Symposium Title: Chicago School Readiness Project: Unpacking the Long-Run Effects of an Early Childhood Intervention

Organizer: Tyler W. Watts, New York University

Symposium Justification (500 words)
 The proposed symposium contains three presentations that each employ a unique approach to understanding how early childhood interventions affect long-term developmental trajectories. Together, the papers address a critical gap in the intervention literature, as few recent studies have examined the long-run impacts of a randomly assigned early intervention. The classic work on the Perry Preschool and Abecedarian programs have shown that early intervention can produce positive effects on development that last through adulthood (e.g., Campbell et al., 2012). But, more recent early interventions have typically yielded positive short-run effects followed by medium-term fadeout during childhood (see Bailey et al., 2016). Despite this discouraging recent evidence, early childhood programs are scaling up at a remarkable rate around the country (e.g., New York), making the need for rigorous early childhood program evaluations with long-run follow-up even more pressing.
 The current symposium contains newly-estimated long-term results drawn from the Chicago School Readiness Project (CSRP), an early childhood program that was implemented in inner-city Head Start centers in 2004-2005. Eighteen centers operating in high poverty Chicago neighborhoods were recruited for study participation, and these centers were grouped into pairs based on observable characteristics. Random assignment then occurred within these blocked pairs.
 The program was designed to improve the quality of inner-city Head Start classrooms. "Treatment" teachers were offered pedagogical development (PD) designed to help them assist students in their behavioral and emotional regulation. Teachers were also provided access to mental health consultants, who supported teachers' efforts to implement the intervention, and they also worked with teachers to relieve stress and burnout. Centers assigned to the control group continued with "business as usual," but control group teachers were provided with part-time teacher aides to balance the student to teacher ratio changes caused in the treatment group.
 Raver and colleagues (2009; 2011) found that the program reduced children's behavior programs and boosted measures of cognitive skills at the end of preschool, but effects largely faded during elementary school. In this symposium, Paper 1 presents newly-estimated impacts on adolescent measures of executive function (EF), emotion regulation and academic achievement, all measured 10 years after the program ended. Preliminary results indicate that the program had surprising positive effects on adolescent measures of EF and GPA. Papers 2 and 3 then offer unique approaches to identifying the mechanisms through which these effects might have occurred. Paper 2 uses a multi-level growth curve modeling approach to investigate whether changes in early childhood skills mediate long-run program effects on achievement. Conversely, Paper 3 investigates the impact of the program on children's school choices, and finds that the program may have compelled children to select into higher quality schools following preschool.
 The symposium will feature comments from a discussant who is a world-renowned expert in program evaluation, and he will speak to the various methodological approaches employed by the papers. Although the findings of these papers are preliminary, they each have

potential to substantially contribute to our knowledge of how early childhood programs can affect long-run development in a policy-relevant context.

References

Bailey, D., Duncan, G., Odgers, C., & Yu, W. (2016). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness*, *10*(1), 7-39.

Campbell, F. A., Pungello, E. P., Burchinal, M., Kainz, K., Pan, Y., Wasik, B. H., ... & Ramey, C. T. (2012). Adult outcomes as a function of an early childhood educational program: an Abecedarian Project follow-up. *Developmental Psychology*, *48*(4), 1033.

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Metzger, M. W., & Solomon, B. (2009). Targeting children's behavior problems in preschool classrooms: a cluster-randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *77*(2), 302.

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Bub, K., & Pressler, E. (2011). CSRP's impact on low-income preschoolers' preacademic skills: self-regulation as a mediating mechanism. *Child Development*, *82*(1), 362-378.

Title: Estimating the Long-Run Impacts of the Chicago School Readiness Project

Authors: Tyler W. Watts, Jill Gandhi, C. Cybele Raver

Words: 995

Background

The current study evaluated the long-run effects of the Chicago School Readiness Project (CSRP), which was tested in a sample of 602 preschoolers enrolled in 18 Head Start centers in Chicago. The program offered targeted pedagogical development sessions to Head Start teachers (see symposium statement for more intervention details). Raver and colleagues (2009; 2011) reported that the program significantly reduced children's end-of-preschool behavioral problems and boosted their cognitive skills, but effects largely faded during elementary school. In the following analyses, we report impacts of the CSRP preschool program on 466 adolescents' measures of neurocognitive, behavioral, emotional, and academic development, all assessed ten years after initial program assignment.

Study Design

The symposium statement provides more details regarding the original study design and data collection procedures. In the current analyses, we rely on data collected at the beginning of preschool (i.e., pre-treatment; fall of 2004) and ten years following the end of the program (i.e., long-run follow-up; 2015-2016 academic year). Our analyses investigated whether assignment to the treatment (i.e., the CSRP preschool program) affected measures of adolescent functioning, thus our models provide intent-to-treat (ITT) estimates.

Because random assignment was conducted through a paired blocking procedure (see symposium statement), we include blocking group fixed effects in all analyses presented. For treatment impact analyses, we began with OLS models in which each respective outcome measure was regressed on treatment status and the set of blocking group dummy variables. We then added a large set of baseline covariates (all shown in Table 1). Robust standard errors were adjusted for site-level clustering using the Huber-White adjustment. The analysis sample was restricted to students who had at least one non-missing outcome measure ($n= 466$), and the 23% attrition rate did not differ between the treatment and control groups ($p = 0.75$).

Measures

Baseline. At preschool entry, study designers collected a host of measures capturing characteristics of study children and Head Start teachers involved with the project. Table 1 shows the full list of baseline measures collected, and measurement details for these variables can be found in Raver et al. (2011).

Ten Year Follow-Up. Adolescent executive function was measured with the Hearts and Flowers task. We used the proportion of correct responses as a measure of accuracy, and the difference between the mixed trials reaction time and hearts-only trials reaction time as a

measure of processing speed. Adolescent-reported behavioral problems were measured using the BIS-Brief, and items were grouped into measures of internalizing and externalizing.

Emotional regulation was measured using the Emotion Go/No Go task (EGNG; Tottenham et al., 2011). The EGNG was given on a computer, and it contained trials in which participants viewed a series of either emotional or neutral faces, and they were asked to either click, or inhibit their inclination to click, on the emotional face. We present results for trials in which adolescents viewed "angry vs neutral" and "sad vs neutral" faces, respectively, and "D-prime" was the primary measure of emotion regulation (see Tottenham et al., 2011). We also present reaction time during "angry" and "sad" trials as a measures of emotional processing speed.

Finally, we also modeled treatment impacts on student-reported GPA (measured on a 4 point scale, with scores of "4" equaling an "A" average). Although we hoped to model outcomes on measures of GPA taken from district offices, administrative data was missing for most students. For the 141 students that had both self-reported GPA and district-reported GPA, these two measures of student grades had a 0.66 correlation.

Results

We began by evaluating baseline equivalence, and as Table 1 reflects, we found evidence of imbalance on several baseline measures. Some of this imbalance seemed to favor the treatment group (e.g., several baseline test scores), while other measures favored the control group (e.g., teacher and class characteristics). These baseline checks implied that despite random assignment, the groups did differ in observable ways, underscoring the need to control for baseline characteristics in subsequent treatment impact models.

Table 2 contains our treatment impact estimates on the set of long-run outcomes, and all outcome measures were standardized so coefficients can be interpreted as effect sizes. Each estimate presented was derived from a separate model, with estimates in Column 1 containing only the blocking group fixed effects. In the next columns, we progressively added baseline characteristics, with Column 4 containing the entire set of characteristics shown in Table 1.

The results in Table 2 further illustrate the baseline imbalance, as results for many of the outcomes depend on the inclusion of the baseline covariates. However, estimates in Column 4 (the fully-controlled model) imply that the program may have had significant long-run effects on measures of adolescent functioning. In particular, we found a positive and marginally statistically significant effect of 0.19 on executive function (i.e., H&F accuracy), and a positive and significant effect on self-reported grades (log odds coefficient =0.403). We also observed statistically significant reductions in reaction time on angry and sad trials for EGNG. Interestingly, we observed a negative effect on emotion regulation (i.e., "Angry D-Prime") in angry trials, and post-hoc tests revealed that this was driven by heightened sensitivity to emotional faces in trials where children were asked to inhibit their response to angry faces. Finally, in most models, we did not detect statistically significant effects on measures of behavioral problems or emotional regulation to sad faces.

Discussion

Our results imply that the CSRP may have had some long-run effects on measures of adolescent functioning assessed ten years after the program ended. We found positive effects on

cognitive and academic outcome, but relatively mixed results for emotional and behavioral measures. Future analyses will focus on mechanisms that explain these outcomes (i.e., mediators) and on possible sources of treatment heterogeneity (e.g., level of poverty at baseline). The presentation will contain a full discussion of the implications of our results for other early childhood interventions, particularly for programs working with inner-city, ethnic-minority, populations.

Table 1 (Panel A)
Child and Parent Characteristics Measured at Baseline (PreK Entry)

	Treatment	Control	P-Value
Child Characteristics			
Female	0.50	0.58	0.002**
Age (years) at PreK Entry	4.92	4.93	0.986
African American	0.71	0.66	0.624
Hispanic	0.25	0.25	0.844
White	0.01	0.04	0.389
Bi-racial or Other	0.50	0.58	0.331
Family/Parent Characteristics			
Income to Needs Ratio	0.03	0.05	0.398
Number of Children in the Home	0.66	0.72	0.003**
Family Size	2.73	2.93	0.225
Years in Current Home	4.51	4.57	0.662
TANF	4.70	4.89	0.109
WIC	0.17	0.12	0.450
Food Stamps	0.28	0.29	0.815
Medicaid/Kidcare	0.53	0.50	0.307
Public Housing	0.68	0.72	0.301
Free/Reduced Price Lunch	0.13	0.17	0.127
SSI Disability	0.56	0.63	0.526
Family Support	0.10	0.10	0.756
Parent or Child is Immigrant	0.18	0.17	0.275
Bio Parent Sometimes Sees Child	0.13	0.18	0.007**
Bio Parent Sees Child Everyday	0.27	0.22	0.001**
Hours Worked per Week	0.43	0.55	0.446
Parent Age	20.58	22.75	0.067+
Parent African American	29.32	30.02	0.612
Parent Hispanic	0.74	0.70	0.922
Living with Partner	0.25	0.25	0.014*
Married/Remarried	0.36	0.44	0.016*
Parent Has Savings	0.17	0.27	0.009**
Parent Full-time Employed	0.66	0.57	0.072+
Parent Unemployed	0.36	0.46	0.942
Mother Graduated H.S.	0.37	0.37	0.522
Mother Attended Some College	0.39	0.42	0.945
Mother Attained B.A. or Higher	0.29	0.30	0.049*
Observations	236	230	

Note. See Panel B for table note.

Table 1 (Panel B)

Child Competencies and Teacher Characteristics Measured at Baseline (PreK Entry)

	Treatment	Control	P-Value
Child Baseline Skills and Behavior			
Executive Functioning	0.11	-0.10	0.013*
Effortful Control	0.05	-0.02	0.595
Attention/Impulse Control	2.30	2.24	0.261
Positive Emotion	2.15	2.15	0.884
Letter Naming	0.25	0.21	0.076+
Math	7.87	7.31	0.085+
PPVT	10.93	10.28	0.003**
Externalizing (Parent Report)	6.55	6.07	0.153
Internalizing (Parent Report)	3.04	3.07	0.607
Externalizing (HS Teacher Report)	6.37	5.42	0.216
Internalizing (HS Teacher Report)	2.53	2.07	0.122
Teacher and Class Characteristics			
Teacher has BA	0.71	0.60	0.628
Teacher Age	37.31	42.43	0.050+
Techer Depression (K6 Score)	3.16	2.10	0.063+
Teacher Job Demand	2.86	2.52	0.002**
Teacher Job Control	3.33	3.14	0.488
Behavioral Management	4.58	5.08	0.001**
Classroom Emotional Climate	15.40	16.49	0.001**
Classroom Overall Quality	4.43	4.88	0.013*
Class Size	16.51	16.17	0.836
Number of Adults in Class	2.50	2.30	0.859
Overall F- Statistic	**22.32**		**0.001***
Observations	236	230	

Note. P-values were generated from a series of regressions in which each respective baseline characteristics was regressed on the dummy indicator for treatment status with blocking group fixed effects included. Thus, there are instances where significant p-values may arise even though mean differences look small (e.g., "living with a partner"), because p-values represent the difference on each characteristics between treatment and control *within each blocking group.* P-values less than 0.001 have been rounded up to 0.001. The F-statistic was generated from a model in which treatment status was regressed on all of the baseline characteristics, and the F-statistic indicates whether the entire set of baseline observables differs between the treatment and control groups.

* $p<0.05$ ** $p<0.01$ *** $p<0.001$

Table 2

Preliminary Treatment Impacts of the CSRP on Adolescent Outcomes

	Blocking Group	Dem/ Family	Baseline Skills/Beh	Class/ Teacher
	(1)	(2)	(3)	(4)
Hearts and Flowers				
Mixed Trials Accuracy	0.137	0.199*	0.154*	0.189+
n= 460	(0.081)	(0.081)	(0.069)	(0.091)
Mixed Trials Reaction Time (adjusted)	0.071	0.073	0.086	0.046
n= 459	(0.057)	(0.061)	(0.066)	(0.066)
Behavior Problems				
Internalizing	0.079	0.123+	0.138+	0.038
n= 461	(0.053)	(0.064)	(0.079)	(0.102)
Externalizing	0.028	0.021	0.046	-0.089
n= 461	(0.098)	(0.099)	(0.108)	(0.103)
Emotion Go/ No Go				
Angry D-Prime	-0.089	-0.128+	-0.157*	-0.189+
n= 447	(0.079)	(0.072)	(0.069)	(0.101)
Angry RT (adjusted)	-0.093	-0.131+	-0.136+	-0.324***
n= 445	(0.075)	(0.071)	(0.076)	(0.075)
Sad D-Prime	-0.028	-0.021	-0.059	-0.067
n= 447	(0.061)	(0.070)	(0.075)	(0.104)
Sad RT (adjusted)	-0.025	-0.022	-0.051	-0.221*
n= 445	(0.029)	(0.042)	(0.047)	(0.077)
Self-reported GPA (ordered logit)	0.125	0.111	0.170	0.403*
n= 418	(0.171)	(0.153)	(0.172)	(0.201)
Baseline Covariates Included				
Blocking Group	Inc.	Inc.	Inc.	Inc.
Demographic, Family and Parent Characteristics		Inc.	Inc.	Inc.
Child Baseline Skills and Behavior			Inc.	Inc.
Classroom/Teacher Characteristics				Inc.

Note. Robust standard errors were adjusted for site-level clustering in preschool and are presented in parentheses. All outcome variables were standardized (with the exception of self-reported GPA), so coefficients can be interpreted as effect sizes. Models using self-reported GPA as the outcome variable were run using ordered logistic regressions, and post-hoc tests revealed a treatment effect of about 12% for Model 4 (i.e., students in the treatment group were 12% more likely to report having a higher grade). Missing data on baseline covariates was accounted for by mean imputation and a set of "missing dummy variables" was also included (e.g., for the baseline executive function score, a missing dummy variable was generated that was equal to one if a student was missing on baseline executive function). Models were also run using FIML to adjust for missing baseline data, and treatment impacts were qualitatively similar and no sign of statistical significance changed.
+ $p<0.10$ * $p< 0.05$ ** $p < 0.01$ *** $p < 0.001$

References

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Metzger, M. W., & Solomon, B. (2009). Targeting children's behavior problems in preschool classrooms: a cluster-randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *77*(2), 302.

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Bub, K., & Pressler, E. (2011). CSRP's impact on low-income preschoolers' preacademic skills: self-regulation as a mediating mechanism. *Child Development*, *82*(1), 362-378.

Tottenham, N., Hare, T. A., & Casey, B. J. (2011). Behavioral assessment of emotion discrimination, emotion regulation, and cognitive control in childhood, adolescence, and adulthood. *Frontiers in Psychology, 2*, 1–9.

Title: Examining Social-emotional Skills as a Mechanism in Explaining the Long-term Impacts of Early Childhood Intervention: Evidence from the Chicago School Readiness Project

Authors: Dana C. McCoy, Kathryn Gonzalez, and Stephanie Jones

Words: 998

Background & Context

Early childhood education (ECE) programs have positive short-term impacts on children's academic and cognitive skills (Camilli, Vargas, Ryan, & Barnett, 2010). Evidence is mixed, however, regarding the long-term efficacy of early interventions. Although several programs (e.g., Perry Preschool, Abecedarian) have shown substantial effects on adult outcomes, the majority of evaluated ECE interventions have shown "fade-out" or "convergence" of cognitive impacts following children's transition to elementary school (Barnett, 2011; Camilli et al., 2010; Campbell et al., 2012; Puma et al., 2012; Schweinhart et al., 2005). Recent work by Bailey, Duncan, Odgers, & Yu (2017) suggests that for interventions to generate sustainable impacts, they must target skills that are: (1) malleable, (2) fundamental, and (3) not addressed in everyday practice (outside of specific interventions). In the present study, we examine skills that fall into these categories (e.g., child self-regulation and behavior) as possible mechanisms for explaining the long-term impact of one ECE intervention on students' academic outcomes (grades) in high school.

Research Question

This study examines whether and to what degree the long-term impacts of the Chicago School Readiness Project (CSRP) – a professional development intervention for Head Start teachers – on students' high school grades are mediated by (1) reductions in children's externalizing and internalizing behavior problems, and/or (2) gains in children's self-regulation and executive function skills in the intervention year.

Setting & Participants

Participants included 418 of the original 602 children who were recruited in 2004 and 2005 from 35 classrooms in 18 Head Start centers. Centers were originally selected for participation based on their location in high-poverty, high-crime neighborhoods in Chicago's West and South sides. Children's mean age at baseline was 49.41 months ($SD = 7.19$). The racial/ethnic composition of full sample was: 67% Black, 27% Latino/a, 3% White, and 4% "Other." Overall, 22% of children came from Spanish-speaking households and 69% came from single-parent families. By virtue of their participation in Head Start, all children were classified as low-income. The average size of classrooms in the Head Start year was 16.03 students ($SD = 2.92$) and approximately 60% of teachers had bachelor's degrees.

Intervention & Research Design

The CSRP evaluation is a cluster randomized control trial in which 18 Head Start centers were matched pairwise based on teacher-, child-, and center-level characteristics and randomly assigned within pairs to treatment or control conditions. The CSRP treatment included five six-hour training sessions for lead and assistant teachers focused on classroom management, emotional support, and stress reduction, as well as weekly, in-class coaching sessions from a mental health consultant. In the spring of Head Start, mental health consultants also delivered one-on-one support to a selected group of individual children identified by teachers as showing particular behavioral difficulties. The control group received no training, but did receive a classroom aid for the same amount of time that the coach supported the treatment classrooms.

Although CSRP was only administered in Head Start, children's cognitive and social-emotional skills were assessed multiple times as part of a longitudinal follow-up. In the present study, we use data from three time points: Head Start Fall (baseline), Head Start Spring (immediately following the conclusion of the intervention), and 10 years after Head Start (when children were in approximately ninth grade).

Measures

Children's grades in ninth grade were captured using self-reports using a letter-grade format (A=3, B=2, C or lower=1). Self-reports were found to be consistent with administratively reported grades based on subset of students for whom both were available. Children's social-emotional skills were measured in the fall and spring of Head Start using teacher-reports of children's internalizing and externalizing symptoms from the Caregiver-Teacher Report Form (CTRF; Achenbach, 1991), and direct assessments of effortful control, executive function, and attention/impulse control from the Preschool Self-Regulation Assessment (PSRA) and its accompanying observational Assessor Report (Smith-Donald, Raver, Hayes, & Richardson, 2007).

Analysis

Preliminary analyses relied on a series of multi-level linear regression analyses used for testing mediation, as illustrated in Figure 1 (Baron & Kenny, 1986). First, we fit a series of models in which center-level CSRP treatment status predicted each of the child-level social-emotional mediators in Head Start Spring (path a), controlling children's scores on these measures in Head Start Fall. Second, we fit a model in which CSRP treatment predicted children's high school grades (path c). Finally, we fit a series of models in which children's high school grades were predicted by their social-emotional scores in Head Start Spring (path b) and their center-level CSRP treatment status (path c'). All models included the same set of child-, classroom-, and center-level covariates and nesting structure used in prior treatment impact analyses (McCoy et al., 2017; Raver et al., 2008; 2011). Sobel's test was used to calculate the significance of the indirect paths between CSRP treatment status and high school grades via each social-emotional mediator.

Results

Results are summarized in Table 1. Consistent with prior work (e.g., Raver et al., 2011), random assignment to CSRP increased children's executive function and decreased their internalizing and externalizing problems in Head Start. CSRP also increased high school grades. Significant indirect paths were not observed via any of the social-emotional mediators. Gains in executive function were, however, significantly predictive of high school grades ($b = 0.125$, $S.E. = 0.061$, $p < .01$).

Conclusions

These results suggest that the Head Start-based CSRP intervention has short-term impacts on children's behavior problems and executive function, as well as long-lasting impacts on students' academic performance in high school. Although significant mediation was not observed in initial analyses, results provide suggestive evidence for executive function as a potential mechanism for explaining long-term academic impacts. This finding is supported by recent empirical work on ECE intervention persistence (e.g., Bailey et al., 2017), which posits early social-emotional competencies as mechanisms that drive sustainable impacts on academic outcomes. For the final presentation, analyses will be replicated using alternative model specifications and structural equation modeling, which will allow for simultaneous exploration of multiple mediating processes. Additional limitations (e.g., concerns regarding causality and mediation) and implications for policy and practice will be discussed.

References

Achenbach, T. M. (1991). *Manual for the Teacher's Report Form and 1991 profile*. Univ Vermont/Department Psychiatry.

Bailey, D., Duncan, G., Odgers, C., & Yu, W. (2017). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness*, *10*(1), 7-39.

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

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, *51*(6), 1173.

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers college record*, *112*(3), 579-620.

Campbell, F. A., Pungello, E. P., Burchinal, M., Kainz, K., Pan, Y., Wasik, B. H., ... & Ramey, C. T. (2012). Adult outcomes as a function of an early childhood educational program: an Abecedarian Project follow-up. *Developmental Psychology*, *48*(4), 1033.

McCoy, D. C., Jones, S., Roy, A., & Raver, C. C. (2017). Classifying social-emotional trajectories through elementary school: Impacts of the Chicago School Readiness Project. *Developmental Psychology, In Press.*

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jeniks, F., ... & Downer, J. (2012). *Third grade follow-up to the Head Start Impact Study final report.* Administration for Children & Families.

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Bub, K., & Pressler, E. (2011). CSRP's

impact on low-income preschoolers' preacademic skills: self-regulation as a mediating mechanism. *Child Development*, *82*(1), 362-378.

Raver, C. C., Jones, S. M., Li-Grining, C. P., Metzger, M., Champion, K. M., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly*, *23*(1), 10-26.

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *Lifetime effects: The High/Scope Perry Preschool study through age 40*. Ypsilanti: High/Scope Press.

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

Figure 1. Conceptual model of CSRP's impact on high school grades via early social-emotional skills
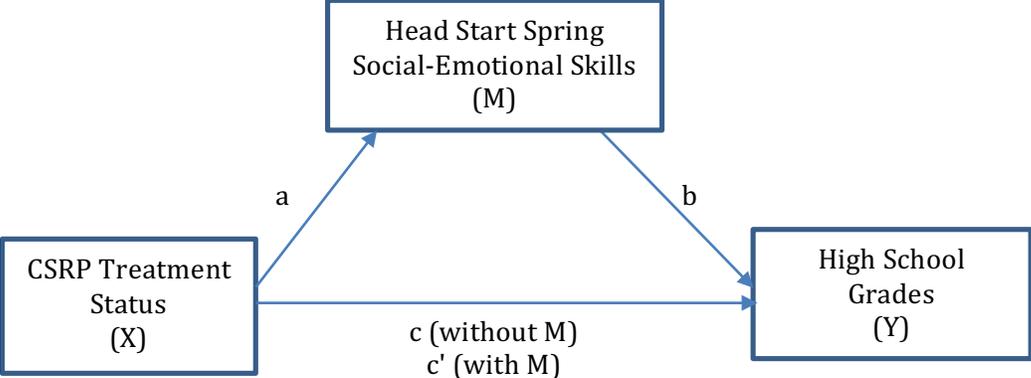
Table 1. Results of multi-level regression models

Outcome	CSRP treatment effect (c)	CSRP treatment effect (a)	Indirect effect on GPA (a*b)
Self-reported GPA	0.294** (0.147)	--	--
Internalizing Behavior (C-TRF)	--	-2.845*** (0.906)	0.032
Externalizing Behavior (C-TRF)	--	-4.195** (1.731)	0.017
Attention/impulsivity (PSRA assessor report)	--	-0.065 (0.091)	-0.011
Executive functioning (PSRA)	--	0.251* (0.138)	0.031
Effortful control (PSRA)	--	-0.192 (0.120)	-0.021

Note: Standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Title: Getting In: The influence of early intervention on later selection into higher-performing schools

Authors: Deanna A. Ibrahim, Tyler W. Watts, Jill Gandhi, Alaa Khadar, C. Cybele Raver

Abstract Text (997 words)

Background:

Given the option of open-enrollment, about half of students in the Chicago Public School System (CPS) decide not to attend their assigned neighborhood school (Cullen et al., 2014). Academic achievement has been shown to vary across districts (Reardon, 2016) and schools (Coca et al., 2012), displaying the weight of school-choice in educational outcomes. In our analyses, we leverage publically available school-level data to examine the influence of the Chicago School Readiness Project (CSRP) on families' school selection over ten years after treatment intervention.

Participants:

We utilized student-level data from the CSRP study's 11-year follow-up, in which 460 students reported the school they attended (96% were enrolled in high school). School-level data was collected from the Illinois School Report Card (ISRC; https://www.illinoisreportcard.com/) and linked to CSRP data by school name (352 students had valid ISRC data). Table 1 presents student characteristics for this analytic sample at the 11-year follow-up. Out of the 352 students whose reported schools had data from the Illinois School Report Card, 53% were female and 47% were male. Most were African American (67%), followed by Hispanic (27%), white (3%), other (1%), and biracial (4%). The mean age of our sample was 16 with an income to needs ratio of 1.0.

Methods:

The ISRC data allowed us to observe several indicators of school performance. First, the ISRC used Illinois state achievement scores to calculate the percentage of students ready for the next level, and the percentage of students meeting expectations both overall and broken down by race and income-level. The Report Card also utilized ACT scores to calculate the percentage of students ready for college. Other school-level demographic data included graduation rate and teacher retention rate.

To increase our understanding of the types of schools CSRP students were attending, we provide descriptive school statistics for those attended by our sample at the 11-year follow-up. Since CPS provides information on open enrollment schools (i.e., schools to which students can apply if they choose to "opt out" of their assigned area school), we provide a breakdown of the distribution of CSRP students across these categories in Table 2. Of the 78% of students enrolled in CPS, 35% attended charter schools, 23% attended College and Career Academies, 20% attended IB High Schools, 14% attended Magnet High Schools, 8% attended Selective Enrollment High Schools, and 4% attended Military High Schools. The remaining 13% attended

schools that did not fall under these categories ("non-categorized"). Graduation rates were highest in Selective Enrollment (94.9%) followed closely by Military (94.5%), then Non-CPS schools in Illinois (i.e., public schools outside of the inner-city Chicago area; 82.8%). This was followed by the non-categorized schools (80.49%), Charter (78.344), College & Career (72.5%), IB (71.5%), and Magnet (66.3%). Selective Enrollment also had the strongest indicators of student achievement (71.4% ready for next level; mean ACT of 25), and the lowest levels of students failing to meet expectations (only 2.6%). On the opposite end of the spectrum were College & Career Academies, with 7.8% ready for the next level (mean ACT of 16), and 36.5% of students failing to meet expectations.

Analysis

As with any treatment impact analyses, we began with models that investigated treatment group balance at baseline. Out of considerations for space, we elected not to give details regarding the balance checks here, because we observed the same patterns that are presented in Table 1 of Paper 1. Because we also found indications of baseline imbalance on a number of student and teacher level measures, we also control for a large set of baseline covariates in most treatment impact models.

We investigated the impact of the CSRP program on selection into schools by estimating treatment impacts on a set of school-level characteristics provided by the ISRC. As such, we began by modeling impacts on school-level graduation rate and school-level teacher retention rate (both measured as percentages), and we created two sample-standardized indexes of school quality. The "college readiness" index was comprised of the percent of students ready for the next level, percent of students who were college ready, and average ACT score. The "underserved groups did not meet expectations" aggregate included the percent of students who did not meet expectations, percent of black students who did not meet expectations, and percent of low-income students who did not meet expectations. For both indexes, we standardized each measure included, then took the average across the set of measures.

Preliminary Results

In Table 3, we began with OLS models that only included blocking group fixed effects (i.e., the unit of random assignment), then we added sets of covariates in each subsequent column. In all models, we adjusted standard errors for school-level clustering. Estimates shown in Column 5 included all baseline covariates, including beginning-of-preschool neighborhood characteristics. As Table 3 reflects, preliminary results suggest that assignment to the program increased the likelihood of attending a high quality high school, as we found positive effects on school level graduation and teacher retention rates. We also found positive impacts on our "college readiness" index, indicating that treatment students attended schools that prepare a higher proportion of their students for college. Finally, we also found that treatment students were more likely to attend schools with lower percentages of Black and low-income students failing to meet state achievement expectations.

Discussion

Our results indicate that certain early childhood interventions could influence eventual selection into high school environments. Specifically, we found evidence that the program influenced students to attend schools with higher markers of achievement over 10 years after the program ended. These results are preliminary, and our future analyses will continue to unpack whether these results were driven by selection into certain *types* of schools (i.e., "College and Career" verses "Selective Enrollment"). We are also working to measure the distance between each child's school of enrollment and their actual neighborhood, so we can better understand how distance may play a role in students' school choices. We believe these results can strongly contribute to our knowledge of how students' early experiences influence their long-run educational experiences.

Table 1. *Student-level demographic information.*

Participant Demographic Information	Mean (SD)
Age (Years)	16 (0.71)
	Column N %
Control	49.7%
Treatment	50.3%
High School Environment	
Enrolled in Chicago Public School	78%
Enrolled in Non-Chicago Public School	22%
Gender	
Female	53.4%
Male	46.6%
Race	
African American	66.5%
Hispanic	26.7%
Bi-Racial	3.7%
White	2.6%
Other	.6%
Parent Demographic Information	Mean(SD)
Mean Income to Needs Ratio	1.0(0.85)
Education Level	Column N%
<12th grade	18.3%
HS diploma or GED	30.0%
Some college	37.6%
BA or higher	14.1%

Table 2. *School-level achievement information by school grouping.*

	Non-CPS IL	CPS						
		College & Career	IB	Magnet	Military	Selective Enrollment	Charter	Non-categorized
Demographics								
Graduation Rate	81.82	72.48	71.53	66.32	94.5	94.87	78.344	80.49
Teacher Retention	89.04	83.19	87.64	89.37	91.9	90.96	62.18	74.34
College Readiness								
Next Level	21.45	7.76	15.75	13.25	20.11	71.41	14.7927	17.46
College Ready	38.29	10.94	19.41	22.83	33.86	79.13	28.57	19.41
ACT Scores	19.56	16.16	17.3	17.62	19.23	24.91	18.371	17.19
Underserved Groups Not Meet Expectations								
School Did Not Meet	28.42	36.51	26.76	26.66	16	2.61	29.55	28.20
Black Did Not Meet	37.15	36.17	33.25	31	13	6.43	31.75	31.00
Low-income Did Not Meet	32.63	33.68	27.15	26.24	16	7.86	29.52	26.83
Student N	77	63	55	38	10	23	96	35
School N	51	23	17	5	4	9	34	16

Note. Unstandardized means are represented in the table below. All values represent percentages with the exception of ACT scores. Students attending non-CPS schools are either enrolled in private schools or non-CPS districts within the state of Illinois. Students attending CPS schools are categorized into six groups. Five of these groups (College & Career, IB, Magnet, Military, and Selective Enrollment) are included in the open-enrollment system, through which students in CPS can apply to schools even if they are outside of their neighborhood boundaries. All of these five groups, with the exception of College & Career, require prequalifying scores for acceptance, with Selective Enrollment schools placing the greatest emphasis on high scores for admission. Students whose schools do not fall under any of the five groups of open-enrollment are classified under the "non-categorized" column. This column includes various types of schools that do not fit into the classifications above, including contract schools, city-wide options alternative education programs, and neighborhood schools.

Table 3. *Associations between treatment and school performance.*

	(1)	(2)	(3)	(4)	(5)
	Blocking Group	Demographics/ Family Characteristics	Baseline Skills/ Behavior	Classroom/ Teacher Characteristics	Neighborhood at Preschool Entry
Graduation Rate	7.266**	5.552**	5.840*	6.725**	7.069*
	(2.304)	(1.758)	(2.048)	(2.314)	(2.697)
Teacher Retention	0.899	2.304	3.028+	2.566	3.113+
	(1.172)	(1.543)	(1.509)	(1.600)	(1.771)
College Readiness	0.118	0.144	0.139	0.192*	0.222*
	(0.116)	(0.089)	(0.088)	(0.084)	(0.088)
Underserved Groups Failing to Meet Expectations	-0.114	-0.159+	-0.178*	-0.213*	-0.203*
	(0.097)	(0.088)	(0.080)	(0.094)	(0.091)
Control Variables					
Blocking Group	Inc.	Inc.	Inc.	Inc.	Inc.
Demographics/Family Characteristics		Inc.	Inc.	Inc.	Inc.
Baseline Skills/Behavior			Inc.	Inc.	Inc.
Classroom/Teacher Characteristics				Inc.	Inc.
Neighborhood at Preschool Entry					Inc.

Note. Standardized coefficients are represented in the table below. Graduation rate and teacher retention rate outcomes are measured as percentage values, whereas college readiness and underserved groups failing to meet expectations outcomes are measured as standardized effects. The first model of this stepwise regression includes the blocking group control variable, which accounts for the group participants were originally assigned to at the beginning of the head start program. The second model controls for blocking group and demographics/family characteristics, which includes descriptors of socioeconomic status, parent education, and other personal factors at head start. Model 3 also controls for baseline skills/behavior, referring to child performance on executive function tasks prior to treatment intervention. Model 4 controls additionally for classroom/teacher characteristics, or qualities of the classroom environment and teacher of each site of the head start program. Finally, the sixth model adds the control variables to account for neighborhood characteristics at preschool entry using variables such as median household income and neighborhood income inequality, which were derived through linking participants' home addresses at preschool entry with census tract data.

References

Coca, Vanessa & Johnson, David & Kelley-Kemple, Thomas & Roderick, Melissa & Moeller, Eliza & Williams, Nicole & Moragne, Kafi. (2012). Working to My Potential: The Postsecondary Experiences of CPS Students in the International Baccalaureate Diploma Programme.

Julie Berry Cullen, Brian A. Jacob, Steven D. Levitt, The impact of school choice on student outcomes: an analysis of the Chicago Public Schools, In Journal of Public Economics, Volume 89, Issues 5–6, 2005, Pages 729-760, ISSN 0047-2727, https://doi.org/10.1016/j.jpubeco.2004.05.001.

"Illinois Report Card." *Illinois Report Card*, www.illinoisreportcard.com/.

Sean F. Reardon, Demetra Kalogrides, Andrew Ho, Ben Shear, Kenneth Shores, Erin Fahle. (2016). *Stanford Education Data Archive* (Version 1.1 File Title). http://purl.stanford.edu/db586ns4974.