DISCUSSION PAPER SERIES

DP18134

NEIGHBORHOOD SPILLOVER EFFECTS OF EARLY CHILDHOOD INTERVENTIONS

John List, Fatemeh Momeni, Michael Vlassopoulos and Yves Zenou

LABOUR ECONOMICS

CEPR

ISSN 0265-8003

NEIGHBORHOOD SPILLOVER EFFECTS OF EARLY CHILDHOOD INTERVENTIONS

John List, Fatemeh Momeni, Michael Vlassopoulos and Yves Zenou

Discussion Paper DP18134
Published 04 May 2023
Submitted 25 April 2023

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

NEIGHBORHOOD SPILLOVER EFFECTS OF EARLY CHILDHOOD INTERVENTIONS

Abstract

This study explores the role of neighborhoods on human capital formation at an early age. We do so by estimating the spillover effects of an early childhood intervention on the educational attainment of a large sample of disadvantaged children in the United States. We document large spillover effects on the cognitive skills of children living near treated children, which amount to approximately 40\% of the direct treatment effects. Interestingly, these spillover effects are localized and decrease with the spatial distance to treated neighbors. We do not find evidence of spillover effects on non-cognitive skills. Perhaps our most novel insight is the underlying mechanisms at work: the spillover effect on cognitive scores is very localized and seems to operate through the child's social network, mostly between treated kids. We do not find evidence that parents' or children's social networks are effective for non-cognitive skills. 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.

JEL Classification: C93, I21, R1

Keywords: Early education, Neighborhood, Field experiment

John List - jlist@uchicago.edu
University Of Chicago

Fatemeh Momeni - fmomeni@uchicago.edu
University Of Chicago

Michael Vlassopoulos - m.vlassopoulos@soton.ac.uk
University of Southampton

Yves Zenou - yves.zenou@monash.edu
Monash University and CEPR

Acknowledgements
The current paper supersedes earlier versions titled: ``The Social Side of Early Human Capital Formation: Using a Field Experiment to Estimate the Causal Impact of Neighborhoods'' and ``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, Steven Durlauf, Justin Holz, Michael Kremer, Thibaut Lamadon, Costas Meghir, Magne Mogstad, Julie Pernaudet, Stephen Raudenbush, Matthias Rodemeier, Juanna Schr{\o}ter 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, Lina Ramirez, and Charles Shi for excellent research assistance.

Neighborhood Spillover Effects of Early Childhood Interventions*

John A. List[†], Fatemeh Momeni[‡], Michael Vlassopoulos[§], Yves Zenou[¶]

April 25, 2023

Abstract

This study explores the role of neighborhoods on human capital formation at an early age. We do so by estimating the spillover effects of an early childhood intervention on the educational attainment of a large sample of disadvantaged children in the United States. We document large spillover effects on the cognitive skills of children living near treated children, which amount to approximately 40% of the direct treatment effects. Interestingly, these spillover effects are localized and decrease with the spatial distance to treated neighbors. We do not find evidence of spillover effects on non-cognitive skills. Perhaps our most novel insight is the underlying mechanisms at work: the spillover effect on cognitive scores is very localized and seems to operate through the child's social network, mostly between treated kids. We do not find evidence that parents' or children's social networks are effective for non-cognitive skills. 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.

Keywords: Early education, social activity, neighborhood, field experiment, spillover effects, non-cognitive skills.
JEL Classification: C93, I21, R1.

*The current paper supersedes earlier versions titled: "The Social Side of Early Human Capital Formation: Using a Field Experiment to Estimate the Causal Impact of Neighborhoods" and "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, Steven Durlauf, 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, Lina Ramirez, and Charles Shi for excellent research assistance.

[†]Department of Economics, University of Chicago, Chicago, IL, USA, and ANU, Canberra, Australia. E-mail: jlist@uchicago.edu

[‡]Department of Economics, University of Chicago, Chicago, IL, USA. E-mail: fmomeni@uchicago.edu

[§]Department of Economics, University of Southampton, UK. E-mail: m.vlassopoulos@soton.ac.uk

[¶]Department of Economics, Monash University, Caulfield, VIC, Australia. E-mail: 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 Jr (1988)

1 Introduction

Human capital theory can be traced back to Mincer (1958), who created the framework to examine the nature and causes of inequality in personal incomes. Empirically, human capital is typically 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 other proxies 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 has been dominated by an empirical and theoretical focus on the individual (Schunk, 2012; 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, 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 achieved 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 (Jackson et al., 2017; Bailey et al., 2018), including those that augment human capital.

With this contribution in mind, our backdrop is a series of early childhood programs that were delivered to low-income families with young children at the Chicago Heights Early Childhood Center

(CHECC) in 2010-12 (Fryer et al., 2015). 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 600 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 explore the interplay between social interactions and human capital formation at an early age. For this, we follow two distinct steps. First, we provide causal evidence of the impact of neighborhoods 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 caused by the random allocation of children into the programs. By doing so, we are able to isolate the role of neighbors on individual outcomes and examine how exogenous variation in treated neighbors' quality affect a child's outcomes. Our second step is 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 cognitive skills from children that have attended the Pre-K program. Our estimates suggest that, on average, each additional treated Pre-K neighbor residing within a 0.5-kilometer radius of a child's home increases that child's cognitive score by 0.058 standard deviations (σ). Given that an average child in our sample has 1.8 Pre-K treated neighbors residing within a 0.5-kilometer radius of her home—and making a (strong) assumption of linearity—we infer that, on average, a child gains 0.10σ in cognitive test scores from spillover effects from their 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 beyond 1km. We do not find evidence of spillover effects on cognitive test scores from children whose parents have attended a Parent Academy program, which is perhaps not too surprising given that we also do not find these programs to have significant direct impact on the children of the treated parents. We also do not find evidence of spillover effects in non-cognitive skills, despite both the Parent Academy and Pre-K interventions having sizeable direct effects on treated children.[1] We discuss further on what these findings indicate about the mechanisms at play.

Fryer et al. (2015) report interesting racial and gender heterogeneity in the treatment effects of the Parent Academy program. For example, through comparing outcomes between treatment and control children, they find that 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 cognitive spillover effects are significantly larger for Hispanics. Focusing on gender, our estimates suggest girls tend to benefit more than boys from cognitive spillovers, although these gender differences are not significant at the conventional levels.

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. Direct exposure to treated children who live in the same neighborhood is a likely mechanism that can generate spatial spillover effects.[2] The second channel is parental interactions. 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; 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.

Comparing the spillover effects from neighbors who were assigned to the parental-education programs with the effects from neighbors who were assigned to the preschool programs allows us to shed light on the mechanisms through which spillover effects operate. Unlike the Pre-K treatments, in the Parent Academies, the goal was to educate parents rather than children; thus, if spillover effects are driven by interactions between parents, we might expect Parent Academy neighbors to generate larger effects

[1] An earlier version of this paper (List et al., 2020) estimated spillover effects from a broader set of CHECC programs using a different identification approach that relied on within individual variation in exposure to treated neighbors caused by the delivery of programs over multiple years. In the current paper, we focus on the first two cohorts of CHECC that underwent the same interventions and randomization approach, rely on experimental (cross-sectional) variation in exposure to treated neighbors for identification of the spillover effects, and consider immediate effects.

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

than Pre-K neighbors.[3] Alternatively, larger spillovers from Pre-K neighbors than from Parent Academy neighbors could imply that the peer-influence channel plays an important role in generating these effects. Our results suggest that cognitive spillover effects mainly operate through children's interactions and mostly between treated kids. For non-cognitive skills, we do not find evidence that either parents' or children's networks are effective in transmitting these skills.

We conclude our analysis by measuring the total impact of the intervention on children's cognitive performance, accounting for the spillover effects. Our estimates suggest that, on average, the intervention increased a child's cognitive test score by 0.36σ. Spillover effects make up a large portion of this total impact: whereas the average direct effect of the intervention on a treated child's cognitive score is 0.26σ, the corresponding indirect effect is 0.10σ. 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.

We view our results as 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.[4] 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 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,b). 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

[3]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. Nevertheless, it does seem reasonable to assume that Parent Academies affect parents *more* than Pre-K treatments do.

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

different approach in that our identification strategy leverages a field experiment that provides 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 children in the neighborhood play an important role in the development of children's cognitive skills.

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; Janssens, 2011; Avitabile, 2012), compliance behavior (Rincke and Traxler, 2011; Boning et al., 2020; Drago et al., 2020), voting behavior (Sinclair et al., 2012; Giné 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. 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; Doepke and Zilibotti, 2017, 2019; Agostinelli et al., 2020; Attanasio et al., 2020b,a; Cotton et al., 2020; Boucher et al., 2023). 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 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, and explore heterogeneity by race and gender. We discuss the mechanisms in section 5. In section 6, 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 7.

2 Program Details

2.1 Overview of treatments

The early childhood interventions that we study in this paper were delivered to low-income families with young children at the Chicago Heights Early Childhood Center (CHECC) between 2010 and 2012. 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, African American 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 the 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, participating families were randomized into either a preschool program (henceforth "Pre-K") or a parental-education program (henceforth "Parent Academy") or a control group.

The Parent Academy was designed to teach parents how 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.[5] 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

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

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.

During the period that we study, besides the Parent Academy, CHECC delivered two preschool programs in which children were treated *directly*. We refer to these programs as Pre-K treatments. These two programs ("Tools", and "Literacy") were nine-month full-day programs delivered during the school year. The curriculum for "Tools" was *Tools of the Mind*, which focuses on improving non-cognitive skills, whereas "Literacy" was based on *Literacy Express*, which focuses on fostering cognitive skills.[6]

2.2 Randomization

Between 2010 and 2012, over 600 children from low-income families in South Side, Chicago, were recruited and randomized into either one of the four treatments or the control group.[7] The randomization took place once per year, at the beginning of each academic year. Table 1 summarizes the randomization schedule for each year of the program. Some children were randomized twice, depending on their age and year of enrollment: those that were aged 3 enrolled in 2010.[8] The yearly randomization schedule created two cohorts of children we refer to by their year of randomization.

2.3 Assessments

Our key outcome measures are children's performances in cognitive and non-cognitive assessments, which were collected 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. Table B.1 in Appendix B presents the assessment schedule for the two cohorts.

[6] For more information on Literacy Express, see `http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=288`. For more information on Tools of the Mind, see `http://toolsofthemind.org`.

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

[8] Of the 333 that appear in Cohort 1 of our analysis sample, 135 were also randomized again in Cohort 2. As a result, some children who were in the control group in 2010 were randomized into a treatment group in 2011. In a few cases, a child who was randomized into a treatment group in 2010 was assigned to a different (or the same) treatment in 2011.

Table 1: Randomization by Year

	Control	Parent Academy	Pre-K	Total
Cohort-1 (2010)	107	97	129	333
Cohort-2 (2011)	129	162	133	424
Total	236	259	262	757
Unique child	220	197	190	607

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

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

3.1 Data

Construction of outcome variables. Our outcome measures are indices generated from standardized test scores on cognitive and non-cognitive assessments.[9] The cognitive assessment included the Peabody Picture Vocabulary Test (PPVT), which assesses verbal ability and receptive vocabulary (Dunn and Dunn, 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

[9]These indices were constructed by Fryer et al. (2015) for the original evaluation of the programs.

et al., 2001).

The non-cognitive component included the Blair and Willoughby Executive Function test (Willoughby et al., 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).

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. To obtain a zero mean and standard deviation of one, these two indices were standardized by the type of assessment (pre-assessment, mid-assessment, etc.), including the entire study population (treated and control) who took that assessment.

Sample. To explore the spatial spillovers on both treatment and control children, we construct an analysis sample that pools observations from children who were randomized into any of the programs, across the two years. Table 2 provides summary statistics on the baseline demographic variables for our analysis sample. Note that the majority (86%) of the children are either African American or Hispanic, and 37% live in families with an annual household income under $35,000.

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.[10] 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 7.7 kilometers (std. dev.= 8.7), and 99.1% 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 of children as *neighbors* if the commuting distance between them is less than "r"

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

Table 2: Baseline Summary Statistics for Analysis Sample

Variable	Share/Mean	Variable	Share/Mean
Gender		Mother's Education	
Female	0.503	Less than high school	.06
Race		Some high school but no diploma	.13
Black	.41	High school diploma	.16
Hispanic	.45	Some college but no degree	.20
White	.13	College degree	.28
Other Race	.01	Other	.05
		Missing Mother's Education	.12
HH Income and Unemployment Benefits		Father's Education	
below 35K	.37	Less than high school	.10
36K-75K	.36	Some high school but no diploma	.13
75K+	.09	High school diploma	.16
Missing Income	.18	Some college but no degree	.14
Receives Unemployment Benefit	.16	College degree	.15
Missing Unemployment Benefit	.21	Other	.07
		Missing Father's Education	.25
Baseline Age (months)	46.2 (7.1)		

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.

kilometers, and we call "r" the *neighborhood radius*. We conduct our analysis for various values of neighborhood radii. We then calculate for each year y of the program the number of treated ($N_{i,y|r}^{Parent}$ and $N_{i,y|r}^{PreK}$) and control ($N_{i,y|r}^{control}$) neighbors of each child i, and define the total number of CHECC neighbors of i as $N_{i,y|r}^{total} = N_{i,y|r}^{Parent} + N_{i,y|r}^{PreK} + N_{i,y|r}^{control}$. Note that to simplify the analysis and retain statistical power, we construct $N_{i,y|r}^{Parent}$ and $N_{i,y|r}^{PreK}$ by pooling neighbors who were assigned to any of the two parent-education programs as *Parent Academy* neighbors, and pooling those who were assigned to any of the two preschool programs as *Pre-K neighbors*.

Table 3 reports summary statistics for $N_{i,y|r}^{treated}, N_{i,y|r}^{Parent}, N_{i,y|r}^{PreK}$ and $N_{i,y|r}^{control}$, and Figure 2 presents histograms of the exposure measure $N_{i,y|r}^{treated}$, for various values of neighborhood radii. Figures C.1 and C.2 In Appendix C present analogous histograms for $N_{i,y|r}^{Parent}$ and $N_{i,y|r}^{PreK}$, respectively.

3.2 Econometric model

To identify the spillover effects of the early childhood interventions that we study, we exploit the experimentally induced variation in our exposure measures. That is, conditional on the total number of

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

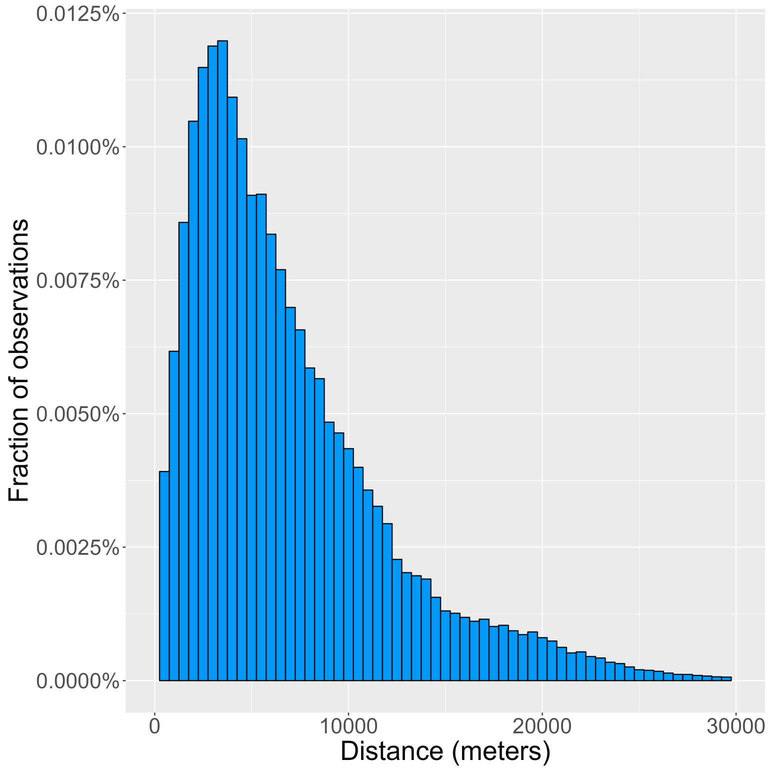


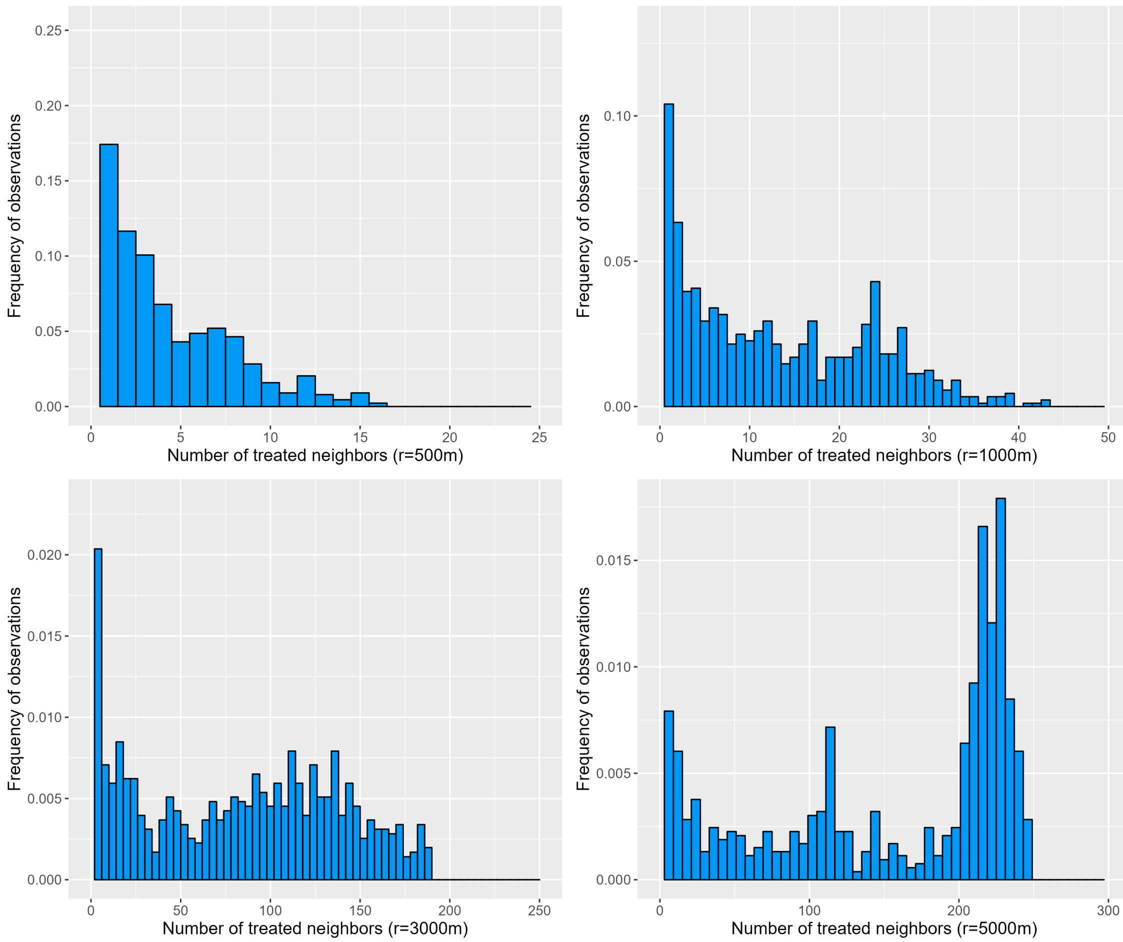
Table 3: Neighbor counts by neighborhood radius

	$r = 0.5$ km	$r = 1$ km	$r = 3$ km	$r = 5$ km	$r = 7$ km
N_r	6.87	23.92	157.15	301.96	414.31
	(6.82)	(21.22)	(115.09)	(171.15)	(183.32)
$N_r^{treated}$	3.31	11.57	76.27	146.88	199.66
	(3.57)	(10.61)	(57.38)	(86.09)	(91.58)
$N_r^{control}$	3.55	12.35	77.14	148.01	205.08
	(4.03)	(12.10)	(63.00)	(95.48)	(106.52)
N_r^{PreK}	1.76	6.11	40.05	76.99	103.34
	(2.16)	(6.20)	(29.40)	(43.47)	(44.52)
N_r^{Parent}	1.56	5.47	36.22	69.89	96.32
	(2.08)	(5.49)	(31.00)	(48.21)	(55.33)

Notes: The average number of neighbors of different treatment status for various definitions of neighborhood radii. Standard deviations are reported in parentheses.

a child's CHECC neighbors at a given point in time ($N_{i,y|r}^{total}$), the number of neighbors who are assigned into treatments ($N_{i,y|r}^{Parent}$ and $N_{i,y|r}^{PreK}$) is determined randomly through the experimental intervention.

Figure 2: Histogram of $N^{treated}_{i,t|r}$ for $r = \{0.5, 1, 3, 5\}$ kilometers.



Formally, we estimate spillover effects using a specification of the following form:

$$Y_{i,a,y} = \beta_0 + \beta_p N^{Parent}_{i,y|r} + \beta_k N^{PreK}_{i,y|r} + \lambda N^{total}_{i,y|r} + \gamma_p T^{Parent}_{i,y} + \gamma_k T^{PreK}_{i,y} + \eta Y_{i,base} + X_i'\delta + \theta_a + \theta_b + \theta_y + \epsilon_{i,a,y}, \quad (1)$$

where $Y_{i,a,y}$ is the standardized cognitive or non-cognitive test score of a child i, on assessment type a, taken in year y. $N^{Parent}_{i,y|r}$ and $N^{PreK}_{i,y|r}$ represent the number of Parent Academy and Pre-K neighbors of a child i who reside within a distance r from i and were randomized into the CHECC program in year y; $N^{total}_{i,y|r}$ represents the total number of i's neighbors who reside within a distance r and who were randomized in the intervention in year y.[11] $T^{Parent}_{i,y}$ and $T^{PreK}_{i,y}$ are indicators capturing whether child i was randomized into a Parent Academy (PA) or Pre-school (Pre-K) treatment in year y. $Y_{i,base}$ is the

[11]As aforementioned, Miguel and Kremer (2004), Giné and Mansuri (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.

baseline test score of child i assessed prior to undertaking treatment, and X_i is a vector of characteristics of the child that includes age, gender, and race. Our specification also includes a series of fixed effects: θ_a for the type of assessment (midline, endline, or summer loss), θ_b the block-group of residence, and θ_y the year of the program (first or second). We also include a dummy variable that takes the value 1 if in year 2 a child is being randomized for the second time. Finally, $\epsilon_{i,a,y}$ is the error term. We cluster standard errors at the cluster-block-group level (139 clusters).

Identification. Random assignment allows us to estimate the direct impact of the intervention on treated children, γ_p and γ_k for the PA and Pre-K programmes, respectively. Our main interest is on identifying spillover effects, meaning that our key parameters of interest in (1) are β_p and β_k, which capture the average effect of moving one of the control neighbors of a child i to treatment group PA and Pre-K, respectively, holding the total number of i's CHECC neighbors constant. Identification rests on the cross-sectional variation in the number of treated neighbors. That is, our identifying assumption is that conditional on the total number of neighbors $N^{total}_{i,y|r}$, variation in $N^{Parent}_{i,y|r}$ and $N^{PreK}_{i,y|r}$ is exogenous. This is because in the experimental design, assignment to the different treatments is random. To provide support for our identifying assumption, we perform a series of balance tests in which we regress our key regressors $N^{Parent}_{i,y|r}$ and $N^{PreK}_{i,y|r}$ on observable characteristics of students (baseline test scores, gender, race, and age) controlling for the total number of neighbors $N^{total}_{i,y|r}$ and cohort fixed effects. In Table D.1 in Appendix D, we report the estimated coefficients on these characteristics, considering exposure to treated neighbors at radii equal to 0.5, 1 and 3 km. For each r, we report 10 tests. For $r = 0.5km$, only one coefficient is significant at 10%, for $r = 1km$ two coefficients at 5%, and for $r = 3kn$ one coefficient is significant at 5%. When we further add block of residence fixed effects (θ_b), then only one out of the 30 tests yields a significant coefficient at 10%. As a further test of our identifying assumption, in Section 4.4 we report a placebo test that is based on permuting the set of treated neighbors and comparing the estimates of the placebo regressions with our baseline estimates.

Channels. Note that a child i may benefit from a PA neighbor k through two channels. The first channel is the parents' social network: k's parents may influence the behavior and decisions 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 PA with i's parents) or preference externalities between parents. The second channel is peer influence: if PA 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. Thus, although $\hat{\beta}_p$ might reflect spillovers through both the parents' and the child's social networks, $\hat{\beta}_k$ is more likely to reflect an effect that is mainly driven by direct interactions between children.

4 Results

4.1 Main findings

Table 4 presents estimated coefficients that capture direct treatment and spillover effects from equation (1) for neighborhood radii of 0.5, 1, and 3 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 results on standardized cognitive (non-cognitive) test scores.

Column (1) reveals a significant positive *direct* treatment effect on the cognitive score of Pre-K of 0.256σ, whereas the effect of PA is smaller (0.048σ) and statistically insignificant. We also find a significant positive *spillover effect on cognitive test scores*: an additional Pre-K treated neighbor within 0.5 km of a child's home increases their cognitive score by 0.058σ ($p < 0.01$). On the other hand, the spillover effect of PA on cognitive scores is smaller and statistically insignificant, which is perhaps not surprising since the direct effect of this treatment on cognitive scores was found to be small and insignificant. In Section 6, we provide an illustration of the size of the Pre-K spillover effect and how it compares to the direct treatment effect.

In columns (2) and (3), we estimate less localized spillovers that can occur in neighborhoods with radii of 1 and 3 kilometers, respectively. What we find is that the Pre-K spillover effects are now much smaller and statistically insignificant.

The effects on *non-cognitive scores* are presented in columns (4)-(6) and display a different picture. We find significant and similar in magnitude direct treatment effects of both Parent Academy and Pre-K interventions (0.22σ). However, we find no significant spillover effects for either type of treatment and any neighborhood size, with the exception of a marginally significant Pre-K effect of 0.016σ at $r = 3km$.

Table 4: Direct & Spillover Effects on Cognitive and Non-cognitive Scores

	Cognitive Scores			Non-cognitive Scores		
	r = 0.5 km	r = 1 km	r = 3 km	r = 0.5 km	r = 1 km	r = 3 km
	(1)	(2)	(3)	(4)	(5)	(6)
N^{Parent}	0.015	-0.005	0.007	-0.009	0.001	-0.015
	(0.024)	(0.016)	(0.007)	(0.031)	(0.016)	(0.011)
N^{PreK}	0.058***	0.009	0.003	0.016	-0.012	0.016*
	(0.016)	(0.007)	(0.005)	(0.017)	(0.008)	(0.009)
T^{Parent}	0.048	0.045	0.053	0.218***	0.217***	0.205***
	(0.072)	(0.070)	(0.067)	(0.065)	(0.063)	(0.065)
T^{PreK}	0.256***	0.264***	0.261***	0.220***	0.209**	0.241***
	(0.078)	(0.083)	(0.082)	(0.082)	(0.085)	(0.083)
Obs.	1,660	1,660	1,660	1,638	1,638	1,638

Notes: Treatment and Spillover effects estimated from equation 1 are presented. Columns 1-3 (4-6) present results 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

It is useful to see whether our results are sensitive to the inclusion of the various controls. In Table F.1 in Appendix F, 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 preferred specification. The results presented in Table F.1 suggest that the estimated coefficients of *cognitive* spillover effects become positive and significant as soon as we add block-group fixed effects (column 2) and remain stable as we include additional controls (columns 2-6). Clearly, in column (9) of Table F.1, we obtain the same results as in column (1) of Table 4.

In sum, we document significant positive spillover effects on cognitive test scores arising from neighbors that have been assigned to a Pre-K program in neighborhood of 0.5km radius. The larger cognitive spillovers from Pre-K neighbors than from PA neighbors suggest that direct social interactions between children (rather than between parents) play an important role in generating the spillover effects. The lack of spillover effects on non-cognitive skills, despite the interventions yielding sizeable direct effects, suggests that these skills are not easily transmitted through social interactions among very young children or among parents. Yet, the richness of the data permits us to explore deeper into both the nature and extent of spillovers.[12]

[12]To evaluate the sensitivity of our findings to different choices of sample and estimation approach, in Appendix E we present some comparative analysis that draws on an earlier version of the paper (List et al., 2020). In that version, spillover

4.2 Spatial fade-out

A closer examination of the estimated spillover effects on cognitive scores 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 0.5 to 1 and then to 3 kilometers. In Figure 3, we further explore the pattern and shed light on the relationship between spillover effects and distance. In particular, we show the estimated β_p's and β_k's for a broad range of r's between 0.5 and 1km.

We find that the spillover effects on cognitive scores from Pre-K neighbors operate very locally. Indeed, as we increase the neighborhood radius, the marginal spillover effects from an additional Pre-K neighbor decrease. Because a larger neighborhood radius corresponds to a longer average distance to neighbors, the negative relationship between the spillover effect and radius r implies that, as the distance between a child and her treated neighbor grows, the spillover effect on cognitive scores weakens. This suggests that the direct interactions in terms of cognitive skills between kids operate at a very local level (up to 1 km) and disappear beyond this distance.

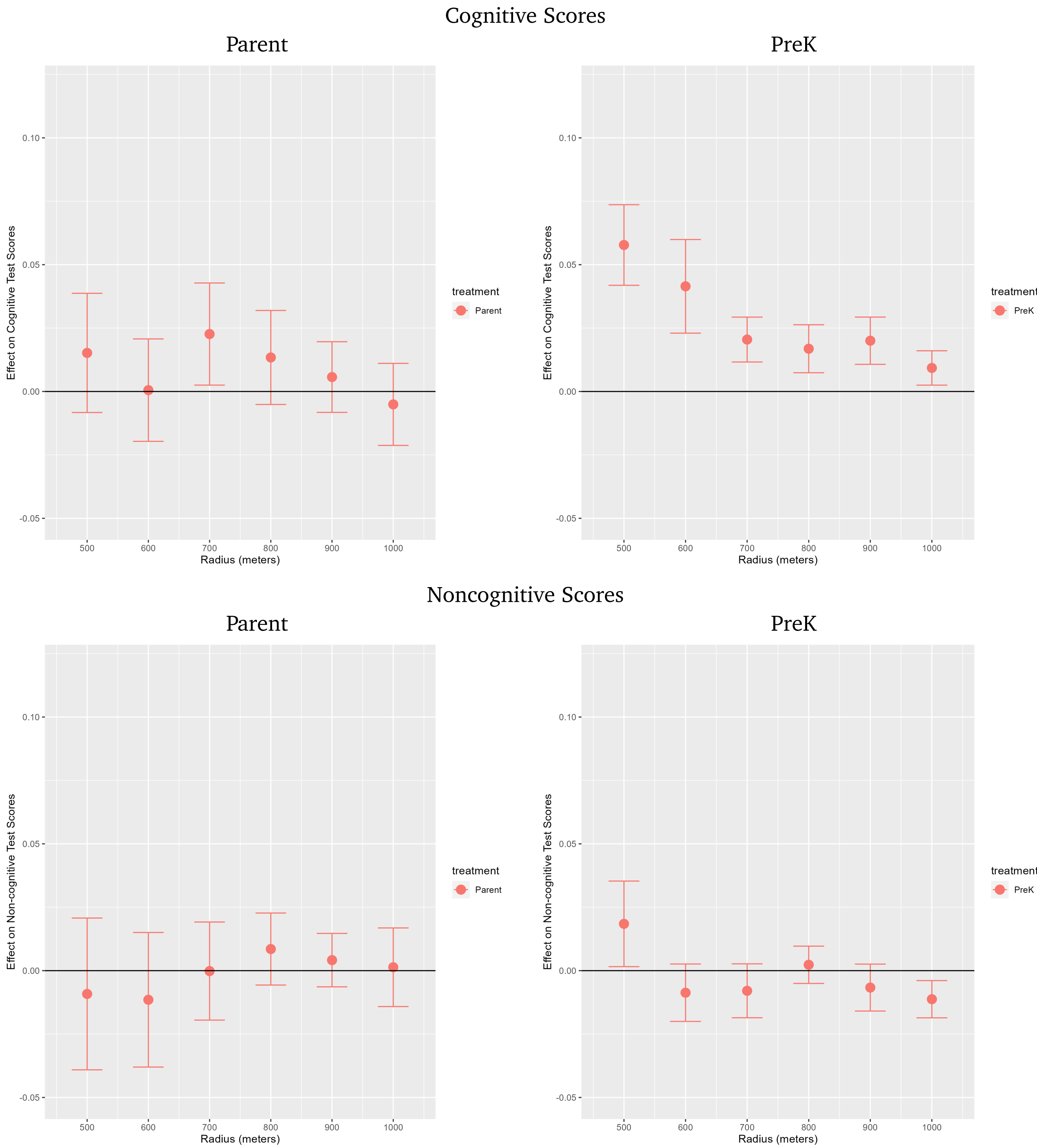
4.3 Heterogeneous effects

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.[13] We do so by estimating equation (1), separately by race and gender. We also examine heterogeneity by baseline ability of the children to assess whether there is an ability gradient to a child's gains from spillovers. In all cases we focus on neighborhoods of r=0.5km, which is the area in which the previous analysis indicates that the spillover effects occur.

effects were estimated from a broader set of CHECC programs, including interventions that were implemented in later years (2012-2014), which are evaluated by Fryer Jr et al. (2020). The previous version used a different econometric approach and sample, and also included assessments that took place after the completion of the CHECC program, when available. The analysis presented in Appendix E suggests that spillover effects estimated in later cohorts of CHECC children that are exposed to different programs are larger and less localized, particularly when available post-program assessments are included. Hence, in this paper we restrict attention to estimating the within-year spillover effects of programs that were implemented in the first phase of CHECC.

[13] Fryer et al. (2015) only evaluated the *direct* effect of the parent-education component of CHECC (Parent Academy) for the 2011 cohort by comparing the outcomes of treatment and control children, under the assumption that treatments 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. This is consistent with the evidence reported in Table 4 using a different specification and sample. 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. This is consistent with the evidence in Table 5, except we do not find a significant direct effect of PA on cognitive scores for Hispanics. Parent Academy was also reported to have slightly larger effects on girls than boys, although the gender differences were not significant.

Figure 3: Spatial fade-out



Notes: Error bars are standard errors.

Gender. Columns (1) and (2) in Table 5 contain the estimated effects by gender. Girls appear to benefit more from spillovers in cognitive scores from Pre-K neighbors than boys (0.064σ vs 0.042σ), though the difference is not statistically significant ($p = 0.53$). On the other hand, for non-cognitive skills, the spillover effects are larger for boys than girls, albeit not statistically significant. For girls, we even estimate a negative spillover effect from neighbors that have been randomised into a parent academy (-0.113; $p = 0.079$). We note also that the direct effect of both treatments is significant for boys, whereas girls seem to only benefit from PA and the effect size is smaller and less precisely estimated than that of boys. That is, it seems that girls' noncognitive test scores are less responsive to the direct and spillover effects of these early childhood programs.

Race. Since African American and Hispanic children make up over 90% of our sample, our analysis of heterogeneity along race focuses on these two groups. Table 5 presents estimation results separately for African American and Hispanic children. Starting with cognitive skills, for African Americans, we find positive direct treatment effects and spillover effects from both type of treated neighbors, however, none of these effects are statistically significant. For Hispanics, we find a large direct effect (0.45σ) and a sizeable spillover effect (0.044σ) of Pre-K, and no significant effects from the PA programs. On noncognitive skills, we find substantial direct treatment effects of both programs for Hispanics, but no spillover effects. For African Americans, we find no significant direct effects and a negative but not very precisely estimated spillover effect from neighbors exposed to PA.

Table 5: Direct and Spillover Effects on Cognitive and Non-cognitive Scores: Heterogeneity by Gender and Race

	Cognitive Scores				Non-cognitive Scores			
	Boys	Girls	Blacks	Hispanics	Boys	Girls	Blacks	Hispanics
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
N^{Parent}	0.005	0.024	0.053	0.009	0.043	-0.044*	-0.113*	0.010
	(0.049)	(0.027)	(0.041)	(0.024)	(0.048)	(0.025)	(0.060)	(0.034)
N^{PreK}	0.042*	0.064**	0.068	0.044**	0.045	0.017	-0.050	0.027
	(0.023)	(0.031)	(0.042)	(0.021)	(0.027)	(0.035)	(0.057)	(0.025)
T^{Parent}	0.166	-0.001	0.081	0.092	0.258**	0.181*	-0.021	0.296***
	(0.114)	(0.103)	(0.149)	(0.109)	(0.118)	(0.100)	(0.152)	(0.061)
T^{PreK}	0.334***	0.259**	0.071	0.449***	0.283**	0.086	-0.098	0.355***
	(0.117)	(0.099)	(0.131)	(0.111)	(0.138)	(0.103)	(0.213)	(0.072)
Obs.	830	830	669	770	815	823	659	761

Notes: Spillover effects from each additional treated neighbor (r=0.5km) estimated from equation (1) are presented. Columns 1-4 (5-8) 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.

Baseline Ability. Finally, we investigate heterogeneity by baseline cognitive and non-cognitive ability. We do so by splitting the sample into two groups on the basis of whether they lie above or below the median of the baseline cognitive and non-cognitive test scores of our sample. Results are presented in Table 6. What emerges is that above median children benefit more from spillovers from Pre-K neighbors than below median children (0.087 versus 0.048) though the difference is not statistically significant ($p = 0.32$). On the other hand, below median children experience larger direct treatment effects from Pre-K treatment than above median children (0.40σ versus 0.23σ), though again the difference is not statistically distinguishable ($p = 0.34$).

Table 6: Direct and Spillover Effects on Cognitive and Non-cognitive Scores: Heterogeneity by Baseline Ability

	Cognitive		Non-Cognitive	
	Below Median	Above Median	Below Median	Above Median
	(1)	(2)	(3)	(4)
N^{Parent}	0.001	0.039	0.005	-0.029
	(0.032)	(0.058)	(0.041)	(0.033)
N^{PreK}	0.048**	0.087**	0.026	0.009
	(0.021)	(0.038)	(0.027)	(0.017)
T^{Parent}	0.113	0.093	0.195	0.194*
	(0.085)	(0.117)	(0.122)	(0.106)
T^{PreK}	0.396***	0.228	0.251	0.213
	(0.108)	(0.137)	(0.174)	(0.141)
Observations	822	838	823	815

Notes: Spillover effects from each additional treated neighbor (r=0.5km) estimated from equation (1) are presented. Columns 1-2 (3-4) 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.

4.4 Additional Results and Robustness Checks

Alternative Specifications. In this section, we present some additional results aiming to explore further the spillover effects at $r = 0.5$ km, the level at which our baseline results indicate that spillovers operate.

First, we estimate alternative specifications that relax the assumption of linearity in the spillover effects imposed in equation (1). These results are presented in Table 7. In columns (1) and (2), we estimate a specification where instead of capturing spillovers *linearly*, we construct indicators D^{Parent} and D^{PreK} that denote whether a child has any treated Parent and Pre-K neighbors, respectively. The coefficients on these dummies reveal whether there is a spillover effect on the extensive margin. What we find is that for cognitive scores, there is no significant effect of having a PA neighbor, whereas having a Pre-K neighbor raises test scores by 0.133σ ($p < 0.1$). This specification offers a convenient way to compare the spillover to the direct effect, which in this specification is estimated to be 0.263σ ($p < 0.01$). This indicates that having at least one Pre-K neighbor has half the impact on cognitive test scores compared to being directly treated by the program. For non-cognitive scores, the estimates are not statistically significant. In columns (3) and (4), we estimate a specification that captures any effects of having multiple neighbors nonparametrically. In particular, we introduce two indicators for each treatment, one which turns on when a child has 1 treated neighbor (N_1^{Parent} and N_1^{PreK}), and one that turns on when a child has at least 2 treated neighbors ($N_{>1}^{Parent}$ and $N_{>1}^{PreK}$). We find that for cognitive skills, the spillover effect from having at least 2 neighbors is larger than having one (0.187σ versus 0.101σ), though the difference between these two coefficients is not statistically significant (p=0.22). These results suggest some tendency for spillover effects on cognitive test scores to be larger in neighborhoods that are more densely populated with treated neighbors. For noncognitive test scores, we do not find this pattern, as the coefficients on N_1^{PreK} and $N_{>1}^{PreK}$ are very similar in value and not statistically significant.

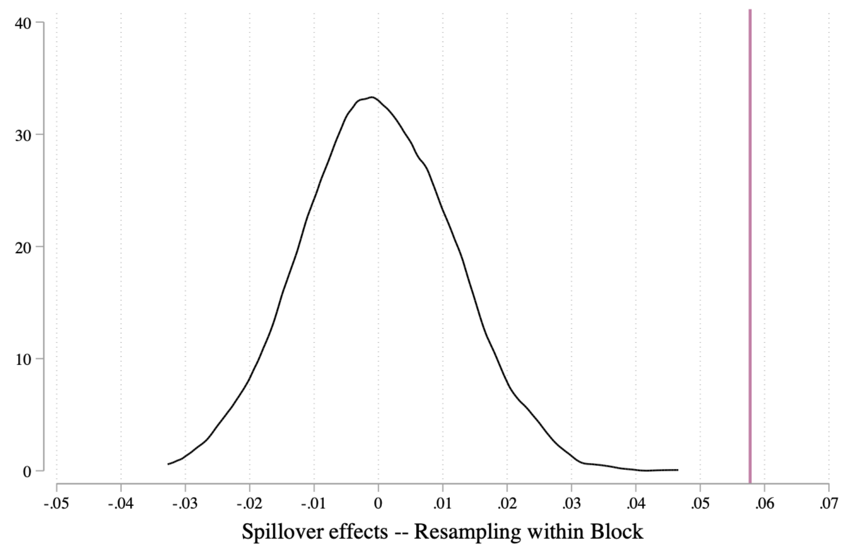
Table 7: Spillover Effects: Alternative Specifications

	Cognitive (1)	Non-Cognitive (2)		Cognitive (3)	Non-Cognitive (4)
D^{Parent}	0.010	-0.070	N_1^{Parent}	0.032	-0.025
	(0.071)	(0.084)		(0.078)	(0.091)
D^{PreK}	0.133*	0.095	$N_{>1}^{Parent}$	-0.004	-0.132
	(0.078)	(0.064)		(0.082)	(0.093)
			N_1^{PreK}	0.101	0.097
				(0.078)	(0.073)
			$N_{>1}^{PreK}$	0.187*	0.095
				(0.098)	(0.077)
Observations	1,660	1,638		1,660	1,638

Notes: D^{Parent} and D^{PreK} are indicators for whether a child has a Parent or Pre-K neighbor (r=0.5km), respectively; N_1^{Parent} and N_1^{PreK} are indicators for having 1 Parent or Pre-K neighbor, respectively; ($N_{>1}^{Parent}$ and $N_{>1}^{PreK}$) are indicators for having at least 2 Parent or Pre-K neighbors, respectively. All columns include controls for treatment, gender, race, age, type of assessment, cohort, and randomization round. Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Placebo test. To provide further support for our identifying assumption that exposure to treated neighbors is quasi-random, we conduct a placebo test. In this test, we permute the value of the number of treated neighbors with values from other children within the same census-block-group. We then estimate equation (1) on 2,000 iterations of this exercise and obtain a set of placebo estimates that we compare with our baseline estimates of β_k reported in Table 4. The outcome of this exercise, shown in Figure 4, indicates that no iteration produced a larger estimate than our baseline estimates, which provides reassurance that our estimates are uncovering spillover effects.

Figure 4: Placebo test



Notes: The graph presents the kernel density of estimated β_ks obtained from estimations of the regression model (1) using permutations of N^{PreK} following the procedure described in Section 4.4. N=2,000 permutations.

Robustness checks. In Appendices G, H, and I we report some further robustness checks. First, in Table G.1, we explore an alternative approach for studying fade-out of spillover effects in which we parameterize the distribution of treated neighbors by creating bands, that is, the number of treated neighbors within 0.5 km, the number of treated neighbors between 0.5 km and 1 km, and the number of treated neighbors between 1 and 5 km, for $N^{Parent}_{i,y|b}$, $N^{PreK}_{i,y|b}$ and $N^{total}_{i,y|b}$ (where $N^{Parent}_{i,y|b}$, $N^{PreK}_{i,y|b}$ and $N^{total}_{i,y|b}$ 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 cognitive skills, the spillover effect is always significant from treated neighbors within the first band (that is, within 0.5 kilometers) and becomes insignificant as we move to treated neighbors who reside in the band that is placed 0.5 km further away from a child's home. For the non-cognitive skills, we find alternating signs in the first two bands that are significant at 10% and of roughly equal magnitude in absolute terms.

Second, in Table H.1, we report another robustness check by examining an alternative way of defining the number of treated neighbors for children who remained in the program for 2 years. For these children, in their second year, we take into account the total number of treated neighbors accumulated over the two years. Our results are robust to this alternative definition of treated neighbors.

Finally, in Table I.1, we check for the robustness of our results to including children that have missing baseline test scores and age. To do this, we impute the missing values to zero and add an indicator variable to capture the missingness. Our results are robust to the inclusion of these children into the analysis.

5 Exploring the Mechanisms

One attractive feature of our field experiment is that it generates data that has a unique variation to explore the underlying mechanisms at work for our observed spillovers. In this section, we leverage randomization in treatment and variation in treatment type, alongside distance to treated and control neighbors to investigate the underlying mechanisms.

Spillover Effects: Neighborhood or Classroom? In Section 4, we documented spillover effects for a child having neighbors who attended the Pre-K program. These spillover effects could arise because of interactions among children that take place at the neighborhood level or in the educational setting where the CHECC program was delivered. To investigate this issue, we parse the spillover effects by treatment assignment. The idea is that since children in the control group did not attend the CHECC programs, then these children can only benefit from spillover effects that occur in the neighborhood. Treated children (or their parents in the case of PA) on the other hand, could benefit either through interactions that happen in the neighborhood or in the settings where the CHECC sessions took place.

In Table 8, we present estimates of our baseline equation (1) by treatment status at neighborhood radius of 0.5 km, since this is where our previous estimates suggested that the spillover effects occurred. These results reveal that only *treated* children experience statistically significant benefits in terms of cognitive scores from living close to Pre-K treated families. In the control group, Pre-K cognitive spillover effects are also positive, but are not precisely estimated; nevertheless, the difference between the two groups is not significant at conventional levels.[14] Columns (3) and (4) report the spillover effects in non-cognitive scores by treatment assignment. These estimates illustrate that both treatment and control children do not benefit from non-cognitive spillovers. The estimated spillover effects on non-cognitive scores on the control group are larger but not statistically significant.

Table 8: Spillover Effects: Treatment versus Control

	Cognitive Scores		Non-cognitive Scores	
	Treated	Control	Treated	Control
	(1)	(2)	(3)	(4)
N^{Parent}	0.007	0.042	-0.016	0.009
	(0.024)	(0.056)	(0.029)	(0.070)
N^{PreK}	0.062***	0.052	-0.005	0.048
	(0.020)	(0.051)	(0.018)	(0.044)
Observations	1,220	440	1,208	430

Notes: Spillover effects from each additional treated neighbor (r=0.5km) estimated from equation (1) are presented. All columns include controls for gender, race, age, type of assessment, cohort, and randomization round. Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

To gain further insight, Figure 5 shows the spillover effect from having an additional treated neighbor on a child's standardized cognitive test score as a function of neighborhood radius, separately for treated and control children. For treated children, we see a clear pattern of diminishing spillover effects as we consider larger neighborhoods, which is consistent with the presence of localized neighborhood effects. For control children, there is a drop in the size of the spillover effect when we look at wider neighborhoods but the effects are more imprecise.

In summary, this evidence suggests that it is likely that the spillover effects we estimate include a significant neighborhood component and that treated children interact mostly with other treated children both in the classroom and in the local neighborhood.

6 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

[14]The p-value of a Wald test of equal Pre-K spillover effects for treatment and control group is 0.84.

Figure 5: Spillover Effects on a Child's Cognitive Test Score: Treatment versus Control

Notes: Error bars are standard errors.

the total impact, we also disentangle the direct and indirect (spillover) effects from the intervention. In what follows, to fix ideas and simplify the presentation of our results, we set the neighborhood radius to 0.5 kilometers.

The total impact of the intervention ($Total$) on a child i who was randomized into one of the treatments (PA 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): $Total = Direct + Spillover$.

We evaluate the total impact of the intervention by using the estimates of equation (1) reported in column (1) of Table 4. Under this specification, γ_p and γ_k represent the average direct effect of the two types of interventions ($Direct$), while β_p and β_k represent the average spillover effects from an additional PA and Pre-K treated neighbor, respectively. Assuming linearity, the average spillover effect from Parent and Pre-K treated neighbors can be expressed as $(\overline{N}_r^{PA} \times \hat{\beta}_p)$ and $(\overline{N}_r^{PreK} \times \hat{\beta}_k)$, where $\overline{N}_r^{PA}$ and $\overline{N}_r^{PreK}$ denote the average number of PA and Pre-K treated neighbors who reside within distance r of a child.

Focusing on the full analysis sample, we find the average direct effect of being randomly assigned to a Pre-K program on a child's standardized cognitive scores is 0.256σ. The average total spillover effects on standardized cognitive scores are estimated to be $1.76 \times 0.058 = 0.102$. Therefore, 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 0.358σ. Note that the spillover effects are about 40% of the direct treatment effects and 28.5% of the overall effects on treated children. This finding implies that if one were to treat a single child in isolation, the average treatment effects would be about 60% of the estimated impacts in the presence of spillovers from other treated children. It also implies that by focusing entirely on the direct effects of the intervention on the treated, we would be neglecting an important additional component, which are the indirect benefits that accrue due to spillovers.

7 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). These findings have heightened debate among academics over the cost effectiveness of Head Start (e.g. Barnett, 2011; Gibbs et al., 2011; 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.[15] 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

[15] 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).

education intervention by causally estimating neighborhood effects. In doing so, we provide unique causal evidence on how neighbors influence children's outcomes. Compared to the previous literature, our approach provides a new glimpse into such effects since we compare children 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 spillovers that are economically significant: on average, they increase cognitive test scores of a child by 0.10σ and benefit both genders. For non-cognitive skills, we do not find evidence of spillover effects, despite the intervention having substantial direct effects on treated children. This suggests that non-cognitive skills are difficult to be enhanced through social interactions at this age. 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 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 preschool population that had broad coverage across Chicago

Heights. 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

Agostinelli, F., Doepke, M., Sorrenti, G., and Zilibotti, F. (2020). It takes a village: the economics of parenting with neighborhood and peer effects. Technical report, National Bureau of Economic Research.

Al-Ubaydli, O., List, J. A., LoRe, D., and Suskind, D. (2017a). Scaling for economists: Lessons from the non-adherence problem in the medical literature. *Journal of Economic Perspectives*, 31(4):125–44.

Al-Ubaydli, O., List, J. A., and Suskind, D. (2020). 2017 Klein lecture: The science of using science: Toward an understanding of the threats to scalability. *International Economic Review*, 61(4):1387–1409.

Al-Ubaydli, O., List, J. A., and Suskind, D. L. (2017b). What can we learn from experiments? Understanding the threats to the scalability of experimental results. *American Economic Review*, 107(5):282–86.

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

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

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

Attanasio, O., Meghir, C., and Nix, E. (2020b). Human capital development and parental investment in India. *Review of Economic Studies*, 87(6):2511–2541.

Avitabile, C. (2012). Spillover effects in healthcare programs: Evidence on social norms and information sharing. Technical report, IDB Working Paper Series.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., and Wong, A. (2018). Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives*, 32(3):259–80.

Barnett, W. S. (2011). Effectiveness of early educational intervention. *Science*, 333(6045):975–978.

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

Boning, W. C., Guyton, J., Hodge, R., and Slemrod, J. (2020). Heard it through the grapevine: The direct and network effects of a tax enforcement field experiment on firms. *Journal of Public Economics*, 190:104261.

Boucher, V., Del Bello, C. L., Panebianco, F., Verdier, T., and Zenou, Y. (2023). Education transmission and network formation. *Journal of Labor Economics*, 41(1):129–173.

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

Chetty, R. and Hendren, N. (2018b). The impacts of neighborhoods on intergenerational mobility ii: County-level estimates. *The Quarterly Journal of Economics*, 133(3):1163–1228.

Chetty, R., Hendren, N., and Katz, L. F. (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.

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

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94:S95–S120.

Cotton, C., Hickman, B. R., List, J. A., Price, J., and Roy, S. (2020). Productivity versus motivation in adolescent human capital production: Evidence from a structurally-motivated field experiment. Technical report, National Bureau of Economic Research.

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

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

Cunha, F., Heckman, J. J., and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.

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

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

Doepke, M. and Zilibotti, F. (2019). *Love, Money, and Parenting: How Economics Explains the Way we Raise our Kids*. Princeton: Princeton University Press.

Drago, F., Mengel, F., and Traxler, C. (2020). Compliance behavior in networks: Evidence from a field experiment. *American Economic Journal: Applied Economics*, 12(2):96–133.

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

Dunn, L. M. and Dunn, L. M. (1965). Peabody picture vocabulary test.

Durlauf, S. N. (2004). Neighborhood effects. *Handbook of Regional and Urban Economics*, 4:2173–2242.

Edin, P.-A., Fredriksson, P., and Åslund, O. (2003). Ethnic enclaves and the economic success of immigrants—evidence from a natural experiment. *The Quarterly Journal of Economics*, 118(1):329–357.

Epple, D. and Romano, R. E. (2011). Peer effects in education: A survey of the theory and evidence. In *Handbook of Social Economics*, volume 1, pages 1053–1163. Elsevier.

Ferracci, M., Jolivet, G., and van den Berg, G. J. (2014). Evidence of treatment spillovers within markets. *Review of Economics and Statistics*, 96(5):812–823.

Fryer, R. G., Levitt, S. D., List, J. A., et al. (2015). Parental incentives and early childhood achievement: A field experiment in chicago heights. Technical report, National Bureau of Economic Research.

Fryer Jr, R. G., Levitt, S. D., List, J. A., and Samek, A. (2020). Introducing cogx: A new preschool education program combining parent and child interventions. Technical report, National Bureau of Economic Research.

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

Gibbs, C., Ludwig, J., and Miller, D. L. (2011). Does head start do any lasting good? Technical report, National Bureau of Economic Research.

Giné, X. and Mansuri, G. (2018). Together we will: Experimental evidence on female voting behavior in Pakistan. *American Economic Journal: Applied Economics*, 10(1):207–35.

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

Hanushek, E. A. (2020). Education production functions. In *The Economics of Education*, pages 161–170. Elsevier.

Heckman, J. J. (2008). Schools, skills, and synapses. *Economic Inquiry*, 46(3):289–324.

Ioannides, Y. M. (2011). Neighborhood effects and housing. In *Handbook of Social Economics*, volume 1, pages 1281–1340. Elsevier.

Jackson, M. O., Rogers, B. W., and Zenou, Y. (2017). The economic consequences of social-network structure. *Journal of Economic Literature*, 55(1):49–95.

Janssens, 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.

Jonassen, D. H. (2004). *Learning to solve problems: An instructional design guide*, volume 6. John Wiley & Sons.

Kalil, A. (2015). Inequality begins at home: The role of parenting in the diverging destinies of rich and poor children. In *Families in an era of increasing inequality*, pages 63–82. Springer.

Katz, L. F., Kling, J. R., and Liebman, J. B. (2001). Moving to opportunity in Boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics*, 116(2):607–654.

Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., and Borghans, L. (2014). Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success.

Kline, P. and Walters, C. R. (2016). Evaluating public programs with close substitutes: The case of head start. *The Quarterly Journal of Economics*, 131(4):1795–1848.

Kling, J. R., Ludwig, J., and Katz, L. F. (2005). Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment. *The Quarterly Journal of Economics*, 120(1):87–130.

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

Leventhal, T. and Brooks-Gunn, J. (2000). The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. *Psychological bulletin*, 126(2):309.

List, J. A. (2020). Non est disputandum de generalizability? A glimpse into the external validity trial. Technical report, National Bureau of Economic Research.

List, J. A., Momeni, F., and Zenou, Y. (2020). The social side of early human capital formation: Using a field experiment to estimate the causal impact of neighborhoods. Technical report, National Bureau of Economic Research.

Lucas Jr, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42.

Miguel, E. and Kremer, M. (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1):159–217.

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

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

Muralidharan, K., Niehaus, P., and Sukhtankar, S. (2017). General equilibrium effects of (improving) public employment programs: Experimental evidence from india. Technical report, National Bureau of Economic Research.

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., Jenkins, F., Fletcher, P., Quinn, L., Friedman, J., et al. (2010). Head start impact study. final report. *Administration for Children & Families*.

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

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the Economics of Education*, volume 3, pages 249–277. Elsevier.

Schunk, D. H. (2012). *Learning theories an educational perspective sixth edition*. pearson.

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

Smith-Donald, R., Raver, C. C., Hayes, T., and 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.

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

Topa, G. and Zenou, Y. (2015). Neighborhood and network effects. In *Handbook of Regional and Urban Economics*, volume 5, pages 561–624. Elsevier.

Waldfogel, J. and Washbrook, E. (2011). Early years policy. *Child development research*, 2011.

Willoughby, M. T., Wirth, R., and Blair, C. B. (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment*, 24(2):418.

Woodcock, R. W., McGrew, K. S., Mather, N., et al. (2001). *Woodcock-Johnson III tests of achievement*. Riverside Publishing Company Itasca, IL.

Neighborhood Spillover Effects of Early Childhood Interventions

Appendix

By John A. List[1], Fatemeh Momeni[2], Michael Vlassopoulos[3], and Yves Zenou[4]

[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, University of Southampton, UK. E-mail: m.vlassopoulos@soton.ac.uk
[4]Department of Economics, Monash University, Caulfield, VIC, Australia, and IFN. Email: yves.zenou@monash.edu.

A Maps

Figure A.1: Map of Home Locations

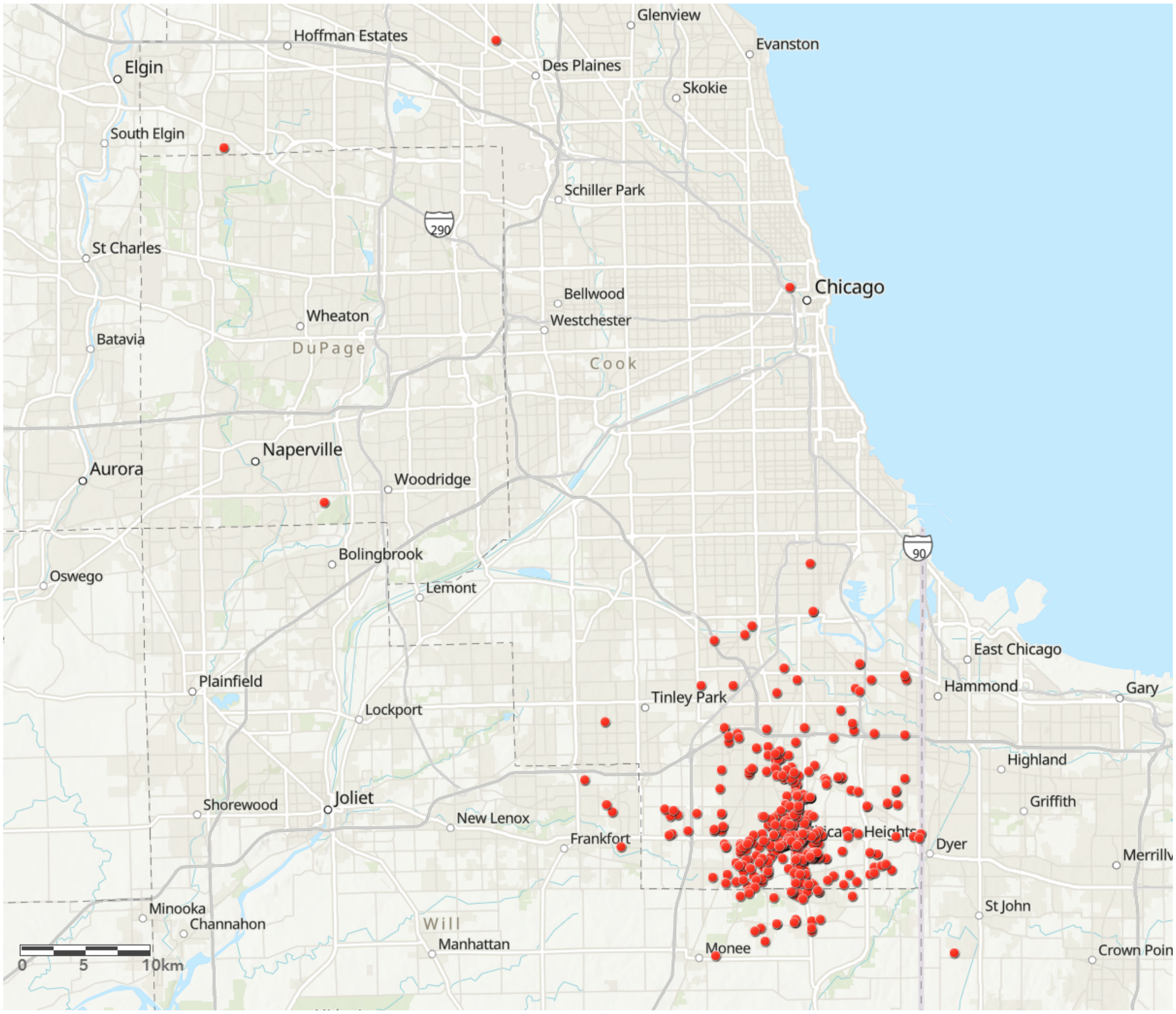
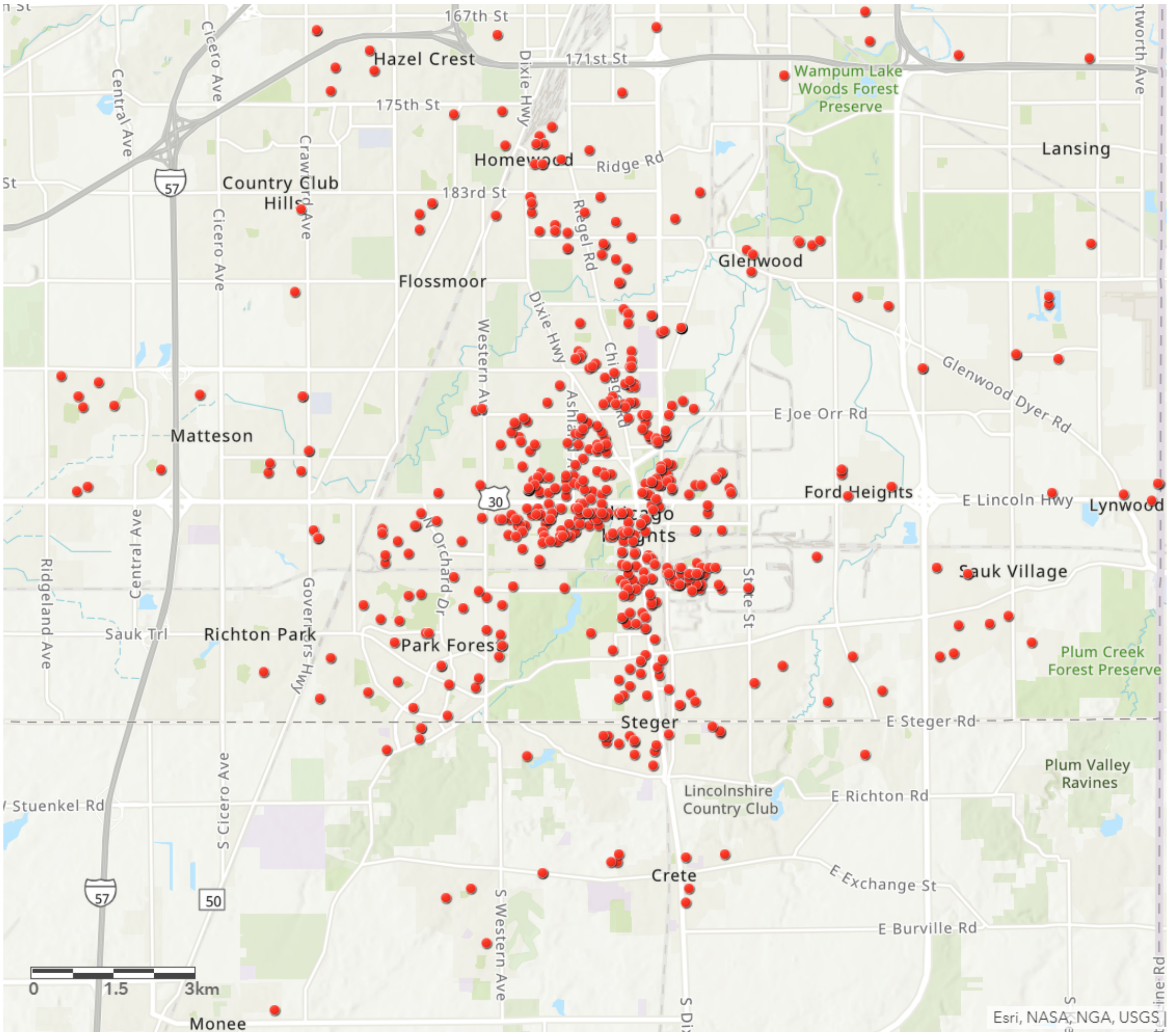


Figure A.2: Map of Home Locations, Zoomed in



B Assessment schedule

Table B.1: Assessment Calendar

	Winter (Jan-Feb)	Late Spring (May)	Late Summer (Jul-Sep)
2010			Pre^{2010}
2011	Mid^{2010}	$Post^{2010}$ Pre^{2011}	SL^{2010}
2012	Mid^{2011}	$Post^{2011}$	SL^{2011}

Notes: Superscripts are cohort identifiers (randomization years).
Pre= pre assessment; Mid= mid assessment; Post= post assessment; SL= summer loss assessment.

C Additional Figures

Figure C.1: Histogram of $N^{PA}_{i,t|r}$ for $r = \{0.5, 1, 3, 5\}$ kilometers.

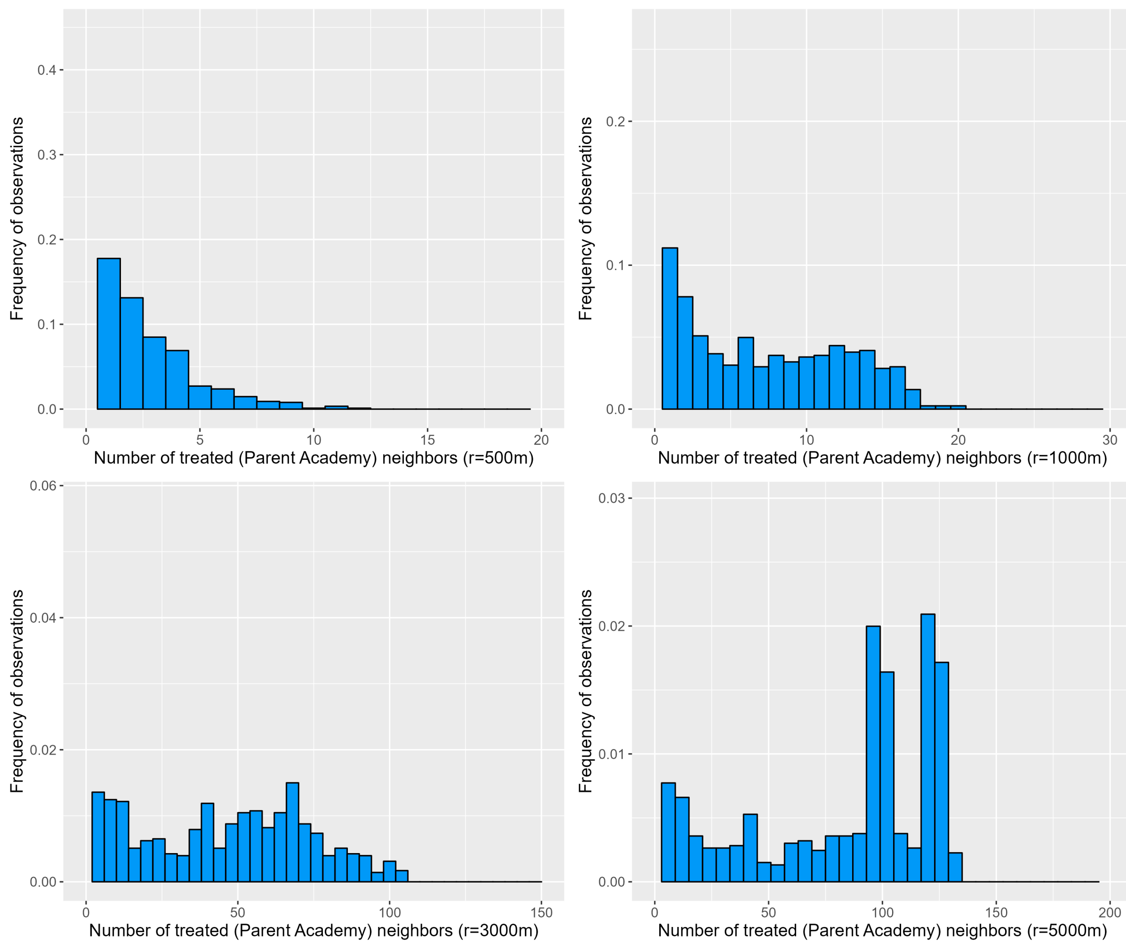
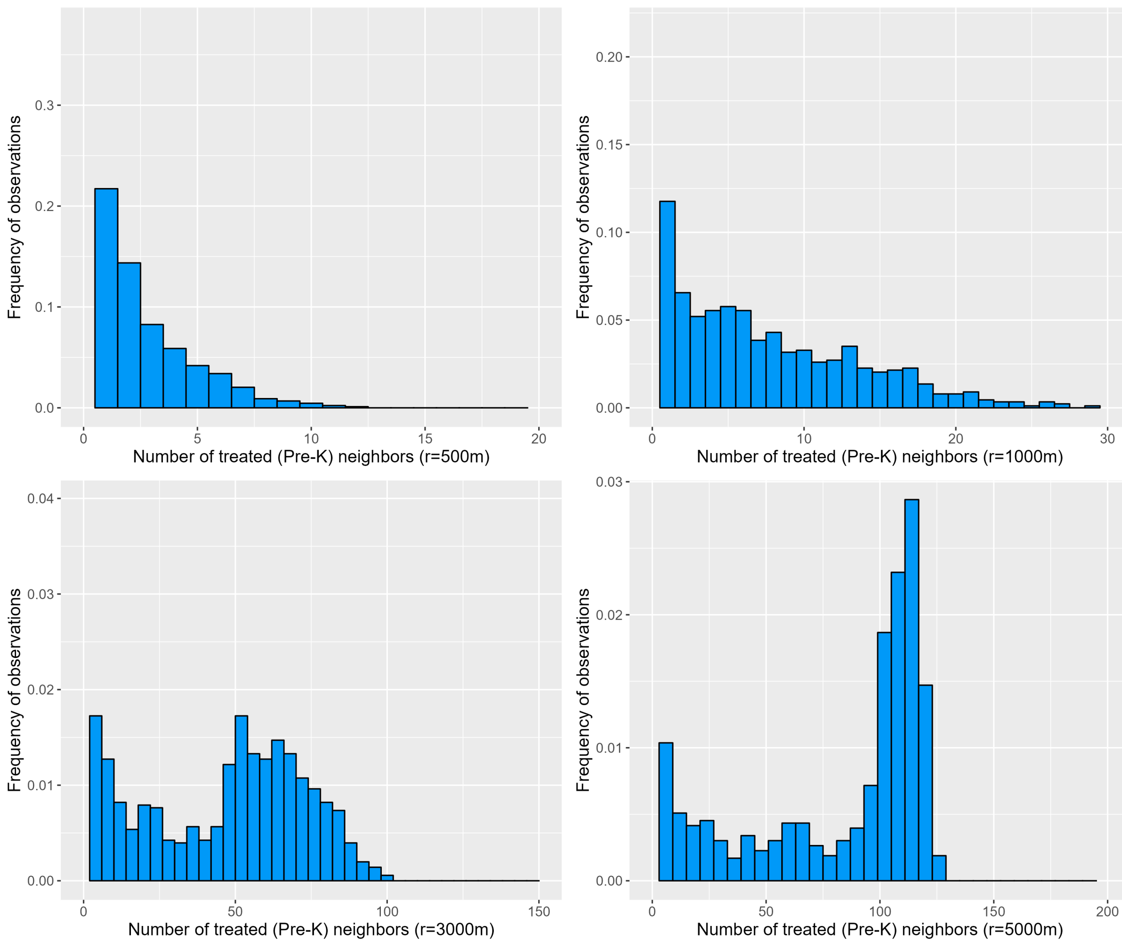


Figure C.2: Histogram of $N_{i,t|r}^{PreK}$ for $r = \{0.5, 1, 3, 5\}$ kilometers.



D Balancing Tests

Table D.1: Balancing Tests

	N^{PA}			N^{PreK}		
	0.5 km	1 km	3 km	0.5 km	1 km	3 km
Y_{base}^{cog}	-0.025	-0.072	-0.056	0.012	-0.140	-0.517**
	(0.042)	(0.076)	(0.126)	(0.044)	(0.098)	(0.203)
Y_{base}^{Noncog}	-0.022	-0.083	-0.158	0.025	0.126	0.060
	(0.042)	(0.080)	(0.111)	(0.046)	(0.093)	(0.193)
Gender	0.080	-0.139	-0.173	0.042	0.241	0.281
	(0.092)	(0.163)	(0.240)	(0.094)	(0.205)	(0.395)
African American	-0.157*	-0.518***	0.184	0.091	0.538**	-0.218
	(0.088)	(0.166)	(0.248)	(0.088)	(0.210)	(0.433)
Age	-0.004	-0.018	-0.041**	0.003	0.018	0.037
	(0.007)	(0.012)	(0.018)	(0.007)	(0.015)	(0.029)

Notes: The Table reports estimated coefficients β on regressions of the following form: $N_r^T = \beta X + \gamma N^{total} + \theta_y + \epsilon$, where $T = \{PA, PreK\}$ and $r = \{0.5, 1, 3\}$. Each row presents a different X; *** p<0.01, ** p<0.05, * p<0.1

E LDV specification and sensitivity analysis

In this section, we present some analysis that draws on an earlier version of the paper (List et al., 2020), which estimated spillover effects from a broader set of CHECC programs using a different identification approach and also considered assessments that took place in later years after the children had completed the program, when available. Specifically, we consider the lagged dependent variable (LDV) specification estimated in this earlier work that takes the form:

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

where $N_{i,t|r}^{treated}$ and $N_{i,t|r}^{total}$ are the cumulative number of treated and total neighbors as of time t and radius r. This specification includes lagged cognitive and non-cognitive test scores ($Y_{i,t-1}$), and 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.

In Table E.1, we report estimated β_1's obtained from estimating equation E.1 under different sample choices to provide some insight into how sensitive the results are to these choices. Note that List et al. (2020) also present a fixed-effects specification that yields similar findings to those of the LDV model. Here, we focus on the LDV specification because it allows us to evaluate the sensitivity of the results to removing age-out assessments, which the fixed-effects specification is unable to deliver.

In particular, column 1 of Table E.1 replicates the results reported in column 1 of Table 6 in List et al. (2020), in which spillover effects on cognitive scores are found to be significant for radii beyond 2km. Column 2 removes the assessments that occur in later years (age-out assessments) to focus on the immediate effects, as we do in the current paper. This leads to spillover effects becoming significant at r=0.5 km. In column 3, we restrict the sample to consider cohorts 1 & 2 of CHECC, as we do in the current paper. In column 4, we restrict to cohorts 1 & 2, and immediate outcomes. As can be seen, the spillover effects are now only significant at r=0.5 km. For comparison, in column (7) we estimate our baseline specification (eq 1) on the same sample as in column (4) and obtain similar patterns. Finally, in columns 5 and 6, we estimate the LDV model only on Cohorts 3 & 4, with (column 5) and without (column 6) age out assessments. As can be seen, in column 5, spillovers are significant at larger radii (r=3km or r=5km) and sensitive to inclusion of the age-out assessments.

This analysis suggests that our baseline results are not sensitive to the choice of specification (LDV or equation (1)). The difference between the earlier results replicated in column 1 of Table E.1 and our baseline results seem to be driven by the fact that spillover effects on cognitive scores for CHECC children that are exposed to later programs are larger and less localized, especially when age-out assessments are included, when available.

Table E.1: LDV specification, Sensitivity Analysis - Cognitive skills

	(1) List et al. (2020)	(2) No age-out	(3) Cohorts 1&2	(4) Cohorts 1&2 No age-out	(5) Cohorts 3&4	(6) Cohorts 3&4 No age-out	(7) Sample Col(4) Eq.1
$N^{treated}$	0.00667	0.01743**	0.02313*	0.03375***	-0.00794	-0.00787	0.05753***
(r=0.5 km)	(0.00745)	(0.00767)	(0.01176)	(0.01223)	(0.00772)	(0.00980)	(0.01753)
$N^{treated}$	0.00312	0.00542	0.00174	0.00156	-0.00350	-0.00145	-0.00010
(r=1 km)	(0.00342)	(0.00355)	(0.00363)	(0.00593)	(0.00443)	(0.00543)	(0.00989)
$N^{treated}$	0.00341**	0.00439*	-0.00648**	-0.00602	0.00310	0.00317	-0.00029.
(r=2 km)	(0.00169)	(0.00229)	(0.00254)	(0.00614)	(0.00250)	(0.00384)	(0.00622)
$N^{treated}$	0.00419***	0.00490***	-0.00231	-0.00241	0.00543***	0.00306	-0.00329
(r=3 km)	(0.00124)	(0.00166)	(0.00269)	(0.00497)	(0.00170)	(0.00273)	(0.00606)
$N^{treated}$	0.00331***	0.00470***	0.00209	0.00642	0.00410***	0.00300*	-0.00008
(r=5 km)	(0.00085)	(0.00107)	(0.00234)	(0.00454)	(0.00121)	(0.00173)	(0.00456)
N	3403	2396	1543	1311	1860	1085	1314

Each row presents estimated β_1's obtained from estimating equation E.1 at different neighborhood radii r.
Column 1 replicates the results reported in column 1 of Table 6 in List et al. (2020).
Column 2 removes the assessments that occur in later years (age-out assessments).
Column 3 restricts the sample to cohorts 1 & 2 of CHECC. Column 4 restricts to cohorts 1 & 2, and within year outcomes.
Columns 5 and 6 restrict to Cohorts 3 & 4, with (column 5) and without (column 6) age out assessments.
Column (7) presents the baseline specification (eq 1) on the same sample as in column (4).
Standard errors clustered at census-block-group level, are presented in parentheses.

For completeness, Table E.2 presents the analogous analysis for non-cognitive scores. We find that our insights regarding the sensitivity of spillover effects to sample choices are consistent with those obtained for cognitive scores. Specifically, we observe that focusing on Cohorts 3 & 4 yields significant and large spillover effects on non-cognitive skills beyond a radius of 2km when age-out assessments are included in the analysis. Focusing on Cohorts 1 & 2 and within-year assessments presents the same picture of no significant spillover effects on non-cognitive scores under both specifications (see columns 4 and 7).

Table E.2: LDV specification, Sensitivity Analysis - Non-cognitive skills

	(1) List et al. (2020)	(2) No age-out	(3) Cohorts 1&2	(4) Cohorts 1&2 No age-out	(5) Cohorts 3&4	(6) Cohorts 3&4 No age-out	(7) Sample Col(4) Eq.1
$N^{treated}$	-0.00355	0.00305	0.00574	0.01233	-0.01289	-0.00310	0.01617
(r=0.5 km)	(0.00697)	(0.00655)	(0.01279)	(0.01862)	(0.01257)	(0.01174)	(0.02175)
$N^{treated}$	0.00445	0.00097	-0.00763*	-0.00774	0.00172	0.00549	-0.00492
(r=1 km)	(0.00287)	(0.00419)	(0.00420)	(0.00789)	(0.00394)	(0.00659)	(0.00910)
$N^{treated}$	0.00952***	0.00529*	-0.00518	-0.00291	0.01291**	0.00877	-0.00206
(r=2 km)	(0.00295)	(0.00317)	(0.00386)	(0.00917)	(0.00521)	(0.00597)	(0.01215)
$N^{treated}$	0.00704***	0.00273	0.00469	0.00889	0.00646**	0.00097	0.00673
(r=3 km)	(0.00152)	(0.00207)	(0.00519)	(0.01178)	(0.00292)	(0.00291)	(0.01135)
$N^{treated}$	0.00592***	0.00116	0.00780*	0.00290	0.00518**	0.00030	0.00110
(r=5 km)	(0.00141)	(0.00137)	(0.00400)	(0.00675)	(0.00213)	(0.00234)	(0.00626)
N	3403	2396	1543	1311	1860	1085	1314

Each row presents estimated β_1's obtained from estimating equation E.1.
Column 1 replicates the results reported in column 1 of Table 6 in List et al. (2020).
Column 2 removes the assessments that occur in later years (age-out assessments).
Column 3 restricts the sample to cohorts 1 & 2 of CHECC. Column 4 restricts to cohorts 1 & 2, and within year outcomes.
Columns 5 and 6 restrict to Cohorts 3 & 4, with (column 5) and without (column 6) age out assessments.
Column (7) presents the baseline specification (eq 1) on the same sample as in column (4).
Standard errors clustered at census-block-group level, are presented in parentheses.

F 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 Table F.1, we start with a stripped-down specification in column (1), which includes no controls, and we gradually add controls until we reach column (9), which corresponds to our main specification displayed in column (1) of Table 4.

Table F.1: Cognitive Test Scores, Main Effects: Gradual Addition of Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
N^{Parent}	-0.006	-0.006	-0.001	0.006	0.007	0.008	0.012	0.017	0.015
	(0.029)	(0.029)	(0.026)	(0.026)	(0.027)	(0.026)	(0.024)	(0.025)	(0.024)
N^{PreK}	-0.012	0.051**	0.039**	0.052***	0.053***	0.052***	0.054***	0.057***	0.058***
	(0.027)	(0.025)	(0.016)	(0.017)	(0.017)	(0.017)	(0.017)	(0.016)	(0.016)
Block-group FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline			Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort				Yes	Yes	Yes	Yes	Yes	Yes
Assessment					Yes	Yes	Yes	Yes	Yes
Gender						Yes	Yes	Yes	Yes
Race							Yes	Yes	Yes
Age								Yes	Yes
Random									Yes
N	1,660	1,660	1,660	1,660	1,660	1,660	1,660	1,660	1,660

Random is an indicator for whether a child has been randomized twice. Standard errors clustered at block-group level, are presented in parentheses.

G Bands

Table G.1: Cognitive Test Scores, Bands

	(1) Cog Score	(2) NonCog Score
$N^{Parent}_{0.5}$	0.019	0.006
	(0.026)	(0.033)
$N^{Parent}_{1-0.5}$	-0.001	0.007
	(0.019)	(0.026)
N^{Parent}_{5-1}	0.006	0.007
	(0.008)	(0.013)
$N^{PreK}_{0.5}$	0.062***	0.027*
	(0.021)	(0.016)
$N^{PreK}_{1-0.5}$	-0.014	-0.022*
	(0.012)	(0.012)
N^{PreK}_{5-1}	0.003	0.001
	(0.004)	(0.005)
$TotalN_{0.5}$	-0.024**	-0.011
	(0.012)	(0.015)
$TotalN_{1-0.5}$	0.009	0.000
	(0.008)	(0.006)
$TotalN_{5-1}$	-0.002	0.000
	(0.002)	(0.003)
Parent Academy	0.051	0.217***
	(0.071)	(0.064)
Pre-K	0.261***	0.201**
	(0.079)	(0.084)
Baseline Cog Score	0.712***	
	(0.035)	
Baseline NonCog Score		0.371***
		(0.039)
Observations	1,660	1,638

Notes: All columns include controls for gender, race, age, type of assessment, cohort, and randomization round. Standard errors clustered at block-group level, are presented in parentheses.

H Cumulative Definition of Neighbors

Table H.1: Spillover Effects: Cumulative

	Cognitive Scores	Non-cognitive Scores
N^{Parent}	0.011	-0.006
	(0.022)	(0.029)
N^{PreK}	0.047***	0.013
	(0.017)	(0.017)
T^{Parent}	0.045	0.217***
	(0.073)	(0.064)
T^{PreK}	0.255***	0.219***
	(0.080)	(0.082)
Observations	1,660	1,638

Notes: All columns include controls for gender, race, age, type of assessment, cohort, and randomization round. Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1

I Imputing missing values

Table I.1: Imputing missing baseline scores and age

	(1) Cognitive Scores	(2) Non-cognitive Scores
N^{Parent}	0.026	0.014
	(0.022)	(0.032)
N^{Pre-K}	0.059***	0.024
	(0.017)	(0.015)
T^{Parent}	0.073	0.235***
	(0.066)	(0.064)
T^{Pre-K}	0.278***	0.213***
	(0.075)	(0.068)
Observations	1,815	1,803

Notes: All columns include controls for gender, race, age, type of assessment, cohort, and randomization round. Robust standard errors, clustered at the census-block-group level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1