Abstract Title Page
Not included in page count.

Title: Comparing the effectiveness of Head Start and state pre-k using a propensity-score matching regression discontinuity design

Authors and Affiliations:

Jade Marcus Jenkins, University of California, Irvine
George Farkas, University of California, Irvine
Greg J. Duncan, University of California, Irvine
Margaret Burchinal, University of North Carolina at Chapel Hill
Deborah Lowe Vandell, University of California, Irvine

Abstract Body

Background / Context:

In light of the evidence that high quality early childhood education programs can improve children's school readiness and future academic success (Duncan & Magnuson, 2013; Gormley, 2008; Magnuson, Ruhm, & Waldfogel, 2007), there are a number of recent efforts at the federal and state levels to expand public preschool programs. These expansions aim to serve not just *more* children, but to also serve *younger* children (e.g. President Obama's "Preschool for All" initiative). Federal Head Start (HS) and state prekindergarten programs (pre-k) are the two most prominent early education policy programs and serve millions of children nationwide. The individual effectiveness of each program in preparing young low-income children for school—in addition to their *relative* effectiveness—is under constant scrutiny.

Evidence suggests that children who participate in pre-k programs have improved early learning outcomes in kindergarten (Gormley & Gayer, 2005; Henry, Gordon, & Rickman, 2006; Wong, Cook, Barnett, & Jung, 2008). Research also shows improvements in early learning skills as a result of HS participation (Gormley, Phillips, Adelstein, & Shaw, 2010; Ludwig & Phillips, 2008; Puma, Bell, Cook, & Heid, 2010). However, few studies directly compare the effectiveness of these two programs with respect to children's readiness for school. Henry and colleagues compare pre-k and HS participants using propensity score matching and find that economically disadvantaged children who participated in state pre-k had slightly higher scores at kindergarten entry relative to similar children who attended HS (2006). Gormley and colleagues do not match children between the two programs, but do find larger program effect sizes for children who attended state pre-k compared to low-income children who attended HS (2010).

Another important policy issue is the dosage of early learning programs. There is little empirical work to date on this topic, but researchers have started to attend to issues of dosage in early learning. While some preschool programs produce large effects in only one year (Gormley, 2008), other early learning interventions that target low-income children like Abecedarian and Perry Preschool provided two to five years of program services (Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001; Schweinhart, 2005). The HS Impact study also found stronger program effects for children who participated in the program at age three relative to those who participated at age four (Puma et al., 2010). The body of evidence suggests that more participation in center-based early learning programs is associated with stronger cognitive outcomes, especially for low-income children (Dearing, McCartney, & Taylor, 2009; Loeb, Fuller, Kagan, & Carrol, 2004; NICHD Early Child Care Research Network, 2000). Still, the research on dosage has not rigorously controlled for selection into treatments and dosage. If longer exposure produces greater outcomes, this would justify increasing access for earlier entry into public programs. However, if the single year approach—that which is most common across pre-k and HS programs—produces similar outcomes, then limited public resources could be directed towards other child development policies.

In our study, we replicate and innovate upon some of the analyses conducted by Gormley and colleagues (2005, 2008, 2010) in their evaluations of the Oklahoma state pre-k and HS programs in Tulsa. We use propensity score matching *in addition* to regression discontinuity to make direct statistical comparisons of the treatment effects of these two major programs on children's

readiness for school. Furthermore, we examine whether the dosage of public early learning programs affects children's pre-academic skills at kindergarten entry by creating treatment groups based on children's age three *and* age four early learning experiences, which we refer to as preschool 'pathways', and make direct comparisons between these exposures.

Purpose / Objective / Research Question / Focus of Study:

The objectives of the study are to compare the overall effects of a high-quality state pre-k program and Head Start, compare the effects of different pathways into pre-k and Head Start, and to test whether two years of participation in public early learning programs (age three and age four participation) are more effective than one year (age four participation only).

Research questions:

1. Are there differences in the effects of a high-quality pre-k program relative to a Head Start program on children's pre-academic skills at school entry?
2. Do the effects of a high-quality pre-k program at age *four* vary by children's exposure to Head Start at age *three* to affect children's pre-academic skills at school entry?
3. Do the effects of Head Start at age *four* vary by children's exposure to Head Start at age-*three* to affect children's pre-academic skills at school entry?
4. Does participation in Head Start at age *three* have a different effect on children's pre-academic skills at school entry depending on the type of preschool exposure at age-*four*?

Setting and Intervention:

This study is a secondary data analysis of data collected on preschool and kindergarten-aged children in Tulsa, Oklahoma in the fall of 2006 to study the effects of school and community-based public early learning programs on children's cognitive and social-emotional outcomes. Children in the study participated in the state-funded pre-k program, in the county HS program, or a combination of both. The study treatment groups are described in further detail below.

Population / Participants / Subjects and Data Collection:

The Tulsa 2006-07 Public Use Data File contains data collected on the Oklahoma's state-funded universal pre-k program administered in Tulsa Public Schools, and the Tulsa County Head Start program administered by local Community Action Project sites. The data include students who were entering: the state pre-K program, the Head Start four-year-old program, or public school kindergarten in the 2006-07 school year. In the prior OK pre-k and HS studies, the children who are entering the two preschool treatments (pre-k or HS) serve as the comparison group for the children who are entering kindergarten and already completed one of the two preschool treatments. The dataset is structured to facilitate this comparison. Within the cohort of children entering kindergarten, there are children who participated in either pre-k, Head Start, or in neither of these public programs in the year prior to kindergarten (age four). The data also include information on whether the child participated in HS at age three. The overall sample includes 1,985 pre-k entrants, 739 HS entrants, and 4,121 kindergarten entrants (1597 pre-k graduates, 474 HS graduates, 2050 who did not attend a public preschool program).
Our study redefines the sample by considering children's public preschool exposures at both age four and age three, creating four groups. These four pathways and their sample sizes are:

participants in pre-k at age four who did *not* participate in HS age three (3,260 total; composed of 1754 kindergarten entrants and 1503 pre-k entrants), participants in pre-k at age four who participated in HS at age three (322 total; 192 kindergarten entrants and 130 pre-k entrants), participants in HS at age four who did *not* participate in HS at age three (733 total; 435 kindergarten entrants, 295 HS entrants), and participants in HS at age four *and* age three (480 total; 300 kindergarten entrants, 179 HS entrants). Child and family characteristics by preschool pathway are presented in Table 1.

The data come from four sources: direct cognitive assessments of children at the beginning of the school year; parent surveys collected at their child's cognitive assessment; social-emotional assessments conducted by each child's teacher; and administrative data from Tulsa Public Schools and Head Start. The three academic subtests used were: the Letter-Word Identification test (pre-reading skills); the Spelling test (pre-writing skills); and the Applied Problems test (pre-math skills) (Woodcock & Johnson, 1989). The same subtests of a comparable Spanish test, the Woodcock-Muñoz Batería, were given to Hispanic students capable of being tested in Spanish.

Research Design and Data Analysis:

The information presented in Table 1 illustrates the differences in observed characteristics between each of the study treatment groups (preschool pathways). We use propensity score matching (PSM) to adjust for observable differences between children in each of the preschool pathways and create balance in covariates across groups where there is evidence of selection into treatment. PSM can mitigate the bias in treatment effect estimates due to this selection on observed variables, assuming that there is no confounding due to unobserved variables (Heckman, Ichimura, & Todd, 1998; Rosenbaum & Rubin, 1983). PSM was implemented using nearest-neighbor matching with replacement and a 0.01 caliper, and we used z-tests, t-tests, and standardized mean differences of characteristics between samples to assess balance (Caliendo & Kopeinig, 2008).

We outlined four treatment contrasts of interest to test our three research questions: 1) age four pre-k vs. age four HS (ignoring any HS treatment at age three), 2) age four pre-k only vs. age four pre-k *and* age three HS, 3) age four HS only vs. age three *and* age four HS, 4) age four pre-k *and* age three HS vs. age four HS *and* age three HS. We created propensity scores by running a logit model predicting one of the two exposures in each contrast, pulling matches from the other exposure, and then estimated an RD model with the matched sample.

Using our propensity-score matched groups, we then implement a regression discontinuity (RD) design to estimate program effects on outcomes at kindergarten entry, and refer to these models as RDPSM. The primary condition for an RD analysis is the use of a quantitative assignment variable and a designated cutoff score to determine treatment status (Imbens & Lemieux, 2008; Shadish, Cook, & Campbell, 2002). This method exploits the fact that the preschool programs implemented a strict age cutoff for participation based on child's birth date, so that children who were born before the cutoff (September 1 of 2005-06 school year) were eligible to participate in the age-four programs, and those born after the cutoff were not. Thus, child age (measured as distance between their birthdate and the cutoff birthdate in days) can be used as the assignment variable in an RD specification. We include an interaction term between the treatment dummy

(born before the cutoff=1) and an indicator for one of the two preschool pathways of interest (e.g. treatment*age three and age four HS) to test for differential effects between the two exposures being contrasted in each analysis. We used a modest bandwidth restriction of 270 days (3/4 year) to ensure exchangeability in observations on either side of the treatment cutoff while also preserving power and precision in smaller treatment groups (Schochet et al., 2010). Outcome models also control for parent's education, child race, sex, free or reduced-price lunch status, exposure to other non-parental care prior to preschool (yes=1), and missing data indicators.

Findings / Results:

The RDPSM results are displayed for each research question in Tables 2-5, with the standard RD for the two comparison groups *without* PSM presented in the column next to the RDPSM estimates for comparison. Coefficients represent changes in raw scale scores. Findings indicate that children who participated in pre-k have stronger reading and language skills than children who participated in HS, but both programs improved children's skills. Children who participated in age three HS in addition to an age four preschool program do not have significantly different pre-academic skills than children who participated in a program at age four only. OK pre-k has a consistently strong positive effect on children's pre-reading, pre-math, and pre-writing skills at the beginning of kindergarten. Results indicate that in many cases the RDPSM makes a difference in the magnitude and significance for some comparisons and for some outcomes, with changes in both directions (i.e. significant to not, increase in magnitude, decrease in magnitude) when compared to a regular RD that includes that does not use PSM.

Conclusions:

While both pre-k and HS are effective at improving children's pre-academic skills at school entry after one year of participation, the high-quality OK pre-k program produces stronger gains in skills at the start of kindergarten than the OK HS program. We find that an additional year of HS at age three does not have a discernible value-added for either age four pre-k or age four HS participation. This is in contrast to findings by Burchinal and colleagues (under review) that children who entered HS at age three and also participated at age four had modestly higher vocabulary scores relative to children who participated in HS at age four only, and other studies that find dosage effects of high-quality early learning programs.

Study limitations include: 1) the OK pre-k program is not representative of most state pre-k programs because of its very high quality standards; 2) children living in Tulsa, OK are not representative of the broader population of children in the U.S.; 3) propensity score matching methods assume there is no unobserved confounding, which not testable; 4) sample sizes for some of the treatment pathways may not provide sufficient power to detect effects, and the age four pre-k only (no age three HS) sample was significantly larger than the other three pathways, and; 5) HS and pre-k have different goals and may often serve different populations. While HS supports child cognitive, emotional, and physical development for very low income children, pre-k programs focus solely on academic activities to prepare children for school entry, and also may be offered to any child who is age-eligible regardless of income or need.

Appendices

Not included in page count.

Appendix A. References

Caliendo, M., & Kopeinig, S. (2008). Some pracical guidance for the implementation of propensity score matching Journal of Economic Surveys, 22(1), 31-72.

Campbell, F. A., Pungello, E., Miller-Johnson, S., Burchinal, M. R., & Ramey, C. T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. Developmental Psychology, 37(2), 231-242.

Dearing, E., McCartney, K., & Taylor, B. A. (2009). Does higher quality early child care promote low-income children's math and reading achievement in middle childhood? Child Development, 80, 1329-1349.

Duncan, G. J., & Magnuson, K. (2013). Investing in Preschool Programs. The Journal of Economic Perspectives, 27(2), 109-132.

Gormley, W. T. (2008). The Effects of Oklahoma's Pre-K Program on Hispanic Children*. Social Science Quarterly, 89(4), 916-936.

Gormley, W. T., & Gayer, T. (2005). Promoting School Readiness in Oklahoma: An Evaluation of Tulsa's Pre-K Program. Journal of Human Resources, XL(3), 533-558.

Gormley, W. T., Phillips, D., Adelstein, S., & Shaw, C. (2010). Head Start's Comparative Advantage: Myth or Reality? Policy Studies Journal, 38(3), 397-418.

Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. The Review of Economic Studies, 65(2), 261-294.

Henry, G. T., Gordon, C. S., & Rickman, D. K. (2006). Early Education Policy Alternatives: Comparing Quality and Outcomes of Head Start and State Prekindergarten. Educational Evaluation and Policy Analysis, 28(1), 77-99.

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. Journal of Econometrics, 142(2), 615-635.

Loeb, S., Fuller, B., Kagan, S. L., & Carrol, B. (2004). Child Care in Poor Communities: Early Learning Effects of Type, Quality, and Stability. Child Development, 75(1), 47-65.

Ludwig, J., & Phillips, D. A. (2008). Long-Term Effects of Head Start on Low-Income Children. Annals of the New York Academy of Sciences, 1136(Reducing the Impact of Poverty on Health and Human Development: Scientific Approaches), 257-268.

Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance? The Economics of Early Childhood Education, 26(1), 33-51.

NICHD Early Child Care Research Network. (2000). The relation of child care to cognitive and language development. Child Development, 71(4), 960-980.

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

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1), 41-55.

Schochet, P., Cook, T. D., Deke, J., Imbens, G. W., Lockwood, J. R., Porter, J., et al. (2010). Standards for regression discontinuity designs: What Works Clearninghouse, Institute for Education Sciences, Department of Education.

Schweinhart, L. (2005). Lifetime Effects: The High/Scope Perry Preschool Study through Age 40. Ypsilanti, M.I.: High/Scope Educational Research Foundation.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin Company.
Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. Journal of Policy Analysis and Management, 27(1), 122-154.
Woodcock, R. W., & Johnson, M. B. (1989). Tests of achievement, standard battery. Chicago, IL: Riverside Publishing.

Appendix B. Tables and Figures

Not included in page count.

Table 1: Descriptive statistics by preschool pathway

Variables	(1) **Age 4 pre-k; nothing age 3**	(2) **Age 4 pre-k; Age 3 HS**	(3) **Age 4 HS; nothing Age 3**	(4) **Age 4 HS; Age 3 HS**
		Mean or proportion		
Reduced-price lunch	0.20	0.10	0.06	0.03
Black	0.32	0.62	0.21	0.44
Hispanic	0.26	0.19	0.59	0.40
Asian/Native/Other	0.11	0.09	0.06	0.07
Below HS	0.23	0.10	0.33	0.22
Some college	0.36	0.38	0.24	0.36
College +	0.04	0.10	0.03	0.06
Child had some non-parental care at age 3	0.54	0.65	0.44	0.53
Internet in the home	0.37	0.39	0.25	0.34
Number of books in the home (scale)	2.42	2.21	1.82	2.22
Parent is foreign born	0.21	0.15	0.50	0.35
English is home language	0.79	0.84	0.52	0.68
Child has health insurance	0.87	0.93	0.77	0.89
Parent respondent is married	0.48	0.31	0.53	0.41
Child assessed in English and Spanish	0.19	0.13	0.47	0.33
Father in home	0.52	0.41	0.57	0.51
Relative age within cohort	188.60	186.60	187.80	190.40
Female	0.48	0.51	0.49	0.54
Health Status of child (scale)	1.50	1.54	1.58	1.53
No medical visit in past year	0.11	0.08	0.12	0.06
Letter-word score	6.45	6.73	4.77	5.92
Applied math score	10.52	9.94	8.65	9.71
Spelling score	6.61	6.22	5.76	6.28
Observations	*1,521*	*178*	*373*	*286*

Table 2. RDPSM and RD results comparing age four HS vs. age four pre-k
Gormley et al. comparisons; ignoring age three HS exposure

	Letter-word		Applied Problems		Spelling	
	B/SE		*B/SE*		*B/SE*	
	RD-PSM	***RD***	***RDPSM***	***RD***	***RDPSM***	***RD***
Pre-k effect (cutoff=1)	**3.31**	**3.86**	0.92	**1.47**	**1.95**	**2.27**
	0.63*	0.35*	0.64	0.39*	0.44*	0.26*
HS effect (HS * cutoff)	**-2.44**	**-2.58**	0.15	-0.61	**-1.29**	**-1.42**
	0.58*	0.45*	0.59	0.49	0.43*	0.32*
Observations	*876*	*2140*	*876*	*2134*	*876*	*2084*
Difference	**2.44**		0.15		**0.80**	
p-value of difference	0.00		0.80		0.00	
Effect size of difference	**0.75**		0.03		**0.31**	

Significant effects and differences are emboldened. Reference group for effect of exposure is age four pre-k. Outcome variable is a raw score. All models use clustered SEs by teacher.

Table 3. RDPSM and RD results comparing pre-k at age four and no HS at age three vs. pre-k at age four and HS at age three

	Letter-word		Applied Problems		Spelling	
	B/SE		*B/SE*		*B/SE*	
	RD-PSM	***RD***	***RDPSM***	***RD***	***RDPSM***	***RD***
Age four pre-k effect (cutoff=1)	**4.17**	**3.94**	**4.46**	**1.50**	**3.62**	**2.40**
	1.16*	0.35*	1.30*	0.40*	0.87*	0.27*
Age four pre-k & age 3 HS effect (exposure*cutoff)	0.04	0.11	**-2.34**	-1.07	-0.14	0.13
	1.11	0.58	1.16*	0.76	0.73	0.42
Observations	*239*	*1628*	*239*	*1625*	*239*	*1581*
Difference	0.04		**2.34**		0.14	
p-value of difference	0.97		0.05		0.85	
Effect size of difference	0.01		**0.85**		0.05	

Significant differences are emboldened. Reference group for effect of exposure is age four pre-k, no age three HS. Outcome variable is a raw score. All models use clustered SEs by teacher.

Table 4. RDPSM and RD results comparing HS at age four and no HS at age three vs. HS at age four and HS at age three

	Letter-word		Applied Problems		Spelling	
	B/SE		*B/SE*		*B/SE*	
	RD-PSM	***RD***	***RDPSM***	***RD***	***RDPSM***	***RD***
Age four HS effect (cutoff=1)	1.14	0.79	2.13	0.85	-0.37	0.38
	1.11	0.72	1.14	0.63	0.60	0.42
Age four HS & age 3 HS effect (exposure*cutoff)	0.20	0.77	0.26	0.35	0.93	0.66
	0.63	0.57	1.02	0.80	0.49	0.44
Observations	*346*	*690*	*346*	*686*	*346*	*673*
Difference	0.20		0.26		0.92	
p-value of difference	0.75		0.80		0.07	
Effect size of difference	0.04		0.05		0.34	

Significant differences are emboldened. Reference group for effect of exposure is age four HS, no age three HS. Outcome variable is a raw score. All models use clustered SEs by teacher.

Table 5. RDPSM and RD results comparing HS at age four and HS at age three vs. pre-k at age four and HS at age three

	Letter-word		Applied Problems		Spelling	
	B/SE		*B/SE*		*B/SE*	
	RD-PSM	***RD***	***RDPSM***	***RD***	***RDPSM***	***RD***
Age four and age three HS effect (cutoff=1)	0.91	**1.83**	1.29	1.43	**1.92**	**1.40**
	0.94	**0.81***	1.22	0.94	0.89*	0.57*
Age four pre-k & age 3 HS effect (exposure*cutoff)	**2.99**	**2.23**	-0.52	-0.79	0.58	0.91
	0.92*	0.78*	1.18	1.10	0.76	0.56
Observations	*286*	*466*	*286*	*462*	*286*	*446*
Difference	**2.99**		0.51		0.58	
p-value of difference	0.01		0.66		0.45	
Effect size of difference	**0.91**		0.11		0.21	

Significant differences are emboldened. Reference group for effect of exposure is age four HS and age three HS. Outcome variable is a raw score. All models use clustered SEs by teacher.