Treatments, Peers, and Treatment Effects in Full-day Kindergarten: Reconciling Experimental and Quasi-experimental Impact Evidence*

Chloe R. Gibbs†

October 2016

*DRAFT. Please do not cite or circulate without permission.

†Contact: Department of Economics, University of Notre Dame, chloe.gibbs@nd.edu. I acknowledge funding support from a National Academy of Education/Spencer Foundation dissertation fellowship in education, a fellowship from the U.S. Department of Education Institute of Education Sciences predoctoral interdisciplinary training grant at the University of Chicago, and predissertation and dissertation awards from the Harris School's Center for Human Potential and Public Policy at the University of Chicago. I thank Kerwin Charles, Jens Ludwig, Ofer Malamud, Steve Raudenbush, Diane Schanzenbach, and seminar participants at Dartmouth College, Georgetown University, University of Notre Dame, Association for Education Finance and Policy, and Association for Public Policy Analysis and Management for helpful feedback. Any remaining errors are my own.

Abstract

As policymakers increasingly focus on the early childhood years as a critical period for establishing school readiness and remediating achievement gaps, one potential policy lever for intervention in early childhood is through greater availability of full-day kindergarten. This paper relies on data from naturally-occurring experiments and quasi-experiments resulting from a policy change in Indiana that greatly expanded resources available for full-day kindergarten. While state grant funding increased availability of and access to full-day kindergarten across the state, the per-pupil funding allocation did not fully fund full-day provision for all interested families and children. Because some school districts employed student assignment procedures based on lotteries and fixed cutpoints on pre-tests, I am able to test the causal impact of full-day assignment on literacy skills at the end of the kindergarten year and to consider the comparability of treatment effects from randomized and regression-discontinuity (RD) designs. Importantly, interpreting these findings in tandem requires consideration of differential effects across the distribution of kindergarten entry literacy skills and by student characteristics, homogeneous versus mixed ability grouping of students, and peer effects under different student assignment policies. While experimental evidence in this context indicates that full-day assignment has a substantial, positive effect (0.31 standard deviations) on early literacy skills, and Hispanic students realize particularly large effects (0.67 s.d.), the RD evidence does not generate similar findings. Importantly, I cannot rule out effects in the RD study of the magnitude realized in the experimental study. The two designs create two different treatment conditions and generate different estimates—though both causal—of the treatment effect. This paper makes use of the experimental data to explore approaches to reconciling, or more suitably comparing, evidence from the two designs.

1 Introduction

The early childhood years have garnered much policy interest as a critical developmental period and opportunity for early intervention, potentially remediating cognitive and noncognitive skill gaps prior to formal schooling and generating high returns on investment in the form of private and social benefits over the life cycle. Several papers have established the early emergence of achievement gaps by race/ethnicity and socioeconomic status (Fryer & Levitt 2004, Fryer & Levitt 2006, Lee & Burkam 2002, Murnane, Willett, Bub & McCartney 2006, Princiotta, Flanagan & Germino Hausken 2006). While there is some dispute about the magnitude of these gaps—and whether and how they can be explained by student characteristics and family background-–there is consistent evidence that test-score gaps by race and socioeconomic status are already sizable at the end of the kindergarten year. In addition, these gaps persist and grow throughout the primary grades.

There is also a growing literature on the importance of kindergarten—both individual skills demonstrated as early as kindergarten and the quality of kindergarten contexts—in predicting later academic and labor market success (Chetty, Friedman, Hilger, Saez, Schanzenbach & Yagan 2011, Duncan, Dowsett, Claessens, Magnuson, Huston, Klebanov, Pagani, Feinstein, Engel, Brooks-Gunn, Sexton & Duckworth 2007, Dynarski, Hyman & Schanzenbach 2011). Notably, President Obama's "Preschool for All" proposal, introduced in the 2013 State of the Union and further articulated on February 13, 2013, included specific mention of full-day kindergarten expansions as a policy tool (The White House, 2013). Nearly all children in the United States attend kindergarten and states have increasingly turned to expansions of full-day kindergarten policies in their efforts to reach children early in their formal schooling.

This paper seeks to contribute to the discussion of the efficacy of interventions early in schooling by providing experimental and quasi-experimental evidence on the impact of full-day kindergarten. Moreover, the evidence informs the decision about how to spend the marginal dollar of investment in early childhood under constraints. The motivating research question is whether or not full-day kindergarten students outperform their half-day kindergarten peers as measured by literacy skills at the end of the kindergarten year, and the paper also discusses how to consider in tandem those estimates of program impact from different study designs.

The experimental results show that full-day kindergarten assignment has a large, positive effect (0.31 standard deviations) on end-of-kindergarten literacy skills when comparing students across treatment conditions within the same school. In particular, I find that Hispanic students realize large full-day kindergarten effects (0.67 s.d.), and notably this impact constitutes 73 percent of the control group's end-of-kindergarten ethnicity gap. The quasi-experimental evidence from a regression-discontinuity (RD) design capitalizes on discontinuities in the probability of treatment assignment around a fixed cutpoint on a pre-test to generate an estimate of the local average treatment effect (LATE), comparing the outcomes of students above and near the cut-score (in half-day kindergarten) with those below and near the cut-score (in full-day kindergarten). Students in full-day kindergarten in this context and near the margin of interest do not outperform their half-day kindergarten counterparts, but the estimates are imprecisely measured. Importantly, interpreting the experimental and quasi-experimental evidence together requires consideration of the different study samples, treatment conditions, and estimates. I leverage the experimental data to explore implications of the RD design and results.

In presenting the evidence on the causal impact of full-day kindergarten, the paper proceeds as follows. This section discusses further the rationale for studying this question as well as previous, related research. Section 2 details the companion experiment and RD studies, including their respective designs, data, empirical approaches, and results. In Section 3, I discuss the lottery and regression discontinuity designs collectively and interpretation of findings from the two studies. I then conclude with reanalyses of the lottery study data to probe possible explanations for inconsistent findings.

1.1 Motivation

Investment in early childhood, as Nobel Laureate James Heckman argues, is one of the rare public policy options in which there are no equity-efficiency tradeoffs. Investing early and in disadvantaged children is both equitable and socially efficient (Heckman & Masterov 2007). The notion that "skills beget skills" describes the role that foundational skill development plays in supporting and augmenting subsequent human capital investments throughout the life cycle, rendering them more efficacious (Heckman 2000, Heckman & Lochner 2000). Moreover, investments in early childhood

have a longer time horizon over which to realize their benefits. As early childhood program and policies seek to compensate for the impoverished developmental environments from which many disadvantaged children come, these interventions address equity concerns while often also generating returns to the individual and society in excess of their costs.

Economists argue that investments in early childhood—that could be viewed as preventative—are less costly than interventions later in life to address crime, substance abuse, or workforce development and training (Currie 2001). The externalities associated with inadequate early education are costly to society and not fully internalized by parents making private decisions about investment in their children's early schooling. In addition to the theoretical argument that preventing social problems now is more cost-effective than remediating them later (Currie 2001), there is also a growing body of empirical evidence that early childhood programs reap long-term effects for participants, generating substantial private and social returns that far outweigh the program costs. Long-term evidence from the Perry Preschool Project, Head Start, and the Project STAR class-size reduction intervention all suggests that interventions in the preschool and early school years can have substantial effects on schooling attainment, labor market success, and other measures of health and well-being into adulthood (Chetty et al. 2011, Deming 2009, Dynarski et al. 2011, Garces, Thomas & Currie 2002, Schweinhart, Montie, Xiang, Barnett, Belfield & Nores 2005).

1.2 Mechanisms

There are a few, entangled mechanisms through which full-day kindergarten specifically may improve students' short- and longer-term outcomes. The education production function, widely used in labor economics to understand the organization and deployment of resources in education, provides a framework for understanding the potential impact of full-day relative to half-day kindergarten. When modeling instructional time in this framework, more instructional minutes potentially result in increased achievement, but likely exhibit diminishing returns (i.e., have a larger impact when the initial instructional minutes are relatively low). The magnitude of the impact varies with other inputs as well, so location on the education production possibilities frontier matters for the effect on individual student outcomes.

Dating back to the Coleman Report (1960) there is considerable debate about the role of re-

sources in improving children's educational outcomes.[1] The body of evidence is mixed and inconclusive, but the quality of the education production function estimates in the literature is also quite varied. The randomized study of Project STAR, an intervention to reduce class size in the early grades, is one of the largest scale and most rigorous tests of the role of increased resources in generating educational output. On average, being in a small class was associated with a 0.22 s.d. effect on test scores (Krueger 2002). Notably, class-size reduction is a resource input that similarly operates on the intensive margin of schooling, and in the case of Project STAR, also occurred early in formal K-12 schooling.

It is important to note that increased instructional time has the additional effect of crowding out other time use activities—a counterfactual condition that may differ by student characteristics, both observed and unobservable—also suggesting the possibility of non-constant treatment effects.

Moreover, full-day kindergarten provides fully subsidized child care in the extended hours, and therefore, constitutes a wealth transfer to parents of full-day kindergarten students. The provision of full-day kindergarten allows parents to adjust to first-grade (and beyond) levels of consumption and investment in advance. Decreasing child care costs increases the effective wage which should lead to greater employment among parents of full-day kindergarten students all else equal. Since all of these mechanisms are operating simultaneously, it is not possible to disentangle these effects from the increased input to the education production function. Importantly, these components of the composite effect would all exist in any implementation of public full-day kindergarten, so the combined impact is the policy relevant parameter to estimate.

1.3 Kindergarten Research

Nearly all students who attend school outside the home participate in kindergarten in the United States and the majority of students in kindergarten are in full-day settings (Walston & West 2004). As displayed in Figure 1, less than 20 percent of kindergarten students were in full-day settings in 1970. Full-day kindergarten enrollment exceeded half-day participation for the first time in 1995 and approximately 70 percent of kindergarteners were in full-day classrooms by 2010.

[1] See Hanushek 1979, Hedges et al. 1994, and Hanushek 1997 for the evolution of the debate, and the corresponding evidence, on whether resources matter for educational achievement and attainment.

Despite its popularity and growing enrollment, research on the benefits of full-day kindergarten, however, is mixed and lacking in rigorous approaches to estimating program impact. Importantly, in this literature full-day kindergarten is compared to half-day kindergarten, rather than no kindergarten attendance at all. This is a notable departure from the literature on other early childhood interventions, particularly pre-kindergarten, and the existing research on kindergarten expansions.

To set the context, a few papers have conducted research relevant to kindergarten that informs the discussion of the full-day kindergarten literature. Two papers explore the impact of kindergarten availability—estimating on the extensive margin—using the plausibly exogenous timing of kindergarten expansions. Cascio (2009) relies on the varied timing of state kindergarten grants to school districts to estimate the long-term effects of kindergarten availability. She finds effects for white children in the form of lower likelihood to drop out of high school and lower institutionalization rates as adults. She does not find similar effects for blacks–though black children experienced similar increases in public kindergarten enrollment–which is likely due to the crowd out of other early childhood interventions available to disadvantaged populations (Cascio 2009). Dhuey (2011) similarly exploits variation in public kindergarten expansions, though she uses significant increases in kindergarten availability within a state for identification. She finds that Hispanic children and those who live in immigrant households, are of low socioeconomic status, and do not speak English experience benefits from the availability of kindergarten with lower likelihood of being below grade for age and higher wages as adults (Dhuey 2011).

Two recent papers explore the intensive margin, including the quality dimension, of kindergarten participation. Using Project STAR data, researchers find that kindergarten test scores are highly correlated with important, long-term outcomes including college attendance, adult earnings, home ownership, and retirement savings (Chetty, Friedman, Hilger, Saez, Schanzenbach & Yagan 2010). They find that kindergarten intensity, as operationalized by small class size, predicts college attendance. In addition, kindergarten quality measured by teacher experience and peer ability relates to college attendance and higher earnings. While they observe fade out of kindergarten quality effects in test scores, the positive effects on sociocognitive measures remain (Chetty et al. 2010). Fitzpatrick, Grissmer, and Hastedt (2011) capitalize on quasi-randomness in dates of test administration to estimate gains to schooling over the course of the kindergarten and first-grade years.

They find that one year of schooling corresponds to 1.2 s.d. on reading tests and 0.9 s.d. on math tests in those early grades, over and above the normal developmental growth children are experiencing. Their results have important implications for thinking about extended school time, particularly early in schooling. While they focus on implications for extending the school year, their work also suggests that substantial learning could take place when doubling the number of instructional hours kindergarten students experience.

The existing literature on full-day kindergarten takes two forms: studies using nationally representative data and district- and school-level evaluations. In observational studies using the ECLS-K, researchers have found significant differences between full- and half-day kindergarten students on literacy and mathematics assessments at the end of the kindergarten year (Cannon et al. 2006, DeCicca 2007, Lee et al. 2006, Votruba-Drzal et al. 2008). These full-day kindergarten advantages failed to persist, however, over the first-grade year. In one study, marginally significant differences were found in the spring of first grade (Cannon et al. 2006). DeCicca (2007) found significant differences in mathematics and reading in the fall of first grade, but only for white children, which faded but continued to be significant in spring literacy performance. No significant differences were found between full- and half-day kindergarten students in the ECLS-K in third grade (Cannon et al. 2006, Votruba-Drzal et al. 2008) or fifth grade (Votruba-Drzal et al. 2008).

Additional smaller-scale evaluations have supported the ECLS-K findings of short-term outcomes in the kindergarten year, but no significant long-term effects (Zvoch, Reynolds & Parker 2008, Hall-Kenyon, Bingham & Korth 2009). In general, findings on the impact of full-day kindergarten relative to half-day kindergarten suggest some positive associations, particularly in the early schooling years. Results related to the impact of full-day kindergarten over time, or the persistence of these positive findings, are more mixed. All of these studies are still subject to concerns about selection bias in that they cannot fully address endogeneity of student assignment to–or school or district provision of–full-day kindergarten. The present studies are the first in the full-day kindergarten literature to use random and fixed cutpoint assignment of students to full- and half-day settings to test the causal impact of full day.

2 The Full-Day Kindergarten Experiment and Quasi-Experiment: Background, Data, Empirical Strategy, and Results

The analysis presented in this paper employs data from seven school districts in Indiana that were unable to provide full-day kindergarten to all interested kindergarten students in the 2007–08 academic year. To allocate the oversubscribed slots, five of these districts used random lotteries to assign entering kindergarten students to either full-day or half-day kindergarten classrooms, creating randomized groups of students in the kindergarten settings. The existence of these lotteries provides a unique opportunity to study rigorously the effects of full-day kindergarten on early literacy skills. In addition, two large school districts used a pre-test measure of academic need to assign students to full- and half-day settings. In particular, the included districts aimed to provide limited full-day kindergarten slots to the most academically needy students, as measured by their literacy skills performance prior to kindergarten entry. Leveraging exogenous variation near the cutpoint of assignment to full- or half-day, this regression-discontinuity design can also obtain unbiased estimates of program impact. The existence of these seven naturally occurring experiments allows for more rigorous analysis of full-day kindergarten impact than previously available in the existing literature and also facilitates comparisons across the two study designs to consider how student assignment to intervention may shape interpretation of results.

2.1 Background

In 2007, the Indiana legislature passed legislation which provided funding to increase access and availability of full-day kindergarten in the 2007–08 school year, with grants targeted directly to school districts and charter schools (Indiana Public Law 234-2007). The stated goal of the bill was to allow, "school corporations [districts] and charter schools to provide full-day kindergarten programs to improve the academic and social development of children in kindergarten." A policy initiative of Governor Mitch Daniels' administration and supported by the Indiana State Board of Education and the Indiana Department of Education, the legislation expanded state grant funds for full-day kindergarten from $8.5 million in the 2006-07 school year to $33.5 million in the 2007-08 school year (Indiana General Assembly 2007). School districts and charter schools, operating as

autonomous school districts in the state, were eligible to apply to the state for full-day kindergarten funding.[1] Grant funding was then dispersed to all applicants on a per-pupil basis, allocated based on kindergarten enrollment in the school or district in the 2007-08 school year.

Full-day kindergarten enrollment in the state increased by 20 percentage points from 41 percent of kindergarten students in 2006-07 to 63 percent in 2007-08. The number of school districts and schools offering full-day kindergarten also increased with a 26 percentage-point increase in the number of school districts enrolling more the vast majority of their kindergarten students in full-day settings, and a 21 percentage-point increase in the number of schools with nearly all of their kindergarten students in full-day settings (Lovell, Kochanek, Mathers & Burke 2009).[2] With the increased funding availability from the state, the full-day kindergarten grant became the primary means for supporting full-day kindergarten enrollment, coupled with federal Title I funds, school district general funds, and parent fees in some cases. School district officials indicated that 91 percent of full-day kindergarten students in the state were funded, at least in part, by the state full-day kindergarten grant monies in 2007-08 (Lovell et al. 2009).

Because grants were provided to all interested district and schools and allocated based on kindergarten enrollment, the per-pupil amount was insufficient to provide full-day kindergarten to all students in the 2007-08 school year in many cases. Districts and schools determined the assignment procedures for allocating oversubscribed slots. Table 1 provides descriptive data on the five participating districts that employed lotteries to determine assignment to oversubscribed slots. The districts are varied in their urbanicity and location. They also vary in their composition of minority and disadvantaged students. The two smallest districts–one is a charter school that operates as an autonomous school district–have only one elementary school while the largest districts have over 500 kindergarten students. Notably, the kindergarten enrollment numbers in Table 1 differ from the study sample in that not all students in the districts were assigned to kindergarten setting by lottery.

Within the five school districts, there are 23 participating schools. Table 2 provides a com-

[1]"Application" for full-day kindergarten funding consisted of indicating interest to the Indiana Department of Education and doing so by a deadline in order to establish enrollment numbers and make grant allocations.

[2]The definition employed for the purposes of reporting change in school district and school provision is that more than 80 percent of enrolled kindergarten students in the district or school participate in full-day kindergarten.

parison of the study schools and all elementary schools in Indiana that provide kindergarten. The comparison of school characteristics demonstrates that study schools are more often suburban and less likely to be rural or located in a town than other Indiana schools. In addition, study schools have a slightly lower prevalence of Title I eligibility and are less likely to be designated for Title I school-wide. The average student composition in study schools is less disadvantaged and comprised of fewer black students and more Hispanic students than the average Indiana elementary school.

Two school districts in Indiana implemented a fixed cutpoint approach to student assignment. While several other districts indicated their intention to use this approach to student assignment, most did not implement the fixed cutpoint based assignment with fidelity and some introduced additional criteria, including parent preferences or principal referral, to the assignment scheme. Table 3 presents the district characteristics for districts that elected to assign oversubscribed full-day kindergarten slots based on academic need, using a pre-kindergarten literacy skills test to assign the lowest performers to full-day settings. As displayed in the table, the two regression-discontinuity (RD) districts are relatively large, both in small cities, and both have relatively high proportions of students who qualify for free or reduced-price lunch. They differ in their minority composition, but have similar numbers of elementary schools and kindergarten enrollment. One district has more full-time equivalent kindergarten teachers, resulting in a higher student-teacher ratio in kindergarten of 27:1 as compared to 22:1 in the other district.

Table 4 presents the 18 study schools as compared to all Indiana elementary schools.[3] As displayed in the table, study schools located in the two regression-discontinuity districts have different characteristics than Indiana public elementary schools at large. In particular, study schools are predominately in city locales (89 percent) while under 30 percent of Indiana elementary schools are in cities. All study schools have Title I eligible students and half are Title I schoolwide schools, as compared to nearly all Indiana elementary schools with Title I eligible students, but only 15 percent with Title I schoolwide funding.

The average student composition in RD study schools is generally more disadvantaged than schools in the state, with average eligibility for free or reduced-price lunch at just under 60 percent.

[3]The data for 1,119 elementary schools in Indiana is restricted to public schools that provided kindergarten in the 2007-08 school year.

Study schools have student populations comprised of one-quarter black students on average compared to the statewide average of 13 percent. There are correspondingly fewer white students in study schools and also a lower percentage of Hispanic students as compared to schools statewide. As displayed in the evidence on district and school characteristics, one threat to the generalizability of results from either design—or the comparability of results across the two designs—is the very different context and composition of schools in the two samples.

2.2 Data

Data for this study were supplied by the seven participating districts, with assistance in collecting the data from the Indiana Department of Education. Districts provided three types of files, which were merged to conduct analysis of program impact. Coupled with student-level identified data in the state's data system, these data provide the necessary information to assess immediate program impact for the purposes of this paper as well as follow-up analyses of longer-term outcomes in subsequent academic years.

Files with lottery results include student names, unique student identifiers, and determination of full-day or half-day kindergarten status based on the randomized lottery. In the RD districts, pre-kindergarten assessment files include student names, unique student identifiers, and scores on the literacy skills assessment administered prior to the beginning of the school year. School district administrators supplied information on the relevant cutpoints for assignment to full- and half-day kindergarten based on those assessment scores. The participating districts used the Kindergarten Readiness Test (KRT) as a diagnostic assessment prior to the start of the kindergarten year. The KRT covers vocabulary, phonemic awareness, letter identification, comprehension and interpretation, visual discrimination, and mathematical knowledge.

School districts also provided administrative records which again include student names and unique student identifiers as well as demographic variables (date of birth, gender, race/ethnicity, and free or reduced-price lunch eligibility). The administrative data include identifiers for the student's teacher, school, and district and an indicator of whether the student was enrolled in full- or half-day kindergarten in the 2007–08 school year. Finally, the districts supplied assessment records that provide results of a end-of-kindergarten year literacy skills assessment.

As required by their participation in the state-funded full-day kindergarten grant, schools in the study districts administered either the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) or the Indiana Reading Diagnostic Assessment (IRDA), scores from which were standardized within sample for use in the study. These assessment records include student names, unique identifiers, scores, and time frame of administration. Both assessments focus on the five essential components of reading, as detailed by the 2000 National Reading Panel: phonemic awareness, phonics, reading comprehension, fluency, and vocabulary. The assessment subparts administered in the kindergarten year emphasize phonics, phoneme segmentation, and letter recognition. The DIBELS and IRDA are formative assessments intended to inform instructional practice and provide teachers and administrators with data for diagnostic purposes. Scores from the assessments were standardized within sample to generate comparable z-scores for inclusion in the analyses.

2.3 Empirical Strategy

The basic ordinary least squares (OLS) model for assessing the impact of full-day kindergarten in lottery sites is:

$$Y_{ik} = \beta_0 + \beta_1 ASSIGN_i + (\beta_2 CH_i) + \gamma_k + \varepsilon_{ik} \tag{1}$$

where Y is the literacy outcome measure for student i in school k. *ASSIGN* is the treatment indicator variable, which takes a value of zero for assignment to half day and one for assignment to full day, regardless of receipt of treatment, resulting in an intention-to-treat estimate (β_1). *CH* is a vector of child-level characteristics for each student i, and γ represents the school fixed effects. In the outcomes model specifications, the child-level characteristics are excluded from one specification and included in one specification for precision. School fixed effects also contribute to precision. Standard errors are heteroskedasticity-robust, clustered at the classroom level.[4]

An analysis accounting for noncompliance with treatment assignment was also conducted in a two-stage least squares (2SLS) framework, employing random assignment to full-day kindergarten

[4]Because of possible concerns about Huber-White standard errors and poor performance in the presence of the relatively small number of clusters (46 classrooms), I employ the Wild cluster-bootstrap percentile-t procedure, imposing the null hypothesis (Cameron, Gelbach & Miller 2008). Inferences are unchanged.

as an instrumental variable (IV) for full-day kindergarten participation. To estimate the impact of enrolling in full-day kindergarten, the IV approach adjusts the effects of being assigned to full-day kindergarten via the lottery to account for the proportions of students assigned to full-day kindergarten who instead attend half-day kindergarten and students assigned to half-day kindergarten who instead attend full-day kindergarten. The 2SLS models are as follows:

$$FDK_{ik} = \alpha_0 + \alpha_1 ASSIGN_i + (\alpha_2 CH_i) + \gamma_k + \mu_{ik} \tag{2}$$

$$Y_{ik} = \delta_0 + \delta_1 \hat{FDK}_i + (\delta_2 CH_i) + \gamma_k + \varepsilon_{ik} \tag{3}$$

where assignment status ($ASSIGN$) is used to predict actual participation (FDK), and the predicted values are employed in the second-stage estimation of treatment impact on outcomes. The IV models produce an estimate of the local average treatment effect, or the LATE, which generalizes to compliers with random assignment. Because of the low frequency of treatment group crossovers, these results do not differ meaningfully from the OLS, or intention-to-treat, results.

Because of the fuzzy regression-discontinuity design (in which the probability of assignment to full-day kindergarten increases discontinuously at the fixed cutpoint), the RD analyses are conducted in a 2SLS framework. In the first stage, the binary assignment variable–based on pre-test score relative to the cutpoint–is used as an instrument for actual participation in full- or half-day kindergarten. The instrumental variable (IV) model is as follows:

$$FDK_{ik} = \alpha_0 + f(PRE_i) + \alpha_1(c_i \leq 0) + (\alpha_2 CH_i) + \gamma_k + \mu_{ik} \tag{4}$$

where FDK is program participation, PRE is the continuous assignment variable, pre-test score for individual student i. Because low performers on the pre-test were assigned to full-day kindergarten, $1(c_i \leq 0)$ takes on a value of one for students with a re-centered pre-test score below the assignment threshold and a zero for students with a re-centered pre-test score above the threshold. CH is a vector of baseline characteristics which are both included and excluded in different model specifications. Table 5 displays the results of the first-stage estimation. In the second stage, the

instrumented values of program participation are included. The outcome model is:

$$Y_{ik} = \delta_0 + g(PRE_i) + \delta_1 F\hat{D}K_i + (\delta_3 CH_i) + \gamma_k + \varepsilon_{ik} \tag{5}$$

School fixed effects (γ_k) are included, leveraging within school variation in kindergarten status to estimate the impact of full- relative to half-day kindergarten. Because the specification of the functional form of the running variable, pre-test score, is critical to the interpretation of regression-discontinuity results, $f(.)$ and $g(.)$ are entered both linearly and quadratically. I also run the models with the inclusion of third and fourth degree polynomials of the running variable, but it did not improve overall model fit and the results do not differ meaningfully from the case with the linear and quadratic pre-test.

To test for differences in slope, I also using the following second-stage model, in which instrumented values of participation as well an instrumented value of the interaction of assignment and pre-test score are entered:

$$Y_{ik} = \delta_0 + \delta_1 F\hat{D}K_i + \delta_2 F\hat{D}K_i * PRE_i + \delta_3 PRE_i + \gamma_k + \varepsilon_{ik} \tag{6}$$

Notably, this model does not include higher-order polynomials of the pre-test because pre-test and the interaction term are entered linearly.

These IV models of full-day kindergarten impact generate estimates of the local average treatment effect (LATE), and thus are notably constrained in their ability to detect an impact and their generalizability. To the first point, the models estimate the effect of full-day kindergarten for compliers at the point in the running variable distribution at which the cutpoint occurs. I can only isolate the presence of full-day kindergarten impact for participants on that margin. In the presence of heterogeneous treatment effects along the running variable, we may or may not expect effects at the measurable margin. Moreover, results from this approach are only applicable to that specific population and not generalizable across the pre-test score distribution. This is particularly troubling when effects for compliers do not constitute an interesting policy parameter and when the margin at which the cutpoint occurs is not policy relevant for broader populations.

2.4 Results

The lottery study results are as follows. As described in the empirical strategy, I present results from the OLS models (2.5) for the outcome of interest, the literacy post-test measure, in Table 6. Model I includes only the indicator variable for full-day kindergarten assignment as a predictor, while Model II also incorporates student characteristics. The models are restricted to the observations included in both specifications, but the results in Model I are robust to inclusion of observations that are missing on the covariates. All models employ school fixed effects and heteroskedasticity-robust standard errors, clustered at the classroom level.

As described in further detail in Gibbs (2014), assignment to full-day kindergarten results in a sizable, statistically significant positive effect (0.31 s.d.) on end-of-kindergarten literacy skills. That finding is unchanged with or without the inclusion of individual student characteristics, which were balanced at baseline. Not surprisingly, since the LATE estimator essentially inflates the intention-to-treat estimate by the compliance rates, the IV results are larger in magnitude (0.34 s.d.). It is reasonable to consider the intention-to-treat estimate, which does not account for noncompliance and treatment crossovers, as a lower bound estimate of the treatment effect. A naïve estimate of the treatment on the treated, estimated from running equation (5) with full-day participation as the treatment indicator on the same sample of 975 observations, results in a similar estimate of the treatment effect (0.32 s.d.). Not surprisingly, these impact estimates are all comparable as compliance with treatment assignment was very high.

To explore heterogeneity in treatment effects, Table 7 presents results from models with the inclusion of interaction terms for poverty by treatment and race by treatment. Exploration of interactions with age and gender did not yield evidence of differential effects. As displayed in Table 7, the interaction of poverty and full-day assignment is not statistically significant at conventional levels in either specification–with and without the inclusion of covariates–but there is suggestive evidence that the treatment effect may be more heavily concentrated among disadvantaged students. The same pattern appears for the interaction of nonwhite race/ethnicity and full-day kindergarten assignment.

Exploration of subgroup effects suggests that disadvantaged students benefit greatly from full-

day kindergarten, as measured by end-of-year literacy skills. Specifically, nonwhite, Hispanic students assigned to full day make sizable gains relative to their half-day kindergarten counterparts at the end of the kindergarten year. The impact estimates from intention-to-treat models are summarized, with standard errors, in Figure 2. Figure 2 displays another important finding related to these subgroup findings. The variation in impact estimates by student characteristics suggests that full-day kindergarten reduces end-of-kindergarten achievement gaps, particularly between Hispanic and non-Hispanic students. This feature of an early childhood program is important when considering universal provision. Not only does full-day kindergarten make all students better off, it simultaneously contributes to closing race/ethnicity gaps.

The RD results do not exhibit a similar pattern of sizable, positive full-day kindergarten effects. To first assess the possible relationship between the treatment of full-day kindergarten and literacy outcomes of interest, I present a figure displaying the relationship between the running variable and baseline characteristics. Figure 4 provides the graphical depiction of the relationship between the assignment variable and the outcome measure. As evidenced by the figure, visual inspection suggests no discontinuity in outcomes at the assignment cutpoint. The lack of visual evidence of a discontinuity in outcomes suggests that full-day kindergarten may have no impact for students on the margin of the pre-test score distribution at which the fixed cutpoint was established. Table 8 presents the results for the full sample, providing both the OLS and IV results. As one would expect, the naïve estimate of the relationship between program participation and end-of-kindergarten literacy skills is negative.

The comparable IV model also shows a negative relationship, but is estimated from the full sample. From left to right, I enrich the IV models with the interaction of treatment status and the assignment variable as well as standardized scores on that literacy pre-test, and then add polynomials of the pre-test to the specification. My preferred specification, in keeping with the literature, includes standardized pre-test scores as well as pre-test polynomials and baseline characteristics. These results correspond to the graphical results suggesting no statistically significant relationship between full-day kindergarten participation and end-of-kindergarten literacy skills.

Table 9 provides results to the same model specifications deployed with a restricted sample. In these models, the included sample of 482 kindergarten students is restricted to those within one

standard deviation of the cutpoint. The results largely mirror those with the full sample. Table 10 again presents the model results with a restricted sample, presenting data from the 233 students within one-half of a standard deviation on either side of the cutpoint, or approximately ten raw score points on either side of the discontinuity. The OLS results switch signs in this sample of students close to the cutpoint, but are not statistically significant. The evidence from the IV models shows no consistent pattern of results, but suggests a negative relationship between full-day participation and the literacy skills outcome. Finally, in Table 11, I present results for a sample very close to the treatment discontinuity, restricting the sample to those 136 kindergarteners within one-quarter of a standard deviation, or approximately five raw score points around the cutpoint. The IV results generally switch signs (relative to the previous restricted sample), but show no systematic pattern in the relationship between treatment and outcomes.

3 Reconciling Findings

Leveraging exogenous variation in student assignment to full- and half-day kindergarten around a cutpoint on an assessment administered prior to kindergarten entry, this RD design capitalizes on these discontinuities in the probability of treatment assignment to assess the impact of full-day kindergarten participation on end-of-year literacy skills. The study generates a LATE estimate, comparing the outcomes of students above and near the cut-score (in half-day kindergarten) with those below and near the cut-score (in full-day kindergarten). While there was a discontinuity in probability of treatment at the cutpoint and no other discontinuities in student characteristics at the same point, I find no evidence that students in full-day kindergarten near the margin of interest outperform those in half-day kindergarten.

In considering the findings on the impact of full-day kindergarten on end-of-kindergarten literacy skills performance in the context of a regression-discontinuity design, a number of important considerations must sit at the forefront. Specifically, issues of homogeneous versus mixed ability grouping of students, peer effects under different student assignment policies, and heterogeneous treatment effects across the distribution of kindergarten entry literacy skills are essential to interpreting this evidence in the context of both the existing evidence on full-day kindergarten as well as the findings

from the lottery study conducted in school districts in the same state at the same time. While individual students are assigned to full- and half-day settings, the length of day is an intervention administered to a whole classroom. Thus, the student assignment policy employed changes the nature of the intervention in different contexts, coupling with the treatment of a lengthened school day the peer group and any resulting peer effects as well as teacher response to the student grouping in his or her classroom.

In the regression-discontinuity study, students who participated in full-day kindergarten were also assigned to a lower-performing peer group. In contrast, students in half-day settings experienced a relatively higher-achieving group of peers in the kindergarten classroom. In addition to the peer effects that may result–and the differing direction of those peer effects in the two settings–the marginal students in full- and half-day contexts also fall in the tails of performance for their classrooms. That is, the marginal student in a full-day setting is the highest performer at kindergarten entry relative to a lower-performing group of students while the marginal student in half-day kindergarten is a low performer as compared to the mean kindergarten-entry literacy skills in a half-day classroom.

Because there are concentrated effects among disadvantaged students in the lottery study, the optimal policy response—in the presence of constrained resources—may be to target full-day kindergarten to specific populations of students at kindergarten entry. The study design, in the lottery case, is individual student assignment with treatment administered at the classroom level, so it is important to note the bundling of treatment with the classroom peer group. Students received the treatment of full-day kindergarten in the context of a varied peer group with mixed ability at kindergarten entry. It remains unknown whether these large, positive treatment effects would generalize to a context in which students participated in full-day kindergarten with a homogeneous peer group.

Finally, in the presence of heterogeneous treatment effects, as realized in the lottery study for nonwhite, predominately Hispanic students, schools and districts that differ in their populations of those subgroups of students would likely experience different effects. If the treatment effects in the lottery study are driven largely by the presence of a large Hispanic student population—and similarly schools with large Hispanic populations generate the largest positive impact estimates

in the school fixed effects model—then schools and districts with less representation of the most affected students would not produce the same gains. The composition of the districts and schools in the regression-discontinuity study is quite different than those in the lottery study, with dramatically lower proportions of Hispanic students.

While the treatment contexts of full-day kindergarten are different in the lottery and regression discontinuity studies, it is important to consider these results in tandem. Reconciling the evidence on full-day kindergarten, including consideration of the existing, observational literature, is most informative for policy development and implementation. To inform this discussion, I use lottery data for further exploration. Since students were also randomly assigned to peer groups in the lottery study, I use variation in the mean classroom literacy skills at kindergarten entry to run a "horse race" between full-day kindergarten and peer quality. These analyses help to understand the important role that peers play in interventions conducted in group settings. Table 12 displays the results of those models with the lottery data, and peers are clearly important. In Table 13, I present similar models with the inclusion of an interaction term to consider the joint relationship of peers and full-day kindergarten in producing individual literacy outcomes. Again, peers are influential, which may suggest that a targeted approach to full-day kindergarten provision—as in the regression-discontinuity study districts–would not produce positive effects of the size shown in the lottery study.

When thinking about the generalizability and scalability of these results, it is also important to consider the demographic makeup of these districts. Table 14 presents the dramatically different composition of the lottery and regression discontinuity samples. In the presence of heterogenous treatment effects for certain subgroups of students, it may be concerning that other student populations would not see the same returns to full-day kindergarten participation. In the lottery sample, all subgroups experience positive gains. Various approaches to cross-cutting the data still uncover benefits for students of different demographic backgrounds and kindergarten entry skill levels. As an additional attempt to address this concern, I reweight the lottery sample data to resemble the regression discontinuity sample (upweighting those students with similar baseline characteristics, and downweighting others). Table 15 displays the results of this analysis and suggests that the full-day kindergarten advantage would look similar even with a differently comprised group of students.

Taken in tandem, the results suggest that peer quality matters, and should be coupled with full-day kindergarten as student assignment policies will allow. When considering full-day kindergarten program expansions, one important policy design consideration is that of targeted versus universal provision. The findings indicate that targeted provision may not prove as effective as full-day participation for all kindergarten students. The presence of concentrated effects for certain subgroups of students, coupled with positive impact for many participants regardless of demographic characteristics and kindergarten entry skills, suggests that universal provision of full-day kindergarten could do much to alleviate early schooling achievement gaps.

Figures

Figure 1: Kindergarten Enrollment (in thousands)

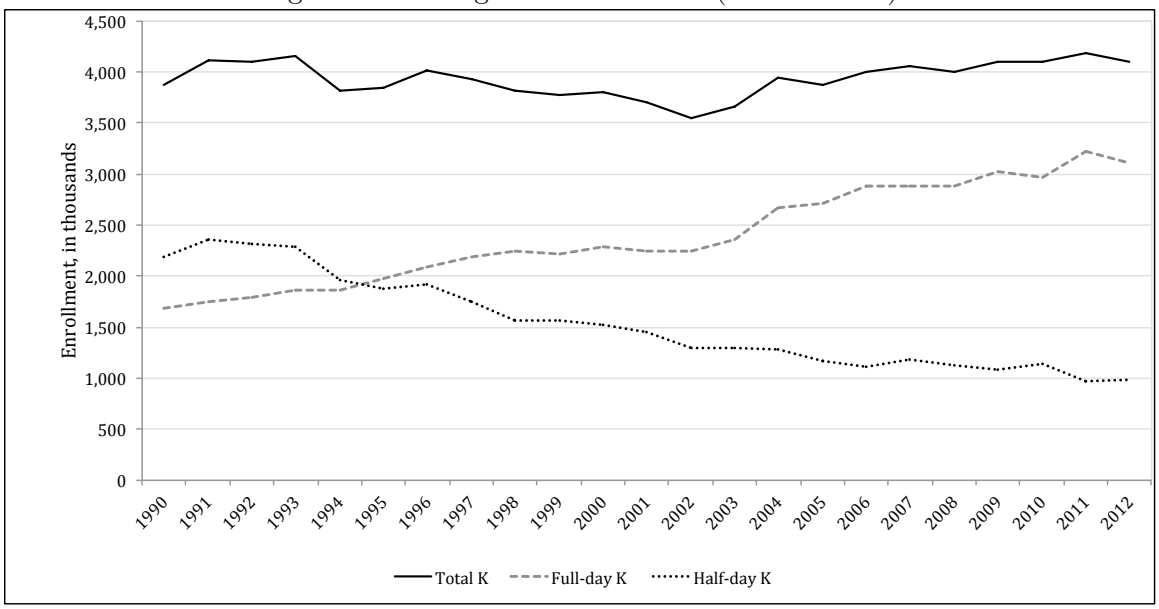


Figure 2: Intention-to-Treat Impact Estimates

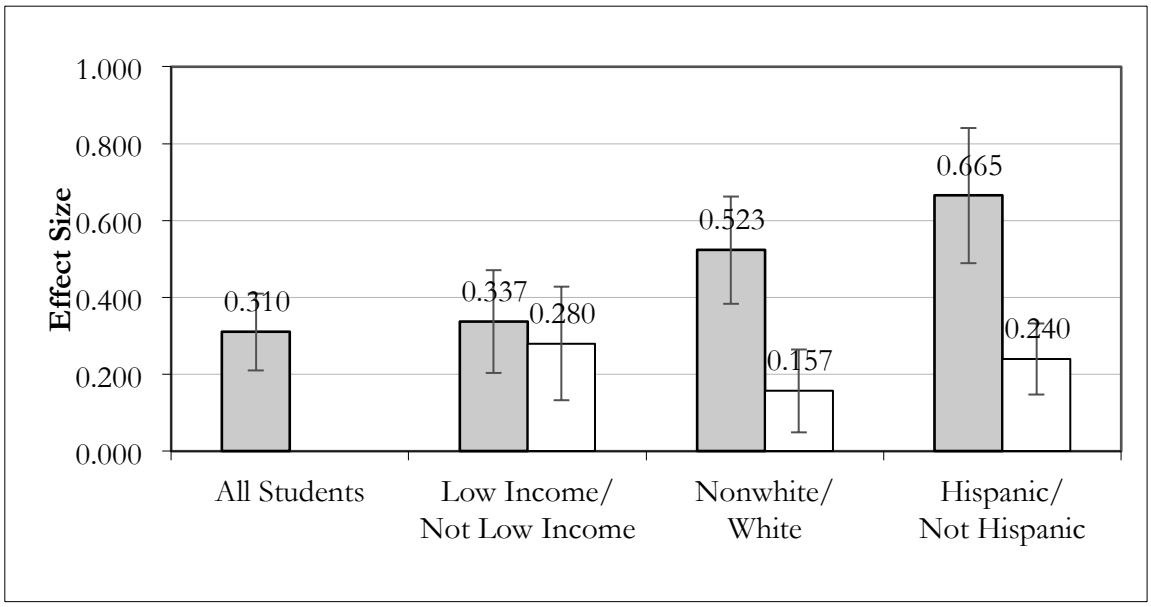


Figure 3: Literacy Skills Gaps in Full- and Half-Day Kindergarten

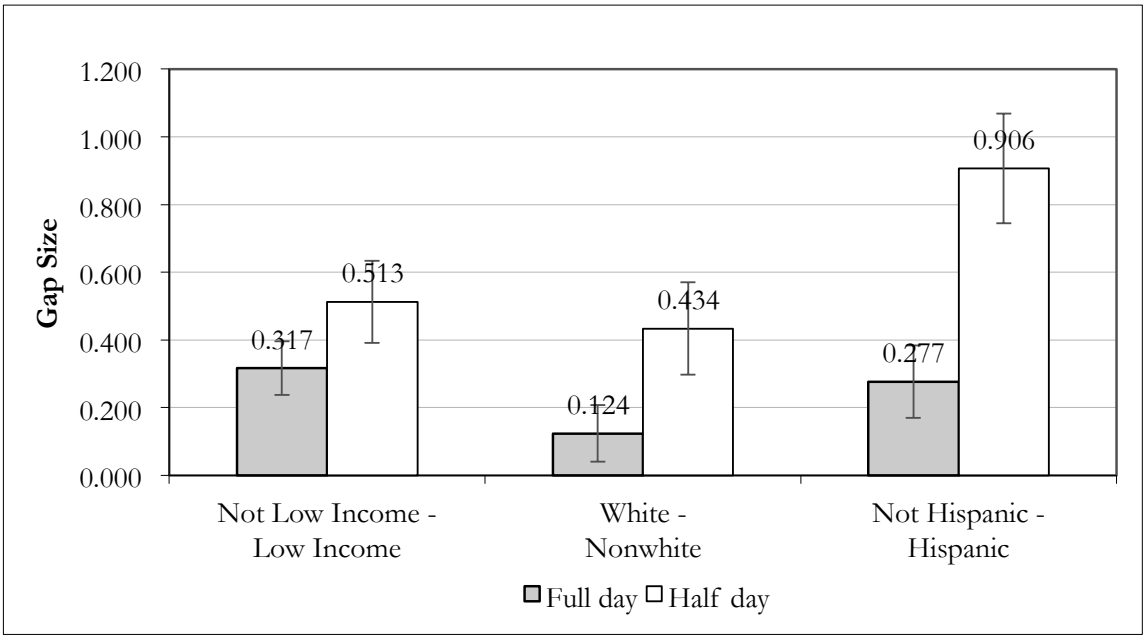
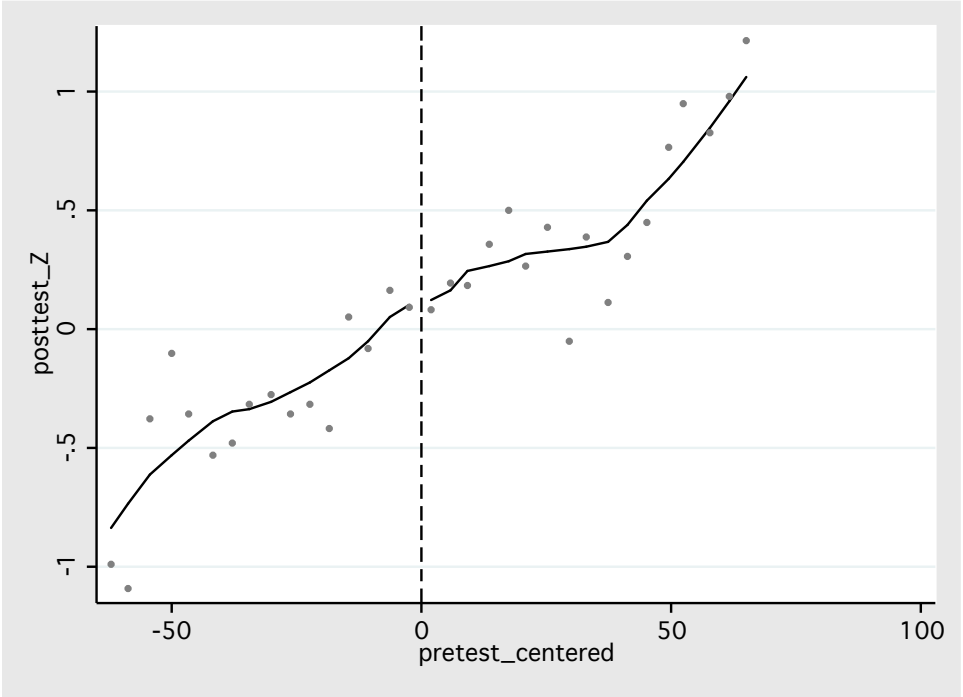


Figure 4: End-of-Kindergarten Standardized Literacy Scores and the Running Variable



Tables

Table 1: Characteristics of Lottery Study Districts

	Locale Type	% Free or Reduced-Price Lunch	% Minority	Elementary Schools	Elementary Enrollment	Kindergarten Enrollment	Kindergarten Teachers
Lottery Districts							
District 1	City, Small	61	48	7	3076	562	18
District 2	City, Midsize	59	26	1	677	119	4
District 3	Suburb, Large	42	36	4	855	116	5
District 4	Suburb, Large	20	16	11	4719	748	18
District 5	Rural, Distant	19	4	1	464	81	3

Source: National Center for Education Statistics, Common Core of Data & Indiana Department of Education, Accountability System for Academic Progress.

Table 2: Comparison of Lottery Study Schools and All Indiana Elementary Schools

	Study Schools *n*=23	All Schools *n*=1,119
School Characteristics		
Location – city (%)	34.78	28.14
Location – rural (%)	21.74	35.75
Location – suburb (%)	43.48	19.80
Location – town (%)	0	16.31
Charter or magnet (%)	4.35	3.40
Title I (%)	86.96	95.62
Title I school-wide (%)	10.00	15.33
Average Student Composition		
Free or reduced-price lunch eligibility (%)	35.06	43.02
Male (%)	50.83	51.58
White (%)	76.42	79.42
Black (%)	3.97	12.88
Hispanic (%)	14.15	6.33
Asian (%)	3.81	1.09

Source: National Center for Education Statistics, Common Core of Data & Indiana Department of Education, Accountability System for Academic Progress.
Note: "All Schools" includes all Indiana elementary schools providing kindergarten in the 2007–08 school year.

Table 3: Characteristics of Regression-Discontinuity Districts

	Locale Type	% Free or Reduced-Price Lunch	% Minority	Elementary Schools	Elementary Enrollment	Kindergarten Enrollment	Kindergarten Teachers
Regression Discontinuity Districts							
District 1	City, Small	57	28	11	3348	580	26
District 2	City, Small	47	66	9	3271	530	19

Source: National Center for Education Statistics, Common Core of Data & Indiana Department of Education, Accountability System for Academic Progress.

Table 4: Comparison of RD Study Schools and All Indiana Elementary Schools

	Study Schools *n*=18	All Schools *n*=1,119
School Characteristics		
Location – city (%)	88.89	28.14
Location – rural (%)	11.11	35.75
Location – suburb (%)	0	19.80
Location – town (%)	0	16.31
Charter or magnet (%)	0	3.40
Title I (%)	100.00	95.62
Title I school-wide (%)	50.00	15.33
Average Student Composition		
Free or reduced-price lunch eligibility (%)	58.66	43.02
Male (%)	51.66	51.58
White (%)	70.22	79.42
Black (%)	24.68	12.88
Hispanic (%)	3.35	6.33
Asian (%)	1.03	1.09

Source: National Center for Education Statistics, Common Core of Data & Indiana Department of Education, Accountability System for Academic Progress.
Note: "All Schools" includes all Indiana elementary schools providing kindergarten in the 2007–08 school year.

Table 5: First-Stage Results in RD Districts

First-stage	Participation Coefficient	F-stat
Full-day assignment full sample	0.439**	161.33
	(.035)	
Full-day assignment within 1 s.d.	0.393**	90.74
	(.041)	
Full-day assignment within 1/2 s.d.	0.274**	20.21
	(.061)	
Full-day assignment within 1/4 s.d.	0.292**	13.49
	(.079)	

Note: Heteroskedastic-consistent standard errors in parentheses.
* p < 0.05, ** p < 0.01

Table 6: Intention-to-Treat Results in Lottery Districts

	Literacy Post-test	
Full Sample	I	II
Assignment to full day	.310**	.328**
	(.100)	(.091)
Age, centered	—	.327**
		(.098)
Age, centered^2	—	-.771**
		(.203)
Female	—	.184*
		(.071)
Poverty	—	-.408**
		(.077)
Nonwhite	—	-.153
		(.100)
School fixed effects	x	x
Observations	975	975

Notes: Heteroskedastic-consistent standard errors in parentheses, clustered at the classroom level.

Because Huber-White standard errors rely on large sample asymptotic theory, the Wild cluster-bootstrap percentile-t procedure, imposing the null hypothesis, was employed (Cameron, Gelbach & Miller 2008). Inferences were unchanged.

* $p < 0.05$, ** $p < 0.01$

Table 7: Interacted Intention-to-Treat Results in Lottery Districts

	Literacy Post-test					
Interaction Effects	I	II	III	IV	V	VI
Full day	.252 (.142)	.246 (.137)	.226 (.118)	.235* (.107)	.232* (.094)	.230* (.089)
Full day*Poverty	.143 (.199)	.166 (.192)	—	—	—	—
Poverty	-.575** (.184)	-.523** (.186)	—	—	—	—
Full day*Nonwhite	—	—	.238 (.189)	.234 (.174)	—	—
Nonwhite	—	—	-.443* (.172)	-.306 (.165)	—	—
Full day*Hispanic	—	—	—	—	.460* (.176)	.483** (.178)
Hispanic	—	—	—	—	-.997** (.104)	-.883** (.113)
Controls		x		x		x
School fixed effects	x	x	x	x	x	x
Observations	975	975	975	975	975	975

Notes: Heteroskedastic-consistent standard errors in parentheses, clustered at the classroom level. Controls include centered age, centered age squared, gender, poverty status, and race/ethnicity.
* $p < 0.05$, ** $p < 0.01$

Table 8: OLS and IV Results in RD Districts - Full Sample

	Literacy Post-test			
	OLS		*IV*	
Full-day participation	-0.172 (.099)	0.544 (.272)	0.582* (.284)	0.547 (.299)
Interaction (Full day*centered pre-test score)	—	—	-0.016 (.009)	—
Literacy pre-test (*z*-scores)	—	0.528** (.070)	0.751** (.160)	0.514** (.063)
School fixed effects	x	x	x	x
Pre-test polynomials				x
Baseline characteristics				x
Observations	621	621	621	621

Note: Heteroskedastic-consistent standard errors in parentheses, clustered at the classroom level.
* p < 0.05, ** p < 0.01

Table 9: OLS and IV Results in RD Districts - Restricted Sample (1 s.d.)

	Literacy Post-test			
	OLS		*IV*	
Full-day participation	-0.116 (.072)	0.667* (.276)	0.471 (.252)	0.748* (.300)
Interaction (Full day*centered pre-test score)	—	—	-0.033* (.015)	—
Literacy pre-test (*z*-scores)	—	0.613** (.095)	0.895** (.162)	0.705** (.125)
School fixed effects	x	x	x	x
Pre-test polynomials				x
Baseline characteristics				x
Observations	482	482	482	482

Note: Heteroskedastic-consistent standard errors in parentheses, clustered at the classroom level.
* p < 0.05, ** p < 0.01

Table 10: OLS and IV Results in RD Districts - Restricted Sample ($\frac{1}{2}$ s.d.)

	Literacy Post-test			
	OLS		*IV*	
Full-day participation	0.118	0.267	-0.569	0.273
	(.097)	(.517)	(.593)	(.564)
Interaction (Full day*centered pre-test score)	—	—	-0.069	—
			(.036)	
Literacy pre-test (*z*-scores)	—	0.262	0.542*	0.346
		(.211)	(.258)	(.271)
School fixed effects	x	x	x	x
Pre-test polynomials				x
Baseline characteristics				x
Observations	233	233	233	233

Note: Heteroskedastic-consistent standard errors in parentheses, clustered at the classroom level.
* p < 0.05, ** p < 0.01

Table 11: OLS and IV Results in RD Districts - Restricted Sample ($\frac{1}{4}$ s.d.)

	Literacy Post-test			
	OLS		*IV*	
Full-day participation	0.036	-0.386	0.315	-0.270
	(.102)	(.595)	(.801)	(.668)
Interaction (Full day*centered pre-test score)	—	—	0.085	—
			(.097)	
Literacy pre-test (*z*-scores)	—	0.424	0.321	1.123
		(.420)	(.460)	(.705)
School fixed effects	x	x	x	x
Pre-test polynomials				x
Baseline characteristics				x
Observations	136	136	136	136

Note: Heteroskedastic-consistent standard errors in parentheses, clustered at the classroom level.
* p < 0.05, ** p < 0.01

Table 12: Intention-to-Treat Model Results with Peer Quality in Lottery Districts

	Literacy Post-test	
Two Treatments	I	II
Assignment to full day	.248*	.273**
	(.092)	(.086)
Class pre-test mean	.502**	.438*
	(.185)	(.181)
Age, centered	—	.328**
		(.097)
Age, centered^2	—	-.766**
		(.205)
Female	—	.186*
		(.072)
Poverty	—	-.401**
		(.077)
Nonwhite	—	-.142
		(.102)
School fixed effects	x	x
Observations	975	975

Notes: Heteroskedastic-consistent standard errors in parentheses, clustered at the classroom level.

* $p < 0.05$, ** $p < 0.01$

Table 13: Intention-to-Treat Model Results with Peer Quality Interactions in Lottery Districts

	Literacy Post-test	
Interacted Models	I	II
Assignment to full day	.311**	.353**
	(.104)	(.091)
Full day * class quality	.516	.647*
	(.287)	(.253)
Class pre-test mean	.137	-.021
	(.256)	(.213)
Age, centered	—	.333**
		(.097)
Age, centered^2	—	-.763**
		(.200)
Female	—	.192*
		(.072)
Poverty	—	-.402**
		(.076)
Nonwhite	—	-.151
		(.101)
School fixed effects	x	x
Observations	975	975

Notes: Heteroskedastic-consistent standard errors in parentheses, clustered at the classroom level.

* p < 0.05, ** p < 0.01

Table 14: Comparison of Lottery and RD Study Samples

	Lottery Study *n*=975	Regression-Discontinuity Study *n*=621
Age (years)	5.647	5.706
Free or reduced-price lunch eligibility (%)	33.95	70.85
Male (%)	48.21	53.78
White (%)	69.74	58.45
Black (%)	3.59	26.25
Hispanic (%)	15.59	4.03
Asian (%)	2.97	0.64
Native American (%)	0.62	0.48
Multiracial (%)	7.49	10.14

Table 15: Reweighted Intention-to-Treat Model Results in Lottery Districts

	Literacy Post-test	
Weighted	I	II
Assignment to full day	.318**	.330**
	(.091)	(.094)
Age, centered	—	.232*
		(.111)
Age, centered^2	—	-.490*
		(.206)
Female	—	.186*
		(.076)
Poverty	—	-.448**
		(.083)
Nonwhite	—	-.005
		(.099)
School fixed effects	x	x
Observations	975	975

Notes: Heteroskedastic-consistent standard errors in parentheses, clustered at the classroom level.

* $p < 0.05$, ** $p < 0.01$

References

[1] Angrist, J.D. & Pischke, J. (2009). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton, NJ: Princeton University Press.

[2] Cannon, J.S., Jacknowitz, A., & Painter, G. (2006). Is full better than half? Examining the longitudinal effects of full-day kindergarten attendance. Journal of Policy Analysis & Management, 25(2), 299-321.

[3] Carneiro, P.M. & Heckman, J.J. (2003). Human capital policy. IZA Discussion Paper No. 821. Available at: http://ssrn.com/abstract=434544.

[4] Cascio, E.U. (2009). Do investments in universal early education pay off? Long-term effects of introducing kindergartens into public schools. NBER Working Paper No. 14951.

[5] Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W. & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. Quarterly Journal of Economics, 126(4), 1593-1660.

[6] Coleman, J.S., Campbell, E.Q., Hobson, C.F., McPartland, J.M., Mood, A.M., York, R.L. & Weinfeld, F.D. (1966). Equality of Educational Opportunity. Washington, DC: National Center for Educational Statistics.

[7] Cunha, F. & Heckman, J.J. (2007). The technology of skill formation. American Economic Review, 97(2), 31-47.

[8] Currie, J. (2001). Early childhood education programs. Journal of Economic Perspectives, 15(2), 213-238.

[9] DeCicca, P. (2007). Does full-day kindergarten matter? Evidence from the first two years of schooling. Economics of Education Review, 26(1), 67-82.

[10] Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. American Economic Journal: Applied Economics, 1(3), 111-134.

[11] Dhuey, E. (2011). Who benefits from kindergarten? Evidence from the introduction of state subsidization. Educational Evaluation and Policy Analysis, 33(1), 3-22.

[12] Duncan, G.J., Dowsett, C.J., Claessens, A., Magnuson, K., Huston, A.C., Klebanov, P., Pagani, L.S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., & Duckworth, K. (2007). School readiness and later achievement. Developmental Psychology, 43(6), 1428-1446.

[13] Dynarski, S., Hyman, J., & Schanzenbach, D.W. (2011). Experimental Evidence on the Effect of Childhood Investments on Postsecondary Attainment and Degree Completion. NBER Working Paper No. 17533.

[14] Elicker, J. & Mathur, S. (1997). What do they do all day? Comprehensive evaluation of a full-day kindergarten. Early Childhood Research Quarterly, 12, 459-480.

[15] Fitzpatrick, M.D., Grissmer, D., & Hastedt, S. (2011). What a difference a day makes: Estimating daily learning gains in kindergarten and first grade using a natural experiment. Economics of Education Review, 30(2), 269-279.

[16] Fryer, R.G. & Levitt, S.D. (2004). Understanding the black-white test score gap in the first two years of school. The Review of Economics and Statistics, 86(2), 447-464.

[17] Fryer, R.G. & Levitt, S.D. (2006). The black-white test score gap through third grade. American Law and Economics Review, 8(2), 249-281.

[18] Fusaro, J.A. (1997). The effect of full-day kindergarten on student achievement: A meta-analysis. Child Study Journal, 27, 269-279.

[19] Garces, E., Thomas, D., & Currie, J. (2002). Longer term effects of Head Start. American Economic Review, 92(4), 999-1012.

[20] Gibbs, C., Ludwig, J. & Miller, D.L. (2011). Does Head Start Do Any Lasting Good? NBER Working Paper No. 17542.

[21] Gibbs, C.R. (2014). Experimental Evidence on Early Intervention: The Impact of Full-day Kindergarten. Batten School Working Paper No. 2014–04.

[22] Hanushek, E. (1979). Conceptual and empirical issues in the estimation of educational production functions. Journal of Human Resources, 14(3), 351-388.

[23] Hanushek, E. (1997). Assessing the effects of school resources on student performance: An update. Educational Evaluation and Policy Analysis, 19(2), 141-164.

[24] Harris, D.N. (2009). Toward policy-relevant benchmarks for interpreting effect sizes: Combining effects with costs. Educational Evaluation and Policy Analysis, 31(1), 3-29.

[25] Heckman, J.J. (2000). Policies to foster human capital. Research in Economics, 54(1), 3-56.

[26] Heckman, J.J. & Lochner, L. (2000). Rethinking myths about education and training: Understanding the sources of skill formation in a modern economy. In Securing the Future: Investing in Children from Birth to College, S. Danziger & J. Waldfogel (Eds.). New York: Russell Sage.

[27] Heckman, J.J. & Masterov, D.V. (2007). The productivity argument for investing in young children. Review of Agricultural Economics, 29(3), 446-493.

[28] Hedges, L., Laine, R. & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. Educational Researcher, 23(3), 5-14.

[29] Imbens, G.W. & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. Journal of Econometrics, 142, 615-635.

[30] Imbens, G.W. & Rubin, D.B. (1997). Estimating outcome distributions for compliers in instrumental variables models. Review of Economic Studies, 64, 555-574.

[31] Indiana General Assembly. (2007). Public Law 234–2007. Indianapolis, IN: Author. Retrieved August 12, 2009 from http://www.in.gov/legislative/pdf/acts_2007.pdf.

[32] Karweit, N. (1989). Effective kindergarten programs and practices for students at risk. In Effective Programs for Students at Risk, R. E. Slavin, N. L. Karweit, & N. A. Madden (Eds.). Boston: Allyn and Bacon.

[33] Kauerz, K. (2005). Full-day kindergarten: A study of state policies in the United States. Denver, CO: Education Commission of the States.

[34] Knudsen, E.I., Heckman, J.J., Cameron, J. & Shonkoff, J.P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. Proceedings of the National Academy of Sciences, 103(27), 10155-10162.

[35] Krueger, A. B. (1999). Experimental estimates of education production functions. Quarterly Journal of Economics, 114(2), 497-532.

[36] Krueger, A.B. (2003). Economic considerations and class size. Economic Journal, 113, F34-F63.

[37] Lee, D.S. & Lemieux, T. (2010). Regression discontinuity designs in economics. Journal of Economic Literature, 48, 281-355.

[38] Lee, V.E. & Burkam, D.T. (2002). Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School. Washington, DC: Economic Policy Institute.

[39] Lee, V.E., Burkam, D.T., Ready, D., Honigman, J., & Meisels, S.J. (2006). Full-day versus half-day kindergarten: In which program do children learn more? American Journal of Education, 112(2), 163-208.

[40] Lovell, R., Kochanek, J., Mathers, C. & Burke, M. (2009). Snapshots of Indiana's full-day kindergarten programs before and after the state's funding increase for the program (REL Technical Brief, REL 2009–No. 013). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest.

[41] Morgan, S.L. & Winship, C. (2007). Counterfactuals and Causal Inference: Methods and Principles for Social Research. New York: Cambridge University Press.

[42] Plucker, J.A., Eaton, J., Rapp, K.E., Lim, W., Nowak, J., Hansen, J., & Bartleson, A. (2004). The effects of full day versus half day kindergarten: Review and analysis of national and Indiana data. Mimeo. Center for Education Evaluation and Policy, Indiana University.

[43] Princiotta, D., Flanagan, K.D., & Germino Hausken, E. (2006). Fifth grade: Findings from the fifth- grade follow-up of the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99. (NCES 2006-038) U.S. Department of Education. Washington, DC: National Center for Education Statistics.

[44] Murnane, R.J., Willett, J.B., Bub, K.L. & McCartney, K. (2006). Understanding trends in the black-white achievement gaps during the first years of school. Brookings-Wharton Papers on Urban Affairs, 97-135.

[45] Schweinhart, L.J., Montie, J., Xiang, Z., Barnett, W.S., Belfield, C.R. & Nores, M. (2005). Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40. Ypsilanti, MI: High/Scope Press.

[46] Shadish, W.R., Cook, T.D. & Campbell, D.T. (2002). Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin Company.

[47] Shonkoff, J.P. & Phillips, D.A., editors. (2000). From Neurons to Neighborhoods: The Science of Early Childhood Development. Washington, DC: National Academy Press.

[48] The White House, Office of the Press Secretary. (2013). Fact sheet President Obama's plan for early education for all Americans [Press release]. Retrieved from http://www.whitehouse.gov/the-press-office/2013/02/13/fact-sheet-president-obama-s-plan-early-education-all-americans.

[49] U.S. Department of Commerce, Bureau of the Census, Current Population Survey (CPS), October supplement, selected years 1977–2001.

[50] U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), Survey of local government finances, school systems (F-33), 2007-08 (FY 2008).

[51] Votruba-Drzal, E., Li-Grining, C.P. & Maldonado-Carreño, C. (2008). A developmental perspective on full- versus part-day kindergarten and children's academic trajectories through fifth grade. Child Development, 79(4), 957-978.

[52] Walston, J. & West, J. (2004). Full-day and half-day kindergarten in the United States: Findings from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99. (NCES 2004-078) U.S. Department of Education. Washington, DC: National Center for Education Statistics.

[53] Yan, W. & Lin, Q. (2005). Effects of class size and length of day on kindergartners' academic achievement: Findings from the Early Childhood Longitudinal Study. Early Education & Development, 16, 49-68.

[54] Zvoch, K., Reynolds, R.E., & Parker, R.P. (2008). Full-day kindergarten and student literacy growth: Does a lengthened school day make a difference? Early Childhood Research Quarterly, 23(1), 94-107.