a muted role.

Panels C and D reveal a different pattern for higher-achieving entrants (at the 70th and 90th percentiles). For these students, peer effects decline sharply with school selectivity: higher-achieving students experience more favorable peer effects in less selective schools. At the same time, for this sample, more selective schools continue to offer higher school value added, with dispersion comparable to that of peer effects. Consequently, for high-achieving students, peer and school effects operate in opposing directions. In particular, our simulations suggest that students at the 70th and 90th percentiles of baseline achievement may achieve higher test-score gains in less selective schools, where favorable peer effects can more than offset lower school effectiveness.

Figure 5 further shows that, under the observed student assignments, the distribution of peer effects is substantially more compressed than the distribution of school value added for low-achieving students. For students at the 10th percentile of baseline achievement, the standard deviation of school value added is 0.22, compared with only 0.01 for peer effects. By contrast, high-achieving students face a more dispersed distribution of peer effects: for students at the 90th percentile, the standard deviation of school value added is 0.05, while the standard deviation of peer effects is 0.07. This pattern highlights that, for high-achieving students, variation in peer composition can rival—or even outweigh—variation in school effectiveness as a source of potential achievement gains.

# 5 Conclusion

This paper develops a unified framework to jointly estimate and separately identify school effects and peer effects on student achievement. Exploiting random tie-breakers in centralized school choice lotteries, we show that widely used assignment mechanisms generate two distinct sources of exogenous variation—variation in school assignments and variation in peer composition within schools—both of which are required for identification. To flexibly model peer effects, we allow student outcomes to depend on the full distribution of

peers' baseline achievement, which we summarize using functional principal component analysis. Endogeneity in both school and peer assignments is addressed by adapting the recentering approach of Borusyak and Hull (2023), which isolates lottery-driven variation in realized assignments.

Applying this framework to New York City middle schools yields three main findings. First, peer effects are predominantly negative and exhibit substantial heterogeneity by students' baseline achievement. For lower-achieving students, peer effects are small, imprecisely estimated, and close to zero on average. In contrast, higher-achieving students experience pronounced and consistently negative peer effects, with magnitudes comparable to—or larger than—school value added. These patterns are masked in full–sample estimates.

Second, accounting for peer effects materially alters the distribution of estimated school value added. Once peer composition is flexibly modeled, the dispersion of school effects reduces for high-achieving students but widens for low-achieving students, indicating that conventional value-added measures may conflate school quality with peer composition. This finding underscores the importance of jointly modeling schools and peers when evaluating school effectiveness in choice-based systems.

Third, a simulation exercise for new entrants—holding incumbent assignments fixed—illustrates how peer and school effects jointly shape the gains from alternative school assignments across the achievement distribution. For lower-achieving students, potential gains are driven primarily by school quality, with peer effects playing a limited role. For higher-achieving students, peer and school effects often operate in opposing directions: less selective schools may offer peer environments with favorable impacts that offset lower school value added. These results highlight that the value of a school assignment depends not only on school quality but also on how peer composition interacts with a student's own achievement.

Taken together, our findings emphasize that peer composition is an economically meaningful input into the education production function and that its effects are highly heterogeneous. Methodologically, the framework developed here provides a tractable way to study rich peer effects in settings with algorithmic assignment mechanisms. Substantively, the results caution against policy evaluations that abstract from peer composition or impose restrictive peer–effect structures, particularly when assessing the consequences of school

choice and assignment reforms.

# References


Abdulkadiroğlu, A., J. Angrist, and P. Pathak (2014): "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," *Econometrica*, 82(1), 137–196.

Abdulkadiroğlu, A., J. D. Angrist, Y. Narita, and P. Pathak (2022): "Breaking Ties: Regression Discontinuity Design Meets Market Design," *Econometrica*, 90(1), 117–151.

Abdulkadiroğlu, A., J. D. Angrist, Y. Narita, and P. A. Pathak (2017): "Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation," *Econometrica*, 85(5), 1373–1432.

Allende, C. (2019): "Competition Under Social Interactions and the Design of Education Policies," p. 104.

Angrist, J., P. Hull, P. A. Pathak, and C. Walters (2024): "Credible School Value-Added with Undersubscribed School Lotteries," *Review of Economics and Statistics*, 106(1), 1–19.

Angrist, J. D. (2014): "The Perils of Peer Effects," *Labour Economics*, 30, 98–108.

Boozer, M., and S. E. Cacciola (2001): "Inside the'Black Box'of Project STAR: Estimation of Peer Effects Using Experimental Data," *Available at SSRN 277009*.

Borusyak, K., and P. Hull (2023): "Nonrandom Exposure to Exogenous Shocks," *Econometrica*, 91(6), 2155–2185.

Carneiro, P., Y. Cruz-Aguayo, F. Salvati, and N. Schady (2025): "The effect of classroom rank on learning throughout elementary school: experimental evidence from Ecuador," *Journal of Labor Economics*, 43(2), 433–466.

Carrell, S. E., R. L. Fullerton, and J. E. West (2009): "Does Your Cohort Matter? Measuring Peer Effects in College Achievement," *Journal of Labor Economics*, 27(3), 439–464.

Carrell, S. E., B. I. Sacerdote, and J. E. West (2013): "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation," *Econometrica*, 81(3), 855–882.

Chang, Y., S. N. Durlauf, S. Lee, and J. Y. Park (2023): "A Trajectories-Based Approach to Measuring Intergenerational Mobility," Discussion paper, National Bureau of Economic Research.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014): "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *American Economic Review*, 104(9), 2593–2632.

Corcoran, S. P., and H. M. Levin (2011): "School Choice and Competition in the New York City Schools," in *Education Reform in New York City: Ambitious Change in the Nation's Most Complex School System*, ed. by J. O'Day, C. Bitter, and L. M. Gomez. Harvard Education Press.

Duflo, E., P. Dupas, and M. Kremer (2011): "Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya," *American economic review*, 101(5), 1739–1774.

Gibbons, S., and S. Telhaj (2016): "Peer effects: Evidence from secondary school transition in England," *Oxford Bulletin of Economics and Statistics*, 78(4), 548–575.

Hoxby, C., and G. Weingarth Salyer (2006): "School Reassignment and the Structure of Peer Effects, Manuscrito," *Universidad de Harvard*.

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

Lyle, D. S. (2009): "The Effects of Peer Group Heterogeneity on the Production of Human Capital at West Point," *American Economic Journal: Applied Economics*, 1(4), 69–84.

Murphy, R., and F. Weinhardt (2020): "Top of the Class: The Importance of Ordinal Rank," *The Review of Economic Studies*, 87(6), 2777–2826.

Rosenbaum, P. R., and D. B. Rubin (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70(1), 41–55.

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

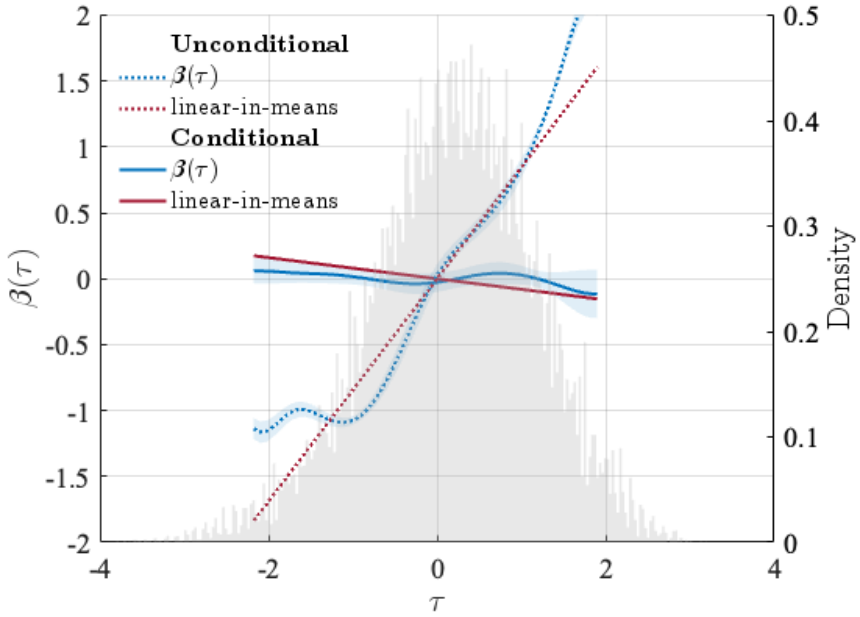
Walters, C. (2024): "Empirical Bayes methods in labor economics," in *Handbook of Labor Economics*, vol. 5, pp. 183–260. Elsevier.

Woodcock, S., J. Friesen, and R. Meilman Lomaz Cohn (2025): "Sorting, Peer Effects and School Effectiveness in Private and Public Schools," .

Zimmerman, D. J. (2003): "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment," *The Review of Economics and Statistics*, 85(1), 9–23.

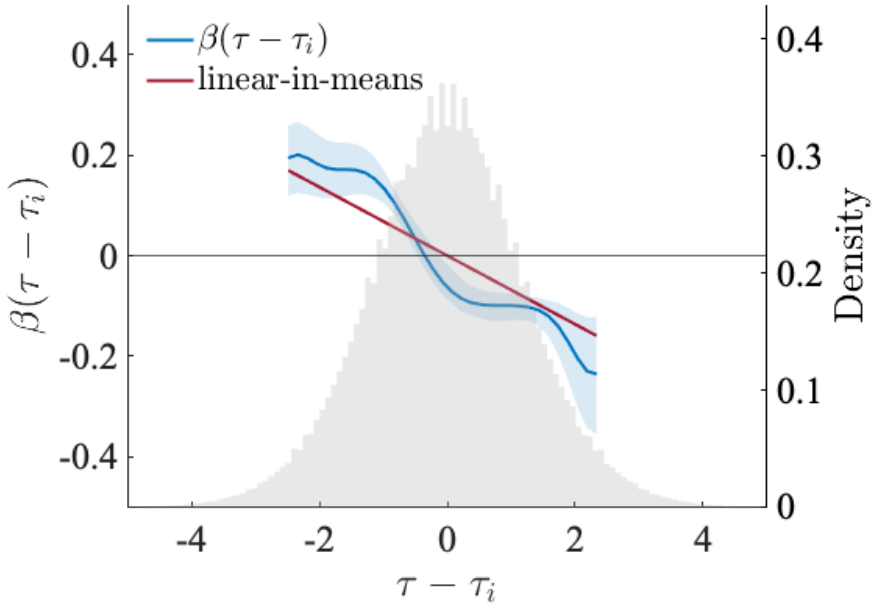
# Figures

Figure 1: Testing the Identifying Assumption



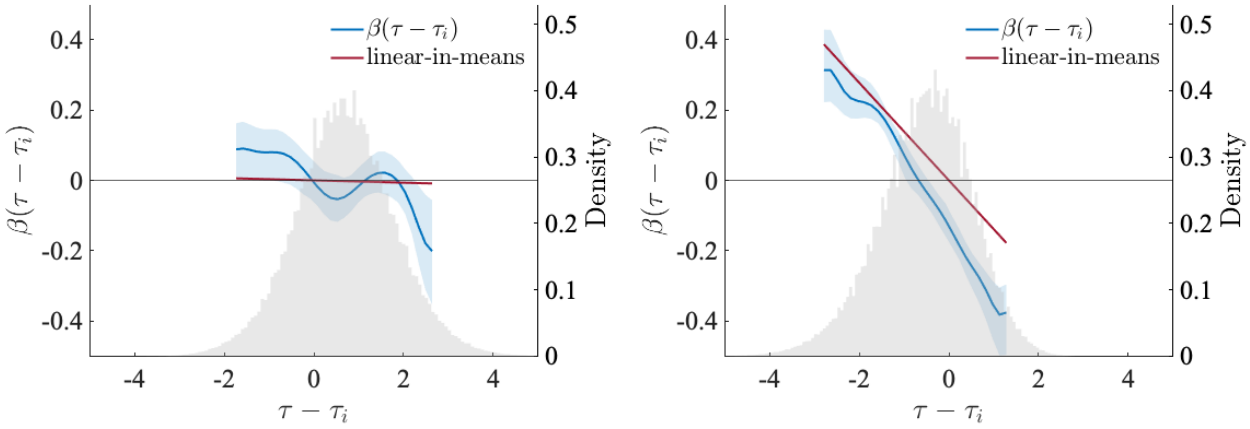

*Notes:* The figure presents results from the placebo test in which students' 4th grade test scores are regressed on the distribution of their future co-enrolled peers' test scores in middle school. The horizontal axis represents peers' 5th grade achievement, ($\tau$). The "Unconditional" plots report estimates from regressions without additional controls. The "Conditional" panels additionally control for the expected co-assigned peer distribution, propensity scores, observed characteristics, and enrolled middle school fixed effects. Shaded bands denote 95% confidence intervals. The shaded gray histogram shows the empirical density of ($\tau$), plotted against the right-side y-axis. The corresponding regression coefficients are reported in columns (3) and (5) of Appendix Tables D1 and D2.

Figure 2: Peer Effect Estimates under aRC-CIA for the Full Sample




*Notes:* The figure plots the estimated peer effect function $\hat{\beta}(\tau - \tau_i)$ under the augmented Conditional Independence Assumption (aRC-CIA). The horizontal axis represents the difference between peers' baseline test scores and the student's own baseline score $(\tau - \tau_i)$. The blue curve shows the non-parametric estimate based on functional principal components with 95% confidence intervals, while the red line represents the linear-in-means specification. The shaded gray area shows the empirical density of $(\tau - \tau_i)$ in the data. Standard errors are computed using a wild bootstrap clustered at the school-market level.

Figure 3: Peer Effect Estimates under aRC-CIA by Baseline Achievement

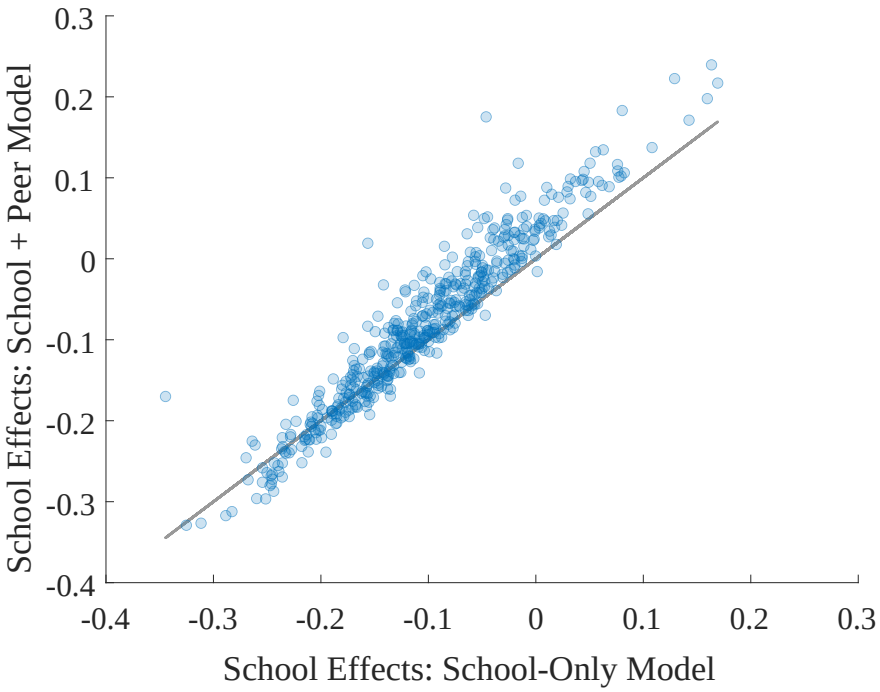



(a) Low Performing Students

(b) High Performing Students

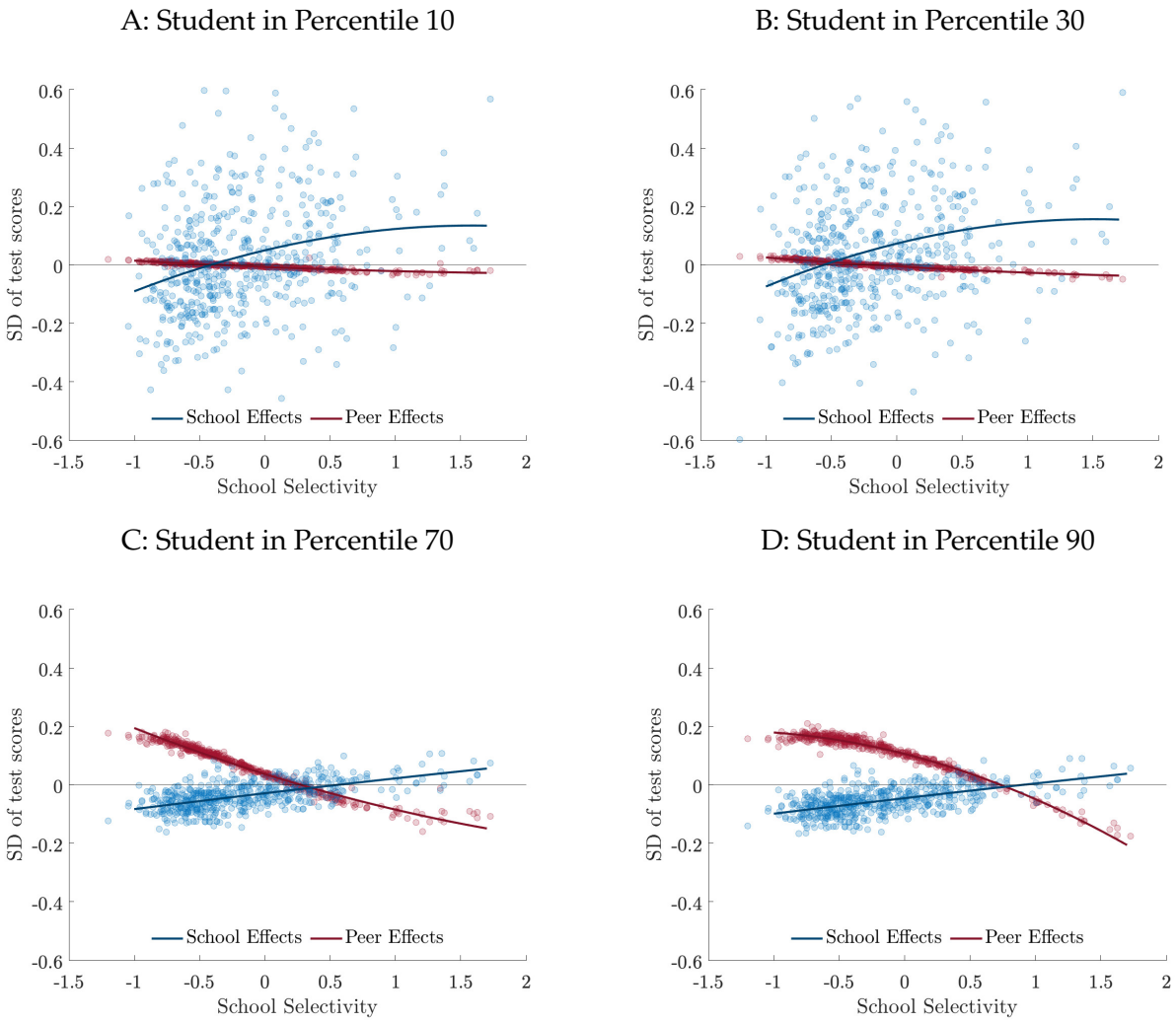
*Notes:* The figure plots the estimated peer effect function $\hat{\beta}(\tau - \tau_i)$ separately for low-performing (left panel) and high-performing (right panel) students, defined by if a student's baseline test score is above median or not. The horizontal axis represents the difference between peers' baseline test scores and the student's own baseline score $(\tau - \tau_i)$. The blue curve shows the non-parametric estimate based on functional principal components with 95% confidence intervals, while the red line represents the linear-in-means specification. The shaded gray area shows the empirical density of $(\tau - \tau_i)$ within each subgroup. Standard errors are computed using a wild bootstrap clustered at the school-market level.

Figure 4: Estimated School Effects: School-Only Model vs. School + Peer Model




*Notes:* The figure plots school-level effect estimates from a model that excludes peer effects (horizontal axis) against estimates from a model that jointly includes peer and school effects (vertical axis). Each point represents a school. Blue markers correspond to estimates obtained for the subsample of lower-achieving students (baseline test scores below the median), while red markers correspond to higher-achieving students (baseline test scores above the median). The solid line denotes the 45-degree line. Deviations from this line reflect changes in estimated school effectiveness once peer composition is accounted for within each subsample.

Figure 5: Simulated Peer and School Effects for a Hypothetical Entrant by Baseline Achievement




*Note:* Each panel plots predicted school value added (blue) and predicted peer effects (red) across schools for a hypothetical incoming student at the 10th, 30th, 70th, or 90th percentile of the baseline achievement distribution. For each percentile type, we evaluate the peer composition and school effectiveness the entrant would face at every school in the city, holding incumbent students' observed assignments fixed. School selectivity is measured as the mean baseline achievement of enrolled students. Solid lines represent fitted cubic relationships between selectivity and the predicted peer and school components. The vertical axis reports gains normalized relative to the corresponding peer effect or school value added at the typical school attended by incumbent students in the same achievement percentile.

# Tables

Table 1: Summary Statistics

| | (1) All | (2) Analysis Sample |
|---|---|---|
| ***Panel A: Student Characteristics*** | | |
| Total Students | 201,602 | 146,972 |
| Students, 2014 | 66,437 | 47,990 |
| Students, 2015 | 67,705 | 49,340 |
| Students, 2016 | 67,460 | 49,596 |
| 5th Grade Score | 0.015 | 0.045 |
| Male | 0.508 | 0.505 |
| Black | 0.237 | 0.220 |
| Hispanic | 0.411 | 0.419 |
| Asian | 0.175 | 0.189 |
| FRL | 0.718 | 0.716 |
| Special Education | 0.208 | 0.202 |
| ELL | 0.111 | 0.111 |
| ***Panel B: School and Program Characteristics*** | | |
| Total Schools | 509 | 507 |
| Schools, 2014 | 505 | 500 |
| Schools, 2015 | 495 | 490 |
| Schools, 2016 | 484 | 483 |
| 6th Grade Cohort Size | 136 | 98 |
| Mean 5th Grade Score | -0.120 | -0.156 |
| Total Programs | 853 | 807 |
| Audition | 0.065 | 0.026 |
| Screened | 0.379 | 0.391 |
| Unscreened | 0.556 | 0.583 |

*Note:* Column (1) reports summary statistics of the population of 2014–15 to 2016–17 applicants. Column (2) reports summary statistics of the FPCA analysis sample. Entries are sample means or shares, except for count rows, which report the number of students, distinct schools, and distinct programs. School rows summarize school-by-cohort cells, and program-method rows report shares of distinct programs. The 5th grade score is the average of students' Math and English standardized test scores. FRL indicates Free- or Reduced-Price Lunch eligibility, and ELL indicates English-Language Learner status.

# Table 2: Lottery-Induced Variation: Switching and Distribution Changes

*Panel A. Assignment variation and TDV exceedance*

| Group (N) | Switch (%) | Auto kernel TDV>0.05 (%) | Manual kernel TDV>0.05 (%) | Binned TDV>0.05 (%) |
|---|---|---|---|---|
| All students (221,225) | 9.16 | 12.41 | 34.58 | 16.27 |
| Non-switchers (200,664) | 0 | 6.09 | 29.72 | 9.36 |
| Switchers (20,561) | 100 | 97.49 | 100.00 | 99.66 |

Auto kernel uses default BW; Manual kernel uses BW = 0.05. Binned uses 20 percentile bins.

**Panel B. Mean vs. FPC changes (SD and TVD-equiv exceedance)**

| | All students | | | Non-switchers | | Switchers | |
|---|---|---|---|---|---|---|---|
| Measure | SD(raw) | Chg/SD (%) | TVD>0.05 (%) | Chg/SD (%) | TVD>0.05 (%) | Chg/SD (%) | TVD>0.05 (%) |
| Mean peer score | 0.387 | 22.43 | – | 4.41 | – | 83.79 | – |
| FPC 1 | 0.252 | 15.33 | 5.69 | 5.23 | 1.38 | 55.09 | 63.74 |
| FPC 2 | 0.171 | 21.10 | 5.83 | 8.55 | 1.60 | 73.86 | 62.80 |
| FPC 3 | 0.086 | 26.97 | 4.77 | 13.76 | 1.36 | 89.16 | 50.59 |
| FPC 4 | 0.062 | 30.49 | 3.57 | 17.63 | 1.00 | 96.15 | 38.12 |

Note: SD(raw) is the cross-sectional standard deviation pooling both lottery draws; SD(Δ) is the standard deviation of lottery-induced changes. Panel B reports Change/SD as SD(Δ)/SD(raw) in percent. For Non-switchers and Switchers, Change/SD uses the All-students SD(raw) as denominator for comparability. TVD-equivalent uses the FPC $L^1$-weighted metric $\left(0.5 \cdot |\Delta c_m| \, \|v_m\|_{L^1}\right)$ with threshold 0.05. Kernel (Auto/Manual) valid $N = 167{,}682$ (Non-switchers $= 156{,}084$, Switchers $= 11{,}598$). Binned TDV valid $N = 174{,}934$ (Non-switchers $= 161{,}551$, Switchers $= 13{,}383$). Multi-year combines 2014–15, 2015–16, 2016–17; DA is run per-year and aggregated.

# A FPCA representation and estimation details

This appendix provides additional details on the FPCA representation used to estimate the peer effects model. Most of the arguments in this appendix closely follow Chang, Durlauf, Lee, and Park (2023). We proceed in three steps: (i) explain why uncorrelated loadings imply no exact multicollinearity in our specification, (ii) explain why truncating the FPCA expansion does not induce omitted-variable bias from approximation error in the linear functional regression setting, and (iii) show why FPCA loadings are uncorrelated.

## A.1 Why uncorrelated loadings imply zero multicollinearity in our setting

In our empirical specification, the regressors for peer composition are the FPCA loadings

$$\pi(f_i) := (\langle v_1, f_i \rangle, \ldots, \langle v_m, f_i \rangle)'.$$

If these loadings are mutually uncorrelated in the population, then their second-moment matrix is diagonal:

$$E[\pi(f_i)\pi(f_i)'] = \text{diag}(\lambda_1, \ldots, \lambda_m),$$

with $\lambda_k > 0$ for retained components. A diagonal second-moment matrix is full rank, so no loading can be written as an exact linear combination of the others. Therefore, the FPCA regressors have no exact multicollinearity in population.

## A.2 Why FPCA truncation does not create omitted-variable bias in the linear model

Consider the linear functional regression model

$$y_i = \langle \beta, f_i \rangle + \varepsilon_i,$$

where $E[\varepsilon_i \mid f_i] = 0$ and $f_i$ is centered. Let $\{v_k\}_{k \geq 1}$ denote the FPCA basis (eigenfunctions of the covariance of $f_i$) and define the loadings (scores) $c_{ik} := \langle v_k, f_i \rangle$ and coefficients

$\beta_k := \langle \beta, v_k \rangle$. By the FPCA expansion,

$$f_i = \sum_{k \geq 1} c_{ik} v_k, \qquad \beta = \sum_{k \geq 1} \beta_k v_k,$$

so that

$$\langle \beta, f_i \rangle = \sum_{k \geq 1} \beta_k c_{ik}.$$

Our empirical specification uses only the first $m$ loadings as regressors, i.e., it replaces $\langle \beta, f_i \rangle$ by its truncated approximation $\sum_{k=1}^{m} \beta_k c_{ik}$. The corresponding approximation residual (the omitted term) is

$$u_i := \sum_{\ell > m} \beta_\ell c_{i\ell}.$$

In a standard finite-dimensional regression, truncation would generate omitted-variable bias if $u_i$ were correlated with the included regressors $\{c_{ik}\}_{k \leq m}$. Under FPCA, however, distinct loadings are uncorrelated: for $k \neq \ell$, $E[c_{ik} c_{i\ell}] = 0$. Therefore, for each $k \leq m$,

$$E[c_{ik} u_i] = \sum_{\ell > m} \beta_\ell \, E[c_{ik} c_{i\ell}] = 0.$$

Thus the approximation residual is orthogonal to the retained components, so truncation does not induce omitted-variable bias in the linear functional regression setting. This argument relies on linearity and does not generally extend to nonlinear functional models; see Chang, Durlauf, Lee, and Park (2023).

## A.3 Why FPCA loadings are uncorrelated

We restrict attention to square-integrable functions and assume that both the functional parameter $\beta(\cdot)$ and the peer distributions $(f_i)_i$ belong to a Hilbert space $H$ endowed with inner product

$$\langle \beta, f \rangle := \int \beta(r) f(r) \, dr.$$

FPCA constructs the basis $\{v_k\}_{k=1}^{\infty}$ as eigenfunctions of the covariance operator of $(f_i)_i$. For centered functions, the loading (score) for component $k$ is $c_{ik} := \langle v_k, f_i \rangle$, and standard FPCA results imply

$$E[c_{ik} c_{i\ell}] = 0 \text{ for } k \neq \ell, \qquad E[c_{ik}^2] = \lambda_k.$$

Hence the loadings are mutually uncorrelated across retained components.

Using the first $m$ components, we approximate

$$f_i \approx \sum_{k=1}^{m} \langle v_k, f_i \rangle v_k.$$

Define the mapping $\pi : H \to \mathbb{R}^m$ by

$$\pi(f) := (\langle v_1, f \rangle, \ldots, \langle v_m, f \rangle)'.$$

Under this representation, the peer-effects term satisfies

$$\langle \beta, f_i \rangle \approx \pi(\beta)' \pi(f_i),$$

which leads to the finite-dimensional approximation of the structural equation. Given an estimate $\widehat{\pi(\beta)}$, the functional parameter is recovered as

$$\hat{\beta}(\cdot) := \sum_{k=1}^{m} \widehat{\pi(\beta)}_k v_k(\cdot).$$

Under regularity conditions, $\hat{\beta}(\cdot)$ consistently estimates the projection of $\beta(\cdot)$ onto the span of $\{v_k\}_{k=1}^m$ (Chang, Durlauf, Lee, and Park, 2023).

Appendix Figure C2 presents Monte Carlo evidence showing that the leading $m = 3$ components provide an accurate approximation of both peer distributions and the peer-effect function $\beta(\cdot)$.

## A.4 Interpretation of $\beta(r)$

The value $\Delta f \times \Delta r \times [\beta(r') - \beta(r)]$ encapsulates the approximate effect of relocating the area $\Delta f \times \Delta r$ in the density $f(\cdot)$ from $r$ to $r'$. We show this using the Riemann sum approximation for $f$:

$$\int \beta(r) f(r) dr \approx \Delta r \sum_{i=1}^{n} \beta(r_i) f(r_i)$$

Now, with the modified density $g$:

$$\begin{aligned} g(r_k) &= f(r_k) + \Delta f \\ g(r_j) &= f(r_j) - \Delta f \\ g(r_i) &= f(r_i) \qquad \text{for all } i \neq j, k \end{aligned}$$

The difference in the Riemann sums for $f$ and $g$ is:

$$\Delta r \left[\beta(r_k) g(r_k) + \beta(r_j) g(r_j) - \beta(r_k) f(r_k) - \beta(r_j) f(r_j)\right] = \Delta f \times \Delta r \times \left[\beta(r_k) - \beta(r_j)\right].$$

# B Details of NYC School Choice Process & Data Cleaning

## B.1 NYC School Admission Methods

Depending on the eligibility criteria, middle schools are classified into three types—district schools, borough schools, and citywide schools. The city is divided into 5 boroughs and 32 community school districts. A student's residence or elementary school decides eligibility at each type of school. In the academic year 2014-2015, 14, 39, and the rest were citywide, borough, and district programs, respectively, among 670 programs.

Middle school programs use a variety of admission methods—Unscreened, Limited Unscreened, Screened, Screened: Language, Zoned, and Talent Test. Unscreened programs admit students by a random lottery number, and Limited Unscreened programs use rules that give priority to those who attend information sessions or open houses. Screened programs as well as Screened: Language programs select students by individually assorted measures such as elementary school GPA, statewide test scores, punctuality, and interviews. Zoned programs guarantee admissions or give priority to students who reside in the school's zone, and Talent Test programs use auditions as the main criteria.

## B.2 Timeline of The Admission Process

The timeline of the admission process is as follows (Corcoran and Levin (2011), Directory of NYC Public High Schools). By December, students are required to submit their ROLs.

By March, DA algorithms are run that determine students' assignments. Students who accept their offer finalize, and if a student rejects an offer she goes to the next round. This describes the main round of the entire system. A majority of students finalize in the main round (about 85% each year). Students who are not assigned in the main round or rejected the assignment go to the Supplementary round, which is organized similarly to the main round and includes schools/programs that did not fill their capacities in the main round, or programs that are newly opened. Finally, there is an administrative round in which students who are not assigned a school even after the second round are administratively assigned to a school.

## B.3 Sample Construction and Data Cleaning

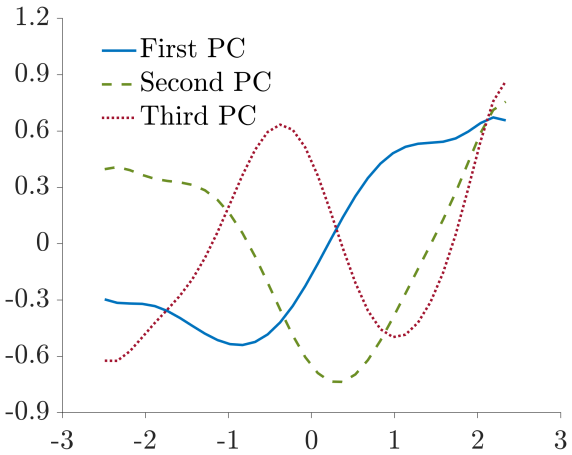
Our starting population consists of 216,525 students who participated in the New York City middle school application process in the 2014–15, 2015–16, and 2016–17 cohorts. We first drop students without baseline characteristics or 5th grade baseline test scores, which leaves 201,602 observations. We then drop students with missing assigned program identifiers or missing enrolled program identifiers, leaving 151,620 observations. Next, we drop five students who are assigned to a school without peers, leaving 151,615 observations. Finally, because our outcome is 6th grade achievement, we drop students without 6th grade test scores, producing a final estimation sample of 146,972 observations. The placebo balance tests use smaller samples, ranging from 137,783 to 141,399 observations, because they additionally require non-missing 4th grade test scores.

# C Appendix Figures

## C.1 FPCA Estimates

Figure C1 displays the first three functional principal components (FPCs) of the co-enrolled peer test score distribution for the full student sample. These components summarize the dominant dimensions of variation in peers' baseline achievement. The first component primarily captures level shifts in peer quality, closely corresponding to changes in mean peer achievement. The second component reflects variation in dispersion, distinguishing between peer groups that are relatively homogeneous and those with greater spread in achievement. The third component captures asymmetry in the distribution, indicating whether peer composition is skewed toward higher- or lower-achieving students. Together, these three components explain more than ninety percent of the total variation in peer distributions, providing a parsimonious yet flexible representation of peer composition.

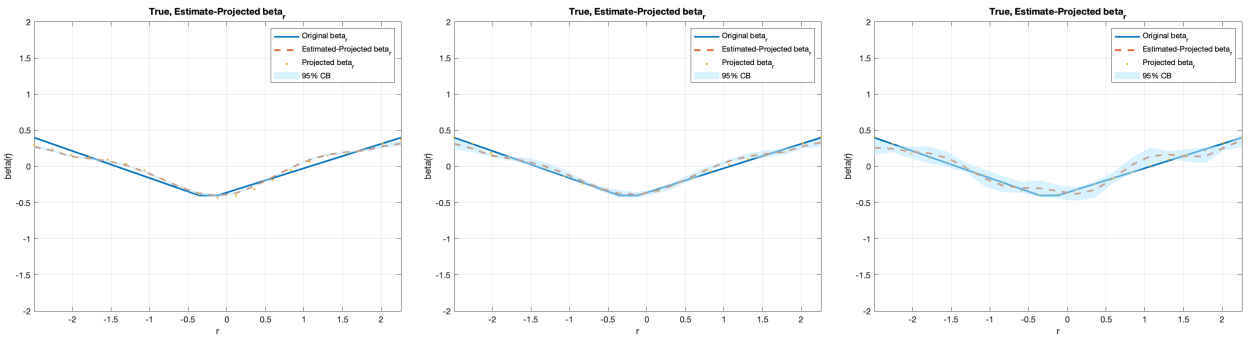
Figure C1: First Three Principal Components




*Notes:* The plot shows the first three functional principal components, $v_j$, used to approximate the peer–effect function $f_j(\tau)$ for the full sample, e.g., when $\mathcal{G}$ contains a single set with all students.

## C.2 Monte-Carlo Simulations

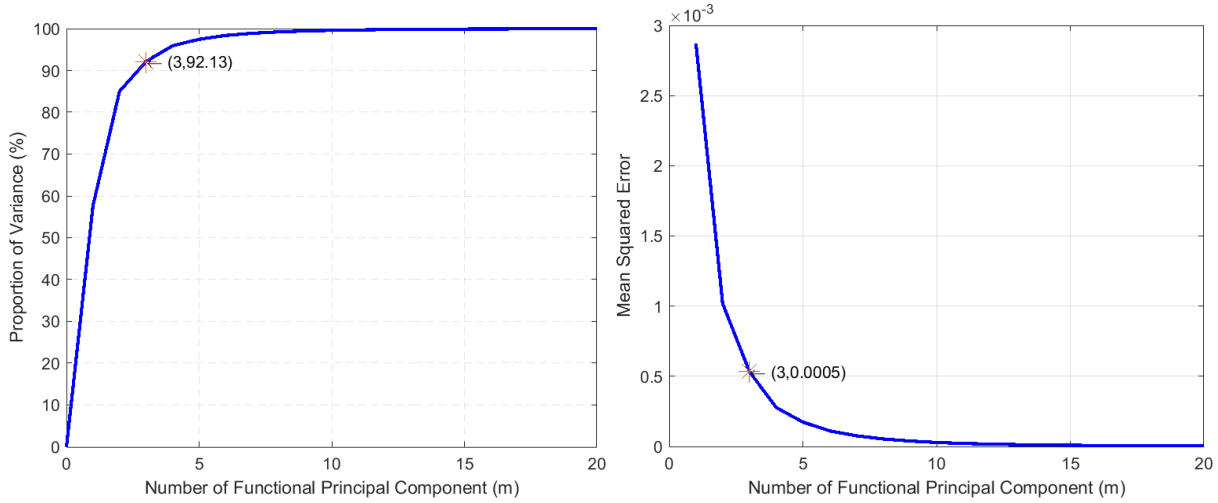
Figure C2: True, Estimated, Projected $\beta(.)$ with $m$ of Functional Principal Components




(a) $m = 3$ (b) $m = 5$ (c) $m = 7$

*Note:* Each figure shows simulated $\beta(\cdot)$, projection of $\beta(\cdot)$ onto the functional principal components of $(f_i)_i$, and the estimate of $\beta(\cdot)$ according to our Monte Carlo analyses. $m$ denotes the number of leading principal components used.

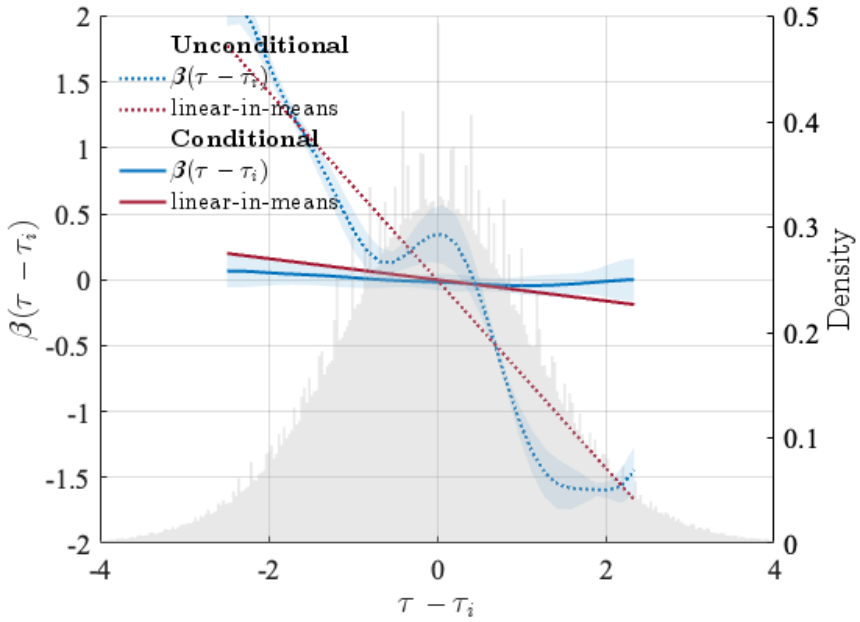
Figure C3: Cross-Validation Errors and Cumulative Screeplot




Note: The left panel is the cumulative scree plot, which presents the total percentage of variation explained by the leading m-functional principal components also as a function of m. In the right panel is the leave-one-out cross-validation (LOOCV) errors as a function of the number m of functional principal components used to approximate parental income trajectory.

## C.3 Placebo Test With Relative Scores

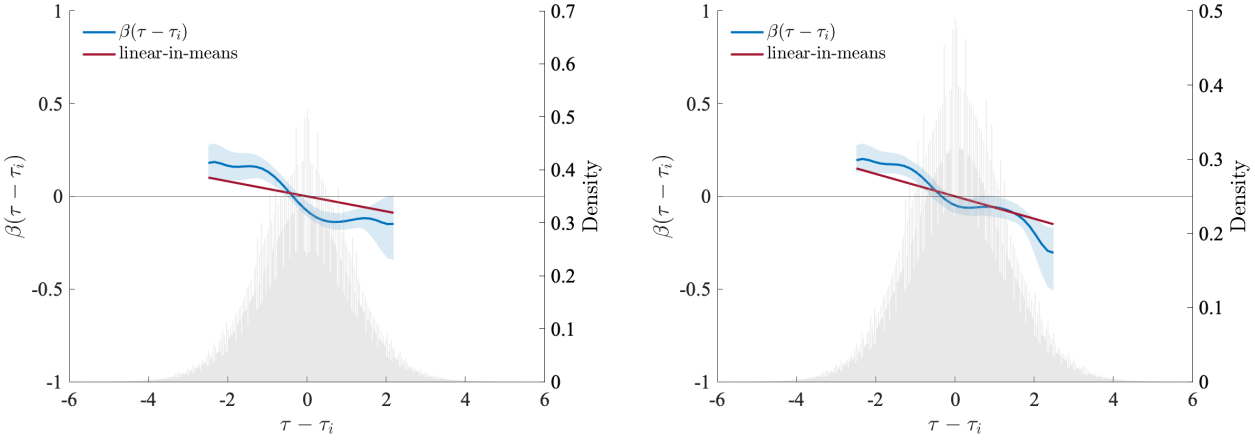
Figure C4: Testing the Identifying Assumption




*Notes:* The figure presents results from the placebo test in which students' 4th grade test scores are regressed on the distribution of their future co-enrolled peers' test scores in middle school. The horizontal axis represents peers' 5th grade achievement relative to the student's own 5th grade score, $(\tau - \tau_i)$. The "Unconditional" plots report estimates from regressions without additional controls. The "Conditional" panels additionally control for the expected co-assigned peer distribution, propensity scores, observed characteristics, and enrolled middle school fixed effects. Shaded bands denote 95% confidence intervals. The shaded gray histogram shows the empirical density of $(\tau - \tau_i)$, plotted against the right-side y-axis. The corresponding regression coefficients are reported in columns (3) and (5) of Appendix Tables D1 and D3.

## C.4 Heterogeneity Analysis

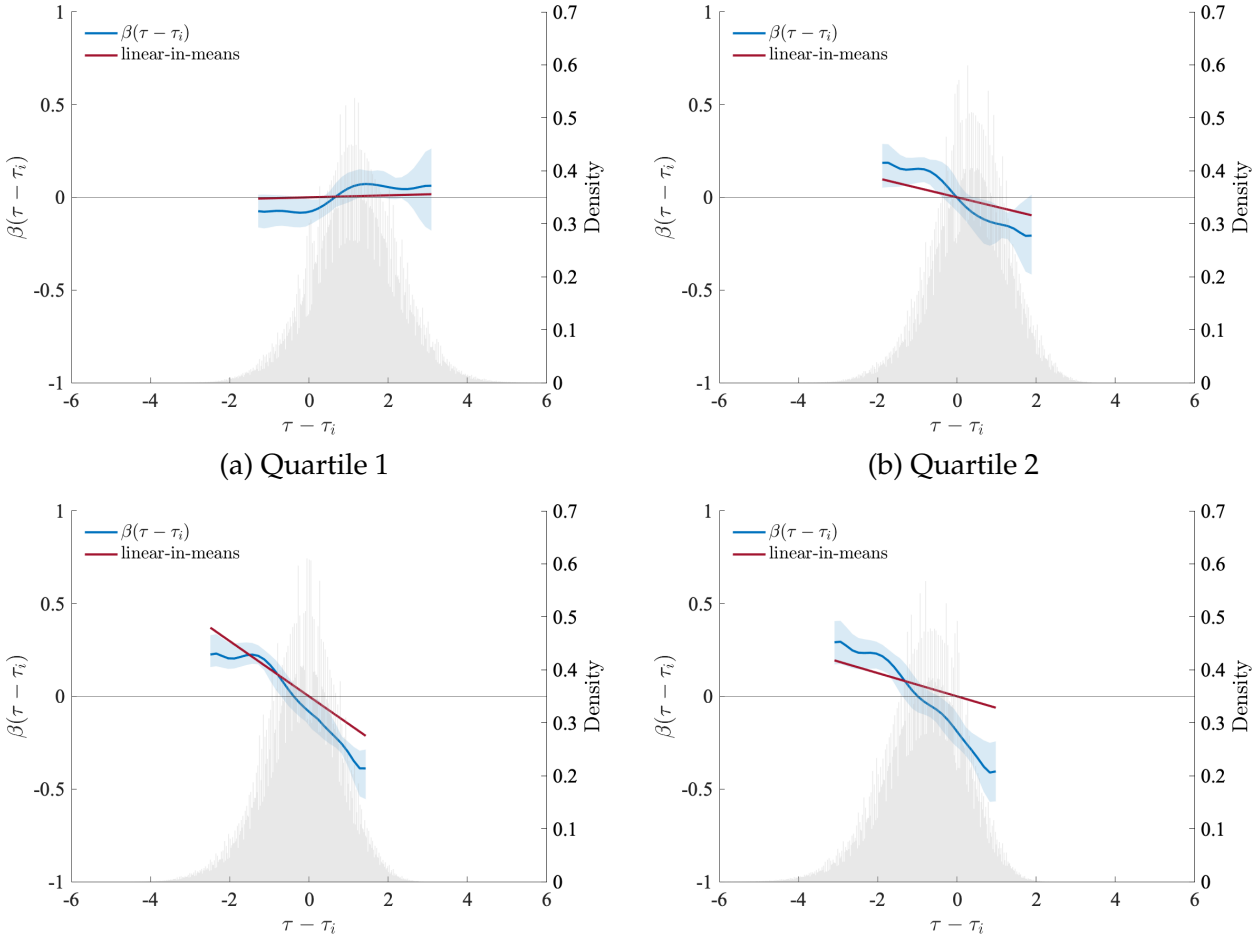
Figure C5: By Gender




(a) Female (b) Male

Note: Peer effects functions $\beta(r - r_i)$ estimated by gender. The blue curves show the non-parametric estimates based on functional principal components with 95% confidence intervals, while the red lines represent the linear-in-means specifications. Gray areas represent the empirical density of $r - r_i$ within each subgroup.

Figure C6: By Baseline Score Quartiles




(a) Quartile 1

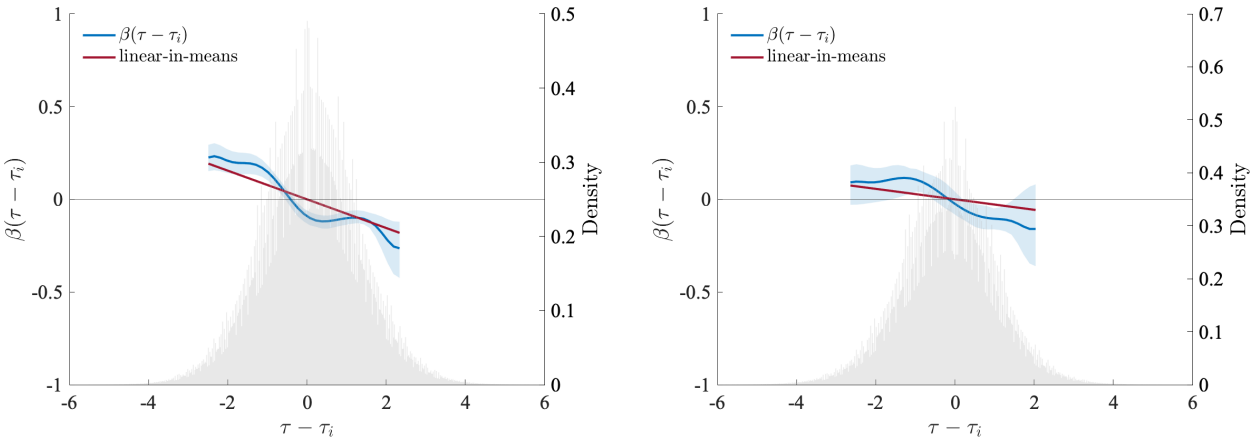
(b) Quartile 2

(c) Quartile 3

(d) Quartile 4

Note: Peer effects functions $\beta(r - r_i)$ estimated by performance quartiles. The blue curves show the non-parametric estimates based on functional principal components with 95% confidence intervals, while the red lines represent the linear-in-means specifications. Gray areas represent the empirical density of $r - r_i$ within each subgroup.

Figure C7: By Socioeconomic Status

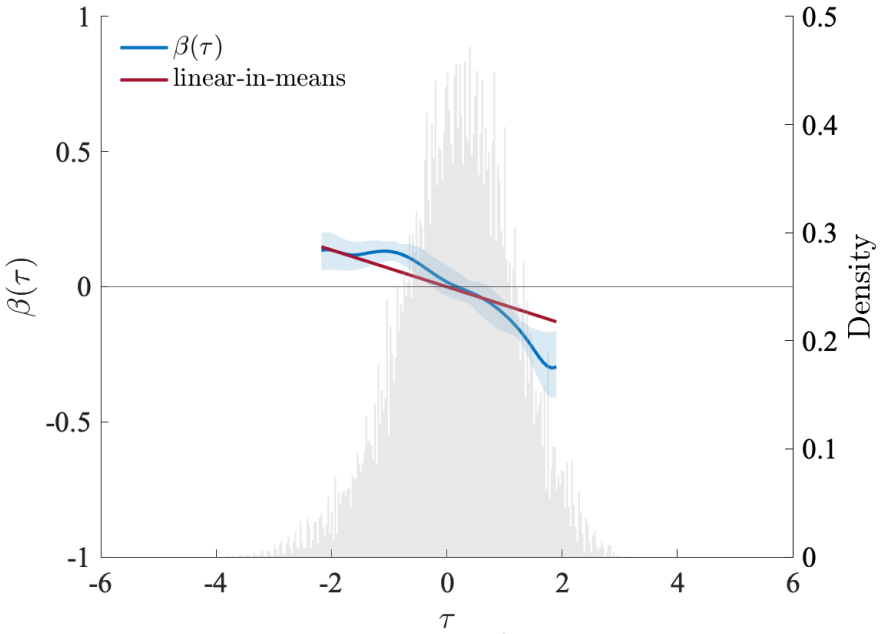



(a) Free/Reduced-Price Lunch (Low SES) (b) Paid Lunch (High SES)

Note: Peer effects functions $\beta(r - r_i)$ estimated by socioeconomic status. The blue curves show the non-parametric estimates based on functional principal components with 95% confidence intervals, while the red lines represent the linear-in-means specifications. Gray areas represent the empirical density of $r - r_i$ within each subgroup.

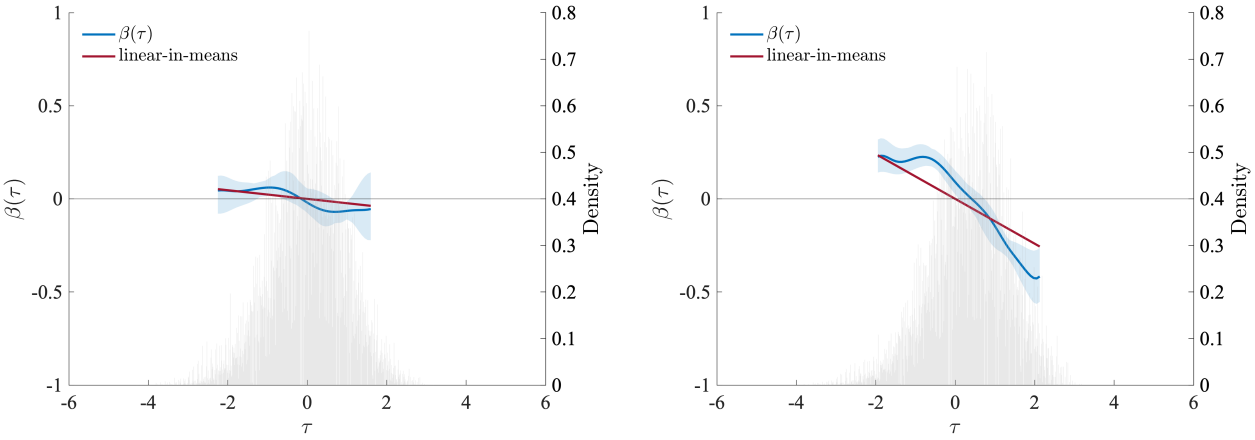
## C.5 Robustness: Absolute Peer Scores

Figure C8: Full Sample (Absolute Peer Scores)




*Note:* The figure plots the estimated peer effect function $\hat{\beta}(\tau)$ using absolute peer scores. The horizontal axis represents peers' absolute baseline test scores $\tau$. The blue curve shows the non-parametric estimate based on functional principal components with 95% confidence intervals, while the red line represents the linear-in-means specification. The shaded gray area shows the empirical density of absolute peer scores $\tau$ in the data.

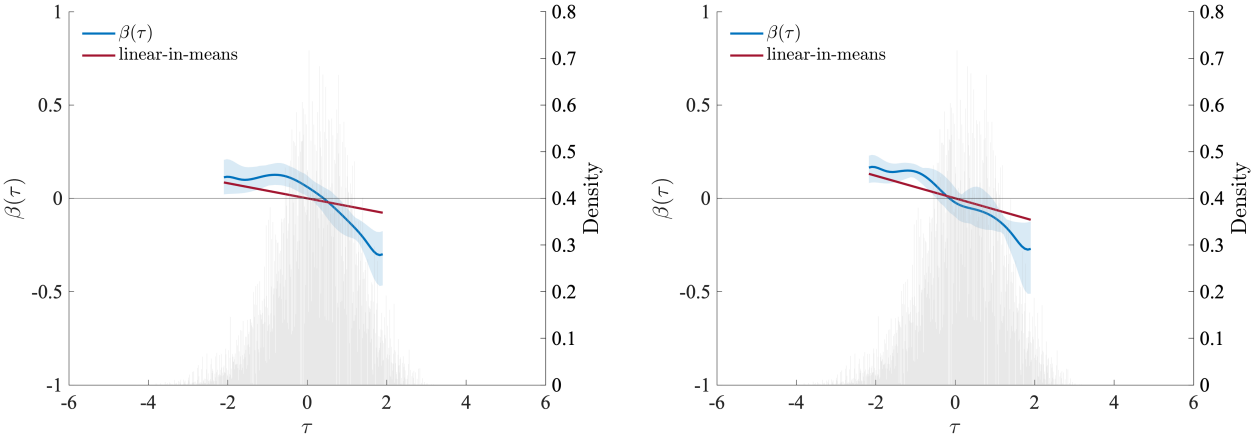
Figure C9: By Baseline Performance (Absolute Peer Scores)




(a) Low-Performing Students (b) High-Performing Students

*Note:* Peer effect functions $\beta(\tau)$ estimated separately for low- and high-performing students, using absolute peer scores. The blue curves show the non-parametric estimates based on functional principal components with 95% confidence intervals, while the red lines represent the linear-in-means specifications. Gray areas represent the empirical density of absolute peer scores $\tau$ within each subgroup.

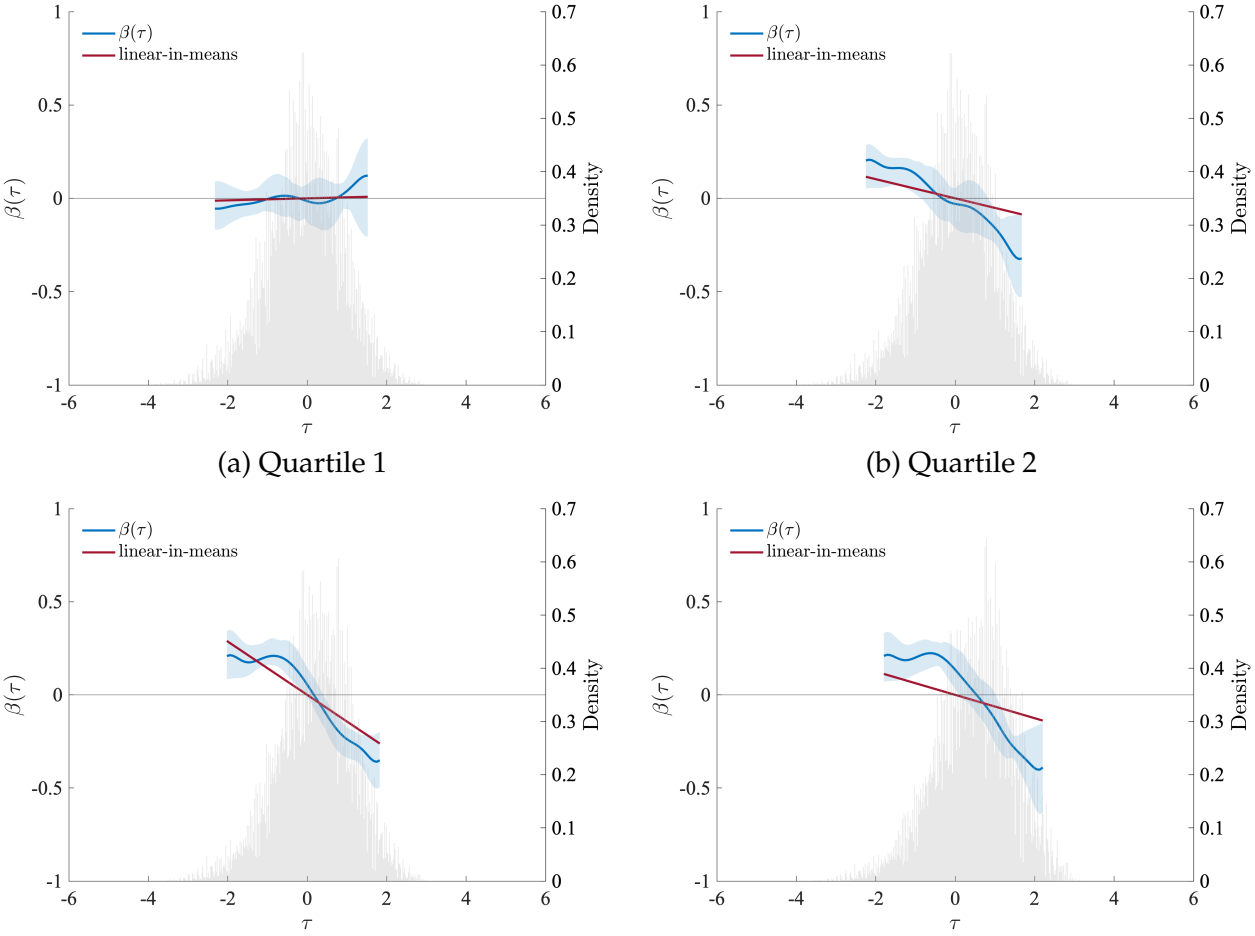
Figure C10: By Gender (Absolute Peer Scores)




(a) Female (b) Male

*Note:* Peer effect functions $\beta(\tau)$ estimated by gender, using absolute peer scores. The blue curves show the non-parametric estimates based on functional principal components with 95% confidence intervals, while the red lines represent the linear-in-means specifications. Gray areas represent the empirical density of absolute peer scores $\tau$ within each subgroup.

Figure C11: By Baseline Score Quartiles (Absolute Peer Scores)




(a) Quartile 1

(b) Quartile 2

(c) Quartile 3

(d) Quartile 4

*Note:* Peer effect functions $\beta(\tau)$ estimated by performance quartiles, using absolute peer scores. The blue curves show the non-parametric estimates based on functional principal components with 95% confidence intervals, while the red lines represent the linear-in-means specifications. Gray areas represent the empirical density of absolute peer scores $\tau$ within each subgroup.

# D Appendix Tables

Table D1: Placebo Test in Linear-in-mean Models

| Tested Assumption | RC-CIA | | aRC-CIA | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| $x_i, p_i, s_i$ | No | No | No | No | Yes |
| ***Panel A: Mean of Peer's Absolute Test Scores*** | | | | | |
| Co-assigned | 0.889** | -0.032 | | | |
| | (0.011) | (0.021) | | | |
| E(Co-assigned) | | 0.944** | | 0.580** | 0.080** |
| | | (0.022) | | (0.016) | (0.028) |
| Co-enrolled | | | 0.845** | 0.353** | -0.080** |
| | | | (0.016) | (0.014) | (0.027) |
| N | 141399 | 137783 | 141398 | 137782 | 137783 |
| $R^2$ | 0.418 | 0.425 | 0.425 | 0.435 | 0.763 |
| ***Panel B: Mean of Peers' Relative Test Scores*** | | | | | |
| Co-assigned | -0.710** | -0.255** | | | |
| | (0.007) | (0.041) | | | |
| E(Co-assigned) | | -0.462** | | -0.492** | 0.080** |
| | | (0.043) | | (0.031) | (0.028) |
| Co-enrolled | | | -0.713** | -0.238** | -0.080** |
| | | | (0.006) | (0.030) | (0.027) |
| N | 141399 | 137783 | 141398 | 137782 | 137783 |
| $R^2$ | 0.498 | 0.497 | 0.609 | 0.615 | 0.763 |

*Notes:* This table reports placebo regressions of students' 4th grade average Math and ELA test scores on peer mean test scores. "Co-assigned" refers to the mean 5th grade test score of peers assigned to the student's middle school. "Co-enrolled" refers to the mean 5th grade test score of peers with whom the student ultimately enrolls. "E(Co-assigned)" denotes the expected mean peer score computed from simulated lottery assignments. Columns (1)–(2) examine co-assigned peer means. Columns (3)–(5) examine co-enrolled peer means. All models include special education status in students' 4th and 6th grades. Column (5) further controls for propensity scores, school fixed effects, district X year X special education status fixed effects, and student observables. Students' observable characteristics include their fifth grade average Math and ELA test scores (baseline score in the main regression), english learner status, female, dummy variables for Asian, Black, and Hispanic, and Free-or-Reduced lunch eligibility. Standard errors in parentheses are clustered at the district X year X special education status level. Significance levels: $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

Table D2: Placebo Test in FPCA Specification with Peer's Absolute Test Scores

| Tested Assumption | RC-CIA | | aRC-CIA | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| $E[z_k^{ij}\|\theta]$ | No | Yes | No | Yes | Yes |
| $x_i, p_i, s_i$ | No | No | No | No | Yes |
| $z_1^{ij}$ | 2.214** | 0.431** | | | |
| | (0.025) | (0.075) | | | |
| $z_2^{ij}$ | 0.246** | -0.045 | | | |
| | (0.042) | (0.082) | | | |
| $z_3^{ij}$ | 1.550** | 0.159 | | | |
| | (0.066) | (0.106) | | | |
| $c_1^{ij}$ | | | 2.001** | 0.812** | -0.039 |
| | | | (0.023) | (0.036) | (0.056) |
| $c_2^{ij}$ | | | -0.182** | -0.123** | -0.025 |
| | | | (0.035) | (0.045) | (0.044) |
| $c_3^{ij}$ | | | 0.755** | 0.346** | -0.081 |
| | | | (0.060) | (0.070) | (0.058) |
| N | 141399 | 141399 | 141399 | 141399 | 141399 |
| $R^2$ | 0.418 | 0.427 | 0.420 | 0.431 | 0.764 |
| *Joint test on coordinates:* | | | | | |
| F-stat | 858.15 | 14.50 | 1395.45 | 152.29 | 0.46 |
| p-value | 0.000 | 0.000 | 0.000 | 0.000 | 0.633 |

*Notes:* This table reports placebo regressions of students' 4th grade average Math and ELA test scores on the functional principal components (FPCs) of peer achievement distributions. $z_k^{ij}$ denotes the $k$-th FPC of the co-assigned peer distribution (instrument), and $c_k^{ij}$ denotes the $k$-th FPC of the co-enrolled peer distribution. $E[z_k^{ij}|\theta]$ denotes the $k$-th FPC of the expected co-assigned peer distribution computed from simulated lottery assignments. Columns (1)–(2) examine the RC-CIA specification using co-assigned peer components. Columns (3)–(5) examine the aRC-CIA specification using co-enrolled peer components. We include the first four FPCs of the co-assigned and co-enrolled peer distributions, and the first ten FPCs of the expected co-assigned peer distribution when included as controls. Student controls ($x_i$) include gender, race/ethnicity, English Learner status, and baseline test scores (omitted when used as the outcome). Propensity scores ($p_i$) capture the probability of assignment to each school. All specifications include market fixed effects (year $\times$ district $\times$ special education status) and special education indicators for grades 4 and 6 where applicable. Standard errors clustered at market level in parentheses. The joint test reports the F-statistic and p-value for the null hypothesis that all four FPC coefficients are jointly equal to zero. Significance levels: $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

Table D3: Placebo Test in FPCA Specification with Peer's Relative Test Scores

| Tested Assumption | RC-CIA | | aRC-CIA | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| $E[z_k^{ij}\|\theta]$ | No | Yes | No | Yes | Yes |
| $x_i, p_i, s_i$ | No | No | No | No | Yes |
| $z_1^{ij}$ | 2.342** | -0.015 | | | |
| | (0.020) | (0.248) | | | |
| $z_2^{ij}$ | 0.280** | 0.042 | | | |
| | (0.072) | (0.169) | | | |
| $z_3^{ij}$ | -2.035** | -0.086 | | | |
| | (0.097) | (0.156) | | | |
| $z_4^{ij}$ | 0.691** | 0.206 | | | |
| | (0.141) | (0.169) | | | |
| $c_1^{ij}$ | | | -2.212** | -0.412** | -0.052 |
| | | | (0.018) | (0.121) | (0.066) |
| $c_2^{ij}$ | | | 0.216** | 0.085 | 0.048$^{+}$ |
| | | | (0.049) | (0.077) | (0.028) |
| $c_3^{ij}$ | | | -0.901** | -0.226* | -0.018 |
| | | | (0.066) | (0.095) | (0.048) |
| $c_4^{ij}$ | | | -1.064** | -0.336** | -0.027 |
| | | | (0.130) | (0.129) | (0.047) |
| N | 141399 | 141399 | 141399 | 141399 | 141399 |
| $R^2$ | 0.488 | 0.512 | 0.496 | 0.513 | 0.764 |
| *Joint test on coordinates:* | | | | | |
| F-stat | 803.97 | 1.06 | 1922.81 | 3.53 | 0.80 |
| p-value | 0.000 | 0.366 | 0.000 | 0.016 | 0.496 |

*Notes:* This table reports placebo regressions of students' 4th grade average Math and ELA test scores on the functional principal components (FPCs) of peer achievement distributions. $z_k^{ij}$ denotes the $k$-th FPC of the co-assigned peer distribution (instrument), and $c_k^{ij}$ denotes the $k$-th FPC of the co-enrolled peer distribution. $E[z_k^{ij}|\theta]$ denotes the $k$-th FPC of the expected co-assigned peer distribution computed from simulated lottery assignments. Columns (1)–(2) examine the RC-CIA specification using co-assigned peer components. Columns (3)–(5) examine the aRC-CIA specification using co-enrolled peer components. We include the first four FPCs of the co-assigned and co-enrolled peer distributions, and the first ten FPCs of the expected co-assigned peer distribution when included as controls. Student controls ($x_i$) include gender, race/ethnicity, English Learner status, and baseline test scores (omitted when used as the outcome). Propensity scores ($p_i$) capture the probability of assignment to each school. All specifications include market fixed effects (year $\times$ district $\times$ special education status) and special education indicators for grades 4 and 6 where applicable. Standard errors clustered at market level in parentheses. The joint test reports the F-statistic and p-value for the null hypothesis that all four FPC coefficients are jointly equal to zero. Significance levels: $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

Table D4: Alternative Non-linear Specification

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| $E$[peer share $\mid \theta$] | No | Yes | Yes | No | Yes | Yes |
| $x_i, p_i, s_i$ | No | No | Yes | No | No | Yes |
| Share peers in bottom quartile | -0.051 | -0.095$^{+}$ | 0.039 | | | |
| | (0.044) | (0.051) | (0.056) | | | |
| Share peers in top quartile | 0.286$^{**}$ | 0.147$^{**}$ | -0.017 | | | |
| | (0.034) | (0.047) | (0.076) | | | |
| Share peers in bottom quartile $\times$ Low-achieving | | | | -0.274$^{**}$ | -0.217$^{**}$ | -0.022 |
| | | | | (0.044) | (0.056) | (0.068) |
| Share peers in bottom quartile $\times$ Middle-achieving | | | | -0.202$^{**}$ | -0.194$^{**}$ | -0.010 |
| | | | | (0.044) | (0.067) | (0.063) |
| Share peers in bottom quartile $\times$ High-achieving | | | | 0.142$^{*}$ | 0.072 | 0.269$^{**}$ |
| | | | | (0.055) | (0.084) | (0.078) |
| Share peers in top quartile $\times$ Low-achieving | | | | -0.123$^{**}$ | -0.049 | -0.128 |
| | | | | (0.039) | (0.069) | (0.088) |
| Share peers in top quartile $\times$ Middle-achieving | | | | 0.104$^{**}$ | 0.078 | -0.062 |
| | | | | (0.034) | (0.062) | (0.080) |
| Share peers in top quartile $\times$ High-achieving | | | | 0.437$^{**}$ | 0.246$^{**}$ | 0.136 |
| | | | | (0.036) | (0.067) | (0.089) |
| N | 138771 | 138771 | 138771 | 138771 | 138771 | 138771 |
| $R^2$ | 0.784 | 0.785 | 0.802 | 0.787 | 0.787 | 0.805 |

*Notes:* This table reports reduced-form regressions that mimic the discrete peer-effects specification in Carrell, Sacerdote, and West (2013). Columns (1)–(3) use the share of co-enrolled peers in the bottom quartile and top quartile of the baseline test-score distribution. Columns (4)–(6) interact those peer shares with the student's own baseline tercile, labeled as low-achieving, middle-achieving, and high-achieving. In columns with expected peer-share controls, we include the corresponding expected co-assigned peer shares computed from simulated lottery assignments. Column (6) additionally controls for propensity scores, enrolled school fixed effects, market fixed effects (year $\times$ district $\times$ special education status), and student observables. Student observables include baseline test scores, English learner status, female, dummy variables for Asian, Black, and Hispanic, and Free-or-Reduced lunch eligibility. All specifications include special education indicators for grades 4 and 6. In column (6), we reject equality of the coefficients on the bottom-quartile peer-share interactions across low-, middle-, and high-achieving students ($F = 7.96$, $p = 0.001$), and likewise reject equality of the coefficients on the top-quartile peer-share interactions ($F = 6.12$, $p = 0.003$). Standard errors clustered at market level are reported in parentheses. Significance levels: $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

Table D5: Alternative Partition Specifications

| | (1) | (2) | (3) |
|---|---|---|---|
| Peer score | Tercile | Quartile | Quintile |
| Own score | Tercile | Quartile | Tercile |
| Peer tercile 1 x own tercile 1 | -0.083 | | |
| | (0.078) | | |
| Peer tercile 3 x own tercile 1 | -0.164$^{+}$ | | |
| | (0.089) | | |
| Peer tercile 1 x own tercile 2 | -0.010 | | |
| | (0.071) | | |
| Peer tercile 3 x own tercile 2 | -0.046 | | |
| | (0.080) | | |
| Peer tercile 1 x own tercile 3 | 0.157$^{+}$ | | |
| | (0.087) | | |
| Peer tercile 3 x own tercile 3 | 0.060 | | |
| | (0.090) | | |
| Peer quartile 1 x own quartile 1 | | -0.033 | |
| | | (0.076) | |
| Peer quartile 4 x own quartile 1 | | -0.156 | |
| | | (0.096) | |
| Peer quartile 1 x own quartile 2 | | 0.022 | |
| | | (0.066) | |
| Peer quartile 4 x own quartile 2 | | 0.032 | |
| | | (0.084) | |
| Peer quartile 1 x own quartile 3 | | 0.082 | |

*Continued on next page*

*Table D5 continued*

| Peer score | Tercile | Quartile | Quintile |
|---|---|---|---|
| Own score | Tercile | Quartile | Tercile |
| | | (0.072) | |
| Peer quartile 4 x own quartile 3 | | 0.006 | |
| | | (0.087) | |
| Peer quartile 1 x own quartile 4 | | 0.244** | |
| | | (0.090) | |
| Peer quartile 4 x own quartile 4 | | 0.100 | |
| | | (0.098) | |
| Peer quintile 1 x own tercile 1 | | | -0.034 |
| | | | (0.093) |
| Peer quintile 2 x own tercile 1 | | | -0.152 |
| | | | (0.115) |
| Peer quintile 4 x own tercile 1 | | | -0.192 |
| | | | (0.122) |
| Peer quintile 5 x own tercile 1 | | | -0.283* |
| | | | (0.116) |
| Peer quintile 1 x own tercile 2 | | | 0.058 |
| | | | (0.082) |
| Peer quintile 2 x own tercile 2 | | | 0.040 |
| | | | (0.085) |
| Peer quintile 4 x own tercile 2 | | | -0.102 |
| | | | (0.093) |
| Peer quintile 5 x own tercile 2 | | | -0.079 |
| | | | (0.095) |

*Continued on next page*

*Table D5 continued*

| Peer score | Tercile | Quartile | Quintile |
|---|---|---|---|
| Own score | Tercile | Quartile | Tercile |
| Peer quintile 1 x own tercile 3 | | | 0.048 |
| | | | (0.102) |
| Peer quintile 2 x own tercile 3 | | | 0.074 |
| | | | (0.116) |
| Peer quintile 4 x own tercile 3 | | | -0.225$^{+}$ |
| | | | (0.125) |
| Peer quintile 5 x own tercile 3 | | | -0.080 |
| | | | (0.112) |
| N | 138771 | 138771 | 138771 |
| $R^2$ | 0.805 | 0.806 | 0.805 |

*Notes:* This table reports full-control reduced-form regressions under three alternative discrete partitions of peer achievement. Column (1) interacts peer terciles with own terciles and omits peer tercile 2. Column (2) interacts peer quartiles with own quartiles and omits peer quartiles 2 and 3. Column (3) interacts peer quintiles with own terciles and omits peer quintile 3. Peer quintiles are constructed by summing adjacent decile shares from the underlying peer-share data. All columns include the corresponding expected peer-share interactions computed from simulated lottery assignments, propensity scores, enrolled school fixed effects, market fixed effects (year $\times$ district $\times$ special education status), student observables, and special education indicators for grades 4 and 6. Standard errors clustered at market level are reported in parentheses. Significance levels: $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.