WORKING PAPER · NO. 2020-187

The Social Side of Early Human Capital Formation: Using a Field Experiment to Estimate the Causal Impact of Neighborhoods

John A. List, Fatemeh Momeni, and Yves Zenou

DECEMBER 2020

Becker Friedman Institute
FOR ECONOMICS AT UCHICAGO

5757 S. University Ave.
Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

THE SOCIAL SIDE OF EARLY HUMAN CAPITAL FORMATION:
USING A FIELD EXPERIMENT TO ESTIMATE THE CAUSAL IMPACT OF NEIGHBORHOODS

John A. List
Fatemeh Momeni
Yves Zenou

This study was previously titled: Are Estimates of Early Education Programs Too Pessimistic? Evidence from a Large-Scale Field Experiment that Causally Measures Neighbor Effects. We thank Alec Brandon, Leonardo Bursztyn, Raj Chetty, Steven Durlauf, Nathaniel Hendren, Justin Holz, Michael Kremer, Thibaut Lamadon, Costas Meghir, Magne Mogstad, Julie Pernaudet, Stephen Raudenbush, Matthias Rodemeier, Juanna Schrøter Joensen, and Daniel Tannenbaum for valuable comments. We received helpful feedback from seminar participants at the University of Chicago, University of Wisconsin Milwaukee, Depaul University, Purdue University, and Monash University. We thank Clark Halliday, Uditi Karna, Alexandr Lenk, Ariel Listo, and Lina Ramirez for excellent research assistance.

The Social Side of Early Human Capital Formation: Using a Field Experiment to Estimate the Causal Impact of Neighborhoods
John A. List, Fatemeh Momeni, and Yves Zenou
December 2020
JEL No. C93,I21,I24,I26,I28,R1

ABSTRACT

The behavioral revolution within economics has been largely driven by psychological insights, with the sister sciences playing a lesser role. This study leverages insights from sociology to explore the role of neighborhoods on human capital formation at an early age. We do so by estimating the spillover effects from a large-scale early childhood intervention on the educational attainment of over 2,000 disadvantaged children in the United States. We document large spillover effects on both treatment and control children who live near treated children. Interestingly, the spillover effects are localized, decreasing with the spatial distance to treated neighbors. Perhaps our most novel insight is the underlying mechanisms at work: the spillover effect on non-cognitive scores operate through the child's social network while parental investment is an important channel through which cognitive spillover effects operate. Overall, our results reveal the importance of public programs and neighborhoods on human capital formation at an early age, highlighting that human capital accumulation is fundamentally a social activity.

John A. List
Department of Economics
University of Chicago
1126 East 59th
Chicago, IL 60637
and NBER
jlist@uchicago.edu

Fatemeh Momeni
Crime and Education Labs
University of Chicago
33 N LaSalle St.
Chicago, IL 60602
fmomeni@uchicago.edu

Yves Zenou
Department of Economics
Monash University
Caulfield VIC 3145
Australia
yves.zenou@monash.edu

"... I will emphasize again and again: that human capital accumulation is a social activity, involving groups of people in a way that has no counterpart in the accumulation of physical capital..." Lucas (1988)

1 Introduction

Human capital theory can be traced to Mincer (1958), who created the framework to examine the nature and causes of inequality in personal incomes. Empirically, human capital is typically operationalized as being measured in years of schooling completed and is commonly tied to labor market outcomes. A key branch of this work explores individual's educational investment decisions and how those choices map into higher future incomes. A related line of work, estimating education production functions, complements the human capital literature by investigating the *determinants* of human capital (Heckman, 2008; Hanushek, 2020; Cotton et al., 2020). In this literature, standardized test scores, or some other proxy for cognitive and executive function skills, are measured and subsequently modeled as individual-specific skills potentially valued by employers. In this manner, the received education production estimates reflect the long-run economic impacts of educational inputs, effectively linking the two literatures (Hanushek, 2020).

To date, this line of economics research and related work in the contemporary psychology of education literature are dominated by an empirical and theoretical focus on the individual (Schunk, 2020; Cotton et al., 2020). This individual-centric approach has served the literatures well, as developing knowledge on issues as varied as the foundations of learning to the causes and consequences of human capital accumulation and skill formation, serve to deepen our understanding and clarify optimal policy solutions. Such insights also have frequently made their way into public policy circles, either through advanced reforms or pedagogical changes in the classroom.

Yet, the Lucas' quote in the epigraph summons a distinctly different line of inquiry, one which includes the wisdom of Sociology to deepen our understanding of human capital accumulation. As Jonassen (2004) notes, Sociology is concerned with many things, but primarily it relates to explaining social phenomena, and this cannot be done if we examine individuals alone. Rather, we must also scrutinize how people interact in group settings, and how those interactions shape individuals and their choices, including those that augment human capital.

With this contribution in mind, our backdrop is that between 2010 and 2014, a series of early childhood programs were delivered to low-income families with young children in the Chicago Heights Early Childhood Center (CHECC; see Fryer et al., 2015; 2018). CHECC was located in Chicago Heights, IL, a neighborhood on Chicago's South Side with characteristics similar to many other low-performing urban school districts. The goals of the intervention were to examine how investing in cognitive and non-cognitive skills of low-income children aged 3 to 4 affects their short- and long-term outcomes, and to evaluate the effectiveness of investing directly in the child's education versus indirectly through the parents. To that end, families of over 2,000 disadvantaged children were randomized into (i) an incentivized parent-education program (Parent Academy), (ii) a high-quality preschool program (Pre-K), or (iii) a control group. The children's cognitive and non-cognitive skills were assessed on a regular basis, starting before the randomization and continuing into the middle and end of the programs. Follow-up assessments were also conducted on a yearly basis.

Making use of these data, we consider insights from Sociology to focus on explorations of group interactions. A useful starting point is Coleman (1988), who introduces social capital to parallel economic concepts (physical capital and human capital) to embody relations among people. Once in place, the effect of social capital is argued to have great import in the formation of human capital, especially in the development of children. The Sociology literature has taken Coleman's work in several directions (Bourdieu, 1985; Putnam, 1993; Schuller, 2000), with critical factors of early child human capital development relating to both parental relationships and the composition of children's peer play groups (Sheldon, 2002). Importantly, the Sociology literature teaches us that detailing group composition at various ages of children is important since there are key age-level interactions that affect human capital development of children (Cochran and Brassard, 1979; Corsaro, 2005).

To explore the interplay between social interactions and human capital formation, we follow two distinct steps. First, we provide causal evidence of the impact of neighborhood on educational outcomes in early childhood. Instead of following the standard approach in economics, which uses residential movers to identify neighborhood effects (see citations below), we exploit a unique form of exogeneity induced by the CHECC intervention: the experimental variation in the spatial exposure to treated families (within and between individuals) caused by the delivery of programs across multiple years. By doing so, we are able to isolate the role of neighbors on individual outcomes and examine how the exogenous changes in treated neighbors' quality affect a child's outcomes. Our second step is to follow the Sociology literature to explore underlying mechanisms at work, both from child to child as well as from parent to parent.

In the first step, we document large and significant spillover effects on both cognitive and non-cognitive skills. We find the non-cognitive spillover effects are about two times larger than the cognitive spillover effects. Our estimates suggest that, on average, *each additional treated neighbor* residing within a three-kilometer radius of a child's home increases that child's cognitive score by 0.0033 to 0.0042 standard deviations (σ), whereas it increases her non-cognitive score by 0.0069σ to 0.0070σ. Given that an average child in our sample has 178 treated neighbors residing within a three-kilometer radius of her home—and making a (strong) assumption of linearity—we infer that, on average, a child gains between 0.6σ to 0.7σ in cognitive test scores and about 1.2σ in non-cognitive test scores in spillover effects from her treated neighbors. As discussed more fully below, the spillover effect is a key component of the total intervention effect. Interestingly, we find that the spillover effects are localized and fall rapidly as the distance to a treated neighbor increases.

Fryer et al. (2015) also report interesting racial and gender heterogeneity in their treatment effects. For example, through comparing outcomes between treatment and control children, they find the Parent Academy significantly increases test scores for Hispanics and Whites, but does not improve outcomes of Black children. These findings prompted us to examine whether such heterogeneities also exist in our estimated spillovers. We find that non-cognitive spillover effects are significantly larger for Blacks than Hispanics. According to our fixed-effects estimates, an additional treated neighbor within a three-kilometer radius increases the non-cognitive test score of a Black child by 0.0100σ, whereas it increases the non-cognitive score of a Hispanic child by only 0.0045σ. We find no significant racial differences in cognitive spillover effects. Focusing on gender, our estimates suggest boys tend to benefit more than girls from cognitive and non-cognitive spillovers, although these gender differences are not significant at the conventional levels. This observation is in the spirit of previous empirical evidence on neighborhood effects, which tend to be larger for boys (Entwisle et al., 1994; Halpern-Felsher et al., 1997; Leventhal and Brooks-Gunn, 2000; Katz et al., 2001; Chetty and Hendren, 2018b).

Turning to Step 2, we recognize that the program effects from CHECC can spill over through two main channels. The first channel is the direct social interactions between children who were randomized during the intervention. Importantly, consonant with the Sociology literature, our analysis includes observations from early childhood (3 to 4 years of age, when peer influence at the neighborhood level starts) to middle childhood (8 to 9 years of age, when social interactions within neighborhoods increase dramatically as children enter school). Therefore, direct exposure to treated children who live in the same neighborhood is a likely mechanism that can generate spatial spillover effects.[1] The second channel is parental interactions. While Sociology presents a useful guide, observational studies in the other sciences have also shown the import of this channel. For example, Psychologists have found that neighborhoods can influence parental behavior and child-rearing practices (Leventhal and Brooks-Gunn, 2000), which play critical roles in early development (Cunha and Heckman, 2007; Waldfogel and Washbrook, 2011; Kautz et al., 2014; Fryer et al., 2015; Kalil, 2015). Because CHECC also offered education programs to parents, treatment effects can spill over through information and preference externalities, generated by parental social interactions.

To shed light on the mechanisms through which spillover effects operate, we start by comparing the effects from neighbors who were assigned to the parental-education programs with the effects from neighbors who were assigned to the preschool programs. Because, unlike in the Pre-K treatments,

[1] See Epple and Romano (2011) and Sacerdote (2011) for recent reviews of the literature on peer effects in Economics.

the focus of Parent Academies was on educating parents rather than children, if spillover effects are driven by interactions between parents, we might expect Parent Academy neighbors to generate larger effects than Pre-K neighbors.[2] Alternatively, larger spillovers from Pre-K neighbors than from Parent Academy neighbors could imply the peer-influence channel plays an important role in generating the effects. Our estimates suggest non-cognitive spillovers are more likely to operate through preschool neighbors. According to our estimates, whereas an additional Parent Academy neighbor within three kilometers of a child's home induces a 0.0017σ to 0.0045σ increase in her non-cognitive score, an additional Pre-K neighbor living within the same distance increases her non-cognitive score by 0.0099σ to 0.0108σ. This finding suggests non-cognitive spillover effects are more likely to operate through children's rather than parents' social networks. We do not find any significant differences in *cognitive* spillover effects from Parent Academy and Pre-K neighbors.

Given our evidence suggesting peer influence at the neighborhood level is a key mechanism in generating non-cognitive spillover effects, we hypothesize that the racial differences in non-cognitive spillovers might be at least partially driven by differences in social interactions within neighborhoods. We explore this idea using data from the National Longitudinal Study of Adolescent Health Survey. Our analysis confirms that African American adolescents are significantly more likely than Hispanics to (i) know most people in their neighborhoods, (ii) stop on the street and talk to someone from the neighborhood, and (iii) use recreation facilities in the neighborhood. Although these results cannot be interpreted as causal evidence, they are consistent with our previous finding that social interactions with peers within neighborhoods is a key channel in generating non-cognitive spillover effects.

Finally, our evidence suggests *cognitive* spillover effects are likely to operate—at least partially—through influencing the parents' decision to enroll their child in a (non-CHECC) preschool program. Using survey data, we show that families with more treated neighbors are significantly more likely to enroll their child in a preschool program (other than the ones offered at CHECC). Our evidence also suggests children whose parents reported enrolling them in an alternative preschool program perform significantly better in cognitive assessments. Therefore, we conclude that influencing parental investment decisions—as measured by the choice to enroll one's child in a preschool program—is a channel through which spillover effects on cognitive test scores operate.

We conclude our analysis by measuring the total impact of the intervention on children's cognitive and non-cognitive performance, accounting for the spillover effects. Our estimates suggest that, on average, the intervention increased a treatment child's cognitive (non-cognitive) test score by 0.82σ (1.32σ). Spillover effects make up a large portion of this total impact: whereas the average direct effect of the intervention on a *treatment* child's cognitive (non-cognitive) score is 0.11σ

[2]This intuition does not rule out possible spillover effects from Pre-K neighbors that are generated through parental interactions. After all, parents of children who received the Pre-K treatments might also be impacted through the Pre-K programs. This intuition merely assumes Parent Academies affect parents *more* than Pre-K treatments do.

(0.05σ), the corresponding indirect effect is 0.71σ (1.27σ). Control children also gain considerably as a result of the intervention: on average, the intervention increased a control child's cognitive (non-cognitive) test score by 0.75σ (1.25σ). If we were to disregard the spillover effects on the control group and had simply based our estimates of the total impact on the outcome differences between the treatment and control children, we would have severely understated the total impact. Specifically, this approach would have indicated that the intervention only improved the cognitive (non-cognitive) test scores of a treatment child by 0.06σ (0.07σ). Ignoring spillover effects would have also led us to underestimate the effects for African American children. Accounting for spillover effects enables us to document a significant and large impact on non-cognitive performance that is significantly larger for African Americans than Hispanics.

We view our results speaking to three distinct strands of research. First, we speak to the various literatures that study the role of *neighborhoods* in shaping children's short- and long-term human capital outcomes. The empirical evidence on how neighborhoods affect children comes mainly from observational studies that document correlations between neighborhood characteristics and children's outcomes, as well as studies that use experimental and quasi-experimental data to disentangle the causal effects of neighborhood from selection effects.[3] We contribute to this literature in two important ways.

Our first contribution to this literature is to provide causal evidence on neighborhood effects by exploiting a unique form of exogeneity, which was induced by our field experiment. The existing experimental and quasi-experimental evidence on how neighborhoods shape children's outcomes identifies neighborhood effects using data from *residential movers* (e.g., Katz et al., 2001; Edin et al., 2003; Kling et al., 2005; Åslund et al., 2010; Damm and Dustmann, 2014; Chetty et al., 2016; Chyn, 2018; Chetty and Hendren, 2018a and 2018b). The identification of neighborhood effects in this literature relies on instruments such as randomly assigned housing vouchers, quasi-random assignment of immigrants to different neighborhoods, or public housing demolitions as sources of exogenous changes in neighborhood quality. We take a different approach in that our identification strategy leverages a field experiment that provides both within and between individual variation in the spatial exposure to treated families.

Our second contribution to this literature is to provide insights on the role of neighbors in generating neighborhood effects and the *mechanism* underlying these effects. Neighborhoods have multiple attributes, which can each influence a child's outcomes, such as school quality, crime rate, neighbors, and so on. Unlike previous estimates on neighborhood effects, we are able to isolate and estimate the effect of *neighbors' quality* as one of the many channels through which neighborhoods can influence children's development. Specifically, our estimates suggest social interactions with other

[3] See Leventhal and Brooks-Gunn (2000), Durlauf (2004), Ioannides and Topa (2010), Ioannides (2011), Topa and Zenou (2015), Minh et al. (2017) and Graham (2018) for reviews of neighborhood effects on children.

children in the neighborhood play an important role in the development of children's non-cognitive skills and that parental interactions influence a complementary aspect of child development.

The second strand of literature our study contributes to is the growing body of work that measures *spillover effects* from programs and policy changes, designed to improve behaviors and outcomes in various domains such as the labor market (Ferracci et al. 2014; Crépon et al., 2013; Lalive et al., 2015; Muralidharan et al., 2017; Gautier et al., 2018), health (Miguel and Kremer 2004; Janssen, 2011; Avitabile, 2012), compliance behavior (Rincke and Traxler, 2011; Boning et al., 2018), voting behavior (Sinclair et al., 2012; Gine and Mansuri, 2018), retirement saving decisions (Duflo and Saez, 2003), and consumption (Angelucci and De Giorgi, 2009). We contribute to this literature by providing the first evidence on spillover effects from a large-scale *early education* intervention, shedding light on mechanisms, and estimating the total program impact when accounting for spillover effects.

Finally, our results provide important insights for academics interested in modeling the formation of early human capital, from economists to psychologists to sociologists. Within economics, for example, a growing body of literature develops dynamic models of skill formation to explore the role of various inputs in the production of cognitive and non-cognitive skills. Through structurally estimating such models, this literature has found inputs such as schools, parental ability, home environment, and parental investments to be important determinants in the formation of future skills (e.g., Todd and Wolpin, 2007; Cunha and Heckman, 2007; Cunha et al., 2010; Attanasio et al., 2015; Doepke and Zilibotti, 2017, 2019; Agostinelli et al., 2020; Attanasio et al., 2020; Boucher et al., 2020; Cotton et al., 2020). We complement this literature by providing empirical evidence for the role of neighbors' influence at young ages. Our estimates suggest neighbors' quality plays an important role in producing cognitive and non-cognitive skills.

The remainder of the paper is structured as follows. Section 2 summarizes key features of our intervention, randomization, and assessments. Section 3 describes our data and presents our estimation strategy. We present our main findings in section 4, where we report our estimates of spillover effects on cognitive and non-cognitive test scores from a fixed-effects model, and explore heterogeneities by race and gender. Section 5 presents our estimates of the spillover effects from a lagged dependent variable (LDV) specification and discusses the robustness of our findings to using this alternative identification strategy. We discuss the mechanisms in section 6. In section 7, we estimate the total impacts of CHECC, break down these estimates into direct and indirect effects, and discuss how ignoring indirect effects would bias our estimates. We discuss policy implications and conclude in section 8.

2 Program Details

2.1 Overview of treatments

Between 2010 and 2014, a series of early childhood interventions were delivered to low-income families with young children in Chicago Heights Early Childhood Center (CHECC). The center was located in Chicago Heights, IL, which is a South Side, Chicago, neighborhood with characteristics similar to many other low-performing urban school districts. According to the 2010 Census, black and Hispanic minorities constituted about 80% of the population of Chicago Heights; its per-capita income was $17,546 per year; and 90% of students attending the Chicago Heights School District were receiving free or reduced-price lunches.

The main goals of this large-scale intervention were (i) to examine how investing in cognitive and non-cognitive skills of low-income children 3 to 4 years of age affects their long-term outcomes, and (ii) to evaluate the effectiveness of investing directly in children's education versus indirectly through their parents. To that end, families of over 2,000 children were randomized into either one of four preschool programs (henceforth "Pre-K") or one of the two parental-education programs (henceforth "Parent Academy") or a control group.

The Parent Academy was designed to teach parents to help their child with cognitive skills, such as counting and spelling, as well as non-cognitive skills, such as working memory and self-control. The curriculum for the Parent Academy was adapted from two effective preschool curricula: *Tools of the Mind*, which focuses on fostering non-cognitive skills, and *Literacy Express*, which focuses on improving cognitive skills.[4] The curriculum was delivered to parents in eighteen, 90-minute sessions, which were held every two weeks over a nine-month period. Parent Academy families had the opportunity to earn up to $7,000 per year and could participate until their child entered kindergarten. Earnings were based on parents' attendance, their performance on homework, and their child's performance on the interim and end-of-year assessments. The two Parent Academy treatments differed only in how they administered incentives. Payments made to families in the "Cash" treatment were made via cash/direct deposits, whereas payments made to families in the "College" treatment were deposited into an account that could only be accessed once the child was enrolled in a full-time post-secondary institution.

Besides the Parent Academy, CHECC delivered four preschool programs in which children were treated *directly*. We refer to these programs as Pre-K treatments. These four treatments differed in their curricula, as well as the duration and intensity of delivery. "Tools," "Literacy," and "Preschool Plus" were nine-month full-day programs delivered during the school year, whereas "Kinderprep" was a two-month half-day program delivered during the summer before a child en-

[4] See Fryer et al. (2015) for more information on curriculum selection.

tered kindergarten.[5] The curriculum for "Tools" was *Tools of the Mind*, which focuses on improving non-cognitive skills, whereas "Literacy" was based on *Literacy Express*,[6] which focuses on fostering cognitive skills.[7] A new curriculum called "Cog-X" was developed for "Preschool Plus" and "Kinderprep", which emphasized both cognitive and non-cognitive skills.[8]

2.2 Randomization

Between 2010 and 2013, 2,185 children from low-income families in South Side, Chicago were recruited and randomized into either one of the six treatments or the control group.[9] The randomization took place once per year, at the beginning of each academic year.[10] Some children were randomized during more than one year, mainly to encourage families who were initially placed in the control group to stay engaged with CHECC for assessments, by offering them a chance to participate in future years.[11] The yearly randomization schedule created four cohorts of children we refer to by their year of randomization.[12] Table 1 summarizes the randomization schedule for each year of the program.

[5] Preschool Plus and Kinderprep also offered a parental component that was much less extensive than the Parent Academies, both in terms of education time and incentives. Parent Academy parents could earn up to $7,000 based on their attendance, their performance on homework, and their child's performance on assessments, whereas Preschool Plus and Kinderprep parents could only earn up to $900 and $200, based merely on their attendance to parental workshops. Preschool Plus and Kinderprep treatments also offered fewer instruction time to parents. Whereas Parent Academy parents could spend 27 hours in parental workshop, Preschool Plus and Kinderprep parents were offered a maximum of 21 and 6 hours of parental education, respectively. The intensity of the preschool component of Preschool Plus was similar to that of "Tools" and "Literacy Express" in terms of instruction time.

[6] For more information on Literacy Express see http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=288.

[7] For more information on Tools of the Mind see http://toolsofthemind.org.

[8] Fryer et al. (2020) evaluated "CogX" under the assumption that the programs did not affect the outcomes of the control group. Through comparing the performance of treatment and control children, the authors found "Cog-X" treatments significantly improved cognitive scores (by about one quarter of a standard deviation), but failed to find any significant effects on non-cognitive scores. For more information on Pre-K programs, see Fryer et al. (2020).

[9] See Appendix A for maps of residential addresses.

[10] The exceptions were years three and four of the intervention during which randomization took place twice per year: In the first randomizations, children were randomized into either the nine-month preschool program, the summer Kinderprep program, or the control group; and in the second, a smaller group of families were recruited and randomized into either the summer kindergarten preparation program or the control group. Table 1 combines the two randomizations.

[11] As a result, some children who were in the control group in an earlier year were randomized into a treatment group in later randomizations. In a few cases, a child who was randomized into a treatment group in an earlier year was assigned to a different (or the same) treatment in later randomizations. Overall, 1,675 children were randomized only once, 509 were randomized twice, and one child was randomized in three years.

[12] Those children who were randomized in multiple years also appear in multiple cohorts.

Table 1: Randomization by Year

	Control	Parent Academies	9-Month Pre-K	Kindergarten Preparation
Cohort-1 (2010)	242	153	172	0
Cohort-2 (2011)	443	216	166	0
Cohort-3 (2012)	422	0	196	107
Cohort-4 (2013)	376	0	104	99
Unique child	1270	317	539	206

Notes: The number of children randomized into each treatment group in each year of the intervention is reported. The bottom row presents the number of unique children in each group, over the course of four years.

2.3 Assessments

Our key outcome measures are children's performances in cognitive and non-cognitive assessments, which were used to evaluate the programs. These assessments consist of a *pre-assessment* administered to all incoming students prior to randomization, a *mid-assessment* between January and February, a *post-assessment*, which occurred in May, immediately after the school year ended, and a *summer assessment* at the end of the summer. Besides the assessments that took place during the program year, graduated children were also assessed annually every April, starting the year after they finished the program. These assessments are referred to as *age-out assessments*. Appendix B presents the assessment schedule for all four cohorts.

Assessments included both cognitive and non-cognitive components and were administered by a team of trained assessors. The cognitive component used a series of nationally normed tests, measuring general intellectual ability and specific cognitive abilities such as receptive vocabulary, verbal ability, oral language, and academic achievements. The non-cognitive component included a combination of subtests measuring executive functions such as working memory, inhibitory control, and attention shifting, as well as a questionnaire completed by assessors, which measured self-regulation in emotional, attentional, and behavioral domains.

3 Data and the Econometric Model

Before describing our empirical approach at a detailed level, we find it potentially useful to provide a roadmap for our exploration. Overall, we identify the spillover effects from CHECC by exploiting certain unique features of our data. First, conditional on the total number of neighbors who signed up to participate, the number of neighbors who were subsequently assigned to treatment is determined exogenously through the randomization process. We leverage this experimental variation in spatial exposure to treatments across children to estimate spillover effects. Second, our main identification strategy also exploits the panel nature of our data and the within-individual variation in exposure to treated neighbors induced by delivery of programs over multiple years. Specifically, by including individual-specific fixed effects, our estimates control for any time-invariant individual, family, and neighborhood unobserved characteristics that might be correlated with spatial exposure to treatments. We also estimate the effects under a second model that relaxes the assumption of time-invariant omitted variables by controlling for the lagged dependent variables (LDV) and dispensing with the fixed effects (Section 5). Whereas our main identification strategy uses *within-individual* variations to estimate the spillover effects, the LDV specification estimates the effects by exploiting both *within-individual* and *between-individual* variations in spatial exposure to treatments. Our findings, presented below, are robust to using this alternative specification.

3.1 Data

3.1.1 Construction of outcome variables

Our outcome measures are indices generated from standardized test scores on cognitive and non-cognitive assessments.[13] The cognitive assessment included the Peabody Picture Vocabulary Test (PPVT), which assesses verbal ability and receptive vocabulary (Dunn et al., 1965), and four subtests of the Woodcock Johnson III Test of Achievement (WJ): (*i*) WJ-Letter and Word Identification (WJL), which measures the ability to identify letters and words; (*ii*) WJ-Spelling (WJS), which measures the ability to correctly write orally presented words; (*iii*) WJ-Applied Problems (WJA), which measures the ability to analyze and solve math problems; and (*iv*) WJ-Quantitative Concepts (WJQ), which assesses the knowledge of mathematical concepts, symbols, and vocabulary (Woodcock, McGrew and Mather, 2001).

The non-cognitive component included the Blair and Willoughby Executive Function test (Willoughby, Wirth and Blair, 2012), which is composed of three subtests assessing attention (Spatial Conflict), working memory (Operation Span), and attention shifting (Same Game) and the Preschool Self-Regulation Assessment (PSRA), which is designed to assess self-regulation in emotional, attentional, and behavioral domains (Smith-Donald et al., 2007).[14]

[13] These indices were constructed by Fryer et al. (2015, 2020) for the original evaluations of the programs.

[14] Because Blair and Willoughby tests are designed for preschool, a new test was added for assessments that were administered to older children (age-out assessments). For children in kindergarten or older, the Same Game test of Blair and Willoughby was replaced with a variant of Wisconsin Card Sort game, which measures attention shifting for children of that age. For more information, see www.parinc.com/Products?pkey=478.

A cognitive index was made up of averaged percentile scores on each cognitive subtest, and a non-cognitive index was made up of average percent-correct scores on each non-cognitive subtest. The two indexes were then standardized by the type of assessment (pre-assessment, mid-assessment, etc.), including the entire study population (treated and control) who took that assessment, to obtain a zero mean and standard deviation of one.

To explore the spatial spillovers on both treatment and control children, we construct three samples: a pooled sample, including observations from both treatment and control groups; a control sample, including data from control children; and a treatment sample, including observations from treated children. Our treatment sample pools observations from children who were randomized into any of the programs. We include observations from the baseline to the fourth age-out assessment. Our final control, treatment, and pooled samples include 2,442, 3,074, and 5,208 observations, respectively.[15] Appendix C presents the details regarding how we construct these three samples.

Table 2 provides summary statistics on the baseline demographic variables for our pooled sample. Note the majority (90%) of the children are either African American or Hispanic, and 53% live in families with an annual household income under $35,000.

3.1.2 Addresses and neighbor counts

To estimate the spatial spillover effects from the intervention, we follow the literature (see, e.g., Miguel and Kremer, 2004 and subsequent work) and calculate the number of treated neighbors of a child at a given time and use it as a measure of spatial *exposure* to treatments. To do so, we start by calculating commuting distances between the home locations of all pairs of children who were randomized during the intervention.[16] Commuting distances are calculated by considering the street network structure and its restrictions (e.g., one-way roads, U-turns, etc.) and finding the closest driving distance between each pair. The average travel distance between a pair of children in our sample is 8.52 kilometers (std. dev.= 8.07), and 99.8% of the sample resides within 60 kilometers of each other. Figure 1 presents a histogram of travel distances between home locations of all children who were randomized during the intervention.

We define a pair as *neighbors* if the commuting distance between the two is less than "r" kilometers, and we call "r" the *neighborhood radius*. We conduct our analysis for various values of neighborhood radii. We then calculate the number of treated ($N_{i,t|r}^{treated}$) and control ($N_{i,t|r}^{control}$) neighbors of each child i at the time of her assessment t, and define the total number of CHECC neighbors of i as

[15] Note that the number of observations in the pooled sample is smaller than the sum of the number of observations in our control and treatment samples. The reason is that in a few cases, when a child was first randomized into the control group and was placed into a treatment group in later randomizations, the pooled sample only includes observations that took place after the child was randomized into treatments. See Appendix C for more information.

[16] Distances were calculated using the ArcGIS OD Cost Matrix Analysis tool.

Table 2: Baseline Summary Statistics for the Pooled Sample

Variable	Share/Mean	Variable	Share/Mean
Gender		Mother's Education	
Male	.51	Less than high school	.08
Race		Some high school but no diploma	.12
Black	.40	High school diploma	.13
Hispanic	.50	Some college but no degree	.17
White	.09	College degree	.18
Other Race	.01	Other	.06
Missing Race	.01	Missing Mother's Education	.25
HH Income and Unemployment Benefits		Father's Education	
below 35K	.53	Less than high school	.09
36K-75K	.14	Some high school but no diploma	.1
75K+	.06	High school diploma	.13
Missing Income	.26	Some college but no degree	.12
Receives Unemployment Benefit	.09	College degree	.08
Missing Unemployment Benefit	.31	Other	.06
		Missing Father's Education	.44
Baseline Age (months)	45.32 (6.91)		

Notes: Summary statistics for baseline demographic variables are presented. For education levels, *Some high school but not diploma* includes parents with a GED or high school attendance without a diploma, *College degree* includes associate's, bachelor's and master's degrees, *Less than high school* includes an education level below 9th grade or no formal schooling, and *Other* includes vocational/technical or other unclassified programs. Standard deviations are reported in parentheses.

$N_{i,t|r}^{total} = N_{i,t|r}^{treated} + N_{i,t|r}^{control}$. Note that as more children are randomized into treatment and control groups over the four years of the intervention, the number of treated and control neighbors vary over time.[17] Table 3 reports the summary statistics for $N_{i,t|r}^{treated}$ and $N_{i,t|r}^{control}$, and Figure 2 presents histograms of the exposure measure $N_{i,t|r}^{treated}$, for various values of neighborhood radii. Whereas for $r = 1$ kilometers, the variation in $N_{i,t|r}^{treated}$ is small, as the neighborhood radius increases to 3 kilometers and beyond, we gain considerable variations in the exposure measure.

3.2 Econometric model

We exploit three unique features of our data to estimate the spillover effects. First, conditional on the total number of a child's CHECC neighbors at a given point in time ($N_{i,t|r}^{total}$), the number of neighbors who are randomized into treatments ($N_{i,t|r}^{treated}$) is determined exogenously through the intervention. Second, the repeated assessment schedule generates a panel, which enables us to track performance over time. Finally, multiple randomizations and the delivery of programs over the four

[17]Because more children were receiving the treatments over the four-year span of the intervention, and a neighbor who was previously in the control group in an earlier randomization might be assigned into a treatment group in later years, $N_{i,t|r}^{control}$ can both increase or decrease over time. However, $N_{i,t|r}^{treated}$ can only increase over time, because no child who was already treated could be assigned into the control group in later years.

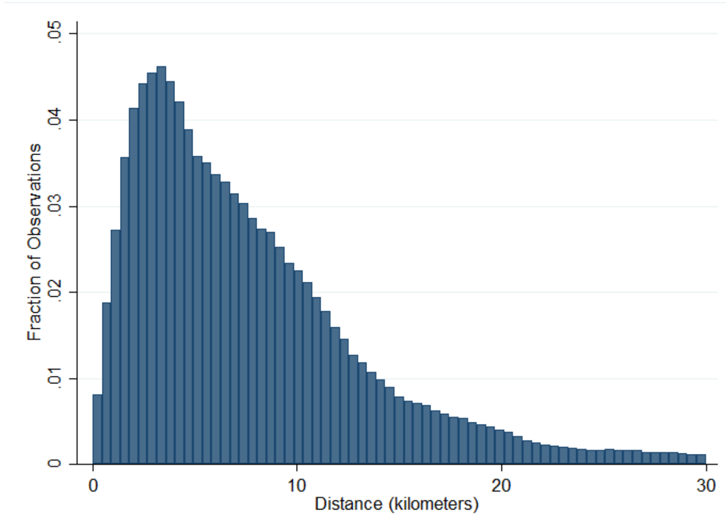


Figure 1: Histogram of distances between children in the study. The horizontal axis is cut at 30 kilometers.

years of the intervention create *within-individual* variations in our exposure measure $N_{i,t|r}^{treated}$.

Although the experimentally induced variation in our exposure measure serves as an important feature, which we exploit for identification, our estimation strategy does not rely on it exclusively. Given our limited sample size and the fact that the intervention was not designed to measure spatial spillovers, our exposure measure could be correlated with individual- or neighborhood-level unobservable characteristics. Therefore, we exploit the panel nature of our data to provide clean estimates of the spillover effects. The above three properties allow us to estimate the spillover effects using *within-individual* variations in our exposure measure ($N_{i,t|r}^{treated}$) through a fixed-effects specification. This technique uses the variations in spatial exposure over time and controls for any unobserved time-invariant individual-, family-, or neighborhood-level characteristics that might be correlated with $N_{i,t|r}^{treated}$.

We estimate spatial spillover effects from CHECC, using an individual fixed-effects specification of the form:

$$Y_{i,t} = \beta_0 + \beta_1 N_{i,t|r}^{treated} + \beta_2 N_{i,t|r}^{total} + \gamma_i + \delta_t + \epsilon_{i,t}, \quad (1)$$

where $Y_{i,t}$ is the standardized cognitive or non-cognitive test score of a child i on test t, $N_{i,t|r}^{treated}$ represents the number of treated neighbors of i at time t as previously defined, and $N_{i,t|r}^{total}$ represents the total number of i's neighbors who were randomized in the intervention by time t.[18] γ_i and δ_t

[18] As aforementioned, Miguel and Kremer (2004), Giné and Mansouri (2018) and Bobba and Gignoux (2019) use similar specifications to estimate spatial spillover effects. Similar to our specification, these studies use the number of treated individuals within a certain neighborhood radius as their measure of spatial exposure to treatments, and control for the total number of neighbors in their regression analyses. Distinct from these studies, which rely exclusively on the experimentally induced variations in the distribution of treated neighbors across individuals, we

Table 3: Neighbor Counts by Neighborhood Radius

	r= 1 km	r = 3 km	r = 5 km	r = 7 km
$N_r^{treated}$	27.89	178.13	325.63	422.81
	(27.73)	(154.21)	(238.75)	(272.08)
$N_r^{control}$	29.49	183.47	333.73	437.79
	(31.84)	(165.34)	(257.95)	(301.11)

Notes: This table presents the average number of treated and control neighbors of a child in our pooled sample, for various definitions of neighborhood radii. The numbers reflect all the observations in our pooled sample for which we observe both cognitive and non-cognitive scores. Standard deviations are reported in parentheses.

are individual and test (time) fixed effects. Under this specification, β_1 represents the average effect of moving one of the control neighbors of a child i to a treatment group, holding the total number of her CHECC neighbors constant. This measure (β_1) provides an intuitive estimate on the spillover effects from the intervention because it enables a policymaker to weigh the benefits against the costs associated with treating an additional child in the neighborhood. Section 5 presents an alternative model, which relaxes the assumption of time invariance for individual effects and exploits both within-individual and between-individual variation in spatial exposure to treated neighbors to estimate spillover effects. As we will further discuss in Section 5, our findings are robust to using this alternative specification.[19]

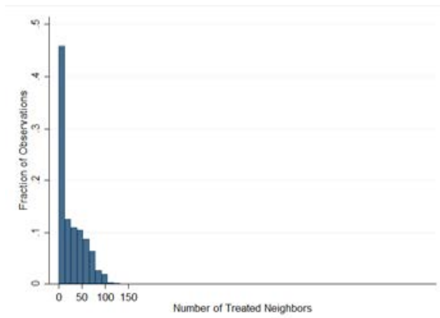
4 Results

4.1 Main findings

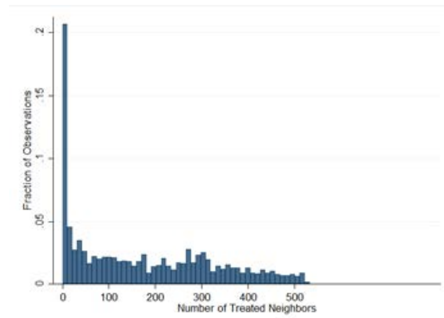
We estimate spillover effects for the neighborhood radii of 3, 5, and 7 kilometers. As Figure 2 suggests, when neighborhood is defined too narrowly, the variation in $N_{i,t|r}^{treated}$ becomes too small, limiting our power to estimate the effects. Therefore, we start with a neighborhood radii of 3 kilometers and larger, which provides us with enough variation in $N_{i,t|r}^{treated}$ to estimate the effects. Arguably, these choices of neighborhood radii are economically relevant. According to the National Household Travel Survey, the average commuting distance to school for a 6 to 12 year-old child

are able to exploit the panel nature of our data and identify the spillover effects using within-individual variations in exposure to treatments and remove all unobserved time-invariant individual-level characteristics that might be correlated with the spatial exposure to treatments.

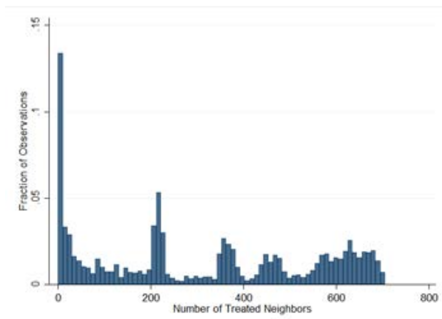
[19] As a robustness check, we also run the regressions without baseline observations and our effects remain similar.



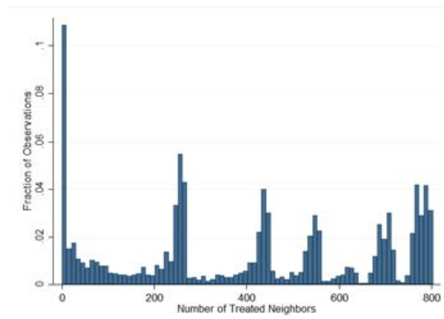
(a) r=1 km



(b) r=3 km



(c) r=5 km



(d) r=7 km

Figure 2: Histogram of $N_{i,t|r}^{treated}$ for $r = \{1, 3, 5, 7\}$ kilometers.

is about 6 kilometers (3.6 miles).[20] The average travel time from home to work for a Chicago Heights resident is estimated to be 26.1 minutes (US Census Bureau statistics), which translates to about 21 kilometers for a speed of 30 miles per hour.[21] Because schools and workplaces provide natural interaction spaces for children and their parents, we can reasonably assume our choices of neighborhood radii are relevant distances within which social interactions can generate spillovers.

Table 4 presents estimated β_1's from equation (1) for neighborhood radii of 3, 5, and 7 kilometers. Standard errors—clustered at the census-block-group level to allow for common error components within geographical units—are reported in parentheses below each point estimate. The left (right) panel presents the average spillover effect from treating an additional neighbor of a child on her standardized cognitive (non-cognitive) test score. Columns (1) and (4) report the pooled effects on both treatment and control children and reveal significant positive spillover effects on both cognitive and non-cognitive test scores. The effects on non-cognitive scores are more than double the effects on cognitive scores: an additional treated neighbor within 3 kilometers of a child's home

[20] https://nhts.ornl.gov/briefs/Travel%20To%20School.pdf

[21] This estimate is consistent with a report by the National Household Travel Survey that suggests the average commuting distance from home to work in the US is about 19 kilometers (11.8 miles). For more information, see: https://www.bts.gov/sites/bts.dot.gov/files/docs/browse-statistical-products-and-data/national-transportation-statistics/220806/ntsentire2018q1.pdf (page 73).

increases her cognitive score by 0.0033σ ($p < 0.01$), whereas it increases her non-cognitive score by 0.0069σ ($p < 0.01$). Empirical differences in cognitive and non-cognitive spillovers are statistically significant.[22]

Columns (2) and (3) parse the effects on cognitive scores by treatment assignment. These estimates reveal that both treatment and control children benefit from living close to treated families. While the control group benefits slightly more than the treatment group from cognitive spillover effects, the difference is not significant at conventional levels.[23] Columns (5) and (6) report the spillover effects in non-cognitive scores by treatment assignment. These estimates illustrate that the treatment and control children both benefit from non-cognitive spillovers. The estimated spillover effects on non-cognitive scores on the control and treatment groups are very similar and are not significantly different across the two groups.[24] These findings are robust to the choice of neighborhood radius, r.[25,26]

In sum, we document significant positive spillover effects on both cognitive and non-cognitive test scores and find the effect sizes are significantly larger for non-cognitive scores versus cognitive scores.[27] Yet, the richness of the data permits us to explore deeper into both the nature and extent of such spillovers.

4.2 Spatial fade-out

A closer examination of the estimated β_1's reported in Table 4 suggests an important spatial pattern: the spillover effect from an additional treated neighbor becomes smaller as we broaden the neighborhood radius from 3 to 7 kilometers. To further explore this pattern and shed light on the relationship between spillover effects and distance, we provide Figure 3, which shows the estimated β_1's for a broader range of r's.[28] Note that the effects on both cognitive and non-cognitive scores operate very locally.

As we increase the neighborhood radius, the marginal spillover effects from an additional treated

[22]The p-values from the Wald test of the null hypothesis H0: $\beta_1^{cog} = \beta_1^{ncog}$ against H1: $\beta_1^{cog} \neq \beta_1^{ncog}$ are 0.001, 0.03, and 0.06 for neighborhood radii of 3, 5, and 7 kilometers, respectively.

[23]The p-values from the Wald test of equal β_1^{cog}'s for the control and treatment group are 0.11, 0.16, and 0.13 for neighborhood radii of 3, 5, and 7 kilometers, respectively.

[24]The p-values from the Wald test of equal β_1^{ncog} for treatment and control group for neighborhood radii of 3K, 5K, and 7K meters are 0.80, 0.84, and 0.78.

[25]Appendix E breaks down these effects by subtests and explores which components of the cognitive/non-cognitive index generate the effects.

[26]Appendix D discusses the robustness of our estimated spillover effects to the exclusion of individual fixed effects from our main specification and to the exclusion of other controls and lagged dependent variables from our alternative (LDV) specification.

[27]In Appendix F, we explore the potential role of sorting by estimating the effects using a subsample of children who attended the majority of assessments. Our evidence suggests that selection is not an important factor in generating our results.

[28]The point estimates are reported in Appendix G.

Table 4: Mean Effect Sizes on Cognitive and Non-cognitive Scores, Fixed-Effects Estimates

	Cognitive Scores			Non-cognitive Scores		
	Pooled	Control	Treatment	Pooled	Control	Treatment
	(1)	(2)	(3)	(4)	(5)	(6)
r = 3 km	0.0033***	0.0038***	0.0016	0.0078***	0.0069***	0.0064***
	(0.0010)	(0.001)	(0.0010)	(0.0013)	(0.0015)	(0.0013)
r = 5 km	0.0021***	0.0023**	0.0010*	0.0043***	0.0037***	0.0034***
	(0.0006)	(0.0008)	(0.0006)	(0.0008)	(0.0011)	(0.0008)
r = 7 km	0.0018***	0.0021***	0.0008*	0.0033***	0.0025***	0.0028***
	(0.0005)	(0.0007)	(0.0005)	(0.0007)	(0.0010)	(0.0007)
Obs.	5,208	2,442	3,074	5,208	2,442	3,074

Notes: Spillover effects from *each* additional treated neighbor ($\hat{\beta}_1$) estimated from equation (1) are presented. Columns 1-3 (4-6) represent the average spillover effects from an additional treated neighbor on a child's standardized cognitive (non-cognitive) score. Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1

neighbor monotonically decrease. Because a larger neighborhood radius corresponds to a longer average distance to neighbors, the negative relationship between $\hat{\beta}_1$ and r implies that, as the distance between a child and her treated neighbor grows, the spillover effect on both cognitive and non-cognitive scores weakens. Specifically, the average spillover effect on a child's non-cognitive score, from a treated neighbor within a 3-kilometer radius is about twice as large as the effect from a treated neighbor who resides within a 7-kilometer radius (0.0069σ vs. 0.0033σ). Similarly, the average effect on the cognitive score from an additional treated neighbor who lives within 3 kilometers of a child is about twice as large as the effect from a treated neighbor who resides within a 7-kilometer radius (0.0033σ vs. 0.0018σ). In summary, we find that the spillover effects on both cognitive and non-cognitive scores are localized and decrease as the distance to a treated neighbor's home increases.[29]

4.3 Heterogeneous effects

Fryer et al. (2015) evaluated the parent-education component of CHECC (Parent Academy) by comparing the outcomes of treatment and control children, under the assumption that treatments

[29] An alternative approach for exploring such fade-out effects would be to parametrize the distribution by creating bands (i.e., number of youth within 3 km, number of youth between 3 and 5 km, number of youth between 5 and 7 km) for both $N_{i,t|b}^{treated}$ and $N_{i,t|b}^{total}$ (where $N_{i,t|b}^{treated}$ and $N_{i,t|b}^{total}$ represent the number of treated neighbors within band b and number of all neighbors within band b), and include these variables in the regression (see, e.g., Miguel and Kremer, 2004). Consistent with the fade-out pattern we document in our main specification, when we use this alternative method, we find that for non-cognitive skills, the spillover effect is always significant from treated neighbors within the first band (within 3 kilometers) and become insignificant as we move to treated neighbors who reside in bands that are placed further away from a child's home. For the cognitive skills, the estimates are sporadically significant from neighbors who reside in the closer bands and never significant from those who reside in bands that are further away.

Pooled

Control

Treated

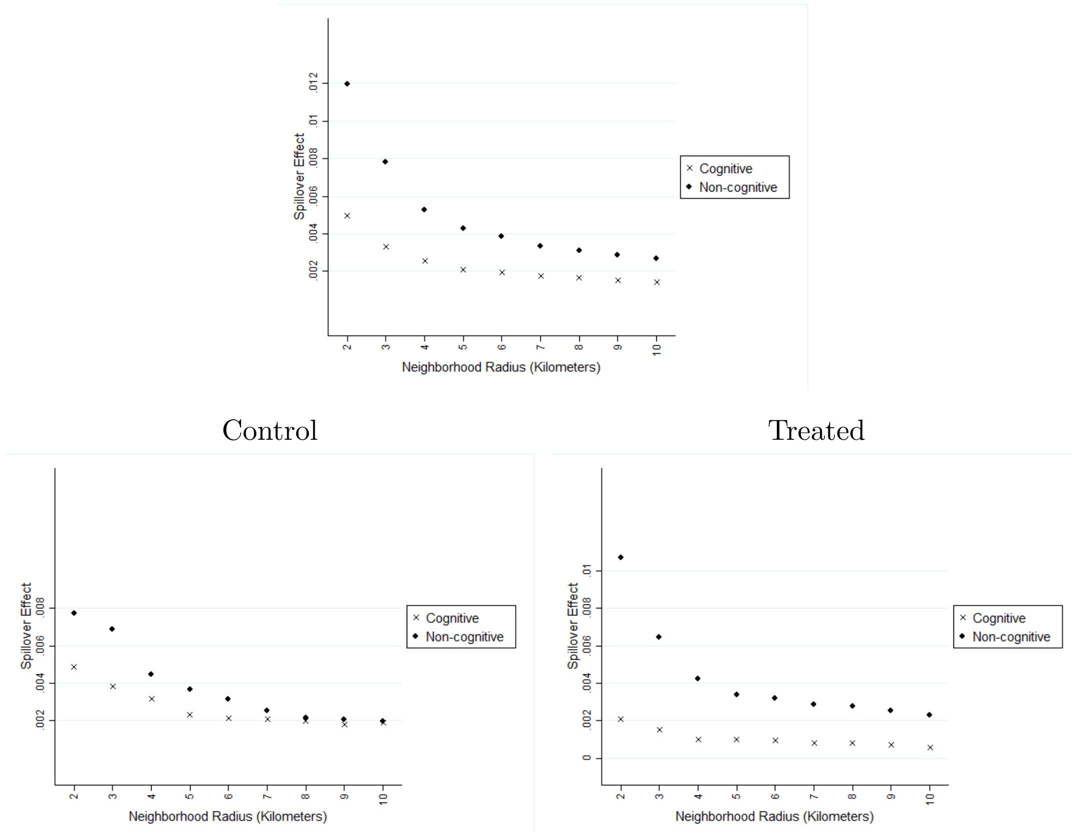


Figure 3: The spillover effect from having an additional treated neighbor on a child's standardized cognitive and non-cognitive scores, as functions of neighborhood radius.

did not induce externalities to the control group. They found that the assignment to Parent Academies increases a child's non-cognitive scores by 0.203σ, but does not significantly impact cognitive scores. Moreover, the authors reported positive treatment effects on cognitive and non-cognitive scores for Hispanic children, but did not find any significant treatment effects on African American children. Parent Academy was also reported to have slightly larger effects on girls than boys, although the gender differences were not significant. Motivated by the heterogeneity in treatment effects from the Parent Academy component of the intervention reported in Fryer et al. (2015), we investigate whether children of different races (or gender) benefit differently from spillover effects. We do so by estimating equation (1), separately by race and gender.

Since African American and Hispanic children make up over 90% of our sample, our analysis on heterogeneity along race focuses on these two groups. Panel (a) of Table 5 and Figure 4 presents $\hat{\beta}_1$'s separately for African American and Hispanic children. Comparing the effects across races, we find no significant differences in cognitive spillover effects between Hispanics and African

Americans. In contrast to the effects on cognitive scores, however, spillovers on non-cognitive scores are significantly larger for African Americans than Hispanics.[30] The empirical estimates indicate that, on average, an additional treated neighbor increases the non-cognitive scores of an African American child by about two to three times more than a Hispanic child. For instance, an additional treated neighbor within a 3-kilometer radius increases the non-cognitive score of a Hispanic child by 0.0045σ, whereas it increases an African American child's non-cognitive score by 0.0100σ.

Table 5: Mean Effect Sizes within Gender and Race Subgroups, Fixed-Effects Estimates

(a) Race

	Cognitive Scores	Non-cognitive Scores
	(1)	(2)
	African American	
r = 3 km	0.0014	0.0100***
	(0.0015)	(0.0024)
r = 5 km	0.0007	0.0055***
	(0.0008)	(0.0015)
r = 7 km	0.0009	0.0042***
	(0.0008)	(0.0012)
Obs.	2,087	2,087
	Hispanic	
r = 3 km	0.0042***	0.0045***
	(0.0013)	(0.0014)
r = 5 km	0.0027***	0.0019**
	(0.0009)	(0.0008)
r = 7 km	0.0023***	0.0016**
	(0.0007)	(0.0008)
Obs.	2,580	2,580

(b) Gender

	Cognitive Scores	Non-cognitive Scores
	(3)	(4)
	Boys	
r = 3 km	0.0048***	0.0088***
	(0.0016)	(0.0019)
r = 5 km	0.0029***	0.0048***
	(0.0009)	(0.0012)
r = 7 km	0.0024***	0.0038***
	(0.0007)	(0.0010)
Obs.	2,583	2,583
	Girls	
r = 3 km	0.0017	0.0068***
	(0.0011)	(0.0019)
r = 5 km	0.0013*	0.0037***
	(0.0007)	(0.0011)
r = 7 km	0.0011	0.0028***
	(0.0006)	(0.0009)
Obs.	2,625	2,625

Notes: The spillover effects from each additional treated neighbor on cognitive and non-cognitive test scores, estimated from the fixed-effects model (equation (1)). Panel (a) presents the effects, separately for African American and Hispanic receiving children. Panel (b) reports the effects separately for boys and girls. Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Panel (b) of Table 5 and Figure 4 presents the estimated effects by gender. Overall, boys benefit more than girls from both cognitive and non-cognitive spillovers. However, these gender differences are not statistically significant at the conventional levels.[31] In section 6, we explore at a deeper level the mechanisms through which spillover effects operate, and discuss a potential source of heterogeneity in spillover effects along race and gender lines.

Note that our estimates cannot be directly compared with the heterogeneity in treatment effects

[30]The p-values from Wald tests of the null of equal β_1 for Hispanics and African American children's cognitive (non-cognitive) test scores (in the pooled sample) are 0.15 (0.07), 0.10 (0.04), and 0.17 (0.07) for neighborhood radii of 3, 5, and 7 kilometers, respectively.

[31]The p-values from Wald tests of the null of equal β_1 for cognitive (non-cognitive) test scores for boys and girls (in the pooled sample) are 0.12 (0.47), 0.13 (0.50), and 0.10 (0.47) for neighborhood radii of 3, 5, and 7 kilometers, respectively.

reported by Fryer et al. (2015) due to two main differences between our samples. First, unlike Fryer et al. (2015) who report heterogeneous effects from the Parent Academy, our analysis uses data from all CHECC programs and considers heterogeneity in spillover effects on all children who were randomized during the intervention. Second, whereas Fryer et al. (2015) base their estimates on observations from the post program assessments (which took place immediately after the end of program year), our estimates use the pre-, mid-, and post-program assessment as well as up to four additional follow-up assessments, which were administered after a program year ended.

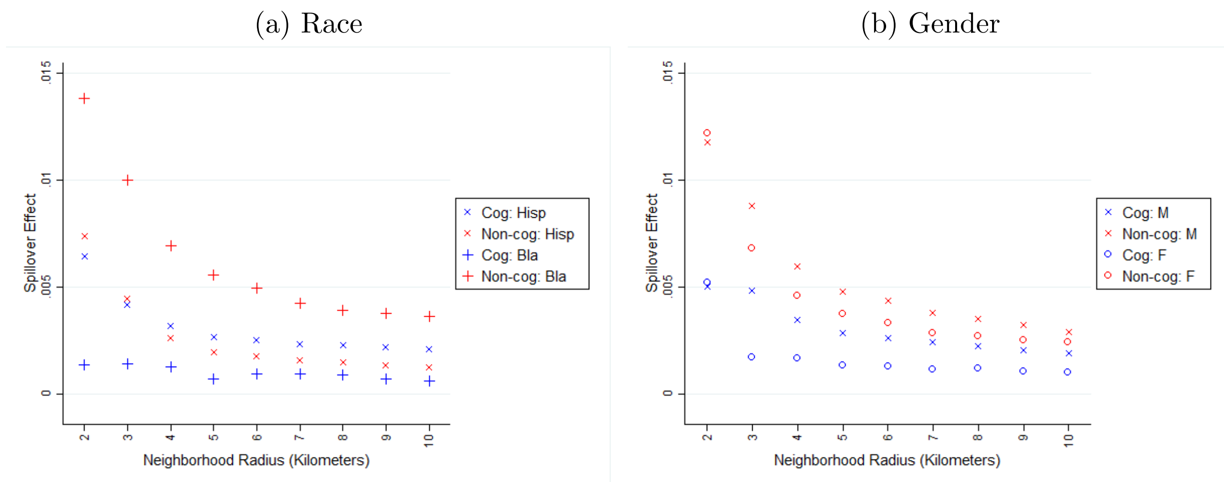


Figure 4: The spillover effect from an additional treated neighbor on a child's standardized cognitive and non-cognitive scores, estimated from the fixed-effects model. Panel (a) presents the effects separately for African American and Hispanic children, and panel (b) presents the effects separately for males and females.

5 Robustness

Our identification strategy, presented in Section 3 (individual fixed-effects model), is based on the assumption that individual effects are *time-invariant* omitted variables. In this section, we relax this assumption by directly controlling for the lagged dependent variables and removing individual fixed effects. Whereas the fixed-effects specification uses within-individual variations in spatial exposure to treatments, the lagged dependent variables (LDV) model exploits both within- and between-individual variations to estimate the spillover effects.

In this spirit, Angrist and Pischke (2008) argue that fixed effects and LDV estimates have a bracketing property such that they provide upper and lower bounds for where the true effect lies. The authors provide an empirical approach to estimate the effects under both specifications and check whether they provide broadly similar results. We formulate this alternative specification as:

$$Y_{i,t} = \beta_0 + \beta_1 N_{i,t|r}^{treated} + \beta_2 N_{i,t|r}^{total} + \eta Y_{i,t-1} + X_i \alpha + \sigma_b + \delta_t + \mu_c + \epsilon_{i,t}, \tag{2}$$

where $Y_{i,t}$, $N_{i,t|r}^{treated}$, and $N_{i,t|r}^{total}$ are defined as previously. We control for the lagged cognitive and non-cognitive test scores through $Y_{i,t-1}$, and include census-block group (σ_b) as well as time and cohort fixed effects (δ_t and μ_c). X_i represents a vector of time-invariant characteristics including gender, race, and age at the time of the baseline assessment. Under this specification, β_1 reflects the average spillover effect from an additional treated neighbor who lives within radius r on a child's standardized cognitive or non-cognitive test score.[32]

5.1 Main effects

Table 6 presents β_1's, estimated from equation (2), for three different neighborhood radii: 3, 5, and 7 kilometers. The standard errors, reported in parentheses, are clustered at the census-block-group level to allow for common error components within geographical units. Consistent with the results presented in section 4.1, we find significant spillover effects on both cognitive and non-cognitive test scores, with larger effects on non-cognitive scores.[33]

Comparing the effects on treated and control children leads to a similar conclusion as we found in our fixed-effects specification: there are not significant differences in spillover effects between children who were assigned to the treatment and control groups.[34] Empirical estimates from the two models are also similar in magnitude: An additional treated neighbor within 3 kilometers of a child is estimated to increase that child's cognitive score by 0.0033σ under the fixed-effects specification and by 0.0042σ under the LDV model. Similarly, an additional treated neighbor residing within 3 kilometers is estimated to increase the child's non-cognitive score by 0.0069σ and 0.0070σ under the fixed-effects and LDV specifications, respectively.[35]

[32] Appendix D discusses the robustness of our estimates to a gradual addition of controls.

[33] The p-values from the Wald test of the null hypothesis of $\beta_1^{cog} = \beta_1^{ncog}$ against the alternative of $\beta_1^{cog} \neq \beta_1^{ncog}$ are 0.045, 0.035, and 0.035 for neighborhood radii of 3, 5 and 7 kilometers, respectively.

[34] The p-values from the Wald test of equal β_1^{cog}'s for the control and treatment group are 0.34, 0.39, and 0.22 for neighborhood radii of 3K, 5K, and 7K meters, respectively. The corresponsing p-values for β_1^{noncog}'s are 0.82, 0.37, and 0.44.

[35] In appendix D, we discuss how these estimates change as we gradually add controls in Equation (2). Our findings suggest that our estimates of spillover effects are not very sensitive to the exclusion of lagged dependent variables or other controls. Specifically, the non-cognitive spillover effects are positive and significant even in our most stripped-down specification, which does not include any controls. That is, we can rely purely on the experimentally induced variations in exposure to estimate the non-cognitive spillover effects. These effects remain stable as we include additional controls to the model. The cognitive spillover effects become stable once we include neighborhood-level (block-group) fixed effects in the model and remain stable afterwards.

Table 6: Mean Effect Sizes on Cognitive and Non-cognitive Scores, LDV Estimates

	Cognitive Scores			Non-cognitive Scores		
	Pooled	Control	Treatment	Pooled	Control	Treatment
	(1)	(2)	(3)	(4)	(5)	(6)
r = 3 km	0.0042***	0.0052**	0.0019	0.0070***	0.0059***	0.0067**
	(0.0012)	(0.0025)	(0.0020)	(0.0015)	(0.0018)	(0.0026)
r = 5 km	0.0033***	0.0038**	0.0021*	0.0059***	0.0037***	0.0060***
	(0.0009)	(0.0016)	(0.0011)	(0.0014)	(0.0011)	(0.0023)
r = 7 km	0.0027**	0.0034***	0.0015	0.0054***	0.0035***	0.0053**
	(0.0007)	(0.0013)	(0.0010)	(0.0014)	(0.0013)	(0.0022)
Obs.	3,403	1,495	2,093	3,403	1,495	2,093

Notes: Spillover effects from *each* additional treated neighbor ($\hat{\beta}_1$) estimated from equation (2) are presented. Columns 1-3 (4-6) represent the effects from an additional treated neighbor on a child's standardized cognitive (non-cognitive) score. Robust standard errors, clustered at the census-block-group level are in parentheses; *** p<0.01, ** p<0.05, * p<0.1

5.2 Spatial fade-out

We can use a similar robustness test for our fade-out estimates. Figure 5 displays the estimated effects from an additional treated neighbor, as a function of neighborhood radius. The point estimates are reported in Appendix G. Similar to Figure 3, the estimated effects from the LDV specification also exhibit a downward-sloping relationship with the neighborhood radius. The spatial fade-out is more salient for the effects on non-cognitive scores. Confirming our finding from estimating the fixed-effects model reported in section 4.2, the LDV estimates show that the spillover effect from a treated neighbor decreases as the spatial distance to that neighbor's home increases.

5.3 Heterogeneous effects

Performing a similar robustness test on the estimated heterogeneous effects, we present Table 7, which reports the spillover effects by race and gender estimated from the LDV specification (equation (2)). Focusing on race (panel (a)), we find no significant differences in spillover on cognitive scores between African Americans and Hispanic children.[36] Similar to the fixed-effects estimates, the LDV estimates of non-cognitive spillover effects reveal a large and significant racial gap, with African American children benefiting about three to four times as much as Hispanics from their treated neighbors.[37]

[36] The p-values from Wald tests of the null hypothesis of equal spillover effects on cognitive scores (β_1) for Hispanics and African Americans in the pooled sample are 0.59, 0.52, and 0.50 for neighborhood radii of 3, 5, and 7 kilometers, respectively.

[37] The p-values from Wald tests of the null hypothesis of equal spillover effects on non-cognitive scores (β_1) for Hispanics and African Americans in the pooled sample are 0.001, 0.005, and 0.020 for neighborhood radii of 3, 5, and 7 kilometers, respectively.

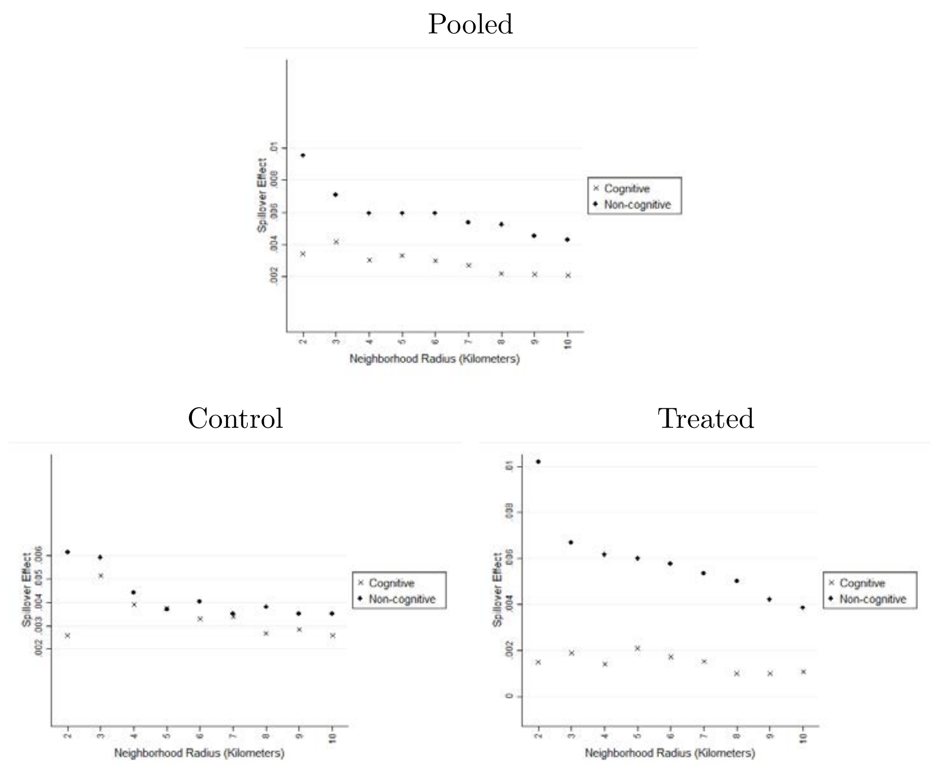


Figure 5: The spillover effects, estimated under the LDV model, from an additional treated neighbor on a child's standardized cognitive and non-cognitive test scores, as functions of neighborhood radius.

Panel (b) of Table 7 reports the effects from the LDV specification by gender, and reveals that in general, boys tend to benefit slightly more than girls from spillover effects. However, the gender differences are not significant at conventional levels.[38]. Overall, both the gender and race effects are largely preserved after performing these robustness tests.

6 Exploring the Mechanisms

One attractive feature of our field experiment is that it generates data that has unique variation to explore the underlying mechanisms at work for our observed spillovers. As the sociology literature argues, early childhood human capital development relates both to parental and child interactions. Whereby Sheldon (2002) and others argue that social activity can be represented by group compositions, and that those critically determine the nature and extent of spillovers, we were unable

[38]The p-values from Wald tests of the null hypothesis of equal spillover effects on cognitive scores (β_1) for boys and girls in the pooled sample are 0.29, 0.36, and 0.16 for neighborhood radii of 3, 5, and 7 kilometers, respectively. The corresponding p-values for spillover effects in non-cognitive scores are 0.88, 0.82, and 0.53.

Table 7: Mean Effect Sizes within Gender and Race Subgroups, LDV Estimates

(a) Race

	Cognitive Scores	Non-cognitive Scores
	(1)	(2)
	African American	
r = 3 km	0.0056***	0.0154***
	(0.0019)	(0.0034)
r = 5 km	0.0037***	0.0108***
	(0.0013)	(0.0025)
r = 7 km	0.0032***	0.0087***
	(0.0012)	(0.0022)
Obs.	1,312	1,312
	Hispanic	
r = 3 km	0.0042**	0.0034**
	(0.0017)	(0.0016)
r = 5 km	0.0026**	0.0034**
	(0.0011)	(0.0014)
r = 7 km	0.0022**	0.0031**
	(0.0009)	(0.0016)
Obs.	1,760	1,760

(b) Gender

	Cognitive Scores	Non-cognitive Scores
	(3)	(4)
	Boys	
r = 3 km	0.0054***	0.0063***
	(0.0026)	(0.0023)
r = 5 km	0.0037***	0.0059***
	(0.0011)	(0.0021)
r = 7 km	0.0033***	0.0060***
	(0.0009)	(0.0021)
Obs.	1,693	1,693
	Girls	
r = 3 km	0.0032**	0.0067***
	(0.0014)	(0.0017)
r = 5 km	0.0027***	0.0053***
	(0.0010)	(0.0019)
r = 7 km	0.0017*	0.0044**
	(0.0009)	(0.0018)
Obs.	1,710	1,710

Notes: Spillover effects from an additional treated neighbor on cognitive and non-cognitive test scores, estimated from LDV specification (equation (2)). Panel (a) presents the effects, separately for African American and Hispanic receiving children. Panel (b) reports the effects separately for boys and girls. Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1

to find literature that uses random assignment to parse the spillover effects. In this manner, using our randomization: variation in treatment and variation in treatment type, alongside distance to treated and control neighbors, gender, and race variables, we can detail the underlying mechanisms proposed to make human capital development inherently a social activity (Coleman, 1988; Cochran and Brassard, 1979; Corsaro, 2005).

6.1 Parent Academy versus Pre-K

Since the intervention offered education programs for both children and parents, one might expect the spillover effects to operate though the social network of children via direct interactions between them, or indirectly through their parents' social networks. To shed light on which channel generates stronger effects, we start by comparing spillovers from treated neighbors who were assigned to the parent-education programs (Parent Academy neighbors) with the effect from those who were assigned to the preschool programs (Pre-K neighbors). If spillovers mainly operate through parents' social network, then we might expect larger effects from Parent Academy neighbors than from Pre-K neighbors, because Parent Academy treatments involved parents more directly and more intensely than Pre-K treatments. Alternatively, larger spillover effects from Pre-K neighbors could imply the

(a) Race (b) Gender

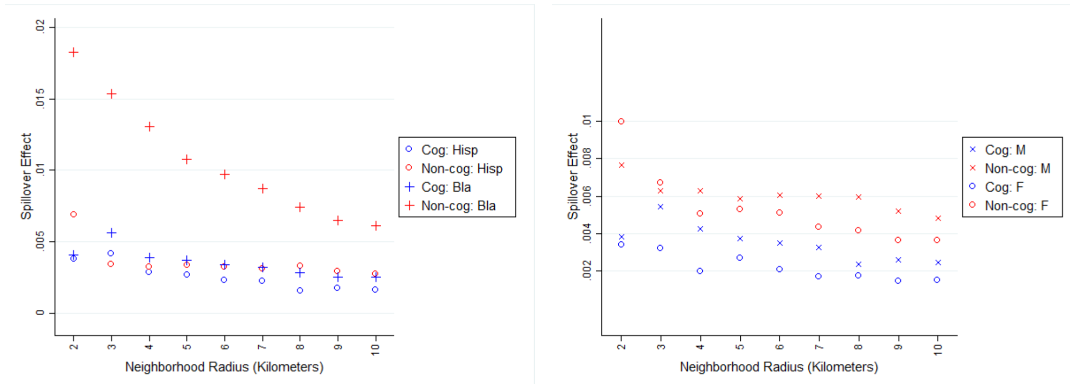


Figure 6: The spillover effect from an additional treated neighbor on a child's standardized cognitive and non-cognitive scores, estimated from the LDV model. Panel (a) presents the effects separately for African American and Hispanics, and panel (b) presents the effects by gender.

direct interactions between children play an important role in generating the effects.

To compare spillovers from Parent Academy and Pre-K neighbors, we estimate the fixed-effects and LDV specifications of the following forms:[39]

$$Y_{i,t} = \beta_0 + \beta_p N^{Parent}_{i,t,r} + \beta_c N^{PreK}_{i,t,r} + \lambda N^{total}_{i,t|r} + \gamma_i + \delta_t + \epsilon_{i,t} \tag{3}$$

$$Y_{i,t} = \beta_0 + \beta_p N^{Parent}_{i,t,r} + \beta_c N^{PreK}_{i,t,r} + \lambda N^{total}_{i,t|r} + \eta Y_{i,t-1} + X_i\alpha + \sigma_b + \mu_c + \delta_t + \epsilon_{i,t}. \tag{4}$$

$N^{Parent}_{i,t,r}$ and $N^{PreK}_{i,t,r}$ represent the number of Parent Academy and Pre-K neighbors of a child i who reside within a distance r from i at the time of her assessment t, and $N^{total}_{i,t|r}$, $Y_{i,t}$, $Y_{i,t-1}$, X_i, γ_i, σ_b, μ_c and δ_t are defined as previously. To simplify the analysis and retain statistical power, we construct $N^{PreK}_{i,t,r}$ and $N^{Parent}_{i,t,r}$ by pooling neighbors who were assigned to any of the preschool programs as *Pre-K neighbors*, and pooling those who were assigned to any of the two parent-education programs as *Parent Academy* neighbors. Under the above specifications, β_p reflects the average spillover effect from an additional Parent Academy neighbor, holding $N^{PreK}_{i,t,r}$ and $N^{total}_{i,t|r}$ constant. In other words, β_p represents the average effect of substituting a *control* neighbor with a *Parent Academy* neighbor. Similarly, β_c represents the average spillover effect from an additional Pre-K neighbor on a child's test scores.

Note that a child i may benefit from a Parent Academy neighbor k through two channels. The first channel is the parents' social networks: k's parents may influence the behavior and decisions

[39] In a few cases in which a treated neighbor k was first assigned to a Parent Academy (Pre-K) treatment and assigned to Pre-K (Parent Academy) in later years, k is counted as a Parent Academy (Pre-K) neighbor for the observations prior to the second randomization, and as a Pre-K (Parent academy) neighbor for the observations following the second randomization.

of i's parents', which may in turn shape i's development. Such effects can occur through information externalities (i.e., k's parents share their acquired knowledge from Parent Academy with i's parents) or preference externalities between parents. The second channel is peer influence: if Parent Academy improves k's outcomes, then child i might benefit from direct interactions with child k. The benefits from a Pre-K neighbor, however, are likely to spill over mainly through direct interactions between children (peer influence) because parents are not the main target of the Pre-K treatments.[40] Thus, although $\hat{\beta}_p$ might reflect spillovers through both the parents' and the child's social networks, $\hat{\beta}_c$ is more likely to reflect an effect that is mainly driven by direct interactions between children.

Table 8 reports estimated β_p and β_c for neighborhood radii of 3, 5, and 7 kilometers from the fixed-effects and LDV specifications. Focusing on *non-cognitive* scores, the estimates from both models suggest larger spillover effects from Pre-K neighbors than Parent Academy neighbors ($\hat{\beta}_p < \hat{\beta}_c$).[41] According to the fixed-effects estimates, an additional Pre-K neighbor within a 3-kilometer radius of a child increases her non-cognitive score by 0.0099σ, whereas an additional Parent Academy neighbor only induces a 0.0045σ increase in the non-cognitive test score. Similarly, the LDV estimates suggest an additional Pre-K neighbor within a 3-kilometer radius of a child increases her non-cognitive test score by 0.0108σ, whereas an extra Parent Academy neighbor within the same radius induces only a 0.0017σ increase in her non-cognitive test score. The larger spillovers from Pre-K neighbors than from Parent Academy neighbors suggests that direct social interactions between children (rather than between parents) plays an important role in generating the non-cognitive effects.

Unlike the effects on non-cognitive scores, we find no significant differences in spillover effects on cognitive scores from Parent Academy and Pre-K neighbors.[42]

6.2 Heterogeneity in neighborhood-level social interactions

Our estimates presented in Sections 4 and 5 suggested that non-cognitive spillover effects are significantly larger for African Americans than Hispanics. We also found spillover effects to be larger for boys than girls, although the gender difference was not significant at conventional levels.

[40] As previously described previously in Section 2, two out of the four Pre-K treatments (Preschool Plus and Kinderprep) had parental-education components that were not incentivised as heavily, and were much less intensive than the one offered to Parent Academy families. Appendix H examines the spillover from Pre-K neighbors who were randomized to preschool programs with and without the parental component and shows that our conclusions are not sensitive to pooling these two treatments together.

[41] The p-values for the fixed-effects estimates from Wald tests of the null hypothesis $\beta_p = \beta_c$ against $\beta_p \neq \beta_c$ for r=3 km, r=5 km and r=7 km equal 0.004, 0.006, and 0.016, respectively. The corresponding p-values from the LDV estimates are 0.004, 0.010, and 0.03.

[42] For the fixed-effects estimates, the p-values from Wald tests of the null hypothesis $\beta_p = \beta_c$ against $\beta_p \neq \beta_c$ for r=3 km, r=5 km and r=7 km are 0.82, 0.43, and 0.34, respectively. The corresponding p-values from the LDV estimates are 0.76, 0.14, and 0.14.

Table 8: Spillover Effects from Parent Academy and Pre-K Neighbors, Pooled Sample

(a) Fixed Effects Estimates

	Cognitive Scores		Non-cognitive Scores	
	β_p	β_c	β_p	β_c
	(1)	(2)	(3)	(4)
r = 3 km	0.0035***	0.0033***	0.0045***	0.0099***
	(0.0012)	(0.0011)	(0.0008)	(0.0018)
r = 5 km	0.0024***	0.0019***	0.0024***	0.0055***
	(0.0008)	(0.0006)	(0.0006)	(0.0011)
r = 7 km	0.0021***	0.0016***	0.0022***	0.0041***
	(0.0005)	(0.0005)	(0.0005)	(0.0009)
Obs.	5,208		5,208	

(b) LDV estimates

	Cognitive Scores		Non-cognitive Scores	
	β_p	β_c	β_p	β_c
	(1)	(2)	(3)	(4)
r = 3 km	0.0045***	0.0040***	0.0017	0.0108***
	(0.0016)	(0.0014)	(0.0017)	(0.0021)
r = 5 km	0.0046***	0.0030***	0.0019	0.0068***
	(0.0013)	(0.0009)	(0.0017)	(0.0015)
r = 7 km	0.0041***	0.0025***	0.0023	0.0058***
	(0.0013)	(0.0007)	(0.0020)	(0.0013)
Obs.	3,403		3,403	

Notes: Columns 1 and 3 (2 and 4) report the average effect of an additional Parent Academy (Pre-K) neighbor who resides within distance r of a child, on her standardized cognitive and non-cognitive scores. Panel (a) presents the estimates from the fixed-effects specifications (equation 3), and panel (b) reports LDV estimates (4). Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Given our evidence suggesting direct interactions with peers at the neighborhood level might be an important mechanism in generating non-cognitive spillover effects, one might hypothesize that the racial and gender differences in non-cognitive spillover effects originate from racial and gender differences in social interactions within neighborhoods.

We explore this idea using data from the National Longitudinal Study of Adolescent Health survey (Add Health),[43] which includes measures of social interactions within neighborhoods. Add Health is a longitudinal study of a nationally representative sample of adolescents in grades 7 to 12 in the United States. Using the public-use data from Add Health, we explore racial and gender differences in variables that reflect children's social interactions within neighborhoods. Table 9 presents our findings. Consistent with our hypothesis, we find African American children are significantly (at $p < 0.001$) more likely than Hispanics, and boys are significantly (at $p < 0.001$) more likely than girls, to (i) know most people in their neighborhood, (ii) stop on the street to talk to someone from the neighborhood, and (iii) use recreation facilities in the neighborhood. Although these data speak to a slightly older age group than our subjects, we believe that these findings provide suggestive evidence that the higher level of neighborhood-level social interactions for African Americans (and for boys) is a possible cause of the larger non-cognitive spillover effects on this subgroup. Yet, more research is necessary as we view this result as merely suggestive.

[43]The AddHealth is a project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

Table 9: Heterogeneity in Neighborhood Level Social Interactions (Add Health Data)

	African American	Hispanic	Male	Female
(1) Know most people in the neighborhood	0.78	0.69	0.79	0.71
(p-value)	(0.00)		(0.00)	
(2) Stop on the street to talk to someone who lives in the neighborhood	0.84	0.78	0.86	0.79
(p-value)	(0.00)		(0.00)	
(3) Use recreation/sports center in the neighborhood	0.25	0.18	0.28	0.17
(p-value)	(0.00)		(0.00)	
Observations	1,121	537	787	871

Notes: The share of respondents who agreed with each statement. The data are restricted to African American and Hispanic respondents. Questions are from Wave II of the Add Health survey. Demographic information was obtained from Wave I of the survey. The precise wording of the survey questions are as follows: (1) You know most of the people in your neighborhood. (0 = false, 1= true); (2) In the past month, you have stopped on the street to talk with someone who lives in your neighborhood. (0 = false, 1 = true); (3) Do you use a physical fitness or recreation center in your neighborhood? (0 = no, 1 = yes). P-values from a two-sample test of proportions under the null hypothesis of equal proportions against the alternative of unequal proportions between each pair of subgroups are reported in parentheses.

6.3 Parental Investment

Spillover effects can also operate through influencing parental decisions, which affect children's development. Parents might learn from their neighbors about returns to investments and the most productive forms of investments in their children and adjust their choices accordingly. Our data include a self-reported measure of investment concerning parents' decision to enroll their child in preschool programs (other than CHECC). This variable was collected through a survey completed by parents, which was administered at the end of each program year (at the time of the children's post-assessment).[44]

We start by exploring whether this self-reported measure of parental investment has any predictive power regarding children's cognitive and non-cognitive performance. That is, whether enrolling one's child in a preschool program (other than CHECC) is associated with the child's cognitive or non-cognitive development. We do so by estimating an LDV model of the following form:[45]

$$Y_{i,t} = \beta_0 + \kappa Z_{i,t} + \eta Y_{i,t-1} + X_i\alpha + \sigma_b + \mu_c + \delta_t + \epsilon_{i,t}, \tag{5}$$

where $Z_{i,t}$ is a binary variable indicating whether parents of a child i reported enrolling i in a preschool program during the school year prior to t.[46] All other terms are defined as previously. In this specification, κ reflects whether a parent's decision to enroll her child in a preschool program

[44] For the last cohort of families who were randomized in the program, this information was also collected from parents at the time of pre-assessment (in addition to the post-assessment).

[45] Since for the vast majority (over 90%) of our sample, we observe the investment measure only once, we cannot estimate their effects on test scores using within-individual variations.

[46] Controlling for $N^{treated}_{i,t|r}$ and $N^{total}_{i,t|r}$ or the child's treatment assignment does not change our point estimates very much.

is correlated with the child's skill development ($Y_{i,t}$).[47]

Table 10 presents the estimated κ from equation (5). The estimates suggest a parent's decision to enroll her child in other (non-CHECC) preschool programs is significantly correlated with cognitive development. Indeed, such enrollment is associated with increases in the child's cognitive test score by 0.134σ, whereas the effect on non-cognitive test-scores is not significant at conventional levels. In the next step, we explore whether spillover effects occur on a parent's decision to enroll her child in other programs.

Table 10: Parent's Decision to Enroll a Child in Preschool Programs and Child's Development

	Cognitive	Non-cognitive
Other Programs	0.134***	0.127
	(0.048)	(0.086)
Observations	655	655
R-squared	0.80	0.58

Notes: Estimated κ's from equation (5) are presented. The point estimates reflect the effect of the parental decision to enroll children in preschool programs, and children's cognitive and non-cognitive test scores. Although we have about 1,000 observations on investment decisions, we observe lagged test scores for only about 60% of these observations. Robust standard errors, clustered at the census-block-group level, are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Our analysis of the spatial spillover effects on cognitive and non-cognitive test scores exploited the panel nature of our outcome variables, allowing us to estimate the causal spillover effects using individual fixed-effects or lagged dependent-variables specifications. Unfortunately, for a large majority (over 90%) of our sample, the data on parental decisions to enroll children in non-CHECC programs were collected only once (at the time of children's post-assessment). Therefore, we cannot use the previous identification strategies to estimate the causal spillover effects on parents' investment decision. Instead, we rely on the following OLS specification to explore this channel:

$$Z_{i,t} = \beta_0 + \beta_1 N^{treated}_{i,t|r} + \lambda N^{total}_{i,t|r} + X_i\alpha + \sigma_b + \mu_c + \delta_t + \epsilon_{i,t}, \tag{6}$$

where $Z_{i,t}$ represents parents' investment decision, and other variables are defined as previously. Under the above specification, β_1 represents the relationship between having an additional treated neighbor residing within distance r and a parent's decision to enroll her child in a non-CHECC preschool program.

Table 11 presents estimated β_1's from equation 6. Interestingly, our estimates suggest positive and

[47]Note that since $Z_{i,t}$ is not assigned randomly, we cannot interpret κ as the causal effect of preschool enrollment on children's development without making additional assumptions. The purpose of this exercise is to merely explore whether this measure of parental investment is associated with child's development rather than to establish a causal relationship.

significant spillover effects on this measure of parents' investment decision. Each additional treated neighbor residing within a 3-kilometer radius of a child's home increases the likelihood of the child's parents enrolling her in a preschool program by 0.55 percentage points. Given our previous finding that enrolling a child in a non-CHECC program significantly improves her cognitive performance, this result provides suggestive evidence that influencing the parental decision to enroll her child in a preschool program is an important channel through which cognitive spillover effects operate.

Table 11: Exposure to Treatments and the Parents' Investment Decision

Dependent Variable	Other Programs
r = 3k	0.0055**
	(0.0023)
r = 5k	0.0066*
	(0.0040)
r = 7k	0.010**
	(0.0041)
Observations	974

Notes: Estimated β_1's from equation (6) are presented. Robust standard errors, clustered at the census-block-group level, are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1

The above finding should be interpreted with caution for two reasons. First, our measure of parental investment decision is self-reported, and therefore is not an objective measure of the actual choices. Second, given that the structure of our survey data prevents us from using individual fixed-effects or LDV methods in estimation, interpreting β_1 as the causal spillover effect on parents' decision would require an additional assumption that (conditional on $N^{total}_{i,t|r}$, X_i, σ_b, μ_c and δ_t) our exposure measure—$N^{treated}_{i,t|r}$—is uncorrelated with unobserved individual-level characteristics. Recognizing these two caveats, we believe our results provide *suggestive* evidence that *cognitive* spillover effects might operate—at least partly—through influencing parents' decisions to enroll their child in a preschool program.

7 Total Impact

After measuring positive spillover effects from the programs delivered at CHECC, we now turn to estimating the total impact of the intervention, accounting for these indirect effects. Beyond estimating the total impact, we also (i) disentangle the direct and indirect (spillover) effects from the intervention, and (ii) estimate the size of bias that would arise if we were to ignore spillovers.

Before presenting our evaluation strategy, we should emphasize three key features of this exercise. First, whereas CHECC offered multiple education programs to both parents and children, the aim of this exercise is not to separately evaluate each program. Instead, we provide an overall evaluation of the intervention as a whole by pooling all treatments and accounting for spillovers. Second, given that we estimate spillovers using panel data over multiple years, our total effect is also based on observations over multiple years, starting at the time of randomization and terminating four years after the program ends for the cohort. Therefore, our analysis provides an average estimate of the total impact of CHECC over time, which includes the immediate as well as the longer-run effects.[48] Finally, to fix ideas and simplify the presentation of our results, we set the neighborhood radii to 3 kilometers. Appendix I presents our estimates of the total impact for a broader range of neighborhood radii and discusses the robustness of our findings to varying r.

The total impact of the intervention ($Total$) on a child i who was randomized into one of the treatments (Parent Academies or Pre-K) can be expressed as the sum of the direct treatment effect (DE) and the spillover effects, which i receives from N other treated individuals in her vicinity (ST^N): $Total = Direct + ST^N$. If an evaluator ignores the spillover effects on control children (SC^N) and follows the standard approach, she naively reports the difference in the mean outcomes between the control and treated groups as the total impact. This standard approach would result in a biased estimate, which we call $Total^{Standard} = Total - SC^N = Direct + ST^N - SC^N$.

We evaluate the total impact of the intervention by estimating the following LDV model, using our pooled sample, which includes observations from both the treated and control children. We focus on the LDV specification for this estimation, because unlike the fixed-effects model, it allows us to exploit between-individual variation to estimate the direct time-invariant treatment effect:

$$Y_{i,t} = \beta_0 + \theta T_i + \beta_1 N_{i,t|r}^{treated} + \lambda(T_i \times N_{i,t|r}^{treated}) + \beta_2 N_{i,t|r}^{total} + \eta Y_{i,t-1} + \alpha X_i + \delta_t + \mu_c + \sigma_b + \epsilon_{i,t}. \quad (7)$$

T_i is a treatment indicator, which equals 1 if i was assigned to a treatment group, and 0 otherwise, and $Y_{i,t}$, $Y_{i,t-1}$, $N_{i,t|r}^{treated}$, $N_{i,t|r}^{total}$, X_i, δ_t, μ_c, and σ_b are defined as previously. We include an interaction term ($T_i \times N_{i,t|r}^{treated}$) to allow for different spillover effects on treatment and control children. Under this specification, θ represents the average direct effect of the intervention on a treatment child ($Direct$), β_1 represents the average spillover effect from an additional treated neighbor on a control child (SC^1), and $\beta_1 + \lambda$ represents the average spillover effect from an additional treated neighbor on a treatment child (ST^1). Assuming linearity, the average total spillover effect from all treated neighbors on a control child (SC^N) can be expressed as $\overline{N_r^{treated}} \times \hat{\beta}_1$, where $\overline{N_r^{treated}}$ represents the average number of treated neighbors who reside within distance r of a child.[49] Likewise, the average total spillover effect from all treated neighbors on a child who was randomized into a treatment (ST^N) can be expressed as $\overline{N_r^{treated}} \times (\hat{\beta}_1 + \hat{\lambda})$. Therefore, the total impact of the intervention on a treatment child ($Total$) is the sum of the direct and the spillover

[48] For this reason, our estimates of the total impact cannot be directly compared to the ones reported in Fryer et al. (2015), which are based on test scores from the assessments administered immediately after the programs ended.

[49] Appendix J shows these findings do not change much if we relax the linearity assumption and allow for quadratic and cubic terms.

effects: $\hat{Total} = \hat{\theta} + \overline{N_r^{treated}} \times (\hat{\beta}_1 + \hat{\lambda})$.

Table 12 reports the estimated coefficients from equation (7) for $r = 3$ kilometers. The first two columns present the estimates for the pooled sample, and the last four report estimates separately for African American and Hispanic children. Focusing on the pooled sample, we find the average direct effects of being randomly assigned to a treatment group on a child's standardized cognitive and non-cognitive scores are 0.11 and 0.05, respectively. The average total spillover effects on a *control child*'s standardized cognitive and non-cognitive scores (SC^N) are estimated to be 0.75 and 1.25. The corresponding spillover effects on a *treatment child* are 0.71σ and 1.27σ. The total impact of being assigned to treatment (including both the direct and spillover effects) on a child's standardized cognitive and non-cognitive test scores is estimated to be 0.82 and 1.32, respectively. Note that the total spillover effects on both cognitive and non-cognitive scores of treatment children are larger than the direct treatment effects, suggesting a large portion of the total impact is due to the network effects that emerge from interactions with other treated individuals. This finding implies that if one were to treat a single child in isolation, the average cognitive and non-cognitive treatment effects would be about ($\frac{0.82}{0.11} \approx$) 7 and ($\frac{1.31}{0.05} \approx$) 26 times smaller than the estimated impacts in the presence of spillovers from other treated children.

Table 12: Total Program Impact

	All		African American		Hispanic	
	Cognitive	Non-cognitive	Cognitive	Non-cognitive	Cognitive	Non-cognitive
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\theta}$ (*Direct*)	0.1070***	0.0456	0.0145	-0.0558	0.276***	0.135
	(0.0393)	(0.0642)	(0.0667)	(0.111)	(0.0671)	(0.0978)
$\hat{\beta}$	0.0042***	0.0070***	0.0056***	0.0154***	0.0044**	0.0034**
	(0.0013)	(0.0015)	(0.0019)	(0.0034)	(0.0018)	(0.0016)
$\hat{\lambda}$	-0.0002*	0.0001	0.0001	0.0005	-0.0007***	-0.0000
	(0.0001)	(0.0002)	(0.0003)	(0.0003)	(0.0002)	(0.0003)
SC^N	0.75***	1.25***	0.72***	1.99***	1.27***	0.81**
	(0.22)	(0.27)	(0.24)	(0.44)	(0.48)	(0.38)
ST^N	0.71***	1.27***	0.74***	2.05***	1.05**	0.80**
	(0.22)	(0.27)	(0.25)	(0.45)	(0.49)	(0.36)
$Total$	0.82***	1.32***	0.75***	1.99***	1.38***	0.93***
	(0.23)	(0.27)	(0.25)	(0.44)	(0.47)	(0.37)
$Total^{Standard}$	0.06***	0.07**	0.03	0.01	0.11***	0.12***
	(0.02)	(0.03)	(0.05)	(0.07)	(0.03)	(0.04)
Observations	3,403	3,403	1,312	1,312	1,760	1,760
R-squared	0.659	0.384	0.684	0.441	0.613	0.362

Notes: Estimated coefficients from equation (7) are presented. Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N includes all observations for which we have lagged cognitive and non-cognitive scores, and other regressors. $\overline{N^{treated}} = 178.13$; $\overline{N^{treated}_{Black}} = 128.83$;$\overline{N^{treated}_{Hispanic}} = 237.38$

The comparison across the last two rows of Table 12 shows that if we were to ignore the spillover effects, we would severely underestimate the total impact of the intervention. For example, if we were to disregard the spillover effects, we would have wrongly concluded the intervention only induced a 0.06σ and 0.07σ increase in cognitive and non-cognitive test scores, respectively.[50] These estimates on cognitive and non-cognitive scores are considerably smaller than our estimated effects, which account for spillovers to the control group ($Total$).

We report our estimates of impacts on African Americans and Hispanics in columns (3)-(6) of Table 12.[51] The first observation is that the direct effects of being randomized into a treatment group ($Direct$) on both cognitive and non-cognitive test scores are larger for Hispanics than African Americans, although the difference is only significant for cognitive skills.[52] While the racial difference in cognitive spillovers are small and insignificant ($p > 0.40$), non-cognitive spillovers are significantly larger for African Americans ($p < 0.05$). Overall, the intervention increases the *cognitive* scores for African American and Hispanic children who were randomized into treatment by 0.75σ and 1.38σ, which are not significantly different from each other ($p = 0.45$).

By contrast, African American children who were offered the chance to participate in one of the programs gain more than their Hispanic counterparts in *non-cognitive* skills as a result of the intervention. The average total program impact ($Total$) on the non-cognitive test score of an African American treatment child is 1.99σ, which is significantly larger than 0.93σ, the corresponding effect on a Hispanic treatment child ($p = 0.07$). Importantly, disregarding spillover effects and evaluating CHECC by naively differencing the outcomes between control and treatment children results in a quite conservative representation of findings. Our estimates, presented in the last row, suggest that ignoring spillovers would have led us to conclude that whereas Hispanic children benefited as a result of the intervention, African Americans gained nothing. Note that this conclusion was in fact the one Fryer et al. (2015) reached in their evaluation of the parent academy intervention.

[50] Note these estimates are considerably smaller than the estimated effects of the Parent Academy arm of the intervention previously reported in Fryer et al. (2015). Fryer et al. (2015)—who ignore spillover effects—report that Parent Academies increased cognitive and non-cognitive test scores by 0.119σ and 0.203σ, by the end of the program year. Our $Total^{Standard}$ estimates of the whole intervention (which ignore spillover effects to the control group) on cognitive and non-cognitive scores are 1.99 (=0.119/0.06) and 2.9 (=0.203/0.07) times smaller than the reported effects by Fryer et al. (2015). Two important factors cause the difference in these two sets of estimates: First, Fryer et al. (2015) focus only on one treatment arm of the intervention (Parent Academies), whereas we estimate the impact of CHECC as a whole, and our sample includes observations from all children who were randomized over the four years of the intervention. Second, Fryer et al. (2015) focus on the short-term treatment effects and compare test scores from assessments that were administered immediately at the end of a program year. Our analysis, on the other hand, uses data from up to four years after a program ends, meaning we rely on both short- and longer-term outcomes to estimate the impact. In fact, the treatment effects wear off and the differences in test scores between the treatment and control groups become smaller over time, which suggests our estimates of $Total^{Standard}$ should be smaller than the effects reported in Fryer et al. (2015).

[51] In our sample, on average, an African American (Hispanic) child has 129 (237) treated neighbors within 3 kilometers from her home.

[52] The corresponding p-values for cognitive and non-cognitive scores are 0.006 and 0.17.

8 Discussion

Evaluations of early childhood programs have played an important role in shaping policy debates on early education. For instance, the Head Start Impact Study (HSIS), a randomized control trial of Head Start, reported small effect sizes that fade considerably over a few years (Puma et al., 2010, 2012). These findings have heightened debate among academics over the cost effectiveness of Head Start (e.g. Barnett, 2011; Gibbs et al., 2013; Kline and Walters, 2016) and have been frequently cited by critics who argue Head Start is ineffective in achieving its mission and should be abandoned or seriously reformed.[53] Given the policy impact of the findings from early education interventions, and more broadly any social intervention, accurate evaluation of the total effect of these programs is crucial.

The standard approach in evaluating social programs is to randomly assign subjects to treatment and control groups. From there, many analysts simply difference the mean outcomes and report the monetized treatment effect within a benefit-cost framework. This approach is based on the assumption that a person's potential outcomes are independent of other participants' treatment assignment; that is, no spillover effects occur. Yet, if one considers that social capital, in the Coleman (1988) sense, has value for a young person's development, then group composition and community structure hold great import in understanding the development of human capital at a young age. For our purposes, this literature opens up the distinct possibility of key spillovers that might arise from such interventions, and guides us where to look for such effects.

In this paper, we provide the first empirical evidence on *spillover* effects from a large-scale early education intervention by causally estimating *neighborhood* effects. In doing so, we provide unique causal evidence on how neighbors influence children's outcomes. By leveraging field experimental variation that has a novel panel data feature, we are allowed a comparison of child i at time t having a certain number of treated neighbors with the same child i at time $t+1$, where the number of treated neighbors has changed. Compared to the previous literature, our approach provides a new glimpse into such effects since we compare the same person at the *same* location with different exposure rates while, in the standard literature using residential movers, one compares the same child at *different* locations with different exposure rates; thus, it is difficult to disentangle the location from the exposure effect.

Overall, we find that ignoring spillover effects results in severe underestimation of the total impact. Indeed, we document spillovers that are economically significant and much larger than we anticipated: on average, the intervention increases cognitive and non-cognitive test scores of a control child by 0.75σ and 1.25σ, respectively. Beyond the main spillover effects, we observe in-

[53]For example, a 2014 report from the House Budget Committee cites the findings of HSIS to conclude Head Start is "failing to prepare children for school" (see: http://budget.house.gov/uploadedfiles/war_on_poverty.pdf).

teresting heterogeneities. For example, we find that non-cognitive spillover effects are larger for African American children than for Hispanics. In addition, our evidence suggests that non-cognitive spillovers are more likely to operate through children's rather than parents' social networks. We also find suggestive evidence that cognitive spillover effects are generated through influencing the parents' decision to enroll their child in an alternative form of treatment–an outside preschool program.

Given the importance of non-cognitive skills in children's future labor market and educational outcomes (Heckman et al., 2006; Borghans et al., 2008; Brunello and Schlotter, 2011), our findings provide practical insights into designing early interventions to better foster such skills. Specifically, our results suggest that interventions that promote social interactions both within participants and between participants and non-participants are likely to generate larger positive externalities on non-cognitive skills.

Our work also speaks to policymakers interested in the science of scaling programs (see, e.g., Al-Ubaydli et al., 2017a,b; 2020). As experimentalists, we have focused almost exclusively on how best to generate data to explore intervention effects and disentangle mechanisms. Yet, what has been lacking is a scientific understanding of how to make best use of the research insights generated. In particular, in what form should we implement the program for policy purposes? And, should we expect the small-scale results to generalize to larger settings? Our findings reveal that traditional measures of early education impacts, which ignore externalities, might be too pessimistic when such programs are taken to scale. In this way, our findings suggest that ignoring the spillover effects can lead to fewer programs being taken to scale than is optimal. Of course, this needs not be a general result, as it is possible that in some cases those treated suppress outcomes of those in the control group. More work is necessary in order to detail the nature and extent of scale-up effects when moving from scientific insight to policy. While evidence-based policy is a useful target, our work highlights that policy-based evidence is necessary to estimate the benefit/cost profile at scale.

More generally, one might wonder if our observed effects are generalizable to other populations of people and situations. For this consideration, we follow the List (2020) SANS conditions. First, in terms of selection, our sample includes a pre-K population that had broad coverage across Chicago Heights (Fryer et al., 2020). Yet, our CHECC program was an opt-in design, so our estimates might be limited to underserved community members who sign-up for early childhood programs. Considering naturalness of the choice task, setting, and time frame, our main identification comes from a natural field experiment, thus our setting is one in which subjects are engaged in a natural task and are not placed on an artificial margin. Finally, since our key results are WAVE1 insights (List, 2020), replications need to be completed to understand if the reduced form direct treatment effect estimates and the spillover estimates manifest in other school districts and over other time horizons.

References

[1] Agostinelli, F., Doepke, M., Sorrenti, G. and Zilibotti, F. (2020), "It takes a village: The economics of parenting with neighborhood and peer effects," NBER Working Paper No. 27050.

[2] Al-Ubaydli, O., List, J.A. and D.L. Suskind (2017a), "What can we learn from experiments? Understanding the threats to the scalability of experimental results," *American Economic Review Papers & Proceedings* 117(5), 282–86.

[3] Al-Ubaydli, O., List, J.A., LoRe, D. and D.L. Suskind (2017b), "Scaling for economists: Lessons from the non-adherence problem in the medical literature," *Journal of Economic Perspectives* 31(4), 125–144.

[4] Al-Ubaydli, O., List, J.A. and D.L. Suskind (2020), "The science of using science: Toward an understanding of the threats to scalability," *International Economic Review* 61(4), 1387–1409.

[5] Angelucci, M., and G. De Giorgi (2009), "Indirect effects of an aid program: How do cash transfers affect ineligibles' consumption?" *American Economic Review* 99(1), 486–508.

[6] Angrist, J.D. and J.S. Pischke (2008), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton: Princeton University Press.

[7] Åslund, O., Östh, J. and Y. Zenou (2010), "How crucial is distance to jobs for ethnic minorities? Old question – Improved answer," *Journal of Economic Geography* 10, 389–422.

[8] Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C. and M. Rubio-Codina (2020), "Estimating the production function for human capital: Results from a randomized control trial in Colombia," *American Economic Review* 110(1), 48–85.

[9] Attanasio, O., Meghir, C. and E. Nix (2015), "Human capital development and parental investment in India," NBER Working Paper No. 21740.

[10] Avitabile, C. (2012), "Spillover effects in healthcare programs: Evidence on social norms and information sharing," IDB Working Paper Series No. 380.

[11] Barnett, W.S. (2011) "Effectiveness of early educational intervention," *Science* 333, 975–978.

[12] Bobba, M. and J. Gignoux (2019), "Neighborhood effects in integrated social policies," *World Bank Economic Review* 33(1), 116–139.

[13] Boning, W.C., Guyton, J., Hodge, R.H., Slemrod, J. and U. Troiano (2018), "Heard it through the grapevine: Direct and network effects of a tax enforcement field experiment," NBER Working Paper No. 24305.

[14] Borghans, L., Duckworth, A.L., Heckman, J.J. and B. ter Weel (2008), “The economics and psychology of personality traits,” *Journal of Human Resources* 43, 972–1059.

[15] Boucher, V., Del Bello, C., Panebianco, F. Verdier, T. and Y. Zenou (2020), “Education transmission and network formation,” CEPR Discussion Paper No. 14997.

[16] Bourdieu, P. (1985), “The social space and the genesis of groups,” *Theory and Society* 14(6), 723–744.

[17] Brunello, G. and M. Schlotter (2011). “Non-cognitive skills and personality traits: Labour market relevance and their development in education and training systems,” IZA Discussion Papers No. 5743.

[18] Chetty, R. and N. Hendren (2018a), “The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects,” *Quarterly Journal of Economics* 133(3), 1107–1162.

[19] Chetty, R. and N. Hendren (2018b), “The impacts of neighborhoods on intergenerational mobility II: County-level estimates,” *Quarterly Journal of Economics* 133(3), 1163-1228.

[20] Chetty, R., Hendren, N. and L.F. Katz (2016), “The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment,” *American Economic Review* 106(4), 855–902.

[21] Chyn, E. (2018), “Moved to opportunity: The long-run effects of public housing demolition on children,”*American Economic Review* 108(10), 3028–56.

[22] Cochran, M.M. and J.A. Brassard (1979), “Child development and personal social networks,”*Child Development* 50(3), 601–616.

[23] Coleman, J.S. (1988), “Social capital in the creation of human capital,”*American Journal of Sociology* 94, S95–S120.

[24] Corsaro, W.A. (2005), *The Sociology of Childhood*, 2nd edition, London: Pine Forge Press.

[25] Cotton, C., Hickman, B., List, J.A., Price, J. and S. Roy (2020), “Productivity versus motivation in adolescent human capital production: Evidence from a structurally-motivated field experiment, ” NBER Working Paper No. 27995.

[26] Crépon, B., Duflo, E., Gurgand, M., Rathelot, R. and P. Zamora (2013), “Do labor market policies have displacement effects? Evidence from a clustered randomized experiment,” *Quarterly Journal of Economics* 128(2), 531–580.

[27] Cunha, F. and J.J. Heckman (2007), “The technology of skill formation,” *American Economic Review* 97(2), 31–47.

[28] Cunha, F., Heckman, J.J. and S.M. Schennach (2010), "Estimating the technology of cognitive and noncognitive skill formation," *Econometrica* 78, 883–931.

[29] Damm, A.P. and C. Dustmann (2014), "Does growing up in a high crime neighborhood affect youth criminal behavior?" *American Economic Review* 104(6), 1806–32.

[30] Doepke, M. and F. Zilibotti (2017), "Parenting with style: Altruism and paternalism in intergenerational preference transmission," *Econometrica* 85(5), 1331–1371.

[31] Doepke, M. and F. Zilibotti (2019), *Love, Money, and Parenting: How Economics Explains the Way We Raise Our Kids*, Princeton: Princeton University Press.

[32] Deaton, A. and N. Cartwright (2018), "Understanding and misunderstanding randomized controlled trials," *Social Science & Medicine* 210, 2–21.

[33] Duflo, E. and E. Saez (2003), "The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment," *Quarterly Journal of Economics* 118(3), 815–842.

[34] Dunn, L.M., Dunn, L.M., Bulheller, S. and H. Häcker (1965), *Peabody Picture Vocabulary Test*, Circle Pines, MN: American Guidance Service.

[35] Durlauf, S.N. (2004), "Neighborhood effects," In: J.V. Henderson and J-F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Vol. 4, Amsterdam: North-Holland, pp. 2173–2242.

[36] Edin, P.-A., Fredriksson, P. and O. Åslund (2003), "Ethnic enclaves and the economic success of immigrants. Evidence from a natural experiment," *Quarterly Journal of Economics* 118, 329–357.

[37] Entwisle, D.R., Alexander, K.L. and L.S. Olson (1994), "The gender gap in math: Its possible origins in neighborhood effects,"*American Sociological Review* 59(6), 822–838.

[38] Epple, D. and R.E. Romano (2011), "Peer effects in education: A survey of the theory and evidence," In: A. Bisin, J. Benhabib and M.O. Jackson (Eds.), *Handbook of Social Economics*, Vol. 1, Amsterdam: North-Holland, pp. 1053–1163.

[39] Ferracci, M., Jolivet, G. and G.J. van den Berg (2014), "Evidence of treatment spillovers within markets,"*Review of Economics and Statistics* 96, 812–823

[40] Fryer Jr, R.G., Levitt, S.D. and J.A. List (2015), "Parental incentives and early childhood achievement: A field experiment in Chicago heights," NBER Working Paper No. 21477.

[41] Fryer, R., Levitt, S., List, J., and A. Samek (2020), "Introducing CogX: A New Preschool Education Program Combining Parent and Child Interventions"˙University of Chicago, Becker Friedman Institute for Economics, Working Paper No. 2020-149, Available at SSRN: https://ssrn.com/abstract=3714022.

[42] Gautier, P., Muller, P., van der Klaauw, B., Rosholm, M. and M. Svarer (2018), "Estimating equilibrium effects of job search assistance," *Journal of Labor Economics* 36(4), 1073–1125.

[43] Gibbs, C., Ludwig, J. and D. Miller (2013), "Does Head Start do any lasting good?" In: M.J. Bailey and S. Danziger (Eds.), *Legacies of the War on Poverty*, New York: Russell Sage Foundation Press, pp. 39–65.

[44] Giné, X. and G. Mansuri (2018), "Together we will: Experimental evidence on female voting behavior in Pakistan," *American Economic Journal: Applied Economics* 10(1), 207–235.

[45] Graham, B.S. (2018), "Identifying and estimating neighborhood effects," *Journal of Economic Literature* 56(2), 450–500.

[46] Halpern-Felsher, B.L., Connell, J.P., Spencer, M.B., Aber, J.L., Duncan, G.J., Clifford, E., Crichlow, W.E., Usinger, P.A., Cole, S.P., Allen, L. and E. Seidman (1997), "Neighborhood and family factors predicting educational risk and attainment in African American and White children and adolescents," *Neighborhood Poverty* 1, 146–173.

[47] Hanushek, E.A. (2020), "Education production functions," In: S. Bradley and C. Green (Eds.), *The Economics of Education, Second Edition*, New York: Academic Press, pp. 161–170.

[48] Heckman, J.J. (2008), "Schools, skills, and synapses," *Economic Inquiry* 46, 289–324.

[49] Heckman, J.J., Stixrud, J. and S. Urzua (2006), "The effects of cognitive and noncognitive abilities on labor-market outcomes and social behavior," *Journal of Labor Economics* 24, 411–482.

[50] Ioannides, Y.M. (2011), "Neighborhood effects and housing," In: J. Benhabib, A. Bisin, and M.O. Jackson (Eds.), *Handbook of Social Economics*, Vol. 1B, Amsterdam: Elsevier Science, 1281–1340.

[51] Ioannides, Y.M. and G. Topa (2010), "Neighborhood effects: Accomplishments and looking beyond them," *Journal of Regional Science* 50, 343–362.

[52] Janssen, W. (2011), "Externalities in program evaluation: The impact of a women's empowerment program on immunization," *Journal of the European Economic Association* 9(6), 1082–1113.

[53] Jonassen, D.H. (2004), *Learning to Solve Problem. An Instructional Design Guide*, San Fransisco: John Wiley Son, Inc.

[54] Kalil, A. (2015), "Inequality begins at home: The role of parenting in the diverging destinies of rich and poor children," In: P.R. Amato, A. Booth, S.M. McHale and J. Van Hook (Eds.), *National Symposium on Family Issues. Families in an Era of Increasing Inequality: Diverging Destinies*, Cham, Switzerland: Springer International Publishing, pp. 63–82.

[55] Katz, L.F., Kling, J.R. and J.B. Liebman (2001), "Moving to opportunity in Boston: Early results of a randomized mobility experiment," *Quarterly Journal of Economics* 116, 607–654.

[56] Kautz, T., Heckman, J.J., Diris, R., Ter Weel, B. and L. Borghans (2014), "Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success," NBER Working Paper No. 20749.

[57] Kline, P. and C.R. Walters (2016), "Evaluating public programs with close substitutes: The case of Head Start," *Quarterly Journal of Economics* 131(4), 1795–1848.

[58] Kling, J.R., Ludwig, J. and L.F. Katz (2005), "Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment," *Quarterly Journal of Economics* 120, 87–130.

[59] Lalive, R., Landais, C., and J. Zweimüller (2015), "Market externalities of large unemployment insurance extension programs," *American Economic Review* 105(12), 3564–96.

[60] Leventhal, T. and J. Brooks-Gunn (2000), "The neighborhoods they live in: The effects of neighborhood residence on child and adolescent outcomes," *Psychological Bulletin* 126(2), 309–337.

[61] List, J.A. (2020), "Non est disputandum de generalizability? A glimpse into the external validity trial," NBER Working Paper No. 27535.

[62] Miguel, E. and M. Kremer (2004), "Worms: Identifying impacts on education and health in the presence of treatment externalities," *Econometrica* 72, 159–217.

[63] Mincer, J. (1958), "Investment in human capital and personal income distribution," *Journal of Political Economy* 66(4), 281–302.

[64] Minh, A., Muhajarine, N., Janus, M., Brownell, M. and M. Guhn (2017), "A review of neighborhood effects and early child development: How, where, and for whom, do neighborhoods matter?" *Health and Place* 46, 155–174.

[65] Muralidharan, K., Niehaus, P. and S. Sukhtankar (2017), "General equilibrium effects of (improving) public employment programs: Experimental evidence from India," NBER Working Paper No. 23838.

[66] Puma, M., Bell, S., Cook, R. and C. Heid (2010), "Head Start Impact Study: Final Report," Washington, DC: U.S. Department of Health and Services, Administration for Children and Families,

[67] Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A. and J. Downer (2012), "Third Grade Follow-up to the Head Start Impact Study Final Report," OPRE Report # 2012-45, Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

[68] Putnam, R.D. (1993), "The prosperous community: Social capital and public life," *The American Prospect* 13, 35–42.

[69] Rincke, J. and C. Traxler (2011), "Enforcement spillovers," *Review of Economics and Statistics* 93(4), 1224–1234.

[70] Sacerdote, B. (2011), "Peer effects in education: How might they work, how big are they and how much do we know thus far?" In: E.A. Hanushek, S. Machin and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Vol. 3, Amsterdam: Elsevier, pp. 249–277.

[71] Schuller, T. (2000), "Social and human capital: The search for appropriate techno methodology," *Policy Studies* 21(1), 1–13.

[72] Schunk, D.H. (2020), *Learning Theories: An Educational Perspective*, 8th Edition, New York: Pearson.

[73] Sheldon, S.B. (2002), "Parents' social networks and beliefs as predictors of parent involvement," *The Elementary School Journal* 102(4), 301–316.

[74] Sinclair, B., McConnell, M. and D.P. Green (2012), "Detecting spillover effects: Design and analysis of multilevel experiments," *American Journal of Political Science* 56(4), 1055–1069.

[75] Smith-Donald, R., Raver, C. C., Hayes, T., Richardson, B. (2007), "Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research," *Early Childhood Research Quarterly*, 22(2), 173–187.

[76] Todd, P.E. and K.I. Wolpin (2007), "The production of cognitive achievement in children: Home, school, and racial test score gaps," *Journal of Human Capital* 1(1), 91–136.

[77] Topa, G. and Y. Zenou (2015), “Neighborhood and network effects,” In: G. Duranton, V. Henderson and W. Strange (Eds.), *Handbook of Regional and Urban Economics*, Vol. 5A, Amsterdam: Elsevier Publisher, pp. 561–624.

[78] Waldfogel, J. and E. Washbrook (2011), “Early years policy,” *Child Development Research* 1–12.

[79] Willoughby, M.T., Wirth, R. J. and C.B. Blair (2012), “Executive function in early childhood: Longitudinal measurement invariance and developmental change,” *Psychological Assessment* 24(2), 418–431.

[80] Woodcock, R.W., McGrew, K. S. and N. Mather (2001), *Woodcock-Johnson Tests of Achievement*, Itasca, IL: Riverside Publishing.

The Social Side of Early Human Capital Formation: Using a Field Experiment to Estimate the Causal Impact of Neighborhoods
Online Appendix

By John A. List[1], Fatemeh Momeni[2] and Yves Zenou[3]

[1]Department of Economics, University of Chicago, Chicago, IL, USA. Email: jlist@uchicago.edu.
[2]Department of Economics, University of Chicago, Chicago, IL, USA. Email: fmomeni@uchicago.edu.
[3]Department of Economics, Monash University, Caulfield, VIC, Australia, and IFN. Email: yves.zenou@monash.edu.

A Maps

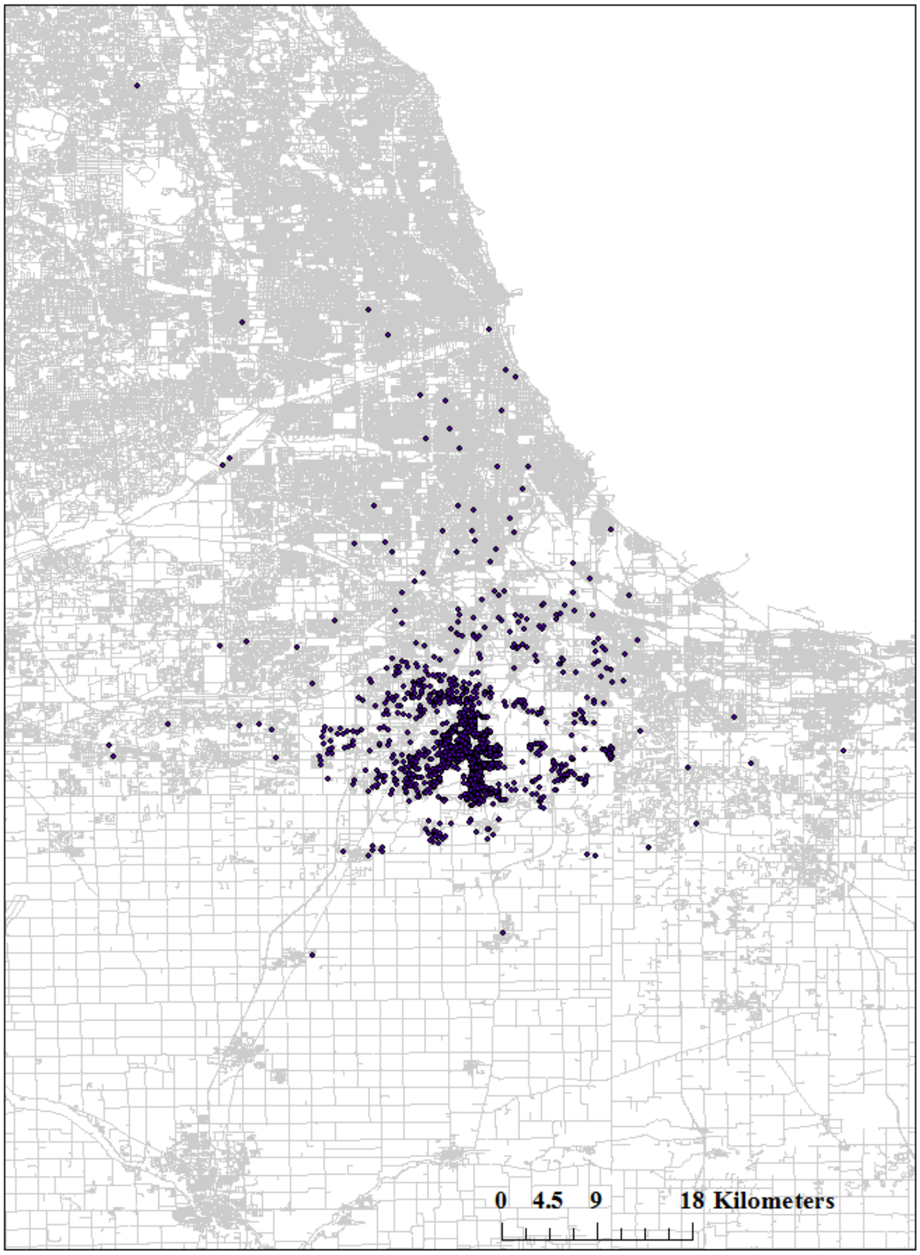


Figure A.1: Map of Home Locations

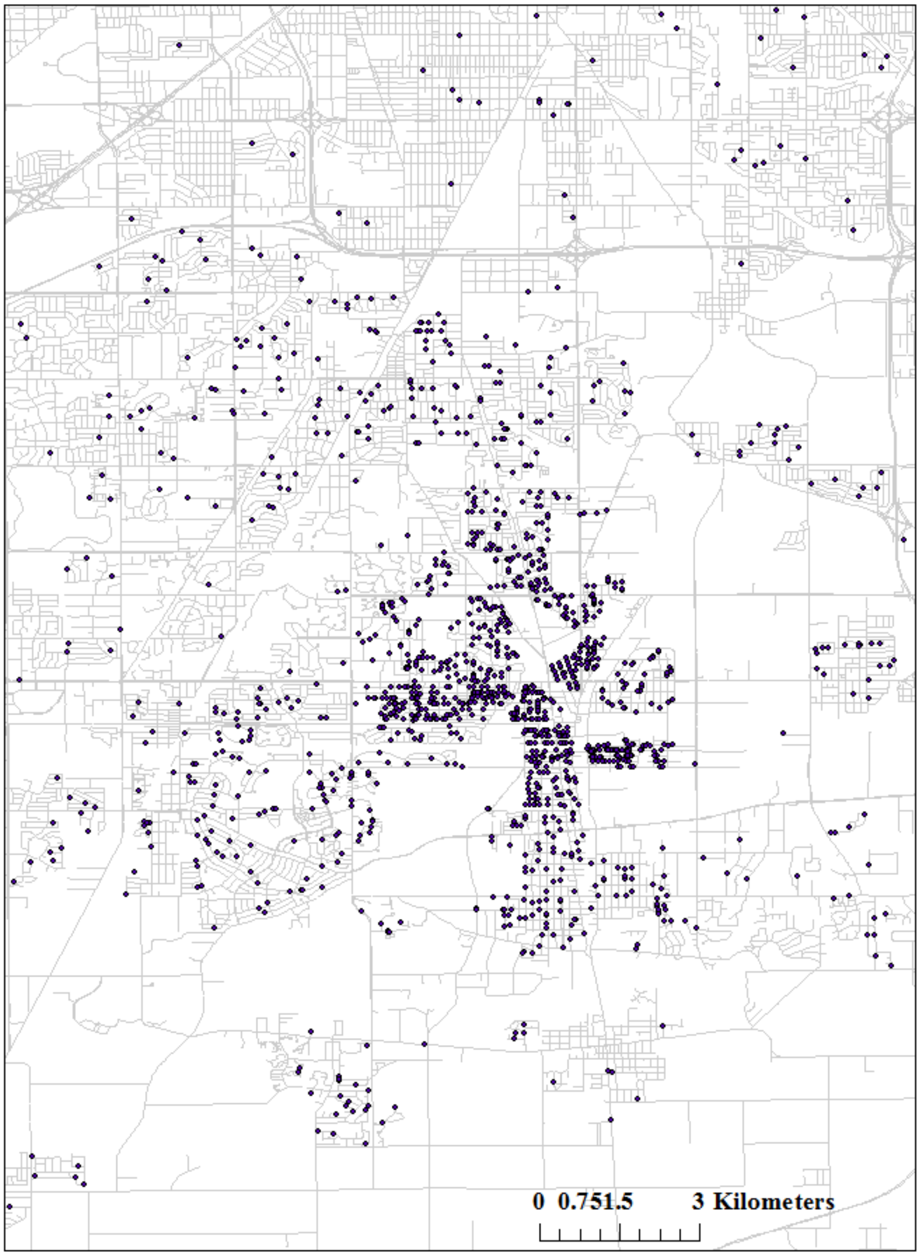


Figure A.2: Map of Home Locations, Zoomed in

B Assessment schedule

Table B.1: Assessment Calendar

	Winter (Jan-Feb)	Early Spring (Apr)	Late Spring (May)	Late Summer (Jul-Sep)
2010				Pre^{2010}
2011	Mid^{2010}		$Post^{2010}$ Pre^{2011}	SL^{2010}
2012	Mid^{2011}	$AO1^{2010}$	$Post^{2011}$ Pre^{2012}	SL^{2011}
2013	Mid^{2012}	$AO1^{2011}$ $AO2^{2010}$	$Post^{2012}$ Pre^{2013}	SL^{2012}
2014	Mid^{2013}	$AO1^{2012}$ $AO2^{2011}$ $AO3^{2010}$	$Post^{2013}$	SL^{2013}
2015		$AO1^{2013}$ $AO2^{2012}$ $AO3^{2011}$ $AO4^{2010}$		
2016		$AO2^{2013}$ $AO3^{2012}$ $AO4^{2011}$ $AO5^{2010}$		

Notes: Superscripts are cohort identifiers (randomization years).
Pre= pre assessment; Mid= mid assessment; Post= post assessment; SL= summer loss assessment; AOx= age-out assessment x years after the treatment ended

C Details on constructing the samples

Here we present how we treat observations from children who were randomized in multiple years in constructing our panel data set for the control, treatment, and pooled samples.

We follow two rules in constructing the control sample: (i) For those control children who were randomized into treatments in later years, we only keep the observations that took place before their treatment started; (ii) for those who were randomized into the control group in more than one year, we only keep the observations corresponding to their first randomization.

In construction of our treatment sample, we follow two rules: (i) For those treatment children who were first randomized into the control and later into a treatment group, we only keep the observations after their treatment started; (ii) for those who were randomized twice and both times into a treatment group, we only keep the observations corresponding to the first randomization.

Finally, we use the following three rules in constructing our pooled sample: (i) For children who were randomized twice and both times into the control group, we only keep the observations that correspond to the first randomization; (ii) for children who were randomized twice and both times into a treatment group, we only keep the observations corresponding to their first randomization; and (iii) for those who were randomized twice, the first time into the control group and the second time into a treatment group, we only keep the observations corresponding to the second randomization.

Three factors can result in missing an observation for a child in our control sample: (*i*) The child was absent on the assessment day; (*ii*) the child was moved to a treatment group in a later randomization and thus her outcomes (for the times after she had entered the treatment groups) are not included in the sample; or (*iii*) the child belongs to later cohorts for which the assessment is taking place at a later date (April 2018 or after). Similarly, an observation from the treatment sample would be missing if (*i*) the child was absent on that assessment day; (*ii*) the child was previously in the control group and thus her outcomes (for the times before she entered the treatment group) are not included in the sample; or (*iii*) corresponding assessment is taking place at a later date (April 2018 or later).

D Gradual addition of controls

In this section, we explore the robustness of our estimated spillover effects to the choice of controls included in our models. In doing so, we study how our estimated spillover effects change as we gradually add controls in our models for the pooled study samples used in the individual fixed-effects and LDV analysis.

Tables D.1 and D.2 present the spillover effects from an additional treated neighbor on the standardized cognitive and non-cognitive scores of a child, as we vary the set of control variables included in the regression, estimated for the sample we used in our LDV analysis. Similarly, Tables D.3 and D.4 present the spillover effects from an additional treated neighbor on the standardized cognitive and non-cognitive scores of a child, as we vary the set of control variables included in the regression, estimated for the sample we used in our individual fixed-effects analysis.[4]

The results presented in Table D.1 suggest that our estimated spillover effects on cognitive scores from our LDV sample become stable as soon as we include neighborhood (block-group) fixed effects (column 2) and remain stable as we continue adding controls to reach our most complete model (column 7). Our estimated effects on non-cognitive scores for our LDV sample, presented in Table D.2, from the most basic model, which does not include any controls (column 1), are very close to

[4]Note that the sample used in our LDV analysis is a subset of the one used in our fixed effects analysis as observations are included in the LDV sample only if we can observe a lagged outcome for the same child.

the ones we get from our preferred model (column 7), suggesting that we can rely exclusively on the experimentally induced variation in exposure to treated neighbors to estimate the non-cognitive spillover effects.

In Tables D.3 and D.4, we perform a similar exercise with the sample we use in our individual-fixed effects analysis. We start with a stripped-down specification in column (1), which includes no controls, and we gradually add controls as we move to column (6), which includes block-group fixed effects, time fixed effects, cohort fixed effects, race fixed effects and gender fixed effects. In column (7), we present the estimated effects from our fixed effects model for the same sample. Note that our individual fixed-effect model only includes individual- and time-fixed effects as the remaining covariates do not vary within observations from the same child.

The results from the fixed effects sample presented in Table D.3 suggest that the estimated coefficients of cognitive spillover effects become positive and significant as soon as we add block-group fixed effect (column 2) and remain stable as we include additional controls (columns 2-6). These estimates are very similar to the ones we get from our individual fixed effects model presented in column (7).

The findings from our fixed effects sample presented in Table D.4 suggest that the estimated coefficients of non-cognitive spillover effects become significant at one percent level in column (4), when we include block-group, time, and cohort fixed effects and remain stable afterwards. Together, these findings suggest that even for the sample used for our individual fixed effects analysis, our results do not solely rely on the inclusion of individual-fixed effects.

E Sub-tests

The cognitive index was constructed by taking the average of a child's percentile scores on five subtests of the Peabody Picture Vocabulary Test (PPVT), WJ Letter-Word, WJ Spelling, WJ Applied Problems, and WJ Quantitative Concepts (the four subtests from the Woodcock Johnson III Test of Achievement). These scores were then standardized by assessment type (pre, mid, post, etc.), including the entire study population. The non-cognitive index was constructed by taking the average of a child's percent-correct scores in the three subtests of the Blair and Willoughby (Spatial Conflict, Operation Span, and Same Game) and the Preschool Self-Regulation Assessment (PSRA) and standardizing with respect to the assessment type.[5]

To explore which components of the cognitive and non-cognitive measures are more important in generating the spillover effects, we estimate the following fixed-effects model:

[5]For children in kindergarten or older, the Same Game test of Blair and Willoughby was replaced with a variant of Wisconsin Card Sort game.

Table D.1: Mean Effect Sizes on Cognitive Scores, LDV Estimates, Pooled Sample: Gradual Addition of Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
r = 3 km	-0.0066***	0.00651**	0.00487*	0.00538**	0.00545**	0.00557**	0.00419***
	(0.00192)	(0.00198)	(0.00202)	(0.00200)	(0.00196)	(0.00192)	(0.00124)
r = 5 km	-0.0041***	0.00508***	0.00363**	0.00347**	0.00353**	0.00360**	0.00331***
	(0.00113)	(0.00102)	(0.00128)	(0.00126)	(0.00122)	(0.00118)	(0.000850)
r = 7 km	-0.0021*	0.00424***	0.00287**	0.00243*	0.00231*	0.00236*	0.00271***
	(0.00095)	(0.000809)	(0.00110)	(0.00109)	(0.00107)	(0.00105)	(0.000712)
Blockgroup FEs	No	Yes	Yes	Yes	Yes	Yes	Yes
Time FEs	No	No	Yes	Yes	Yes	Yes	Yes
Cohort FEs	No	No	No	Yes	Yes	Yes	Yes
Race FEs	No	No	No	No	Yes	Yes	Yes
Gender FEs	No	No	No	No	No	Yes	Yes
Age at baseline	No	No	No	No	No	No	Yes
LDVs	No	No	No	No	No	No	Yes
Adj. R^2 at 3km	0.0375	0.1811	0.1888	0.1907	0.2053	0.2084	0.6359
Adj. R^2 at 5km	0.0290	0.1825	0.1889	0.1916	0.2062	0.2093	0.6367
Adj. R^2 at 7km	0.0141	0.1818	0.1886	0.1924	0.2070	0.2100	0.6369
Observations	3403	3403	3403	3403	3403	3403	3403

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All regressions include a constant.
Standard errors in parentheses and clustered at the blockgroup level.

Table D.2: Mean Effect Sizes on Non-Cognitive Scores, LDV Estimates, Pooled Sample: Gradual Addition of Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
r = 3 km	0.00700***	0.00992***	0.0126***	0.0111***	0.0108***	0.0109***	0.00704***
	(0.00201)	(0.00166)	(0.00213)	(0.00210)	(0.00203)	(0.00199)	(0.00152)
r = 5 km	0.00522***	0.00631***	0.00943***	0.00860***	0.00823***	0.00827***	0.00592***
	(0.00111)	(0.000872)	(0.00132)	(0.00139)	(0.00137)	(0.00135)	(0.00141)
r = 7 km	0.0044***	0.00513***	0.00795***	0.00751***	0.00741***	0.00744***	0.00538***
	(0.000982)	(0.000829)	(0.00134)	(0.00150)	(0.00146)	(0.00144)	(0.00144)
Blockgroup FEs	No	Yes	Yes	Yes	Yes	Yes	Yes
Time FEs	No	No	Yes	Yes	Yes	Yes	Yes
Cohort FEs	No	No	No	Yes	Yes	Yes	Yes
Race FEs	No	No	No	No	Yes	Yes	Yes
Gender FEs	No	No	No	No	No	Yes	Yes
Age at baseline	No	No	No	No	No	No	Yes
LDVs	No	No	No	No	No	No	Yes
Adj. R^2 at 3km	0.0064	0.0692	0.0733	0.0900	0.1052	0.1061	0.3417
Adj. R^2 at 5km	0.0093	0.0701	0.0759	0.0086	0.1070	0.1079	0.3447
Adj. R^2 at 7km	0.0092	0.0704	0.0762	0.0922	0.1080	0.1088	0.3453
Observations	3403	3403	3403	3403	3403	3403	3403

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All regressions include a constant.
Standard errors in parentheses and clustered at the blockgroup level.

Table D.3: Mean Effect Sizes on Cognitive Scores, Fixed-Effects Estimates, Pooled Sample: Gradual Addition of Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
r = 3 km	-0.00530***	0.00513***	0.00443**	0.00470***	0.00474***	0.00482***	0.00334**
	(0.00155)	(0.00144)	(0.00137)	(0.00129)	(0.00121)	(0.00119)	(0.00101)
r = 5 km	-0.00320**	0.00373***	0.00280**	0.00263**	0.00265***	0.00271**	0.00209**
	(0.000999)	(0.000792)	(0.000874)	(0.000822)	(0.000782)	(0.000764)	(0.000638)
r = 7 km	-0.00130	0.00307***	0.00212**	0.00173*	0.00164*	0.00170***	0.00177***
	(0.000830)	(0.000662)	(0.000752)	(0.000690)	(0.000667)	(0.000667)	(0.000529)
Blockgroup FEs	No	Yes	Yes	Yes	Yes	Yes	No
Time FEs	No	No	Yes	Yes	Yes	Yes	Yes
Cohort FEs	No	No	No	Yes	Yes	Yes	No
Race FEs	No	No	No	No	Yes	Yes	No
Gender FEs	No	No	No	No	No	Yes	No
Individual FEs	No	No	No	No	No	No	Yes
Adj. R^2 at 3km	0.0246	0.1734	0.1823	0.1833	0.1960	0.2014	0.1121
Adj. R^2 at 5km	0.0173	0.1740	0.1821	0.1835	0.1963	0.2017	0.1136
Adj. R^2 at 7km	0.0067	0.1734	0.1816	0.1836	0.1963	0.2017	0.1138
Observations	5208	5208	5208	5208	5208	5208	5208

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All regressions include a constant.
Standard errors in parentheses and clustered at the blockgroup level.

Table D.4: Mean Effect Sizes on Non-Cognitive Scores, Fixed-Effects Estimates, Pooled Sample: Gradual Addition of Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
r = 3 km	0.000793	0.00248	0.00222	0.00510**	0.00485**	0.00490**	0.00780***
	(0.00157)	(0.00154)	(0.00167)	(0.00156)	(0.00154)	(0.00152)	(0.00126)
r = 5 km	0.00123	0.00156*	0.00106	0.00342***	0.00317***	0.00321***	0.00428***
	(0.000807)	(0.000688)	(0.000794)	(0.000842)	(0.000829)	(0.000829)	(0.000778)
r = 7 km	0.00125	0.00144*	0.000883	0.00319***	0.00311***	0.00314***	0.00331***
	(0.000698)	(0.000560)	(0.000690)	(0.000790)	(0.000768)	(0.000764)	(0.000667)
Blockgroup FEs	No	Yes	Yes	Yes	Yes	Yes	No
Time FEs	No	No	Yes	Yes	Yes	Yes	Yes
Cohort FEs	No	No	No	Yes	Yes	Yes	No
Race FEs	No	No	No	No	Yes	Yes	No
Gender FEs	No	No	No	No	No	Yes	No
Individual FEs	No	No	No	No	No	No	Yes
Adj. R^2 at 3km	0.0002	0.0520	0.0518	0.0757	0.0869	0.0887	0.0283
Adj. R^2 at 5km	0.0003	0.0519	0.0523	0.0745	0.0858	0.0877	0.0243
Adj. R^2 at 7km	0.0006	0.0522	0.0530	0.0744	0.0863	0.0881	0.0220
Observations	5208	5208	5208	5208	5208	5208	5208

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All regressions include a constant.
Standard errors in parentheses and clustered at the blockgroup level.

$$Y_{i,t}^k = \beta_0 + \beta_1 N_{i,t|r}^{treated} + \beta_2 N_{i,t|r}^{total} + \gamma_i + \delta_t + \epsilon_{i,t}, \tag{E.1}$$

where $Y_{i,t}^k$ is the standardized score of a child i at time t on subtest k, and $N_{i,t|r}^{treated}$ and $N_{i,t|r}^{total}$ represent the number of treated neighbors and total number of neighbors, as previously defined. We include the time and individual fixed effects, and cluster standard errors at the census-block-group level. Under this specification, β_1 represents the average spillover effect from an additional treated neighbor who resides within a radius r of a child.

Table E.1: Spillover Effects by Subtests

	Cognitive Sub-scores					Non-cognitive Sub-scores				
r (meters)	PPVT	WJL	WJA	WJS	WJQ	OSP	SPAT	PSRA	SAME	CARD
3000	.0025*	.0044***	-.0057***	.0044***	.0018	.0093***	.0142***	.0002	-.0052	.0082***
	(.0014)	(.001)	(.0011)	(.0011)	(.0019)	(.0026)	(.0031)	(.0021)	(.0128)	(.0023)
5000	.0014	.0026***	-.003***	.0031***	.0009	.0055***	.0089***	.0003	-.0105	.0051***
	(.0009)	(.0007)	(.0009)	(.0007)	(.001)	(.0014)	(.0019)	(.0013)	(.009)	(.0013)
7000	.0012*	.002***	-.0026***	.0025***	.0007	.0041***	.0071***	.0001	-.0077	.0045***
	(.0007)	(.0006)	(.0008)	(.0006)	(.0009)	(.0011)	(.0015)	(.001)	(.0078)	(.0011)

Notes: This table presents point estimates for the spillover effects from an additional treated neighbor on a child's standardized scores on each subtest, estimated from equation (E.1). Robust standard errors, clustered at the census-block-group level, are in parentheses;
*** p<0.01, ** p<0.05, * p<0.1

As point estimates for β_1 presented in Table E.1 suggest, the spillovers on cognitive skills are mainly driven by the effects on receptive vocabulary (PPVT), spelling ability (WJS), and the ability to identify letters and words (WJL). The spillovers on non-cognitive skills are driven by the effects on attention (Spatial Conflict), working memory (Operation Span), and attention shifting in children who are kindergarten age or older.

F Estimated effects on the restricted sample

Sections 4 and 5 present the estimated spillover effects using all observations. One potential concern is the role of sorting and whether selection into taking assessments is the factor deriving our findings. We address this concern by estimating the main effects from equations 1 and 2 for a subset of our sample who attended at least five out of the eight possible assessments. Our data includes 1,792 observations from 313 children who attend a minimum of five assessments. Note these children represent under 20% of of the total number of children in our pooled sample. Table F.1 presents the point estimates. Our estimates are similar in magnitude to the ones from the whole sample. The estimates on non-cognitive spillover effects are especially close to the ones presented in Tables

4 and 6. The fixed-effects estimates of β_1 on cognitive spillover effects from the whole sample for neighborhood radii of 3, 5, and 7 kilometers are 0.0033σ, 0.0021σ, and 0.0018σ, respectively, whereas the corresponding estimates from the restricted sample are 0.0040σ, 0.0021σ, and 0.0018σ. Likewise, the LDV estimates of β_1 on cognitive spillover effects from the whole sample for neighborhood radii of 3, 5, and 7 kilometers are 0.0042σ, 0.0033σ, and 0.0027σ. The corresponding estimates for the restricted subsample are 0.0029σ, 0.0017σ, and 0.0010σ. The fixed effects estimates on non-cognitive spillover effects from the whole sample for neighborhood radii of 3, 5 and 7 kilometers were 0.0069σ, 0.0043σ, and 0.0033σ, whereas the fixed-effects estimates from the restricted sample are 0.0070σ, 0.0036σ, and 0.0027σ. Similarly, the LDV estimates on non-cognitive spillover effects from the whole sample for neighborhood radii of 3, 5, and 7 kilometers are 0.0070σ, 0.0059σ, and 0.0054σ, whereas the corresponding fixed-effects estimates from the restricted sample are 0.0064σ, 0.0050σ, and 0.0045σ.

Table F.1: Estimates of Spillover Effects for the Restricted Sample

	Fixed Effects		LDV	
	Cognitive	Non-cognitive	Cognitive	Non-cognitive
r =3 km	0.0040**	0.0070***	0.0029**	0.0064***
	(0.0011)	(0.0018)	(0.0012)	(0.0019)
r = 5 km	0.0021***	0.0036***	0.0017*	0.0050***
	(0.0007)	(0.0013)	(0.0009)	(0.0015)
r =7 km	0.0018***	0.0027***	0.0010	0.0045***
	(0.0006)	(0.0010)	(0.0007)	(0.0013)
Obs.	1,792	1,792	1400	1400
Unique child	313	313	313	313

Notes: Estimated spillover effects from equations (1) and (2) for a subsample of observations from children who attended a minimum of five out of eight assessments. Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1

G Spatial fade-out

Table G.1: Mean Effect Sizes on Cognitive and Non-cognitive Scores, Fixed-Effects Estimates

	Cognitive Scores			Non-cognitive Scores		
	Pooled	Control	Treatment	Pooled	Control	Treatment
	(1)	(2)	(3)	(4)	(5)	(6)
r= 2 km	0.0050***	0.0049**	0.0021	0.0120***	0.0077***	0.0107***
	(0.0019)	(0.0024)	(0.0016)	(0.0025)	(0.0030)	(0.0023)
r= 3 km	0.0033***	0.0038***	0.0016	0.0078***	0.0069***	0.0064***
	(0.0010)	(0.0012)	(0.0010)	(0.0013)	(0.0015)	(0.0013)
r= 4 km	0.0026***	0.0032***	0.0010*	0.0053***	0.0045***	0.0042***
	(0.0007)	(0.0011)	(0.0006)	(0.0009)	(0.0013)	(0.0010)
r= 5 km	0.0021***	0.0023***	0.0010*	0.0043***	0.0037***	0.0034***
	(0.0006)	(0.0008)	(0.0006)	(0.0008)	(0.0011)	(0.0008)
r= 6 km	0.0019***	0.0022***	0.0010*	0.0038***	0.0031***	0.0032***
	(0.0006)	(0.0008)	(0.0005)	(0.0007)	(0.0010)	(0.0008)
r= 7 km	0.0018***	0.0021***	0.0008*	0.0033***	0.0025***	0.0028***
	(0.0005)	(0.0007)	(0.0005)	(0.0007)	(0.0010)	(0.0007)
r= 8 km	0.0017***	0.0020***	0.0008*	0.0031***	0.0021**	0.0028***
	(0.0005)	(0.0007)	(0.0005)	(0.0006)	(0.0009)	(0.0007)
r= 9 km	0.0015***	0.0018***	0.0007	0.0029***	0.0021**	0.0025***
	(0.0005)	(0.0007)	(0.0005)	(0.00060)	(0.0009)	(0.0007)
r= 10 km	0.0014***	0.0019***	0.0006	0.0027***	0.0020**	0.0023***
	(0.0005)	(0.0006)	(0.0005)	(0.0006)	(0.0009)	(0.0006)
	5,208	2,442	3,074	5,208	2,442	3,074

Notes: Columns 1-3 (4-6) represent the effect of an additional treated neighbor on a child's standardized cognitive (non-cognitive) score. Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table G.2: Mean Effect Sizes on Cognitive and Non-cognitive Scores, LDV Estimates

	Cognitive Scores			Non-cognitive Scores		
r	Pooled	Control	Treatment	Pooled	Control	Treatment
(meters)	(1)	(2)	(3)	(4)	(5)	(6)
r = 2 km	0.0034**	0.0026	0.0015	0.0095***	0.0061**	0.0102**
	(0.0017)	(0.0039)	(0.0028)	(0.0030)	(0.0031)	(0.0044)
r = 3 km	0.0042***	0.0052**	0.0019	0.0070***	0.0059***	0.0067**
	(0.0012)	(0.0025)	(0.0020)	(0.0015)	(0.0018)	(0.0026)
r = 4 km	0.0030***	0.0039**	0.0014	0.0059***	0.0037***	0.0062***
	(0.0007)	(0.0018)	(0.0012)	(0.0013)	(0.0011)	(0.0023)
r = 5 km	0.0033***	0.0038**	0.0021*	0.0059***	0.0040***	0.0060***
	(0.0009)	(0.0016)	(0.0011)	(0.0014)	(0.0013)	(0.0023)
r = 6 km	0.0030***	0.0033**	0.0017	0.0060***	0.0035***	0.0058**
	(0.0007)	(0.0013)	(0.0011)	(0.0016)	(0.0013)	(0.0023)
r = 7 km	0.0027***	0.0034***	0.0015	0.0054***	0.0038***	0.0053**
	(0.0007)	(0.0013)	(0.0010)	(0.0014)	(0.0012)	(0.0022)
r = 8 km	0.0022***	0.0027**	0.0010	0.0052***	0.0035***	0.0050**
	(0.0006)	(0.0011)	(0.0010)	(0.0014)	(0.0012)	(0.0020)
r = 9 km	0.0021***	0.0028**	0.0010	0.0045***	0.0035***	0.0042**
	(0.0006)	(0.0011)	(0.0008)	(0.0013)	(0.0011)	(0.0019)
r = 10 km	0.0021***	0.0026**	0.0011	0.0043***	0.0031***	0.0039**
	(0.0006)	(0.0010)	(0.0008)	(0.0012)	(0.0011)	(0.0018)
Obs.	3,403	1,495	2,093	3,403	1,495	2,093

Notes: Columns 1-3 (4-6) represent the effect from an additional treated neighbor on a child's standardized cognitive (non-cognitive) score. Robust standard errors, clustered at the census-block-group, are level in parentheses; *** p<0.01, ** p<0.05, * p<0.1

H Spillovers from Parent Academy, Preschool, and Cog-X treatments

In section 6, we compared the spillover effects from the Parent Academy treatments, which exclusively offered education program for parents, to the four Pre-K treatments, which offered pre-school programs to children. Two of the four Pre-K treatment groups (Preschool-Plus and Kinderprep), also included parental components, which, compared to Parent Academies, were shorter and not as heavily incentivized. We refer to these two treatments as Cog-X treatments. In this section, instead of estimating the overall spillover effects from Pre-K treatments, we separately estimate the effects from Cog-X to the ones from the preschool treatments, which exclusively targeted children, using the following specification:

$$Y_{i,t} = \beta_0+\beta_{parent}N_{i,t|r}^{Parent}+\beta_{parent\_child}N_{i,t|r}^{Cogx}+\beta_{child}N_{i,t|r}^{Preschool}+\lambda N_{i,t|r}^{total}+\eta Y_{i,t-1}+X_i\alpha+\sigma_b+\mu_c+\delta_t+\epsilon_{i,t}, \tag{H.1}$$

where $N_{i,t,r}^{Parent}$, $N_{i,t,r}^{Cogx}$, and $N_{i,t,r}^{Preschool}$ represent the number of neighbors residing within distance r of a child i who were assigned to Parent Academy, Cog-X , and the two preschool treatments that exclusively targeted children. All other arguments are defined as in section 6. Under the above specifications, β_{parent}, β_{parent_child}, and β_{child} represent the spillover effects from an additional treated neighbor who was assigned to Parent Academies, Cog-X, or the two preschool treatments with no parental components. Table H.1 presents the estimates of $\hat{\beta}_{parent}$, $\hat{\beta}_{parent\_child}$ and $\hat{\beta}_{child}$.

Table H.1: Mean Effect Sizes on Cognitive and Non-cognitive Scores

	Cognitive Scores			Non-cognitive Scores		
	β_{parent} (1)	β_{parent_child} (2)	β_{child} (3)	β_{parent} (4)	β_{parent_child} (5)	β_{child} (6)
r = 3 km	0.0045**	0.0040***	0.0040	-0.0010	0.0105***	0.0132***
	(0.0022)	(0.0013)	(0.0026)	(0.0035)	(0.0023)	(0.0029)
r = 5 km	0.0074***	0.0034***	0.0005	0.0007	0.0067***	0.0079*
	(0.0024)	(0.0009)	(0.0024)	(0.0045)	(0.0015)	(0.0044)
r = 7 km	0.0074***	0.0029***	-0.0006	0.0006	0.0056***	0.0074
	(0.0026)	(0.0007)	(0.0025)	(0.0040)	(0.0013)	(0.0045)
Obs.	3,403	3,403	3,403	3,403	3,403	3,403

Notes: Columns 1-3 (4-6) represent the effect of an additional treated neighbor of each type, on a child's standardized cognitive (non-cognitive) score. Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Our estimates suggest the programs that included parental-education components are likely to generate larger cognitive spillovers than those that exclusively targeted children.[6] Focusing on non-cognitive spillovers, our estimates confirm our previous findings that programs that directly targeted children (Cog-X and preschool treatments) generate significantly larger non-cognitive spillovers than the ones that exclusively targeted parents ($p < 0.10$).

I Total impact evaluation and neighborhood radius

In section 7, we estimated the impact of the intervention under the neighborhood radius of 3 kilometers. Table I.1 presents the corresponding effects for two other neighborhood radii: r=5 and r=7 kilometers. Our estimates of the direct program effects on treatment children ($\hat{\theta}$) are not significantly different across various neighborhood radii.[7]

As we broaden the definition of the neighborhood, the estimated total spillover effect on both the control (SC^N) and treatment (ST^N) children increases. This finding is intuitive, because broadening the neighborhood would allow for neighbors who live farther away to also impact a

[6] The differences in spillover effects are insignificant for $r = 3$ km, but as we broaden neighborhood radii to 5 and 7 kilometers, the difference in spillovers on cognitive skills from Cog-X and Parent Academy to the ones from the preschool treatments become significant ($p < 0.10$).

[7] P-values from the Wald tests of $\hat{\theta}^{Cog}_{r1} = \hat{\theta}^{Cog}_{r2}$ against $\hat{\theta}^{Cog}_{r1} \neq \hat{\theta}^{Cog}_{r2}$ for r1=3K and r2=5K; r1=3K and r2=7K; and r1=5K and r2=7K are 0.28, 0.12, and 0.08. The corresponding p-values for non-cognitive scores ($\hat{\theta}^{Ncog}$) are 0.75, 0.85, and 1.00.

Table I.1: Total Program Impact under r=3, 5, and 7 kilometers, LDV Estimates

r	5 km		5 km		7 km	
$\overline{N_d^{treated}}$	178.13		325.63		422.81	
	Cognitive	Non-cognitive	Cognitive	Non-cognitive	Cognitive	Non-cognitive
$\hat{\theta}$ $(Direct)$	0.108***	0.0456	0.126***	0.0540	0.150***	0.0541
	(0.0393)	(0.0642)	(0.0458)	(0.0731)	(0.0528)	(0.0837)
$\hat{\beta}$	0.0042***	0.0070***	0.0034***	0.0058***	0.0028***	0.0053***
	(0.0017)	(0.0015)	(0.0009)	(0.0014)	(0.0007)	(0.0014)
$\hat{\lambda}$	-0.0002*	0.0001	-0.0002*	0.0001	-0.0002*	0.0000
	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0002)
SC^N	0.75***	1.25***	1.09***	1.90***	1.17***	2.25***
	(0.22)	(0.27)	(0.28)	(0.46)	(0.31)	(0.61)
ST^N	0.71***	1.27***	1.03***	1.91***	1.09***	2.27***
	(0.22)	(0.27)	(0.27)	(0.46)	(0.30)	(0.61)
$Total$	0.82***	1.32***	1.16***	1.97***	1.24***	2.32***
	(0.23)	(0.27)	(0.28)	(0.45)	(0.31)	(0.60)
$Total^{Standard}$	0.06***	0.07**	0.06***	0.07**	0.07***	0.07**
	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
Observations	3,403	3,403	3,403	3,403	3,403	3,403
R-squared	0.659	0.384	0.660	0.387	0.660	0.387

Notes: Estimated coefficients from equation (7) are presented;. N includes all the observations for which we have the lagged cognitive and non-cognitive scores, and other regressors. $\overline{N_{3km}^{treated}} = 178.13$; $\overline{N_{5km}^{treated}} = 325.63$; $\overline{N_{7km}^{treated}} = 422.81$; Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1

child's outcomes. The increases in the total spillover effects result in larger estimates of the total impacts ($Total$) as we increase the neighborhood radius. The average estimated total impact of the intervention ($Total$) on a treatment child's cognitive scores is 0.82σ for $r = 3$ km and increases to 1.16σ and 1.24σ for $r = 5$ km and $r = 7$ km. Likewise, the average estimated total impact of the intervention ($Total$) on the non-cognitive score of a child increases from 1.32σ to 1.97σ and 2.32σ as we increase the radius from 3 to 5 and 7 kilometers. Finally, our estimates for program impacts if we were to ignore spillovers to control children ($Total^{Standard}$) are very similar and not significantly different across various radii.[8]

J Exploring non-linearities in measuring the total spillover effects

We calculated our estimation of the total spillover effects under the assumption of linearity. In this section, we explore whether and how allowing for nonlinearities affects our estimates. We explore

[8] P-values from the Wald tests of $\hat{Total}_{r1}^{Standard-Cog} = \hat{Total}_{r2}^{Standard-Cog}$ against $\hat{Total}_{r1}^{Standard-Cog} \neq \hat{Total}_{r2}^{Standard-Cog}$ for r1=3 km and r2=5 km; r1=3 km and r2=7 km; and r1=5 km and r2=7 km are 0.71, 0.35, and 0.18. The corresponding p-values for non-cognitive scores are 0.99, 0.93, and 0.85.

non-linearities by considering polynomial functional forms of up to degree 3.

$$\begin{aligned} Y_{i,t} &= \alpha_0 + \theta T_i + \alpha_1 N^{treated}_{i,t|r} + \alpha_2 (N^{treated}_{i,t|r})^2 + ... + \\ &\lambda_1 T_i \times N^{treated}_{i,t|r} + \lambda_2 T_i \times (N^{treated}_{i,t|r})^2 + ... + \beta_2 N^{total}_{i,t|r} + \eta Y_{i,t-1} + X_i \alpha + \sigma_b + \delta_t + \mu_c + \epsilon_{i,t} \end{aligned} \tag{J.1}$$

While under the linear specification, the marginal spillover effect from the j-th treated neighbor is given by α_1 for a control child and by $\alpha_1 + \lambda_1$ for a treated child, the corresponding effects under polynomials of degrees 2 and 3 are given by $\alpha_1 + 2\alpha_2 j$ and $\alpha_1 + 2\alpha_2 j + 3\alpha_3 j^2$ for a control child, and $\alpha_1 + \lambda_1 + 2(\alpha_2 + \lambda_2)j$ and $\alpha_1 + 2(\alpha_2 + \lambda_2)j + 3(\alpha_3 + \lambda_3)j^2$ for a treated child. Therefore, the average spillover effects on a control child from all neighbors, using polynomials of degrees 1, 2, and 3, can be calculated as follows:[9]

$$\overline{SC^N_1} = \alpha_1 \overline{N^{tr}}$$

$$\overline{SC^N_2} = \alpha_1 \overline{N^{tr}} + 2\alpha_2 \sum_{j=1}^{\overline{N^{tr}}} j = \alpha_1 \overline{N^{tr}} + \alpha_2 \overline{N^{tr}}(\overline{N^{tr}} + 1)$$

$$\overline{SC^N_3} = \alpha_1 \overline{N^{tr}} + 2\alpha_2 \sum_{j=1}^{\overline{N^{tr}}} j + 3\alpha_3 \sum_{j=1}^{\overline{N^{tr}}} j^2 = \alpha_1 \overline{N^{tr}} + \alpha_2 \overline{N^{tr}}(\overline{N^{tr}} + 1) + 0.5\alpha_3(\overline{N^{tr}} + 1)(2\overline{N^{tr}} + 1).$$

Table J.1 presents the estimated coefficients from equation J.1 for neighborhood radius of $r = 3$ km. Note the coefficients of the quadratic and cubic terms are all insignificant, suggesting our linear specification is an appropriate representation. Our estimates of the average cognitive spillover effects from all treated neighbors become slightly smaller as we add quadratic and cubic terms, but the changes are small. The estimated non-cognitive spillover effects become larger as we move away from the linear specification. However, these increases are small. Overall, we find no strong evidence suggesting the spillover effect from an additional treated neighbor has a non-linear relationship with the number of treated neighbors.

[9] Replacing α_i with $\alpha_i + \lambda_i$ ($\forall i \in \{1,2,3\}$) would give us the corresponding average spillover effects on a treatment child.

Table J.1: Estimates of Spillover Effects under Non-linear Specifications

	Cognitive			Non-cognitive		
	(1)	(2)	(3)	(4)	(5)	(6)
T_i	0.11***	0.08	0.05	0.05	0.12	0.08
	(0.04)	(0.06)	(0.07)	(0.06)	(0.08)	(0.11)
$N_{i,t\|r}^{treated}$	0.0042***	0.0031	0.0023	0.0070***	0.0080***	0.0096***
	(0.0013)	(0.0019)	(0.0019)	(0.0015)	(0.0021)	(0.0028)
$(N_{i,t\|r}^{treated})^2$		1.58e-06	6.10e-06		-1.42e-06	-1.08e-05
		(1.52e-06)	(7.42e-06)		(1.39e-06)	(9.33e-06)
$(N_{i,t\|r}^{treated})^3$			-5.44e-09			1.04e-08
			(9.01e-09)			(1.07e-08)
$T_i \times N_{i,t\|r}^{treated}$	-0.0002*	0.0001	0.0007	0.0001	-0.0008	0.0004
	(0.0001)	(0.0006)	(0.0014)	(0.0002)	(0.0007)	(0.0021)
$T_i \times (N_{i,t\|r}^{treated})^2$		-7.08e-07	-3.34e-06		1.89e-06	-4.41e-06
		(1.12e-06)	(6.72e-06)		(1.33e-06)	(1.04e-05)
$T_i \times (N_{i,t\|r}^{treated})^3$			3.36e-09			8.57e-09
			(8.53e-09)			(1.34e-08)
SC^N	0.75***	0.60*	0.56**	1.24***	1.37***	1.41***
	(0.22)	(0.31)	(0.29)	(0.27)	(0.35)	(0.37)
ST^N	0.71***	0.55*	0.60**	1.27***	1.34***	1.39***
	(0.22)	(0.32)	(0.29)	(0.27)	(0.38)	(0.38)
Total	0.82***	0.63*	0.65**	1.32***	1.45***	1.47***
	(0.23)	(0.34)	(0.28)	(0.27)	(0.38)	(0.37)
Constant	0.469***	0.478***	0.491***	-2.431***	-2.500***	-2.449***
	(0.113)	(0.114)	(0.123)	(0.180)	(0.186)	(0.188)
Observations	3,403	3,403	3,403	3,403	3,403	3,403
R-squared	0.659	0.659	0.660	0.384	0.384	0.385

Notes: Estimated coefficients from equation (J.1) for neighborhood radius of $r = 3$ km. Columns 1 and 4 correspond to the linear specification; 2 and 5 correspond to polynomials of degree 2; columns 3 and 6 correspond to polynomials of degree 3. Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1