Extended abstract for Paper Session Consideration:

Does preschool attendance matter to children's long-term well-being in China?

Xin Gong, Teachers College, Columbia University
Di Xu, Teachers College, Columbia University
Wen-Jui Han, New York University & NYU-ECNU Institute for Social Development at NYU Shanghai

The number of programs and the enrollment in formal preschool education has been rapidly growing in China in the past three decades (China National Statistical Bureau, 2011). In addition to the potential benefits of formal childhood education on child cognitive and non-cognitive development documented in the existing literature worldwide (e.g. Camilli et al., 2010; Heckman, 2006; Nores & Barnett, 2010), parents in China have been particularly attentive to using preschool as a means to cultivating children's academic and social skills (Tobin, Hsueh, & Karasawa, 2009). During the drafting stage of the *2010-2020 Medium and Long-term Education Development and Reform Outline*, a quarter of the feedback from polling the public opinion was about preschool education (Zeng, Fan, & Zhou, 2011).

However, evidence regarding the effects of preschool attendance in China is limited, though emerging. Only a few studies so far have examined the effects of preschool attendance on child development, most of which reported positive benefits (Luo et al., 2012; Peng, 2011; Rao et al., 2013; Wu et al., 2012; Zhang, 2013).[1] For example, Rao et al. (2013) examined a total of 207 first-grade children from a rural county in Guizhou province in the local primary schools in the 2007-2008 academic year. They found positive effects of attending developmentally appropriate preschool (including "kindergarten"/*you er yuan* and "separate pre-primary class") on school readiness scores based on ANCOVA (Analysis of covariance). Yet, for all the studies that directly examined the effects of preschool attendance in China, the generalizability of the findings is limited due to small sample size in a single province or even a single county. In addition, despite the importance to understand the long-term effects of preschool attendance (Here "long-term" means after age 10, as defined in Camilli et al., 2010), none of the existing studies in China were able to include outcome measures beyond grade 2 of primary school.

Using data from Chinese Family Panel Studies (CFPS), a national survey that follows about 16,000 households in 25 provinces of China since 2010, we provide the first estimates of the long-term effects of preschool attendance on child development in rural and urban China for a national sample of 2,857 teenagers (11-15 year olds). The preschool attendance rate for this sample was 57.93%, with a considerably large gap between the rural and urban areas (42.34% vs. 84.27%). To provide a comprehensive understanding of the effects of preschool attendance on child development, we explore a wide range of outcome measures including cognitive (e.g. self-reported and expert-assessed Chinese and math performance) and non-cognitive skills (e.g. self-

[1] Among the five studies, Luo et al. (2012) and Wu et al. (2012) didn't take preschool effect as a focus of the study: it was either a supplemental product, or was analyzed by treating preschool attendance as one of the predictors of child development.

reported social competence and attention skills), which have shown to be important to later labor market outcomes and life prospects (Almlund et al., 2011; Duncan & Dunifon, 2012; Levin, 2013). Among them, cognitive skills include: (1) self-reported and parent-reported academic performance; (2) language and math skills, either directly measured during the survey or rated by the interviewer. Non-cognitive skills include: (1) self-reported and surveyor-rated social competence, measured by four items including "surveyor-rated interpersonal communication skills", "whether having any friends", "the number of friends", and "whether being a student cadre at school"; and (2) self-reported attention skills, measured by four survey items including "ability to concentrate on learning", "to comply with school disciplines", "to organize stuff in an orderly manner", "to control oneself from playing during study". The descriptive statistics of the outcome measures used in this study is presented in Table 1, for the full sample as well as separately for the urban sample and the rural sample.

To rigorously estimate the long-term effects of preschool attendance, we use multiple empirical strategies. First, OLS is conducted as a baseline model, where each of the outcome variables is regressed upon the preschool attendance along with a set of control variables (e.g. demographic characteristics, family socioeconomic indicators, and county-level variables). Second, we use random effect multi-level modeling (MLM) to account for correlation of child development outcomes for children in the same county attributable to unobserved county factors. Third, we further use propensity score matching to account for the observed differences between people who did and did not attend preschools. Propensities derived from the multilevel model that accounts for county-level variations in preschool attendance are used to match preschool attendees and non-attendees, followed by a weight-adjusted regression on the matched sample to derive the average treatment effect on the treated children (ATT). Using only matched observations through this strategy help derive more precise and robust estimates than the regular regression analysis and serves as a robust check for the results derived from the OLS and MLM (Conniffe, Gash, & O'Connell, 2000; Rubin & Thomas, 2000).

Preliminary results based on the full sample are presented in column 1 and 3 in Table 2. Estimates based on OLS indicate that preschool attendance has significant positive effects on cognitive development and social skills. Specifically, compared to children who did not attend preschools, children who attended preschools had significantly better academic scores and performed better in language and math assessment. In addition, they were more likely to have friends and be cadre at school. No significant associations were found for attention skills. These results indicate that preschool in China, on average, may matter to children's cognitive development and interpersonal interaction skills that hold important implications to individuals' later life prospects.

In addition to the results on full sample, we also present results using only the rural subsample based on two reasons: understanding the impact of preschool attendance on rural children is of particularly policy interest, considering the great disparity in educational resources between rural and urban areas and that much of the current education reform efforts in China aiming to improve the educational outcomes for rural children; methodologically, we could not achieve balanced samples for urban children given more than 85% of urban children attended preschools and as a result, most of the urban preschool attendees failed to find a match among the urban non-attendees). For the subsample of rural children (column 2, 4, and 5), the positive effects of

preschool on cognitive skills are consistently smaller across all model specifications, and become non-significant with propensity score matching strategy. Such attenuation in preschool effects may be due to poorer quality of preschools in rural areas. Yet, the positive results of preschool attendance on student social competence is much more promising for the rural sample: the estimates across various measures are consistently significant, where the coefficients based on the propensity score matching are considerably larger than those based on straightforward OLS and MLM estimates. This robust finding about preschool attendance on social skills for rural children suggests that preschool attendance may assume an important role in allowing rural children to develop social skills when rural areas compared to urban areas tend to have limited access to social interaction resources.

Table 1. Descriptive Statistics for the Outcomes and the Key Predictor

Outcome measures/key predictor	Full sample		Rural sample		Urban sample		T-test: rural-urban difference
	mean	s.d.	mean	s.d.	mean	s.d.	
(1)Cognitive skills: parent and self-reported							
Parent-reported performance in most recent Chinese examination	2.73	0.93	2.62	0.94	2.86	0.90	-0.24**
Parent-reported performance in most recent math examination	2.66	1.01	2.54	1.00	2.80	1.00	-0.26**
Self-reported scores in most recent Chinese examination	80.87	17.03	77.96	17.56	84.35	15.68	-6.38**
Self-reported scores in most recent math examination	79.18	21.99	75.66	22.27	83.32	20.91	-7.66**
Whether still at school (for the sub sample of 14-15 year olds in 2010)	0.78	0.42	0.73	0.45	0.84	0.37	-0.11**
Whether at an academic high school (for the subsample of 14-15 year olds)	0.65	0.48	0.61	0.49	0.69	0.46	-0.08
(2) Cognitive skills: language and math skills							
Scores on Chinese test	21.95	8.56	20.43	9.30	23.81	7.13	-3.38**
Scores on math test	11.47	5.41	10.73	5.89	12.37	4.61	-1.64**
Surveyor rated comprehension skills	5.75	1.12	5.61	1.15	5.92	1.04	-0.31**
Surveyor rated IQ	5.63	1.08	5.50	1.08	5.80	1.06	-0.30**
Surveyor rated expression skills	5.51	1.13	5.39	1.17	5.64	1.07	-0.25*
Surveyor rated interpersonal interaction skills	5.50	1.11	5.42	1.16	5.59	1.04	-0.18*
(3) Non-cognitive skills: social competence							
Whether having any friends	0.92	0.27	0.89	0.32	0.95	0.21	-0.07**
Number of friends	6.79	8.31	6.23	7.66	7.48	9.00	-1.25**
Whether being a student cadre	0.36	0.48	0.33	0.47	0.40	0.49	-0.07**
(4) Non-cognitive skills: attention skills							
Ability to concentrate	2.31	2.08	2.29	2.14	2.34	2.01	-0.06
Ability to comply with school disciplines	2.01	1.66	2.04	2.16	1.98	0.69	0.06
Ability to be organized	2.17	3.37	2.16	2.94	2.18	3.83	-0.02
Ability to control oneself from playing	2.21	3.79	2.24	3.83	2.18	3.75	0.06
Prior preschool attendance	0.58	0.49	0.42	0.49	0.84	0.36	-0.32**

** $p < 0.01$ * $p < 0.05$ † $p < 0.1$

Note. All descriptive statistics in this table are weighted by sample weight. For the t-test, in order to apply sample weight, we used regression: each of the variables was regressed upon a rural dummy as the only predictor, and the t statistic showed the statistical significance of the mean difference between the rural and urban samples.

Table 2. MLM and Multi-level PSM estimates for the preschool effects

Outcome measures	OLS		MLM		MLM-PSM	
sample	Full	Rural	Full	Rural	Full	Rural
(1)Cognitive skills: parent and self-reported						
Parent-reported performance in most	0.26**	0.19**	0.23**	0.13*	0.34**	-0.13
recent Chinese examination	(0.05)	(0.06)	(0.05)	(0.06)	(0.11)	(0.13)
Parent-reported performance in most	0.19**	0.17*	0.17**	0.13*	0.41***	0.15
recent math examination	(0.06)	(0.07)	(0.15)	(0.06)	(0.12)	(0.16)
Self-reported scores in most recent	3.83**	3.04**	2.56**	1.39	0.76	-0.11
Chinese examination	(1.07)	(1.32)	(0.84)	(1.07)	(1.84)	(3.45)
Self-reported scores in most recent math	3.79**	4.68**	2.34*	3.02*	0.97	-2.44
examination	(1.42)	(1.72)	(1.16)	(1.42)	(1.89)	(2.97)
Whether still at school (for the sub sample	-0.04	-0.05	-0.05	-0.05	-0.07†	-0.21**
of 14-15 year olds in 2010)	(0.04)	(0.05)	(0.03)	(0.05)	(0.04)	(0.05)
Whether at an academic high school (for	0.09	0.03	0.08	-0.02	0.03	-0.19*
the subsample of 14-15 year olds in 2010)	(0.07)	(0.08)	(0.06)	(0.07)	(0.10)	(0.10)
(2) Cognitive skills: language and math skills						
Scores on Chinese test	1.54**	1.49**	1.63**	1.23*	2.30	-0.92
	(0.52)	(0.55)	(0.39)	(0.50)	(1.71)	(0.88)
Scores on math test	0.76*	0.60†	0.81**	0.53†	1.79	-1.00**
	(0.31)	(0.32)	(0.24)	(0.31)	(1.14)	(0.40)
Surveyor rated comprehension skills	0.23†	0.20	0.27**	0.13	-0.08	0.33
	(0.12)	(0.14)	(0.09)	(0.12)	(0.20)	(0.23)
Surveyor rated IQ	0.25*	0.10	0.29**	0.10	0.02	0.22
	(0.11)	(0.13)	(0.09)	(0.11)	(0.16)	(0.18)
Surveyor rated expression skills	0.13	0.11	0.17*	0.13	-0.34†	0.15
	(0.11)	(0.14)	(0.10)	(0.13)	(0.17)	(0.23)
Surveyor rated interpersonal interaction	0.27*	0.17	0.34**	0.18	0.33	1.11*
skills	(0.13)	(0.15)	(0.10)	(0.13)	(0.31)	(0.55)
(3) Non-cognitive skills: social competence						
Whether having any friends	0.05**	0.05*	0.06**	0.06**	0.09*	0.04
	(0.02)	(0.02)	(0.01)	(0.02)	(0.05)	(0.04)
Number of friends	1.64**	1.33**	1.94**	1.63**	2.05***	2.27**
	(0.46)	(0.48)	(0.43)	(0.47)	(0.55)	(0.66)
Whether being a student cadre	0.10*	0.12**	0.11**	0.13**	0.16***	0.18*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.08)
(4) Non-cognitive skills: attention skills						
Ability to concentrate	0.03	0.10	-0.08	-0.28	-0.007	0.01
	(0.10)	(0.14)	(0.20)	(0.23)	(0.12)	(0.16)
Ability to comply with school disciplines	0.07	0.13*	0.08	0.12	0.06	0.07
	(0.05)	(0.06)	(0.09)	(0.13)	(0.06)	(0.16)
Ability to be organized	0.04	0.11	0.07	0.11	0.14	0.20*
	(0.08)	(0.09)	(0.11)	(0.13)	(0.10)	(0.09)
Ability to control oneself from playing	-0.21	-0.41	0.05	0.08	-0.07	-0.31
	(0.30)	(0.44)	(0.14)	(0.18)	(0.15)	(0.37)

** $p < 0.01$ * $p < 0.05$ † $p < 0.1$

Notes. (1) All the models include sampling weight as well as control for the following covariates: age, gender, whether having an agricultural Hukou, ethnicity, whether disabled, mother's education, father's education, parental marriage status, location, number of months when the child started speaking, and per capita household income. (2) As explained in the text, we focus on the MLM-PSM results of the rural sample, because of the relative comparability between rural attendees with rural non-attendees, as compared to the matching of an urban attendee with a rural non-attendee that didn't account for the prevalent rural-urban contextual differences; at the same time, we present MLM-PSM results of the full sample for the reader's interest, in the second to the last column.