Does Having a Preschool Teacher with a Bachelor's Degree Matter for Children's Development Outcomes?

Xin Gong

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2015

ABSTRACT

Does Having a Preschool Teacher with a Bachelor's Degree Matter for Children's Development Outcomes?

Xin Gong

As part of the complex but intriguing question of what defines a highly qualified early childhood teacher (Kagan, Kauerz, & Tarrant, 2008), there has been a heated policy debate over whether to make a bachelor's degree (B.A.) the minimum education requirement for preschools' lead teachers in publicly funded programs (Zigler, Gilliam, & Barnett, 2011). A mixed and non-causal research base of the effect of a B.A. on preschool-teacher performance and child development outcomes is a partial source of the controversy (Kelley & Camilli, 2007; Early et al., 2007). Particularly, no experimental or quasi-experimental studies have been conducted for this topic (Barnett, 2011b). To fulfill the need for better causal inference, this dissertation first uses a nationally representative sample of American children born in 2001 who attended a preschool in 2005, to estimate the effect of having a lead teacher with a B.A. in preschool on the children's development outcomes assessed at aged 4, based on data from the Early Childhood Longitudinal Survey, Birth Cohort (ECLS-B). The term preschool is an umbrella term for all types of center-based programs.

Based on three rigorous quantitative methods, including ordinary least squares with rich controls (OLS with rich controls) and two quasi-experimental methods (propensity score matching (PSM) and instrumental variables (IV)), this study finds: (1) In the model of OLS with rich controls, four of the eight comprehensive child development outcome constructs at age 4 are affected by teachers' B.A. status. Children with B.A. teachers are shown to exhibit higher early reading and math skills and fewer parent-reported internalizing behavior problems than children

with non-B.A. teachers. No effects are found for story-telling skills, color recognition, parent-reported externalizing behavior problems or approaches to learning skills. Yet the children in the treatment group are reported by parents to have lower social competence. In PSM, B.A. positively predicts math skills and negatively affects social competence. In the IV estimates, a B.A. effect is only found for reducing parent-reported externalizing behavior. Comparatively, the PSM and IV estimates tend to be less statistically significant than the OLS estimates. This difference may be attributed to either bias or heterogeneity, given that the PSM and IV estimate may have removed some endogeneity of the treatment in a better way than OLS but they cannot represent the whole sample---the PSM estimate is for those matched and the IV estimate is only local to compliers. Further, when comparing teachers who have just a B.A. (as opposed to a B.A. or higher) with teachers who have an associate's degree (A.A.), the B.A. is found to have fewer statistically significant effects in the model of OLS with rich controls. Significant effects are found for two outcomes: Having a teacher with a B.A. increases math skills and reduces internalizing behavior problems. (2) There has not been much evidence of differential effects by preschool type, and the B.A. effects are no larger for children from low socioeconomic status (SES) families. Neither does the specialized education in early childhood education (ECE), as measured by whether a teacher has a degree in ECE or a related field and the number of college courses in ECE, interplay with the B.A. effect. (3) The supplemental analysis that uses two steps regression to link B.A., teacher-child interactions and child outcomes also returns some interesting findings. The treatment B.A. is found to increase the frequency of several classroom activities and the quality of teacher-child interactions (i.e., being more sensitive, less harsh, less detached and less permissive); but the two steps of the analysis only provide slight evidence for the mediating role of teacher-child interactions.

Overall, there is some positive evidence of B.A. effects on children's early reading, math, the reduction of parent-reported internalizing behavior problems, the reduction of internalizing behavior problems and positive teacher behavior for the center-attending children in the ECLS-B dataset. Still, the evidence is not very strong given the inconsistency of findings across models and the negative effect of B.A. on parent-reported social competence. Such findings identified by rigorous methods in this study speak directly to the B.A. debate by adding a new piece of empirical information for a new generation of children and teachers; it adds some positive evidence to the pro side. Still, for future research and practice that aim to elevate quality, a full picture of cost-effectiveness and cost-benefit of the B.A. threshold policy is recommended and other teacher quality components should be considered.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

This dissertation cannot be done to this level without the guidance, encouragement, help and support by my professors, colleagues, family and friends.

First of all, I want to express my infinite and sincere gratitude to my advisor and dissertation sponsor Professor Mun C. Tsang. I am very grateful to his generous guidance of this study even since its initial literature review stage. I benefit a lot from his careful reading and constructive feedbacks of the three drafts of the chapters. His support also strengthens my passion to apply economics in early childhood education. I have learnt a lot from him through the journey of my doctoral study, from being a graduate assistant, a teaching assistant, a research assistant, a conference presenter to a journal paper author. His sponsorship of the whole doctoral process is very important. All of the above influences make me grow both academically and personally. I am so lucky to be his student and will cherish it for a lifetime.

I also own enormous thanks to Professor Sharon Lynn Kagan, who spared time to comment on my dissertation proposal and findings from a busy schedule. Her expertise on early childhood education research and policy improves the design of the whole study as well as the understanding of the findings and policy implications. I like her class on *Early Childhood Development and Education I & II* very much.

Special thanks go to Professor Peter Bergman, Professor Kevin Dougherty and Professor Jane Waldfogel for their service in my dissertation committee and their valuable feedbacks on this study. The discussion was so inspiring.

In addition, I benefit greatly from Professor Henry Levin, especially in term of the conceptualizing of the non-cognitive outcomes and the suggestion on peer context for my study.

His passion for lifelong teaching and research has always been a lighthouse for my career. To Professor Jeffery Henig, I want to give thanks for his comments on an early draft of my dissertation proposal and two presentations in the dissertation seminar of our department. I am also grateful to Professor Thomas Bailey, Professor Judith Scott-Clayton and Professor Roy Joydeep, for their helpful comments and suggestions in the *Economics and Education* workshop, and for what I learnt from their courses. I also want to thank Professor Wen-Jui Han and again Professor Jane Waldfogel, who co-taught *the Advanced Methods in Policy Analysis* class. I enjoyed the course very much. I also own thanks to Professor Jeanne Brooks-Gunn for her help in screening my dissertation topics, and to the discussant and audience of Section 6.05 in the American Education Finance and Policy 2014 annual conference in San Antonio. To Professor Rivera-Batiz, I would like to express thanks for his teaching of the *Advanced Microeconomics: Application in Education* course and other courses. His summary of four important elements of success in his retirement gathering have always been the soup for my soul---hard work, persistence, help from others and love.

Particularly, I appreciate the Institute of Education Science of National Center on Education Statistics for the provision of the ECLS-B dataset and the help from Nick Spanos, Douglas Hill, and Xiang Li of the Data Products branch in the office of American Community Survey for the instrumental variable data.

My thanks go to my colleagues and friends in and outside Teachers College. Their company or encouragement intensified my hard work and persistence with this big project. Their valuable friendship has partially defined my life as a Ph.D. student. I appreciate the discussions with Dr. You You, Yilin Pan, Dr. Yifei Li, Dr. Henan Cheng, Dr. Di Xu, Dr. Fei Guo, Dr. Li Yu, Dr. Emma Garcia and Jing Li, which enhanced my understanding of the research questions. I also

want to thank Yuan Wang, Shuangshuang Liu, Xiaoying Qi, Vivian Liu, Xiaotao Ran, Chien-yu Liu, Haogen Yao, Na Feng, Yan Shi, Dakai Li, Feng Hu, Jihye Kim, Robert Shand, Dr. Na Li, Kaidan Chen, Ze Geng, Min Li, Li Chen, Dr. Yu Zhang, Dr. Hui Zhang, Professor Jingli Zhai, Professor Huimei Zhou, Professor Bin Tang, Professor Chun Fu, Professor Lijing Jiang, Dr. Nii Nartey, Barbara Tanner, Shibani Khanna, Christof Spaeth, Wenyan Ye, Dr. Ronghua Guo, Dr. Xiaonan Qi, Hui Mai, Xuan Zhang, Shunan Wu and Dr. Ruohan Wu. Thanks also go to the participants of the *Doctoral Research Seminar* hosted by Professor Mun C. Tsang, including Professor Lei Zheng, Huanhuan Xia and Yipeng Tang, and other friends of Center on Chinese Education.

I also want to thank the Project Hope of China, China Youth Development Foundation-Jiangxi and the donors for the sponsoring of my study from middle school to college. The help from strangers has been very touching. I appreciate the sponsorship of China Scholarship Council for the living expense of the first four years of my doctoral study as well as other scholarship and assistantship from Teachers College, Center on Chinese Education and Mingyuan Foundation.

Finally, I cannot express sufficient gratitude to my beloved parents (Minjun Gong and Hong Zhou) and sister (Jiaping Gong). Their unconditional love and care have always been the drive of my efforts, ever since my childhood. At times of tiredness and hardship, they are my hope. Without their support, I wouldn’t have gone far and achieved so much. Even though they don’t know my academic work, their interest in it is not less. This dissertation is dedicated to them and other family members yet to come.

Chapter 1

Introduction

Amidst the complex but intriguing debate regarding what defines a highly qualified preschool teacher (Kagan, Kauerz, & Tarrant, 2008), there is an important policy discussion about whether to set a bachelor's degree (B.A.) threshold or standard for lead teachers in center-based preschool programs, especially for the public funded programs. Alongside, an accountability movement also draws more attention to teacher's impact on classroom and student outcomes. In this context, this dissertation aims to conduct an updated and rigorous empirical study regarding the effect of experiencing a B.A. teacher in preschool on children's cognitive development, social-emotional development and approaches to learning skills, based on a nationally representative sample of children born in 2001, followed at 9 months, age 2, age 4, and their kindergarten year (2006 or 2007). The operational sample includes children who were enrolled in center-based preschool programs (to be called "preschools" throughout this dissertation) at age 4.

This chapter introduces the background in Section 1.1 and states the key issues under examination in Section 1.2. Section 1.3 presents the definition of key terms used in this dissertation study.

1.1 Background

This study has context in several historical events in the early childhood education (ECE) field. These include: a significant expansion of preschool enrollment in this country, increased public investment in early childhood education, a need for identifying the essential teacher

quality components for elevating quality through rigorous empirical research, and a heated policy debate regarding using a B.A. as a threshold for preschool teachers.

The expansion of preschool enrollment. Based on data from the National Household Education Survey (NHES) for the year of 2005, the majority of 3- to 4-year-olds in the United States (U.S.) attend a preschool program: Approximately 75% of 4-year-olds and 50% of 3-year-olds (Barnett, 2010). According to National Center for Education Statistics (NCES), for the 2005-2006 year, 57.4% of the 4-year-olds were enrolled in center-based care, which includes public programs such as Head Start programs that serve 3- and 4-year-olds, state pre-kindergartens (state pre-Ks) that currently serve mainly 4 year olds[1] and private preschools.[2] The expansion of preschool enrolments are the results of increased female labor force participation (Bassok, Fitzpatrick, Loeb, & Paglayan, 2013) and, more importantly, an enhanced belief in the beneficial effects of preschool education on child development, which were revealed in a rich set of rigid empirical studies conducted in various disciplines, psychology and economics in particular (Currie, 2001; Barnett, 2008; Camilli et al., 2010; Heckman, 2011; Nores & Barnett, 2010; Yoshikawa et al., 2013). Domains of benefits include short-term benefits such as school readiness (including cognitive development often measured by early reading and math achievement at kindergarten entry, as well as social and emotional skills such as social competence and self-control), medium-term effect such as school progress (reduction of special education replacement, grade retention and school dropout), and long-term benefits like increased college attendance, earnings, and the reduction of teenage pregnancy and crime.

[1] Data indicates that about 67% of the state preKs are located in public settings while the rest are located in private settings (Epstein & Barnett, 2012).

[2] In NCES (2012)'s Fast Facts table, "center based care" is defined to be care provided in places such as early learning centers, nursery schools, and preschools. See https://nces.ed.gov/fastfacts/display.asp?id=4

Most of the interventions studies focused on economically disadvantaged children (Barnett, 2011a; Camilli, Vargas, Ryan, & Barnett, 2010, p.599). In studies that included a diversified child population, benefits are found to be larger for children from low income families or other disadvantaged backgrounds (Barnett, 2008; Currie, 2001; Yoshikawa et al., 2013), especially in the cognitive domain (Brooks-Gunn, Gross, Kraemer, Spiker, & Shapiro, 1992; Gormley, Phillips, & Dawson, 2005; Melhuish et al., 2008). Given that both cognitive (e.g., early literacy and math) and non-cognitive early skills (e.g., social competence and attention skills) affect later development (Duncan et al., 2007; Duncan & Magnuson, 2011), as "skills beget skills" (Cunha, Heckman, Lochner, & Masterov, 2006), preschool education interventions might have the potential to close school readiness gaps, later achievement gaps, and social inequality among children from different socioeconomic status (SES) groups (Duncan & Magnuson, 2005; Duncan et al., 2007;[3] National Institute of Child Health and Human Development (NICHD), 2000).

Government's interest in promoting school readiness. Aware of the importance of early learning on future academic success (Ackerman & Barnett, 2005), the government has an interest in getting all students ready for school, which was outlined as one of the six National Education Goals (NEG) originating in the *Goals 2000: Educate America Act* enacted by the Clinton administration in 1994. While there is still not a consensus on what school readiness means, the National Association for the Education of Young Children (NAEYC), a national authority on early childhood education, pointed out that "readiness is more than basic knowledge of language and math", other domains such as physical, cognitive, social-emotional competence and positive attitudes toward learning should also be taken into account when talking about ready children (NAEYC, 2009). The NAEYC's view is toward the whole child approach (Bishop-Josef &

[3] According to Duncan et al. (2007), the strongest predictors of later achievement (grade 3 or 5) are school-entry math, reading, and attention skills.

Zigler, 2011), and is consistent with the definition of the National Educational Goals Panel (NEGP) (Kagan, Moore, & Bredekamp, 1998), where five dimensions for early learning and development are identified: physical well-being and motor development, social and emotional development, approaches toward learning (initiative & curiosity, engagement & persistence, reasoning & problem solving), language development, cognition and general knowledge. [4]

Increased public investments in early childhood education and the need for elevating quality. Along with the interest in promoting school readiness, public investments have been flowing into early childhood education to boost children's school readiness. The total public spending on early care and education was estimated to approach $40 billion in 2011 (W. Steven Barnett & Hustedt, 2011), much more than decades ago when federal spending for the well-known Head Start program was less than $100 million in 1960s (Currie & Neidell, 2007) and when state funding was less than $25 million pre 1970s (Katherine A. Magnuson, Ruhm, & Waldfogel, 2007; Mitchell, 2001). Also, the new century has seen booming state funded pre-K initiatives (i.e., state pre-Ks) (Epstein & Barnett, 2012).

Quality is an important issue for the previously mentioned benefits of preschool education to be delivered to children (NICHD, 2003; Barnett, 2011a; National Research Council, 2012). This is especially important given that the majority of children are enrolled in ECE programs of low quality (Kamerman & Gatenio-Gabel, 2007; Kagan, 2009). The aim for "quality improvement" returned to the policy forum in President Obama's push of $10 billion award in

[4] Two points are worth noting: (1) of the five domains, "approaches to learning" has been identified as the least understood, the least researched, and perhaps the most important dimension (Kagan, Moore, & Bredekamp, 1998); (2) when language and cognition are often taken as cognitive development, social-emotional development and approaches toward learning are frequently categorized under the umbrella term "non-cognitive skills". It is a term often used in economics literature and was meant to embrace all development outcomes that are not "cognitive" (Almlund, Duckworth, Heckman & Kautz, 2013; Levin, 2013). Non-cognitive skills (or non-academic skills, character skills, virtue, etc.) are important not only for their influence on cognitive skills, but also for their intrinsic value for personal development and their long-term effect on personal and social productivity (Levin, 2013; Carneiro et al., 2007). To some degree, physical development may also be taken as "non-cognitive", but it will not be included in this study.

the challenge grants (the Race to the Top Early Learning Challenge) to states (Fuller, 2011), and this includes the goal of supporting a great early childhood education workforce. Teacher quality is the first preschool quality ingredient or mark that comes to mind for many researchers and policymakers when considering how to improve preschool (Fuller, 2011; Pianta, 2011; Kagan et al., 2008; Kagan & Gomez, 2011; NAEYC, 2007; Whitebook, Gomby, Bellm, Sakai, & Kipnis, 2009). The bio-ecological theory of human development states that everyday interactions between adults and children are proximal processes that serve as the primary mechanism through which children develop (Bronfenbrenner & Morris, 2006); and during their time in preschool, the most important adult a preschooler interacts with is his/her teacher. Empirical studies show that teachers are the key determinant of positive preschool classroom experience, and that teacher behavior especially high-quality teacher-child interactions is important in promoting children's academic skills and social-emotional competence (Burchinal et al., 2008; Curby, Rimm-Kaufman, & Ponitz, 2009; Hamre & Pianta, 2007; Hamre, Hatfield, Pianta, & Jamil, 2013; Howes et al., 2008; Mashburn et al., 2008; O'Connor & McCartney, 2007; Vandell & Wolfe, 2000; Whitebook, 2003; Yoshikawa et al., 2013), including a positive association to children's attention skills (Peisner-Feinberg & Burchinal, 1997).

Persistent efforts in understanding teacher quality and its components. Like K-12 education, especially in the background of an accountability movement, policymakers and scholars in the early childhood education field are eager to know what makes a good preschool teacher. What are the essential components of teacher quality that empirically correlate with teacher effects on children's development? This is an intriguing yet complex question for ECE (Phillips, 1987; Kagan et al., 2008). It is intriguing because policymakers want to know how to recruit, train, and support highly qualified teachers for ECE programs and how to set standards for regulating and

monitoring teacher quality, including setting minimal entry requirement. Meanwhile, it is complex in the sense that many of the teacher attributes are not easily observable and/or are costly to measure (e.g., general academic ability, personality traits, belief, professional commitment, knowledge, and interaction behavior in the classroom). In the terminology of economics of information, this creates a lack of information for preschool directors, parents and policymakers (Staiger & Rockoff, 2010; Temin, 2003). While studies on teacher behaviors continue to emerge (e.g., Hamre et al., 2014), teachers' general academic ability and personality traits are least understood.[5]

The policy debate over using a B.A. as a threshold. The initial policy parameters designed to improve teacher quality target regulable features, e.g., teacher degree and certification along with regulations on non-teacher quality factors such as child-staff ratio and group size (Clarke-Stewart, Vandell, Burchinal, O'Brien, & McCartney, 2002). A heated debate over teacher qualifications has taken place in policy discussions regarding early childhood education, with a particular focus on whether a B.A. should be the minimum educational requirement for a teacher in publicly funded preschool programs.[6] The book entitled *Pre-K Debates* provides an excellent collection of opinions on both sides of the debate over "credentials versus competence and support", in which the debate over the B.A. plays a part (Zigler, Gilliam, & Barnett, 2011). Table 1-1 summarizes the major arguments for the B.A. debate from both sides.

[5] Although personality traits began to be empirically examined in the field of K-12 education (Bastian, 2013; Rockoff, Jacob, Kane, & Staiger, 2011), studies are much more limited in early childhood education because of the lack of measures in existing surveys (Thomason, 2011).

[6] Teacher degree and qualifications were once hot topics in the re-authorization of Head Start (e.g., H.R. 1429, 2007); and similar debates took place for state pre-Ks. This debate had its counterpart in the history of K-12 education, before B.A. was accepted as a uniform entry-level requirement in K-12.

Table 1-1 Summary of the B.A. threshold debate for publicly funded programs

Arguments	Points/behaviors for supporting the B.A.	Points for having reservations about B.A.
1	B.A. teachers have better knowledge and skills and are more able to connect research to practices (Goffin & Washington, 2007)	As described in Bowman (2011), many used to think that early childhood teachers do not generally teach academic subjects and therefore need little training (p.54)
2	Supporters cite empirical studies reporting the positive effects of B.A. (Barnett, 2011b; Bowman, 2011); acknowledging that pay increase should go with a B.A. but benefit surpasses cost (Barnett, 2011b)	mixed findings not valid for supporting the policy (Fuller, 2011, p.58)
3	The use of a B.A. threshold make ECE easier to be aligned to the K-12 system (Early et al., 2007; Bogard et al, 2008)	cost: increased compensation (Fuller, 2011) + reduction of racial diversity of the existing ECE workforce (Fuller, Livas, & Bridges, 2006)
4	Degree qualifications of the current workforce is low: less than 50%	too narrow to judge teacher quality just by degree (Kagan & Gomez, 2011; Pianta, 2011)
5	-	Two reasons for opposing the use of a "universal B.A." requirement: 1) there is lack of precedent; 2) there is "variability in state certificate requirements" (Kagan & Gomez, 2011)

Sources: (1) Author's self-conducted summary based on Zigler, Gilliam, & Barnett's (2011) *Pre-K Debates*, chapter 8–14. The leading authors are Steven Barnett, Barbara Bowman, Bruce Fuller, Robert Pianta, Sharon Kagan, Margaret Burchinal and Barbara Willer. Researchers such as Steven Barnett showed positive support for B.A.s, while Bruce Fuller held more reservations.[7]

Supporters cite evidence showing positive B.A. effects from the literature, and believe that a B.A., especially a B.A. with specialization in ECE or a related field, or in other words, B.A. in ECE, contributes to better knowledge and skills in early childhood education, and in turn leads to richer language and more sensitive interactions with the children in the classroom (Barnett,

[7] Fuller (2011) turned his focus on "caring" instead of "college credentials". Many support B.A.s while acknowledging it is not sufficient. For example, Bowman (2011) stated that a "B.A. is necessary but not sufficient". Robert Pianta pointed out that, "a degree is not enough" (Pianta, 2011, p.64). Kagan and Gomez (2011) suggested using "some combination of B.A. plus other strategies" (p.73). Burchinal et al. (2011) suggested focusing on "the content and quality of the degree-granting institute, the context the early educator is teaching in, and the supports the educator receives in the teaching setting" (p.77). Willier et al. (2011) mentioned that a B.A. in ECE was stated as a preference for teacher qualification by the NAEYC (p.80) while their focus is on "teacher preparation" as a whole. (2) Author's literature review regarding the research base for the effect of teacher education in center-based programs for preschoolers.

2011b).[8] Relatedly, some supporters doubt that minimally educated teachers without a bachelor's degree can connect new scientific research about early education with their teaching practice (Goffin & Washington, 2007), especially when the essential knowledge for teaching young children has expanded (Bowman, 2011).[9] Similarly, as pointed out by Bryant et al. (2010), B.A. teachers may respond better to a quality improvement approach or policy, as implied in the National Early Reading First Evaluation (Jackson, 2007). Another argument is that the B.A. standard is required in the K-12 grades (Early et al., 2007; Bogard, Traylor, & Takanishi, 2008).

One more point relates to the fact that the level of educational qualifications of the ECE teaching workforce is low (Kagan & Gomez, 2011; GAO, 2012). The most recent figure from the Bureau of Labor Statistics Current Population Survey (CPS) shows that only 28.4% of the early child care and education workers had a B.A. in 2010; furthermore, if these figures are broken down by different types of ECE, the figures are 23.7% for center-based workers defined mainly as those in center-based programs excluding schools[10] and 61.2% for "school-based" workers as defined by "prekindergarten and kindergarten teachers" or "early childhood teacher assistants" employed in the "elementary and secondary schools" industry (Bassok, Fitzpatrick, Loeb, & Paglayan, 2013).[11] It started only in 2011 and 2012 that more than 50% of all Head Start center-based preschool teachers had a B.A. or higher in early childhood education, or in a related

[8] Also, having well-educated teachers with a B.A. or higher is one of the shared features of high quality ECE programs such as Perry Preschool, Abecedarian, and the Chicago Child-Parent Centers (Barnett, 2011b, pp.52–53) that have proved to be effective by randomized control trials.

[9] This includes new development in the child development research, cultural awareness and the ability to tailor teaching to different groups of children, including children with risk factors such as low-income, home language other than English, and immigration.

[10] Specifically in their paper, "center-based" workers are those who are not self-employed and work in the "child day care service" industry, or have child care occupations (i.e., "child care workers", "pre-kindergarten and kindergarten teachers" or "early childhood teacher assistants").

[11] Although it is impossible to differentiate the percentage for lead teachers and other staff in center-based programs based on the current available data, due to measure inconsistency and a lack of accuracy (National Research Council, 2012), this figure indicates low degree levels of the workers in general.

field with experience, following the mandate in the Head Start Act.[12] For state pre-Kindergarten programs, according to NIEER's state preschool yearbook 2012 (for 2011-2012), only a little more than half (58%) adopted the B.A. as a threshold for a lead teacher, and some of them specifically required a B.A. in ECE. In many cases, the standards apply in all settings in which the state Pre-K program is located. However, some only implemented a B.A. threshold for teachers in public-school settings whereas Associate's degree (A.A.) or Child Development Associate (CDA)[13] is required for teachers in non-public school settings (Barnett, Carolan, Fitzgerald, & Squires, 2012). The state of Rhode Island even extended the B.A. standard to teachers in non-public preschool programs that are least regulated (Child Care Aware of American, 2013). In a nutshell, the overall status of degree achievement for preschool education in this country is different from K-12 education, where a B.A. is the threshold and all teachers hold at least a bachelor's degree.

Explicit opponents or doubters like Bruce Fuller have concerns about the mixed nature of the existing evidence (Fuller, 2011). He pointed out the government interest behind the strategic use of "an empirically futile remedy" (p. 58). Another point against a B.A. remedy is that early childhood teachers do not generally teach academic subjects and therefore need little training, as mentioned by Bowman (2011). An additional concern lies in the following question: whether the

[12] The percentage was 62% in 2012 and 57% in 2011. Data source: http://eclkc.ohs.acf.hhs.gov/hslc/mr/factsheets/2012-hs-program-factsheet.html; information about the Head Start Act can be retrieved from: http://eclkc.ohs.acf.hhs.gov/hslc/standards/Head%20Start%20Act

[13] CDA is a widely recognized credential in early childhood education administrated by the *Council for Professional Recognition* , which requires: (1) 480 hours (about 1 year full time) of experience working with children within the past 5 years; (2) 120 clock hours of formal education/training/coursework (12 credits of education) within 5 years of the application date. More information can be obtained from http://www.cdacouncil.org/the-cda-credential/how-to-earn-a-cda and http://www.earlychildhoodnyc.org/education/CDA.cfm

cost for implementing this policy can be justified by its benefit (Fuller, 2011).[14] Most researchers, however, realize that it is necessary, but too narrow and insufficient to judge the quality of the teacher based simply on the degree (Burchinal, Hyson, & Zaslow, 2011; Bowman, 2011; Kagan & Gomez, 2011; Pianta, 2011; Willier et al., 2011). For example, Pianta (2011) advocated focusing on building professional development supports for teachers to enhance their capability to execute real-time high-quality teacher-child interactions and to make contributions to child development, while using coursework, credits and degrees as the vehicle or incentives. Kagan and Gomez (2011) suggested using "some combination of B.A. plus other strategies" like mentoring and coaching, allowing for state variations (p. 73).

More elaboration on the evidence base of the B.A. effect on child development. Apparently researchers from both sides of the debate have come to different conclusions about the evidence base of the B.A. effect. Therefore a careful review of the empirical studies is necessary. According to my updated literature review, as presented in Chapter 2, the empirical picture for B.A.'s effects on child development outcomes is mixed and the effect of B.A. on some intermediate outcomes such as teacher belief, and teacher-child interactions is mostly positive (Barnett, 2011b; Early et al., 2007; Whitebook & Ryan, 2011). For example, a meta-analysis on a sample of 32 studies conducted by Kelley and Camilli (2007) concluded a "small but significant effect size" (0.16) of B.A. on measures of classroom quality and child outcomes in center-based ECE settings (0.14-0.50 for child cognitive outcomes; 0.03-0.17 for social outcomes; 0.21-0.54 for teacher-child interactions). However, Early et al. (2007), an important study of seven well-known datasets, concluded that the link between the B.A. and the academic skills of the children is mostly insignificant (only two out of six studies reported significant

[14] Examples of implementation cost include: (1) the associated wage cost increase for B.A. achievers; (2) the need for replacement of nonqualified teachers; and (3) a potential reduction in racial and cultural diversity of the workforce.

differences for pre-reading skills, and five out of seven found no association between B.A. and early math skills). Furthermore, an insignificant association was found between B.A. and classroom quality for most of the seven studies examined by Early et al. (2007).[15]

Notably, beyond inconsistency in findings, there are several weaknesses with this literature. First, the lack of experimental or quasi-experimental studies is the foremost weakness of this literature. The effect size and the credibility of the mostly descriptive or non-causal regression studies are still in doubt, as such models were not able to address the estimation biases very well (Kelly & Camilli, 2007; Barnett, 2011b).[16] There are at least four sources of identification challenges that may cause bias to the estimate of the effect: (1) failure to account for prior developmental status of the children; (2) family selection: children of different family background and child characteristics may select into preschools of different quality and thus teachers with differential degrees (Lamb, 1998; NICHD & Duncan, 2003); (3) omitting in the estimation model other confounding teacher quality attributes (e.g., teacher's career motivation); and (4) omitting center quality attributes that might be associated with teachers (e.g., a supportive work environment). These threats to causality have not been addressed or insufficiently addressed in the traditionally non-causal studies. For instance, the quality of parenting may be both highly correlated with child development and teacher's degree qualification. If the estimation model doesn't hold constant parenting quality, the estimated B.A. effect may not be pure: it may also contain some effect of the parenting quality or the ideology of parenting behind it. Another example is, if a teacher's employment motivation is not included in

[15] Specifically, only two out seven found positive associations, 0.44-0.65 in effect size; one study found a negative effect size, i.e., -0.26.

[16] For example, studies underlying the small overall effect in Kelley and Camilli's (2007) meta-analysis paper were correlational in nature, and "stronger causal claims are not possible" (p. 31). There are at best "structural models that model nonrandom assignment of students to teachers and the ways in which teachers may affect classroom experiences" (Barnett, 2011b, p.50).

the estimation model, the estimated coefficient of B.A. attainment may reflect the effect of the teacher's deep motivation to work for children (e.g., viewing it as a calling rather than merely seeing it as a job for income) on children's high quality experiences and development outcomes rather than the effect of the B.A. per se, given that a teacher who is motivated to work for the children with a sense of calling is also more likely to get a bachelor's degree (Kagan, Kaurez, & Tarrant, 2008; Tout, Zaslow, & Berry, 2006).

Second, more than half of the empirical studies used other measures of teacher education instead of the B.A. dummy, such as years of education, which precludes a threshold interpretation. We only know more education is better but this is not helpful in deciding which level of education to be set as a threshold or minimum requirement for teaching in preschools. Third, there are other limitations in the empirical literature. To name a few, most studies comprise on a small sample scale and lack national representativeness; the number of studies for child development outcomes is smaller than the studies for intermediate outcomes such as teacher-child interactions and classroom quality; the old studies often had a narrow focus of child outcomes toward academic skills; and few looked at the heterogeneous effect by program type or by family background. All these weaknesses speak of the importance of conducting a quasi-experimental study of the B.A. effect based on a newer dataset with broader development outcome measures that involve multiple domains (cognitive, social emotional and learning approaches). The need for using quasi-experimental designs with population data is also confirmed by Fuller (2011).

Overall, as the B.A. threshold debate continues in the unreconstructed early childhood education system, rigorous empirical research of the B.A. effect on teacher-child interactions and development outcomes is still very important.

1.2 Problem statement

The complex role a teacher plays in ensuring quality and positive outcomes for children needs nuanced research. To help clarify the above-mentioned policy debate as well as the confusing and inconsistent evidence, and to enrich the literature with better methods, I conduct this empirical study to ascertain the effect of experiencing a B.A. teacher in preschool on young children's cognitive and socio-emotional development, based on a nationally representative survey of children born in 2001, which followed the children, their families, and child care or school settings at 9 month, age 2, age 4, and kindergarten (age 5 or 6). This dissertation study is the first quasi-experimental study for estimating the B.A. effect on child developmental outcomes, aimed at obtaining the B.A. treatment effect through two comparisons, each between two groups of children whose teachers had different degree attainment (i.e., B.A. versus non-B.A., and B.A. versus A.A.). Three methods are used to estimate the B.A. effect: (1) Ordinary Least Squares (OLS) with rich controls; and (2) two quasi-experimental methods, i.e., propensity score matching (PSM) and the instrumental variable (IV) approach. All three models will control for children's prior outcomes at 2 years old, so the models are "value-added" to some degree (Early et al., 2007). Estimating the effect of the B.A. on child development outcomes is the first and major line of inquiry for this dissertation study.

The data come from the Early Childhood Longitudinal Survey, Birth Cohort (ECLS-B), which is 10 years newer than the well-known NICHD dataset that has followed children since 1991. To ensure comparability of teacher measures,[17] I restrict the sample to children who were enrolled in center-based settings at the age of 4, excluding others whose primary child care arrangement was home-based care or parental care at that time. Short-term child development outcomes measured at age 4 (the preschool year) and at age 5 are available for analysis.

[17] Such measures are not comparable between center-based care settings and home-based care settings.

Additionally, accounting for the fact that studies on pre-K tends to report insignificant B.A. effect, additional analyses were conducted to see if the impact differs by types of preschool programs (Head Start, state-funded pre-kindergartens, partially publicly funded private preschools, and exclusively private preschools). Note that private preschools normally belong to "child care centers" in legal terms. Similarly, with an equity concern and accounting for the current policy interest in closing the empirically documented school readiness gap (Lee & Burkam, 2002; McLanaham, Haskins, & Paxson, 2005), this dissertation also tests whether the effect of having a B.A. teacher is larger for children from low SES backgrounds.

Also, this dissertation tests whether "specialized education in early childhood education[18] (ECE) " interplay with bachelor's degree (B.A.) in delivering effects, in other words, whether having a degree in early childhood education (ECE) or a related field (or the number of college-level ECE courses) increases the effect of having a bachelor's degree. Due to limitations in the questions designed for the teacher survey, specialized education in ECE is approximately measured by two indicators: "having any degree in ECE or related field" and "the number of college courses in ECE". ECE courses are those that focus on early childhood education and child development. The ECE related fields include nursing, psychology, elementary education, social work, pathology or special education. Many believe a B.A. in ECE would be a better threshold than a B.A. alone (Barnett, 2004; Bueno, Darling-hammond, & Gonzales, 2010), whereas the empirical evidence on its importance is limited (Early et al., 2007). Therefore, findings from this inquiry may help understand whether specialized education in ECE is a necessity for the B.A. effect to take place, and whether it matters more than the length of formal education.

[18] This term was also used by Epstein and Barnett (2012, p.14).

In a second and supplemental line of inquiry, this dissertation explores the linkages between a B.A. and an important process quality variable, namely, the quality of teacher-child interactions, and tests its empirical relation to child development outcomes. This study uses several frequency and quality measures of teacher-child interactions, including the well-established Caregiver Interaction Scale (CIS) for a sub-sample of children and teachers (Arnett, 1989). A two-step OLS estimation connecting the three factors (B.A., teacher-child interactions, and child development outcomes) is conducted. The first step tests the effect of B.A. on teacher-child interactions and the second step on the effect of teacher-child interactions on child development outcome. The first step examines the likelihood of B.A. education on teacher behaviour in the classroom. The two steps together will shed light on whether the measured teacher-child interactions mediate a B.A.'s effect on the relevant child development outcomes.

1.3 Definition of key terms

Several key terms are used in this dissertation.

(1) "**Preschool**" refers to the center-based programs serving preschoolers aged three and four (or more accurately, between the ages of three and five) in the United States, including publicly funded programs such as Head Start and state pre-kindergartens, other public preschool programs (e.g., those funded locally), and child care centers.[19] It is part of the birth-to-eight early childhood education system. Among these preschools, "Head Start" is a federal initiative that provides comprehensive early childhood education, health, nutrition, and parent involvement services to low-income children and their families. The major form of provision is center-based

[19] According to Barnett et al. (2009), whose findings were cited in Epstein and Barnett (2012, p.10), "excluding children in state PreK center based program, Head Start, or special education programs, approximately 33% of the 3-year-olds, and 35% of 4-year-olds were either in a local public school program, a private child care program or a center-based program during the 2008-2009 school year." Also, in Kagan, Kaurez and Tarrant (2008), "center based programs" were referred to public and private child care centers, Head Start and prekindergarten programs" (p. 23).

and it remains "the federal government's major center-based early childhood initiative" after more than four decades since its inception (Epstein & Barnett, 2012, p.50). "State pre-kindergartens" are mostly funded and administrated by state government, and they follow specific state regulations. An often used definition define them as "school or center-based programs that serve 4-year olds, have an explicit goal of improving school readiness, and are funded fully or partially by the state" (Early et al., 2006, p.178). About two thirds of state pre-Ks are located in public schools, with others in Head Start or private settings (Epstein & Barnett, 2012). Child care centers, day care centers or nursery schools are mostly non-public programs serving young children under the age of 5, which has been the major form of early childhood care and education provision before the introduction of government programs. Note that some of the private child care centers may receive partial funding from the public. Also, Head Start and pre-kindergartens sometimes contract with private providers (Epstein & Barnett, 2012). In a nutshell, throughout this study, "preschool" is a term used to encompass all types of center-based programs for 3- to 5-year-old children, and for the operational sample in this study, it exclusively refers to the center-based programs the sampled child attended when there were at age 4. This is consistent with the definition in ECLS-B dataset's survey instruments, in which child care centers, nursery schools, pre-kindergarten programs, and center-based Head Start are taken as center based. Therefore the operational definition of "preschool" for data analysis includes all center-based non-Head Start programs and center-based Head Start programs selected as the children's primary care arrangement.[20] The umbrella term center-based is used in comparison to "home-based", as widely done so in the early childhood education field. Also, for the analysis of

[20] This is similar to the "center based program" inclusions in National Household Education Survey (NHES)'s survey instruments for 2005 and 2013. In these instruments, "center based program" includes "a day care center, preschool, prekindergarten, or program". The survey instruments can be downloaded from: http://nces.ed.gov/nhes/questionnaires.asp

the B.A. effect by preschool type, four types of preschools (Head Start, state preK, partially publicly funded private preschool and exclusively private preschool) are categorized, and their operational definitions are provided in Section 3.6 of Chapter 3 for research question 1.2.

(2) "**Preschool teacher**" refers to a teacher in a preschool, and in this dissertation, exclusively refers to the lead teacher (or "head teacher", a term that is used interchangeably in the other research and practice). The survey does not include information on assistant teachers.[21] A lead teacher is, as defined in NAEYC accreditation[22], the "teacher", namely, "the adult with primary responsibility for a group of children" who "must spend the vast majority of time" with the designated children.

(3) "**Child development**" refers to "the sequence of physical, cognitive, psychological and social changes that children undergo as they grow older" (Cole, Cole, Lightfoot, & Lightfoot, 2005, p. 2). A whole child development definition includes all aspects of child development: physical, cognitive, and social emotional (Bishop-Josep & Zigler, 2011, pp. 83-88); and they are inter-correlated. In the empirical literature for teacher effectiveness in preschool education, two domains of child development outcomes are often examined: (a) cognitive outcomes, such as early language, reading skills and math skills, in other words, early learning outcomes; and (b) social-emotional outcomes, such as social skills that is often used interchangeably with "social competence" and behavioral problems. Relatedly yet less studied, a third domain contains indicators such as sustained attention, self-regulation and eagerness to learn, and sometimes appears in the name of "approaches to learning" (ATL).[23] Both the cognitive and non-cognitive outcomes are important for the child's future (Carneiro, Crawford, & Goodman, 2007; Levin,

[21] In the survey instrument for early care and education provider (ECEP), "primary teacher" was defined as "the person who spends the most time" taking care of the child in center based programs.
[22] Source: http://www.naeyc.org/academy/teachingstaffdef
[23] This was included as one of the domains of the National Education Goals Panel (Kagan, Moore, & Bredekamp, 1998) and also as one domain in Head Start's definition of school readiness.

2013). Physical development, for example, motor skills, has not been examined in the literature, perhaps because it is less responsive to teachers' skills. This dissertation study will not examine physical development.

(4) "**Teacher quality"** and **"teacher quality component**". "Teacher quality" is defined as "the positive actions and behaviors of teachers" (Kagan, Kaurez, & Tarrant, 2008, p. 41).[24] In this study, a teacher quality component, or an indicator of teacher quality, is defined as a quality-contributing characteristic or behavior of a teacher and, in this dissertation, of a preschool teacher. If such a characteristic or behavior contributes to children's developmental outcomes, either directly or indirectly, it can be defined as a teacher quality component. Such component could be either "structural" or "process" in nature.[25] Based on the empirical literature and practice, such a component may include: (a) Degree qualification. This is a structural feature of a teacher and is often studied in the literature. For this dissertation, the ownership a bachelor's degree is used as an indicator of a college-level education. It is also a policy parameter or policy tool being considered by policymakers who wish to improve teacher quality. When B.A. is used in this dissertation, if not with other terms, it is used without specifying the ECE content, meaning that the degree can be in any major. (b) Specialized education or training. Here specialization means having specific education or training regarding early childhood education.

[24] For clarification, it is important to note the differences among the three related terms, including "teacher quality": (1) teacher qualifications: education, training and credential of teachers; (2) teacher quality: positive actions and behaviors of teachers; (3) teacher effectiveness: whether the teacher improves students/children's outcomes (outcome-based, value-added).

[25] This use of "structural" or "process" to describe various teacher components (including B.A. and teacher-child interactions) is enlightened by researchers' differentiation between "process quality" and "structure quality" for the definition of program quality (Bryant, Zaslow, & Burchinal, 2010, p. 48; Peisner-Feinberg & Yazejian, 2010). According to Peisner-Feinberg and Yazejian (2010), "process quality" refers to the direct experiences of children in early childhood programs, and it is more dynamic and qualitative in nature, requiring more in-depth observation. Structural features are "those aspects of early education programs that can be regulated tend to be more quantitative in nature and are readily observable and/or measurable" (p.21). Similarly, according to Helburn et al.'s (1995) definition (in the Glossary section), process quality refers to the general environmental and social interactions. Definition of process quality is also provided in Pianta et al. (2005, p.145).

A difference between "education" and "training" is noted here.[26] "Education" is often associated with formal education (in a degree program) in the pre-service sense, whereas "training" refers to "professional development activities outside the formal education system" (e.g., workshop, seminar, mentoring), often is informal and in the in-service sense (Maxwell, Field, & Clifford, 2006, p.23). Therefore, specialized education or training can be decomposed into two parts: pre-service education and in-service training. Regarding pre-service education, this dissertation will analyze "specialized education in ECE" in the formal college education settings and in pre-service sense, to be measured by the status of having a college major in ECE or not, and number of college courses in ECE. This term will be further explained in item 6. Another part is in-service training, that is, specialized training in the workplace. (c) Others: such as credential (including CDA and other certificates), professional commitment and teacher-child interactions. Among them, a behavior-related teacher quality component, i.e., teacher-child interactions, is often seen as one of the most important teacher quality components.

(5) "**Teacher-child interactions**", as a concept for teacher's interactive behavior with young children, is defined as a process dimension of teacher quality, mostly used as a crucial mediator for the impacts of some structural components of teacher quality (such as degree, in-service training and experience) on child development outcomes. A comprehensive definition of teacher-child interactions involves three dimensions: instructional support, emotional support and classroom organization, according to their Classroom Assessment Scoring System (CLASS, Hamre & Pianta, 2007; Hamre et al., 2013). Developmentally appropriate, warm, and responsive caregiving plus stimulating instruction are usually seen as high-quality teacher-child interactions (Arnett, 1989; Burchinal et al., 2008; Yoshikawa et al., 2013). This is the quality aspect of

[26] This differentiation is consistent with Maxwell et al. (2006)'s identification of the three components of professional development: education, training and credential (p.23).

"teacher child interactions". In another aspect, the frequency of teacher-child interactions can be measured in terms of the number of times teachers conduct activities with children (e.g., reading, singing, playing games) within a time span (e.g., a week). It is related to the "instruction" domain of CLASS. Both the frequency and quality aspects of teacher-child interactions defined here are consistent with NAEYC's guidelines for Developmentally Appropriate Practice (DAP), which emphasizes three core dimensions of classroom practices with young children: provision of appropriate materials and activities; effective teaching; and teacher–child relationships (Bredekamp & Copple, 1997). For empirical research and monitoring practice, CLASS is the most comprehensive measure for teacher-child interactions, while the CIS developed by Arnett (1989) and the sub component "teaching and interactions" of the Early Childhood Environment Rating Scale-Revised (ECERS-R)'s (Harms, Clifford, & Cryer, 1998) are partial or comparatively narrower measures of teacher-child interactions. This dissertation uses the activity frequency measure and the CIS measure of interaction quality that "largely focuses on caregiver sensitivity and responsiveness" (Burchinal, 2010, p.6), which are the most relevant child specific measures available from the ECLS-B dataset.

(6) "**Specialized education in ECE**". This refers to the specialized education for early childhood teachers acquired through two-, four-, or five-year teacher education programs, including "a broad base of theoretical and practical knowledge of child development to prepare them to work with young children and their families" (Snider & Fu, 1990, p.70). Because of this definition, it is by nature pre-service education, and may be called alternatively "specialized pre-service training in ECE" in other occasions. Note also that this is a term specifically for formal education (A.A. or B.A. programs); it doesn't include CDA that is also pre-service specialized professional development but is granted by organizations that are not the same as the degree

programs (Maxwell et al., 2006). In other words, CDA belongs to “credential”, not “education” and “training” in rigid terms; and the effects of CDA and specialized education in ECE defined here will be analyzed separately in this study. This dissertation uses two operational measures for “specialized education in ECE”: having a degree in ECE or a related field; and the number of college courses in ECE. Due to data limitation, the specialized college education here is not necessarily associated with the B.A.; teachers with a claim of a degree in ECE may have been obtained specialized education from her associate’s degree, if having one.

Chapter 2

Literature Review

Researchers in early childhood education have always been interested in how various teacher quality components may affect children's development, including components related to formal education, training, certification, and teacher-child interactions. This chapter focuses on the B.A. as a teacher quality component and reviews studies on the relationship between the B.A. status and child development outcomes. Section 2.1 gives an overview of the overall situation of the literature and how the B.A. studies are situated in the broad research field on teacher quality. Section 2.2 and Section 2.3 examine the relevant empirical studies on the B.A. effect and classify them into two groups based on the proximity of the outcome measures. Section 2.4 reviews studies regarding the relationship between B.A. and specialized training in ECE. Section 2.5 synthesizes findings of the teacher education effect based on other measures of teacher education. Section 2.6 summarizes the findings of the literature review and identifies the knowledge gaps.

2.1 Overview of relevant studies

In the literature on the effect of preschool teacher's formal education on child development outcomes, there are several literature review papers (Kagan, Kauerz, & Tarrant, 2008; Kelley & Camilli, 2007; Tout et al., 2007; Whitebook, 2002), lots of policy reports (e.g., Barnett, 2003; Whitebook et al., 2009), and a summary study of empirical studies based on seven datasets (Early et al., 2007). Building on existing reviews, I conducted a comprehensive and updated

literature review on the effect of preschool teachers' B.A. attainment on children's cognitive and social-emotional development[27] (Gong, 2013). I regrouped empirical studies on the effect of formal education into two categories by the measures of teacher education: the B.A. dummy measure and other measures (e.g., a continuous measure named years of education and an ordered measure for level of education). My synthesis of the literature is based on all studies regarding "teacher education", but the focus for the presentation of the findings is the "B.A. studies," that is, studies that directly compare preschool teachers with a bachelor's degree (B.A.) to preschool teachers without a B.A.[28]

Empirical studies have examined the relationship between teacher education and child development outcomes, the relationship between teacher education and quality of teacher-child interactions, and the relationship between teacher education and teacher belief in Developmentally Appropriate Practice (DAP). Given the needs to present clearly various relationships among various teacher attributes and behavior measures (or in other words "teacher quality components") and to offer a framework for thinking about potentially important but currently unexamined teacher attributes or behavior, I put B.A. (and teacher education) in a broader conceptual framework for teacher quality, as shown in Figure 2-1.[29] In this framework,

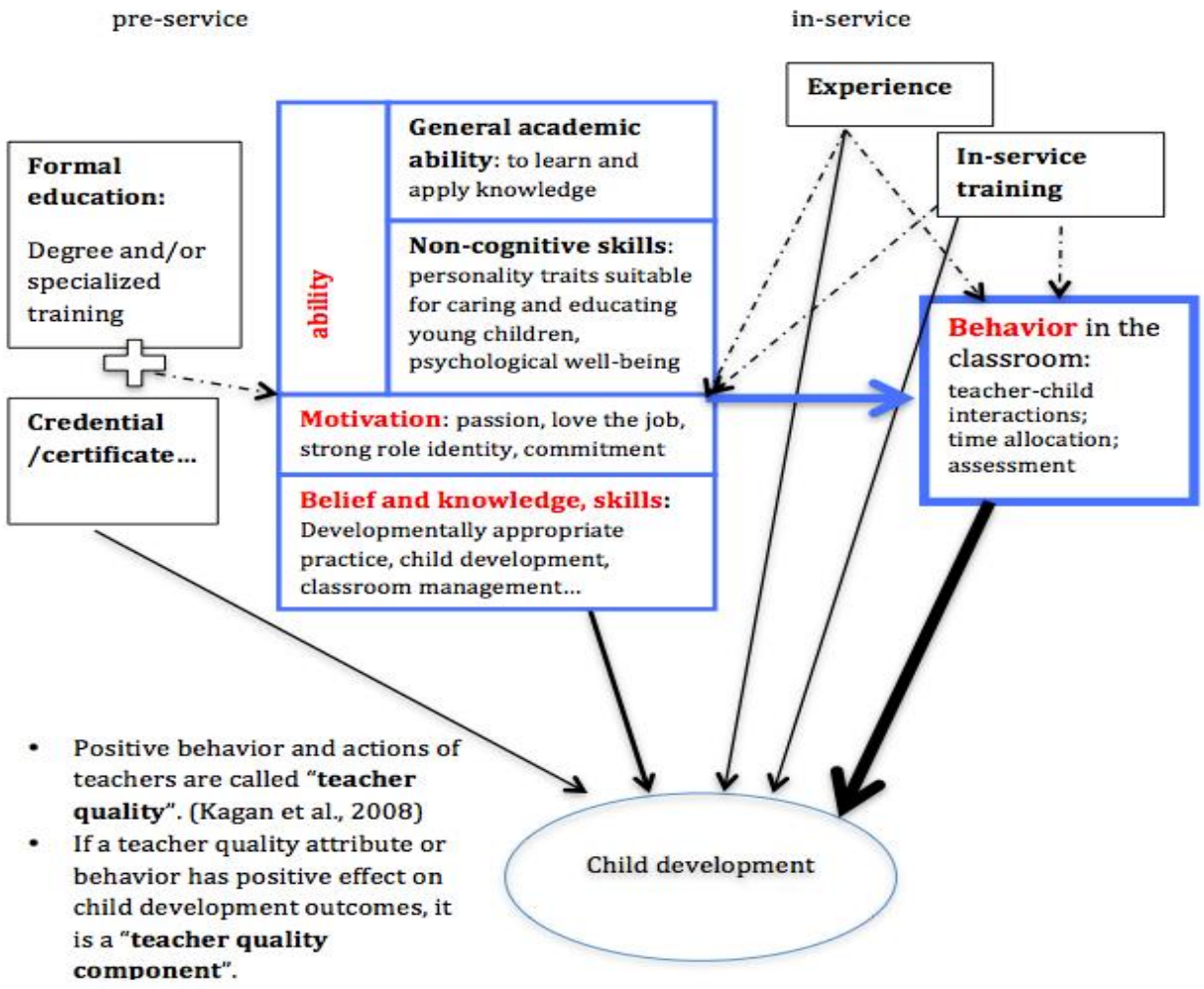
[27] Because the focused early childhood education settings of this dissertation are center-based settings (preschools), my literature review excludes home-based settings. However, studies based on datasets with both home based setting and center-based settings are included.

[28] The B.A. dummy measure of teacher education measures the attainment of a specific minimum level of education. In the context of a research base that has not systematically examined formal education thresholds (Tout et al., 2006), the B.A. studies are helpful in finding answers to an important policy question about threshold, that is, the effect of B.A. versus A.A. as a type of cutoff point analysis (Whitebook, 2003). Note also that Kelley and Camilli (2007) did not single out B.A. studies, although they did differentiate between a continuous measure of teacher education and a categorical measure that allows comparison. Those studies that used a categorical measure of teacher education are called "comparative studies" in their paper.

[29] This way of conceptualizing teacher quality was constructed by the author of this dissertation, accrediting the following insights: (1) efforts at bringing the teacher quality components together (Kagan, Kauerz, & Tarrant, 2008; Early et al., 2007); (2) recent recognition of the importance of incorporating subjective features of teachers in the definition (Colker, 2008; Rohacek, Adams, & Kisker, 2010; Thomason, 2011); (3) the corresponding policy tools used for improving teacher quality.

B.A. is seen as a component of teacher quality. This framework provides a conceptual guide for my literature review.

Figure 2-1 Conceptualizing teacher quality



Note. This framework outlines teacher quality components and their positions. The lightest black arrow (full line) links the most distal teacher quality components, the middle-level heavy line links the teacher quality components that are somewhat close to child development outcomes, and the darkest arrow links the teacher quality component that is most proximate to child development outcomes.

According to Figure 2-1, apart from B.A., there are other teacher quality components, such as credential, in-service training, teacher's general academic ability, non-cognitive skills, motivation (or professional commitment), belief, knowledge about developmentally appropriate

practice, and behavior, which serve as intermediate outcomes between B.A. and child development outcomes. These various teacher factors were identified either as proximal or distal factors in the research (Pianta et al., 2005) and different positions of the components in the conceptual framework define whether a particular component is a "proximal" component or a "distal" component. B.A. is one of the most distal teacher contributors for child development outcomes. Since there is no direct effect of B.A. as a diploma itself on child development,[30] the B.A. attainment of a preschool teacher is more likely to go through some intermediate process (motivation, belief, knowledge and quality of teacher-child interactions) to affect child development outcomes.

With this framework in mind, I reorganize the empirically tested outcomes of the B.A. effect into two sets. One set of studies examines B.A. effect on child development outcomes. The other set of studies examines the B.A. effect on certain intermediate outcomes, such as teacher-child interactions, belief on DAP, knowledge and overall classroom or program quality. The intermediate outcomes are often referred as "ECE quality" (Kelley & Camilli, 2007) or "child care quality" (Vandell & Wolfe, 2000). Among them, teacher-child interactions as a teacher behavior dimension can be taken as "teacher quality", which is defined to be positive actions of teachers in Kagan et al. (2008). Note that some studies examined both sets of outcomes (child development outcomes and intermediate outcomes), although there are not many. Early et al. (2007) is an example.

Comparatively, there are fewer studies regarding the impact of teacher education on child development outcomes than on intermediate outcomes: 24 versus 47. Specifically for the 24 studies on child development outcomes, 12 of them used the B.A. dummy measure. 30 out of the 47 studies on intermediate outcomes have focused on teacher-child interactions, and 14 of them

[30] A related factor, the signaling of the diploma on a teacher candidate's ability is an "indirect" path too.

used the B.A. dummy measure for teacher education. Citations for the 12 B.A. studies for child outcomes and 14 B.A. studies on teacher-child interactions are given in Appendix A. Correspondingly in Section 2.2 and Section 2.3, information regarding data and samples, child outcome measures, measures for teacher education, findings and interpretation of findings will be presented for each set of studies, to reveal the methodological weakness and strength of the existing studies and indicate how this dissertation study can contribute to the literature. In Section 2.4, the empirical understanding about the connection between B.A. and specialized training (including specialized pre-service education in ECE, and specialized in-service training in ECE such as workshops, professional meetings and mentoring) is also summarized.

2.2 Studies for the effects of B.A. on child development outcomes

The 24 studies on child developmental outcomes were published from 1979 to 2011. 12 of them used the B.A. dummy measure, and the key features of the studies are shown in Table 2-1. Most of the outcomes were measured in the preschool years, and therefore concurrent outcomes; only a few were in the longer term, such as the kindergarten years or later, as shown in studies with longitudinal data (Vandell, Henderson, & Wilson, 1988; NICHD & Duncan, 2003).

Table 2-1 Summary of the 12 studies for the B.A. effect on child development outcomes

Study	Data and sample	Age group	Program type	Measures of teacher education	Outcome measures	Outcome dimension	Method	Findings (statistically significant if without further indication)
Vandell and Powers (1983)	55 children from six different daycare centers were observed (location not mentioned)	3-4	Center	Effect of B.A. dummy re-estimated by Kelley and Camilli (2007)	Positive & negative behavior w/ adults, total adult-directed behavior (caregiver interaction quality)	Social	ANOVA	Comparative, effect size is 0.18 for social outcomes when compared to High school diploma
Vandell et al. (1988)	Re-examined the above dataset: initial observation + follow-up observation) in 6 daycare centers, 20 white middle-class children (half girls)	Mean of 51m for initial observation and 100m for later	Daycare centers	Staff training: college degree. Not also: centers were classified as good or poor quality; staff training was taken as one of the center characteristics.	(1) Checklist (interaction: friendly, solitary play); (2) observer-rated (e.g., positive affect, social competence); (3) peer nomination; (4) parent-rated	Social	Hierarchical regression	Comparative, 0.21 for social outcomes (same as above) when compared to high school diploma
Layzer et al. (1993)	1989; 119 classrooms & 5 sites; for each classroom selected, lead teacher (119) interviewed and observed; assistance also interviewed, 1 for each classroom observed	4	Childcare centers	"Teacher education"	Child's task engagement and use of higher level social strategies with other children (children's behavior as defined to proxy child development outcomes)	Cog and social	Correlation --> multiple regression	Comparative: 0.29 for cog and 0.00 for social when compared to High school diploma

Helburn (1995)	Cost, Quality and Outcomes Study (CQO): a four-state sample of 100 daycare centers in one region of each of four states in 1993; for this study, it included a subsample of 401 centers and 826 preschool-aged children	Preschool aged children	Child care centers (evenly distributed for for-profit and non-profit)	"Teacher education": one of the teacher education measures is categorical: "BA/BS"; the other is "years of education"	Receptive language, pre-academic skills (math, reading), attitude/perceptions, positive aspect in T-C relationship, positive behavior, sociability and problem behavior (table 4.3)	Cog and social	Two steps: (1) mean comparison, quality production function equation, hierarchical regression (p.35); (2) mean comparison	(1) teacher education (BA/BS) is important for center quality; (2) center quality is positively related to language and some social outcomes
Burchinal et al. (2000)	89 African American infants attending community-based childcare centers participated in this follow-up study (see Burchinal et al., 1996 for complete details)	Infant, toddler, 2-year-olds and 3-year-olds, mixed	Childcare centers	Years of education; B.A. as supplemental analysis (page 351, at the end of the Results Section)	Overall cognitive development measured by Bayley MDI +3 measures of language development	Cog	HLM (the same as in Table 7)	Correlational: 0.15 is the correlation coefficient, for cognitive outcomes
Burchinal, Howes, and Kontos (2002)	CQO	Infant/toddler /preschooler;2 classrooms in each center	Center: 50 profit and 51 non-profit centers in each state	(1) The highest level of formal education; (2) B.A. in ECE or related field; A.A. in ECE or CDA; completion of ECE courses at a college; workshops only or no formal training	PPVT-R (Dunn& Dunn, 1981): having them point to the picture that matches the word spoken by the examiner	Cog	HLM	Comparative: 0.20 for B.A. (as compared to HS); 0.38 for A.A. (as compared with HS), for language skills

Henry et al. (2004)	A probability sample of children who participated in a full year of Georgia's Pre-K Program during the 2001-02 school year: identified by four-stage sampling	Georgia UPK (Universal PreK): born in 1997-2001 (4 - to 5-year -olds)	State pre-K	Teachers' qualifications: bachelor's degrees vs. associates' degrees or technical diplomas; 80% have B.A. degree ("degreed") or higher ("certified") (table 2.2)	Math and academic skills, and behavior (including curiosity)	Cog, social	Multiple regression	No statistically significant difference, for all 3 types of outcomes (table 4.2 and 4.3)
Early et al. (2006)	NCEDL: 326 children, nearly evenly split between boys and girls (Table 2.1).	1 classroom serving primarily 4-year-olds was selected at random	State pre-K	(1) Years of education; (2) highest degree; (3) bachelor's versus no bachelor's, which is actually a combination measure of education and credentials/major (B.A. and higher in EC or EC teaching credentials)	Academic gains: direct child assessments	Cog	ANCOVA	Overall, not linked (significant only for gains in math skills)
Early et al. (2007)	7 studies (datasets: EHS, FACES, GECS, MAF, NCEDL, NICHD, PCER), 5 of 7 statistically representative, based on randomly selected sample for a population	4-year-old: mean child age, 37m to 56m (Table 1)	Should be center based, including Head Start and state preK	1) Highest level of education (4 categories); 2) bachelor's degree: high school degree/GED, A.A., B.A., and graduate degree; 3) major: ECD/CD; any other education major; non-education major.	Academic skills (receptive language and prereading skills) in the year before kindergarten entry, 4 outcomes: classroom quality, receptive vocabulary, prereading skills, and early math skills	Cog	Value-added specification for HLM	Weak and inconsistent, for academic outcomes

Howes et al. (2008)	From 2 datasets (NCEDL and SWEEP): nearly 3000 children randomly selected, 4 per classroom, from about 700 randomly selected state-funded pre-k classrooms in 11 states (n=2800)	Most classrooms served only 4-year-olds, with a minority serving 3- & 4-year-olds	State-funded pre-K	Teachers completed questionnaires about their educational background and study children: B.A. and EC certification	Pre-academic achievement: growth in school-related learning and social skills over the pre-k year	Cog+ some social	Hierarchical linear model (HLM)	Not related for cog; only marginally significant for social competence (Table 5)
Choi (2011)	Self-administrated survey: 34 teachers and 132 3-5-year-old preschool children in 20 preschools in West Lafayette and Lafayette, Indiana	Preschooler	Preschool of all types	Categorical, BA as one category==>in the final HLM, use the B.A. dummy (have a B.A. or higher or not), see page 30 (Section 4.2.1.1) and Table 5	Math outcomes (Test of Early Mathematics Ability)	Cog	HLM	Not significant for growth in math scores (residualized gains), Table 5
Kim, Chang, and Kim (2011)	Head Start FACES, 2003: 2457 3- and 4r-year-old children from 2003 when they entered Head Start programs to 2005 or 2006 when they graduated from kindergarten	3- and 4-year-olds	Head Start	Combined with teacher training and certification: 6 categories[31] (Section 3.2)	Focused on children's cognitive development in the mathematical domain	Cog	HLM	Positive for math skills (ELL children in Head Start programs)

Note. (a) Only those with the B.A. dummy measure are presented here; and the 12 studies were arranged by the publication year. (b) HS stands for "High School", and A.A. stands for "Associate's degree". (c) "Effect size" in the *Findings* column (last column) is mostly collected from Kelley and Camilli (2007). (d) Studies using teacher education measures other than the B.A. dummies are not presented here but are available upon request.

[31] The six categories are: (1) teachers with high education (bachelor's degree or higher), no preschool certificate, no professional training; (2) Teachers with high education, no certificate, with training; (3) Teachers with high education, with a certificate, no training; (4) Teachers with low education, no certificate, no training; (5) Teachers with low education, with a certificate, no training; and (5) Teachers with low education, no certificate, with training.

Data and samples. Two types of observation data pertaining to sampling differences were employed to empirically test the linkages between teacher education and child development outcomes: site-based data and child-follow up data. The majority used site-based data, which usually select centers from certain cities of certain states, randomly draw one or several classrooms from each of the sampled centers, and then sample all children in the selected classroom or a certain number of children from the selected classroom, or even one child per classroom. The other type is the child follow-up dataset, which usually samples newborn children and their mothers from hospitals, following them at intervals. Examples are the Early Head Start follow-up data, as one of the seven datasets used in Early et al. (2007), and the National Institute of Child Health and Human Development (NICHD)[32] Study of Early Child Care dataset used by three studies (NICHD, 2002; NICHD & Duncan, 2003; Early et al., 2007).

Some samples from the empirical studies have large sample sizes and reflect varying contexts, such as the Cost, Quality and Outcomes Study (CQO) (Helburn, 1995)[33] and some of the datasets used in Early et al. (2007) (e.g., NICHD and GECS/Georgia Early Childhood Study[34]). Some studies, however, have limited sample size for a special context. For example, Vandell and Powers (1983)'s study included 55 3- to 4-year-old middle-class white children in six centers. Also, some focused exclusively on state pre-kindergartens (Early et al., 2006; Howes et al., 2008) or Head Start programs (Kim,

[32] NICHD has tracked the development of 1300 children in 10 sites of 9 states since 1991 from the time they were one month of age, including four phases (birth through three years of age, 54 months through first grade, second through sixth grades, and seventh through ninth grades). Such a longitudinal data allows for the examination of the impact over a longer span. However, the two NICHD studies only used a continuous measure for teacher education, namely, "years of education".

[33] CQO covers 745 children in 100 representative centers of 4 states (California, Colorado, Connecticut, and North Carolina) with complete child assessments, classroom observations, teacher surveys, and parent surveys conducted in 1993 (Helburn, 1995; Peisner-Feinberg et al., 1999).

[34] GECS, as used in Early et al. (2007), contains 630 children from 128 classrooms; it covers "full-day, full-year preschool participants in Georgia including Head Start, Georgia Pre-K and private programs" (p. 563).

Chang, & Kim, 2011). In this regard, my dissertation study based on a national dataset for young children will maintain the strength of examining the diverse context and may be more appealing by its national representativeness.

Child development outcome measures. First, measures for cognitive skills are mostly based on well-established assessment tests such as Peabody Picture Vocabulary Test-Revised (PPVT-R) for receptive language, Woodcock-Johnson tests for early math skills, and Bayley Mental Development Index (MDI) for overall cognitive development, whereas social-emotional outcomes are usually derived from parents or teachers. Some established tests for social emotional tests were used as well, such as the Classroom Behavior Inventory (CBI)-sociability scale for preschool, as used in Dunn (1993). Second, many studies only examined the cognitive domain of development (Helburn, 1995; Burchinal et al., 2000; Early et al., 2006; Early et al., 2007; Howes et al., 2008; Choi, 2011; Kim et al., 2011[35]). Only two studies included both cognitive and social-developmental outcomes (Henry et al., 2004; Mashburn et al., 2008), when they also included outcomes like task persistence and curiosity. That's why the literature has been criticized for using narrow measures of child outcomes towards the academic skills (Bogard et al., 2008; Early et al., 2008). Also, in the B.A. studies, 4 studies did investigate the learning approach domain but did not emphasize it. [36] New studies need to examine the effect of B.A. on all the important domains of development outcomes available, as defined in Section 1.3 of Chapter 1.

[35] They only looked at early math skills.
[36] These include: (1) Vandell et al. (1988), with a measure for "task orientation"; (2) Layzer et al. (1993), with a measure for "task engagement"; (3) Henry et al. (2004), with measures for "task persistence" and "curiosity"; (4) Howes et al. (2008), with a measure for "motivational skills"(social and motivational skills).

Measures of teacher education. Among all of the 24 teacher education studies on child development outcomes, only 12 of them directly estimated the B.A. impact. For example, Early et al. (2006) and Early et al. (2007) are the most relevant studies that directly estimated the B.A. effect. Alternatively, Henry et al. (2004) used a "teacher qualification" measure that was called "teacher credentialing," which had three categories: certified, degreed, and an associate's degree (A.A.). Burchinal et al. (2000) and Howes et al. (2008) estimated the B.A. effect as a byproduct of their analysis on the effect of "child care quality".

Analytical methods. A variety of statistical methods have been used to estimate the B.A. impact, from basic observational non-causal methods to more advanced yet still non-causal methods, such as Ordinary Least Squares (OLS) with a rich set of variables and a change model that controls for prior child development outcomes. Specifically, the simplest way is to calculate the correlation coefficient (Dunn, 1993; Goelman, 1988[37]; Early et al, 2006). Other non-regression methods include Analysis of Variance (ANOVA) (Vandell & Powers, 1983) and Analysis of Covariance (ANCOVA) (Early et al., 2006). ANOVA is a special case of linear regression (Gelman, 2005). It is suitable for detecting an association between a continuous outcome variable and a categorical variable, and is often used by psychologists (Myers, Well, & Lorch, 2010). When categorical covariates are added, ANCOVA is used instead. [38] More complex inference tools that are variations of regression analysis are also employed, including single regression or multiple regression (several regression models modeling different but similar outcomes separately) (Layzer & Others, 1993; Henry et al., 2004), hierarchical regression (Vandell,

[37] This study used teacher education measures other than the direct B.A. dummy.
[38] It is a general linear model that blends ANOVA and regression.

Henderson, & Wilson, 1988), hierarchical linear model (HLM) (Howes et al., 2008; Burchinal et al., 2000, 2002; Choi, 2011; Kim et al., 2011), and OLS with controls for prior outcomes.

However, most of the studies are correlational, in the sense that selection biases are not seriously tackled, precluding a causality finding of the B.A. effect on child development outcomes. Better designs with richer controls were found in NICHD and Duncan (2003[39]), Early et al. (2007) and Howes et al. (2008). All of them controlled for child and family variables so that some of the selection bias in terms of which pre-K program parents selected was eliminated; at the same time, children's entry-level skills were controlled for by using a gain-score model or by including entry scores in the level model. Even for these three studies with some of the identification challenges addressed, they cannot be described as "causal" studies.

Findings. The findings are mixed in that both significant positive effects and insignificant or inconsistent effects on child development outcomes in the preschool years or at kindergarten entry are found. More specifically, according to my review, among the 12 studies that directly estimated the coefficient of B.A. (versus non B.A.), five studies reported insignificant findings (Henry et al., 2004; Early et al., 2006, 2007; Howes et al., 2008; Choi, 2011). Accounting for the fact that Early et al. (2007) reanalyzed 7 large datasets, this study should be given more weight, and the case for mixed picture is stronger (9 out 18 samples are with insignificant findings). The mixed picture of findings also applies specifically to cognitive outcomes and social outcomes when they were considered separately. My conclusion about the relative magnitude of the

[39] Note that their study focused on the effect of overall child care quality. Still, it analyzed the effect of teacher education measured by "total years of education". More importantly, it is very informative in terms of methodology, for modeling the relationship between B.A. and child development outcomes.

effect for these two domains of child development outcomes is that the impact seems to be larger for cognitive outcomes than for social outcomes, which is consistent with Kelley and Camilli's (2007)'s summary Table 4 in their paper (0.50 vs. 0.17). When the B.A. is compared with high school diplomas, according to four studies included in Kelley and Camilli (2007)'s meta-analysis result in their Table 2, significant coefficients have effect sizes ranging from 0.18 (Vandell & Powers, 1983, for social skills), 0.29 (Layzer et al., 1993 for cognitive skills), 0.38 (Burchinal, Cryer, Clifford, & Howes, 2002, for cognitive skills), and to 0.82 (Berk, 1985, for cognitive skills). The only study that allows for the comparison of B.A. with A.A. is Burchinal et al. (2002), and its effect size is 0.18 for cognitive skills.

Worth noting is that, if we further examine those studies that are considered to be relatively better executed, they still showed mixed findings. For instance, the above-mentioned study by Layzer et al. (1993) used data from 119 classrooms and five sites (San Francisco and Richmond counties in CA; Bexar county, Texas; Dade and Broward counties, Florida; Union, Hudson and Essex counties, NJ; Oakland, Wayne, and Washtenaw counties in Michigan). They found a B.A. effect size of 0.29 for cognitive outcomes and 0.00 for social outcomes (Kelley & Camilli, 2007). In another study, Early et al. (2007), the value added model for seven studies with samples from varying contexts (from state pre-K to Head Start, to NICHD's child follow-up sample with all types of care settings), most of the impact estimations of the B.A. on children's academic skills are insignificant and therefore the overall effect was concluded to be none. This study has been taken as the "most methodologically rigorous analysis of this topic", as it used a value added specification for HLM consistent across 7 datasets to account for the nested

data structure where children were nested in classrooms, and as it also accounted for missing data[40] (Bassok et al., 2013). [41]

Additionally, a recent study by Walters (2014) reported statistically insignificant association of the B.A. percentage among all teachers (lead teachers and assistant teachers) in Head Start centers with the average outcomes of children in that center. Although this finding is only a small component of that paper and the analysis is at the center level rather than individual child level, his study still adds some evidence for the insignificant side.

Interpretation of the mixed findings. There are several potential reasons underlying the inconsistency of findings. First, contextual differences may lead to effect differences. One interesting observation is that studies with samples from the state pre-kindergarten (preK) settings tend to document insignificant impact (Henry et al., 2004; Early et al, 2006, 2007; Howes et al., 2008). The impact is not strong in Head Start programs either (e.g., ACF, 2003, as mentioned in Gerde & Powell, 2009). This observation is consistent with Burchinal (2010) saying that "caregiver education modestly predicted better outcomes in community child care, but was not related to either observed quality or child outcomes in Head Start or pre-kindergartens in seven large studies of preschool center-based care (Early et al., 2007)" (p. 4). The lack of impact for state pre-kindergartens may be related to three factors: (1) While B.A. teachers may do a better job in promoting child development, there might be supplementary investments (e.g., greater levels of supports,

[40] The problem caused by nested data structure is that a high value of intra-class correlation (ICC) and in this case the variation-covariance structure of the model needs to be adjusted using hierarchical models.
[41] Another methodologically rigid study, NICHD and Duncan (2003), utilized three types of models (multiple regression models of 54-month cognitive development and achievement scores, longitudinal models of 24-54-month child outcomes, and residualized change models) to explore the robustness of the impact of childcare quality. According to their Table 5, some positive impacts were detected for teachers' years of education on 54-month achievement.

supervision and monitoring in publicly funded programs) to the existing non-B.A. teachers provided by the school or local education bureau to help the non-B.A. teachers become better teachers, especially when accounting for the fact that it is more costly to hire a new B.A. teacher from the applicant pool (Vu, Jeon, & Howes, 2008). (2) The percentage of B.A. holders in the existing workforce in a typical state pre-kindergarten is already very high; therefore, the difference between a B.A. teacher and non-B.A. teacher is not obvious (Tout et al., 2006; Vu et al., 2008), or in other words, B.A. is less useful as a marker of talent (Gormely, 2007). A further understanding for this inference is that non-B.A. teachers might benefit from B.A. teachers, as they could learn from the B.A. teachers the best teacher behavior through self-initiated learning or mentoring. (3) The third reason for not finding an effect was provided in Early et al. (2007). They talked about the differential compensation system between K-12 education and early childhood education systems, which on the one hand attracted high quality non-B.A. teachers to teach older children and on the other hand retained low-quality B.A. teachers in the ECE setting. These two contracting forces diminished the link between teacher education and quality in ECE settings.

Second, sample size differences can lead to differences in terms of statistical power of the estimated coefficient. A positive effect may not be detected if there is not enough variation in teacher education for the sampled teachers.

Third, differences in methods and control variables may be the cause of inconsistent significance of the effect and/or inconsistent size of the effect. If the selection bias is positive, studies that attempt to remove selection bias are expected to have a smaller effect size. Therefore, as Blau (1999) mentioned, “the nature of the data and the methods

used to analyze them may strongly affect the results obtained when analyzing the effect of child care inputs" (p.814). Although the Blau (1999) study is about the effect of specialized education or training in ECE[42] without using a B.A. dummy measure for teacher education, the logic is similar when we think why results for the B.A. effect differ across studies and across models. This highlights the need for better-designed empirical studies that follow the same methodological standard to facilitate across-study comparison.

Fourth, cohort differences and sampling differences may work together to explain part of the mixed findings. Earlier studies (Vandell & Powell, 1983; Vandell et al., 1988; Helburn, 1995) tend to report positive relationship while later studies (Henry et al., 2004; Early et al., 2007; Choi, 2011) show no relationship. On the one hand, the importance of B.A. as a signal for talent is weaker when more people get college education in more recent decades (Gormley, 2007).[43] On the other hand, the content and quality of the B.A. granting programs examined in certain studies that sampled particular child cohorts and therefore particular teacher cohorts graduated from certain B.A. programs in the 1970-1980s may be better than that of the particular B.A. programs for teacher cohorts in the1990s covered in some later studies, and this may lead to inconsistent findings.[44] Or the importance of a B.A. degree may have declined over time.

[42] In this study, the measure of "caregiver training" was not specified in detail, except that "received any education or training specifically related to children such as early childhood education, special education or childhood psychology" (p. 797).

[43] As pointed out by Gormley (2007), "when B.A. degrees were relatively scarce, those who held them were arguably more intelligent, more motivated, or otherwise better equipped. Now that B.A. degrees are more common, they are less useful as a marker of talent" (p.662).

[44] While it is true that the quality of program in later years is higher than before, sampling difference adds another possibility of larger effect for certain studies in the 1980s or 1990s.

2.3 Studies for the effects of B.A. on teacher-child interactions and other intermediate outcomes

Similar to but slightly different from Kelly and Camilli (2007),[45] I reorganized the empirically tested intermediate outcomes into three categories: (1) teacher belief and knowledge; (2) teacher-child interactions; and (3) overall classroom quality or program quality, acknowledging that the first category is correlated to the third and the second category is often one component of the overall classroom quality or program quality. For example, teacher-child interaction ("activities and interaction") is an important "process quality" component for the Early Childhood Environment Rating Scale-Revised (ECERS-R) measure of program quality (Bryant, Zaslow, & Burchinal, 2010, p.48).

Overall, the impact of the B.A. on measures of teacher-child interactions is mostly positive. There are 30 studies that related teacher education to teacher-child interactions and 14 of them used the B.A. dummy measure (B.A. versus non-B.A.) for teacher education. Brief summary of the studies are also provided in Table 2-2 of this chapter. Ten out of the 14 B.A. studies on the quality of teacher-child interactions reported positive findings. The remaining four studies either found insignificant (Blau, 1997; Early et al., 2006) or mixed findings (Pianta et al., 2005; Vu et al., 2008[46]). Note that the average effect size calculated in Kelley and Camilli's (2007) meta-analysis is 0.54 (with a standard error of 0.07) when the two authors compared B.A. with a high school diploma or A.A. by redoing the analysis for the data from their 32 reviewed studies. Their

[45] The six categories used in their study are: (1) class quality; (2) teacher-child interactions; (3) teacher beliefs; (4) child cognitive skills; (5) child social skills; (6) instructional activities. My review wants to differentiate "final outcome" (child development) and "intermediate outcome" (all other categories). For the two categories regarding teacher behavior (category (2) and category (6) in Kelly and Camilli (2007), I use the term "teacher-child interactions" in a in a broader sense, meaning that my definition of "teacher-child interactions" involves instructional activities. It also has two perspectives: frequency and quality.

[46] The effect is only significantly positive in private non-profit centers.

conclusion is mainly based on effect sizes calculated for the comparisons between B.A. and the high school diploma, which is not the policy question of "B.A. vs. A.A.". This renders most studies in the meta-analysis in Kelley and Camilli (2007) not so helpful for policy consultation. Even so, when they compared B.A. with A.A., the average effect size for their reviewed studies is around 0.26, based on Kelley and Camilli (2007)'s table 2 results and my amended calculations for 2 of the 4 studies pertaining teacher-child interactions (0.29 for Honig & Hirallal, 1998; 0.30 for Howes et al., 2003; about 0.31 for Burchinal et al., 2002; about 0.14 for Whitebook, Howes, & Phillips, 1989).[47] These effect sizes indicate that the effect of B.A. is moderately large; and is slightly larger than the effect on child cognitive skills and much larger than the effect on social-emotional outcomes. The above conclusions are unchanged when I append additional studies from my updated literature review.[48]

[47] If B.A. is compared to some college, effect size is 1.46 for Arnett (1989), 0.23 for Whitebook et al. (1989), and 0.14 for Phillipsen et al. (1997).

[48] There are two more observations regarding the findings: (1) Interestingly, some studies with the same dataset but different methods returned different results. For example, Howes (1997) and Blau (1997) both used the NCCSS data but their findings are different. (2) State pre-kindergarten samples (Pianta et al., 2005; Early et al., 2006) tend to be associated with insignificant B.A. effect on teacher-child interactions.

Table 2-2 Summary of the 14 studies for the B.A. effect on teacher-child interactions

Study	Data and sample	Age group	Mean statistics of teacher education for that sample	Measures of teacher-child interactions	Method	Findings: effect size from Kelley and Camilli (2007)
Arnett (1989)	Bermuda: 59 preschool teachers from child care centers in Bermuda in 22 day-care centers	preschooler	100% B.A., has specified the content of training/education (Tout et al., 2006)	(1) 26-item Caregiver Interaction Scale (CIS) developed by the author; (2) childrearing attiitude from Parental Modernity Scale (Schaefer & Edgerton, 1981)	ANCOVA analyses with a priori contrasts	BA vs. HS: 1.18; BA vs. SC, 1.46, pretty large (show positive association with interactions, and less detached, less punitive)
Whitebook et al. (1989)	National Child Care Staffing Study: 83 centers selected from 5 representative metropolitan areas (Atlanta, Boston, Detroit, Phoenix, and Seattle)	infant, toddler and preschooler	B.A./B.S. or more: 12%, 25% and 19% for Black, White, and other minorities (Table 8), 893 teachers (lead and assistant)	(1) appropriate caregiving, from ECERS; (2) teacher behavior (sensitive, harsh, detached), the latter of which derived from Arnett's scale of teacher sensitivity and AIS (Adult Involvement Scale)	(1) multiple regression; (2) ANOVA: analysis of variance (Table 14 & Table 17)	BA vs. HS: 0.41 (larger compared to the other two: SC vs. HS, 0.18; AA vs. HS, 0.27)
Howes et al. (1992)	National Child Care Staffing Study: representative of full time, full year, center based care", 1309 teachers, 45 child care centers	infant, toddler, and preschooler, analyzed separately	30.7% for lead teacher; 18.5% for assistant teacher	a specific index of teacher sensitivity (Arnett, 1989)	(1) tabulation; (2) one predictor regression; (3) +ANOVA (F statistics)	positive and significant for preschoolers (table 4 and table 6); Formal education was a better predictor than specialized training
Layzer et al. (1993) (Observational Study of Early Childhood Programs)	119 classrooms, disadvantaged preschoolers, 5 sites	4 year olds	55%r with a college degrees or higher	(1) interaction with children (Abt Associates' adult focused observation); (2) global rating scale of caregiver behavior (CIS)	correlation+ multiple regression	0.40 for Interaction (more responsive to children, use positive techniques more often, and spend more time interacting & teaching)

Howes (1997)	(1) CQO; (2)Florida Child Care Quality Improvement Study: 150 child care programs in four different Florida counties, 3 classrooms for each	(1) CQO: mean age of 4.25; (2) Florida: 3 classroom in each (infant, toddler and preschooler)	not specified	(1) CQO: CIS, AIS; (2) Florida: the same	(1) CQO: descriptive statistics, F values, and post hoc tests for teacher background differences; (2) Florida: ANOVA	B.A. teachers engaged in more language play and positive management than teachers with a high school education; less sensitive than BA.
Pillipsen et al. (1997)	CQO: "large multi state project", 228 infant/toddler rooms	infant, toddler and/or preschoolers	(1) for infant/toddler classrooms, 8% to 42% by profit/non-profit by state (table 1); (2) for preschool classrooms, 21% to 58% (table 2)	CIS and the Teacher Involvement Scale (TIS)	correlations + hierarchical multiple-regression models/hierarchical regression (table 4 & 5)	for preschool classrooms, mostly positive
Blau (1997)	NCCSS: 1,309 teachers and 664 classrooms in 227 day care center selected through a stratified random sampling process			(1) behavior of the teachers and the appropriateness of the classroom environment ECERS and ITERS; (2) CIS	regression + fixed effects	a positive, statistically significant impact on SENSITIVE and CARE; become much smaller after including center fixed effects, and only post college education remains statistically significant for SENSITIVE
Honig & Hirallal (1998)	81 caregivers in 24 urban centers	three to five-year-old preschool children	25 out of 81 have both high level of training and education; 26 had a BA degree	n domains of positive and negative socioemotional inputs, language facilitation, concept promotion, and caregiving and cleaning up based on ABC (Adult Behaviors in Caregiving) (Honig & Lally, 1973, 1975)	Hierarchical stepwise regressions and ANOVAs	0.29 for BA compared to AA/SC (explained about 15.1% of the variance)

De Kruif, McWilliam, Ridley, & Wakely (2000)	63 child care lead teachers (all women) working in child care centers in central North Carolina	child care centers, 12-36 months: child care centers	8 out of 63 lead teachers have bachelor's degree (table 1)	eight interaction behaviors taken from the Teaching Style	comparison of educational composition by quality clusters	positive ("teachers demonstrating a positive interaction style were more likely to have a college degree ")
Burchinal et al. (2002)	CQO: child care centers in four regions	Both: in each center, 2 selected rooms serving infant toddlers or preschoolers	30% have BA-BS(no higher scale) (table 1)	CIS	ANCOVA+ comparison of adjusted means and SD (table 4)	0.33 for interaction
Howes et al. (2003)	The Best Practices (BP) and Partnership Projects (BPPP): 80 primarily African-American and Latino teachers working in child care programs serving low-income children (78%)	Possibly preschoolers	24% had BA or MA degrees (n=19) (1/6 teacher and 5/6 assistants)	(1) Responsive involvement; (2) Adult engagement(positive initiations, responds positively, positive management, facilitate peer play, language play); (3) learning activities (creative, language arts)	Descriptive statistics for the measured pathways and turns	comparative 0.30 for Interaction and 0.37 for instructional activities (BA compared to AA)
Pianta et al. (2005)	National Center for Early Development and Learning (NCEDL)'s Multi-State Pre-Kindergarten Study: a sample of 238 classrooms representing 6 states' pre-kindergarten programs	4 year old	48% have BA (table 1)	(1)The Teaching and Interactions scale and the Provisions for Learning scale based on CLASS and ECERS; (2) activity setting (free choice, whole-group, routines)	multivariate analysis	mixed (some impact on emotional measure of CLASS but not on interactions measure of ECERS)

Early et al. (2006)	NCDEL's Multi-State Study, well-established state funded pre-K IN 6 states: 237 pre-K classrooms, over 800 children	Mostly 4 year old	59% with B.A. and higher	ECERS-R + CLASS	ANCOVA; a general linear mixed model failed due to small within-class correlation	largely "non associations"
Vu et al. (2008), AIS measure for interaction included	California:279 classrooms (231 finally used), sponsored by 122 different agencies (29 in private, for profit agencies, 48 in private, non-profit agencies, 25 HS classrooms, and 177 California DOE sponsored programs)	4 year old children eligible for kindergarten in the next academic year	less than one quarter (18.0%) of lead teachers held B.A.s	(1) teaching and interaction, and provisions for learning based on ECERS-R; (2) concept development and quality of feedback based on CLASS; (3) AIS (adult involvement)	ANOVA & HLM(intercept and slope, qualification of program director as the only agency(program) level predictor)	significant only in private nonprofit programs and HS/general child care programs

Note. All 14 studies used site-based data.

Measures of teacher-child interactions. Several of the B.A. studies on teacher-child interactions used Caregiver Interaction Scale (CIS, Arnett, 1989) to measure teacher-child interactions (Arnett, 1989; Blau, 1997; Burchinal et al., 2002; Howes et al., 1992; Howes, 1997; Layzer et al., 1993; Phillipsen et al., 1997; Whitebook, Howes, & Phillips, 1989). It contains 26 items aggregated to four scales: sensitive, harsh, detached, and permissive. Simultaneously, some studies used Adult Involvement Scale (AIS, developed by Howes and Stewart (1987) (Howes, 1997; Vu et al., 2008; Whitebook et al., 1989). Both measures highlight the method of interaction, which is mostly emotional in content. Notably, as one of the measures for the quality of teacher-child interactions, teacher sensitivity sometimes substitutes for the concept of teacher-child interactions (Howes, Whitebook, & Phillips, 1992). The most recent measure, however, tends to define the concept broadly by specifying the type of interaction in terms of the nature and content of interaction. In the Classroom Assessment Scoring System (CLASS, Hamre & Pianta, 2007; Pianta, LaParo, & Hamre, 2007), three dimensions of teacher-child interactions are defined: emotional support, classroom organization, and instructional support. Among them, instructional support and emotional support are the two often-used dimensions (Mashburn et al., 2008). Some of the B.A. studies on teacher-child interactions used the "emotional support" and "instructional support" components of CLASS and/or the "interaction" component from ECERS-R (Early et al., 2006; Pianta et al., 2005; Vu et al., 2008). For example, Early et al. (2006) used components both from CLASS and ECERS-R based on factor analysis.[49] Vu et al. (2008) selected ECERS-R subscales ("teaching and interaction" and "provisions for learning") and

[49] For ECERS-R, the relevant factor identified through factor analysis with varimax rotation is labeled "teaching and interactions", and it involves of 11 indicators: greeting/departing, encouraging children to communicate, using language to develop reasoning skills, informal use of language, supervision of gross motor activities, general supervision of children, discipline, staff-child interactions, interactions among children, free play, and group time (alpha=0.88). For CLASS, two factors are generated: emotional climate and instructional climate. This is consistent with a study of kindergarten classrooms (Pianta et al., 2002).

CLASS subscales ("concept development" and "quality of feedback"), and AIS ("adult involvement"). Other measures used in the B.A. studies include TIS (Teacher Involvement Scale) (Phillipsen et al., 1997), ABC (Adult Behaviors in Caregiving) scale developed by Honig and Lally (1973, 1975) (Honig & Hirallal, 1998), Abt Associates' Adult focused observation (Layzer et al., 1993) , interaction behaviors from TSRS (Teaching Styles Rating Scale, McWilliam, Scarborough, Bagby, & Sweeney, 1998) (De Kruif, McWilliam, Ridley, & Wakely, 2000), and self-designed variables that measure responsive involvement, adult involvement, and learning activities (Howes et al. 2003). What is noticeable here is that the measures do differ in some way. For example, CIS focuses on the quality of emotional support, and CLASS adds in the instructional perspective. ECERS-R's interaction component includes break-down quality in specific activities and interactions. Some use proportion of time in a particular type of behavior to define "adult engagement", or use "time spent in learning activities" to define a type of behavior (Howes et al., 2003).

Findings. Among the studies for the B.A. effect on teacher-child interactions, a comparatively important one was Burchinal et al. (2002). This study used the "B.A. in ECE training" measure for teacher education and examined the effects of education and training simultaneously. Their analysis was based on the CQO data, and their ANCOVA analysis found that "care-givers with formal education in early childhood or who attended workshops were rated as more sensitive in interactions with children and as providing higher quality care than other caregivers, even after adjusting for the caregivers' experience and differences related to state, adult–child ratios, and type of classroom" (p.2). The effect size of B.A. was 0.33 when compared to high school diploma and about 0.31 when compared to A.A., according to Kelley and Camilli

(2007)'s table 2[50]. Given, however, that infant/toddler classrooms and preschool classrooms were combined in the estimation (though with dummy indicators), it is difficult to isolate the separate impact for the preschool classrooms. Also using ANCOVA, Early et al.'s (2006) focused exploration of the impact of B.A. based on a large sample of state preK classrooms (237), however, it reported mostly insignificant findings.

Teacher's developmentally appropriate practice (DAP) beliefs were found to be strongly correlated with practices (McMullen, 1999), but empirical studies linking B.A. to teacher belief are small in number and most of them reported positive findings (Berk, 1985; Cassidy & Lawrence, 2000; McMullen & Alat, 2002; McMullen, 2003; Cassidy et al., 1995). For example, McMullen and Alat (2002) found that professionals with a bachelor's degree or higher more strongly adopted[51] as a philosophy overall than colleagues with less education, and coursework specific to working with young children was found to be significant on producing beliefs related to child-initiated learning. Also, Cassidy and Lawrence (2000) found that early childhood teachers with B.A. teachers were better able to articulate their beliefs concerning their practices with young children as compared to A.A. teachers; B.A. teachers were also twice as likely as to provide "cognitively focused" rationales for their curriculum choices than teachers with less education. The B.A. impact was modest but not significant in Bryant et al. (1994). As for knowledge, Snider and Fu (1990) documented positive effect of degree in ECE (as compared to degree in other areas) on DAP knowledge; and more importantly, when B.A. holders are compared to high school diploma holders by Kelley and Camilli (2007), the effect size is 0.56.

[50] The corresponding effect size for A.A. as compared to high school is 0.02. Therefor the effect size for B.A. as compared to A.A. is about 0.31.

[51] B.A. was also found to contribute to more developmentally appropriate practice (McMullen & Alat, 2002), although the impact was modest but non-significant in Bryant et al. (1994).

This paragraph provides only a brief review of B.A.'s effect on teacher belief and knowledge because such effect will not be analyzed in this dissertation study.

Evidence of the B.A. effect on the global measures for overall classroom quality or program quality, which has been heavily studied, is mixed (Blau, 2000; Buysse, Wesley, Bryant, & Gardner, 1998; Early et al., 2007; Whitebook, Sakai, Gerber, & Howes, 2001). Note also that program quality's effect on child outcomes, is mostly positive, with small to moderate effect size (Committee on Family and Work Policies, 2003; Vandell, 2004; Peisner-Feinberg et al., 1999; NICHD, 1999). The effect is not strong in state preK studies (Howes et al., 2008; Mashburn et al., 2008) and mixed for Head Start programs (Zill et al., 2001; Zill et al., 2006).[52] The mixed picture of B.A. effect on program or classroom quality is understandable given that the effect of B.A. on teacher-child interactions is mostly positive whereas teacher quality is just one portion of overall program quality. The often-used overall quality measures include ECERS (Early Childhood Environment Rating Scale) or ECERS-R (ECERS's revised version) for preschool classrooms/programs, ORCE (Observation Record of the Caregiving Environment), CLASS (Classroom Assessment Scoring System, Pianta et al., 2008), and High/Scope quality assessment (Epstein, 1999). These comprehensive measures of the quality of a classroom or a program in general cover both process features (e.g., emotional environment) and structural features (e.g., adult-child ratio, group size, teacher education, classroom structure, space, materials and facilities).[53] Because components like curriculum, facilities, and director quality are different from teacher quality (though correlated), this dissertation study will not analyze the effect of

[52] The effect was also found to last through kindergarten and even into second grade for many skills (Peisner-Feinberg and Yazejian, 2010).

[53] For example, ECERS-R is designed to assess the quality of programs for preschool-kindergarten aged children from 2 through 5 years of age. It consists of 43 items in which "interaction and activities" is a part (de Kruif et al., 2000).

B.A. on overall classroom quality or program quality. Thus I have kept the summary for this literature short.

2.4 The interplay between B.A. and specialized education in ECE

For the B.A. effect literature, when it comes to the impact estimates of B.A., it is often entangled with another teacher quality component examined by researchers, i.e., specialized education or training in ECE, for which positive effect was found for teacher quality but not so for child development outcomes (Snider & Fu, 1989[54]; Fukkink & Lont, 2007; Blau, 1999).

"Specialized training" defines training or education that is specialized for a certain area, and it is about how to care and educate young children properly for the early childhood education field. In the literature, many times "specialized training" is used interchangeably with "training" perhaps because "general training" is not the research interest [55]. As pointed out by Whitebook (2003), some studies also lack specificity regarding what constitutes ECE or child-related training. Some include only college-level education, and some include informal workshops or high school and vocational school training. To summarize, the widely used meanings of specialized training involve two senses depending on when and where it happens: (1) In the pre-service sense, it refers to specialized education or training in ECE or a related filed before entering the profession. Such specialized pre-service education for teachers could be received in 2-year or 4-year degree granting institute, or in an integrated program of specialized training, such as Child Development Associate (CDA) or other credential programs. (2) In the in-service

[54] Snider and Fu (1990) shows that knowledge of appropriate practice is dependent upon academic training in ECE, supervised practical experiences and the number of content areas (such as general education, child growth and development, curriculum and supervised practicum) covered in ECE courses.

[55] Similarly, teacher education (usually "general education") and training (usually "in-service training") are often put under the umbrella of "professional development." For example, in NAEYC (2011)'s glossary for "professional development", professional development encompasses education, training, and technical assistance. Still, pre-service training and in-service training are included in their definition for "training" and NAEYC admitted the possibility of crossover between pre-service training and in-service training.

sense, it means specialized in-service training in the workplace, in the form of workshops, professional meetings, or mentoring. Particularly, in this dissertation study, specialized pre-service training obtained in college degree programs belongs to the first sense, and is called "specialized education in ECE". In some empirical studies, it was measured by major in ECE for the highest degree (Early et al., 2006, 2007) or college courses in ECE (Blau, 1997).

There are two major reasons for the mixture of formal education and specialized training when we actually need to estimate the effect of one factor: measurement difficulties and identification difficulties that result from a high intercorrelation between formal education and specialized training (Whiebook, 2003; Tout et al., 2006). With regards to high correlation, "it is difficult to study a sufficient number of child-care teachers for whom training (specialized training) and formal education is not confounded within the person" (Howes, 1997, p.406).[56] Different B.A. studies have addressed the issue of correlation between teacher education and specialized training differently, either estimating them in separate models, estimating the effect of one while controlling for the other, or combining the two variables into a single variable.[57]

Two observations from this literature are relevant to my dissertation study. First, because the term "specialized training" is used without specification of the timing (pre-service vs. in-service) and form (formal vs. informal) of training, findings regarding the link between B.A. and specialized training in general is helpful in revealing how B.A. and specialized education in ECE

[56] As pointed out by Whitebook (2003), "many in the field who have a BA, for example, have also completed high levels of other training in early childhood education, and this training, when analyzed as a separate variable from college education, may in fact reflect differences in education as well"(p.4).

[57] In my literature review paper (Gong, 2013), based on Tout et al. (2006), I grouped some important studies of the B.A. impact into six types of analysis regarding the way of incorporating specialized training in the empirical estimation. First, some studied the impact of teacher education only, not analyzing or controlling for the influence of specialized training (Type 1), some of them did so because of the concern that it is confounded with education (Layzer et al., 1993). Some added specialized training as a covariate and estimate the impact of it separately, that is, they included both types of education measures (Type 2). Also, some chose to estimate education and training in different models (Type 3). There are also studies combining measures of education and specialized training: either fully combined, in terms of "education with ECE content", or partially combined using a "combined ladder" (Kathryn Tout et al., 2006). These two are Type 4 and Type 5, respectively. Finally, some studies analyzed education and training both in a separate way and in a combined way, and they are classified into Type 6.

(i.e. the specialized in-service training) interact with each other, and the relative importance of them for teacher quality and child development outcomes. According to previous research review papers in the early 2000s that addressed the relationship of teacher background and quality (Barnett, 2003; Whitebook, 2003), the presence of B.A. teachers with specialized training in ECE leads to better outcomes for young children.[58] However, in Kelley and Camilli (2007) meta-analysis, a small difference was found between B.A. with specialized training in ECE and B.A. without, but "the difference was within the range of sampling fluctuation" (p.20), indicative of a lack of incremental effect of specialized training to a bachelor's degree.

Additionally, regarding the relative importance of the B.A. and specialized training on outcomes, the existing empirical literature is still not conclusive. Some conclusions can be derived or inferred from studies that include both the B.A. dummy measure and training in the estimation and in this regard, the relativity is unclear. For example, according to three relevant studies on intermediate outcomes, the effect of teacher education seemed to be stronger than specialized training or equally important in some contexts (Howes et al., 1992; Burchinal et al., 2002), but weaker in other contexts (Honig & Hirallal, 1998).[59] Given the number of studies linking B.A.[60] and specialized training is still limited, more empirical research on the connection between them and their relative roles are needed.

Second, among the B.A. studies with a good measure for "specialized education in ECE", the effect of B.A. with specialized education in ECE is different for different comparisons.

[58] Besides, empirical studies regarding CDA support the potential importance of a bachelor's degree with major in ECE (B.A. in ECE).
[59] In these studies, specialized training was measured by having an ECE major or coursework plus in-service training.
[60] As in other places of this proposal, B.A. means the B.A dummy that is measured in a way to allow for comparison between B.A. and non-B.A., and between B.A. and A.A.

Consider two studies on this subject,[61] higher quality of teacher-child interactions was found in the teachers with B.A. in ECE than teachers without the B.A. in ECE training (Arnett, 1989; Pianta et al., 2005).[62] However, a positive association was not supported in Early et al. (2006) and Early et al. (2007) when they analyzed whether ECE major (as compared to non-ECE major) matter among those whose highest degree is a bachelor's degree or above, which is a policy relevant question. The two studies are the only two studies that actually answered this policy question. New studies are needed to follow Early et al. (2007)'s analysis and test the relative effect of B.A. in ECE versus B.A. in other majors.

2.5 Findings from the other measures of teacher education

As mentioned before, not all teacher education studies used the B.A. dummy measure to estimate the effect of teacher education; instead, some studies used measures like years of education and level of education. Are the findings from these studies similar to the picture for the B.A. studies, i.e., mixed for child development outcomes and mostly positive for teacher-child interactions? Generally speaking, the answer is yes. For child development outcomes, 3 out of 10 studies with teacher education measures other than B.A. reported positive effects. For teacher-child interactions, 11 of the 16 studies with other type of teacher education measures indicated positive effects of teacher education on appropriate teacher-child interactions.[63]

[61] Pianta et al. (2005) used a measure of level of ECE training of which "B.A. in ECE" is a category whereas Arnett (1989)'s measure for specialized education in ECE is B.A. with certification so that the comparison group is those who didn't attend this program.

[62] For Arnett (1989), the comparison group refers to those without training in the B.A. program of Bermuda college; for Pianta et al. (2005), it is A.A. training or no training.

[63] For example, Berk (1985) found that caregivers with at least 2 years of college were more likely than less educated caregivers to display encouragement, teacher direction, and promotion of verbal skills. They were also lower in restrictive behavior. College educated caregivers with a child-related major exhibited more indirect guidance, less restriction, as well as more encouragement of children's self-initiations and verbal expression.

2.6 Summary of the literature review and knowledge gaps

Studies aimed to link B.A. to child development outcomes and intermediate outcomes used varying samples and methods (such as ANOVA, ANCOVA, value added OLS and HLM). We also know that almost all the empirical studies aimed to link B.A. and child development are about short-term outcomes. Comparatively, there are a larger number of studies that link B.A. to the often-studied intermediate outcomes, i.e., teacher-child interactions.

Findings from the studies can be summarized as follows: (1) The findings on the effect of preschool teacher's B.A. attainment (comparing B.A. and non-B.A. teachers) on child development outcomes is mixed. (2) The impact of B.A. on teacher-child interactions is more consistently positive and statistically significant. (3) Comparatively, the link between B.A. and high-quality teacher-child interactions is stronger than the link between the B.A. and children's development outcomes, both in terms of a larger effect size and a more consistent research base (Bryant, Zaslow, & Burchinal, 2010; Commodari, 2013; NICHD, 2003; Pianta, Karen, Paro, & Hamre, 2008). It is reasonable because the B.A. is closer to the intermediate outcome than the final outcome. (4) Some of the studies that attempted to link B.A. with specialized education in ECE showed positive impacts, but the only two studies that examined the policy relevant comparison between B.A. in ECE and B.A. without ECE reported insignificant effects (Early et al., 2006, 2007).

At the same time, there are several important knowledge gaps, some of which can be addressed by this dissertation study.

(1) For the effects of B.A. (or teacher education in general) on child development and on the intermediate outcomes, the major flaw is that there haven't been any quasi-experimental or experimental studies. Therefore we need stronger research designs to strengthen B.A.'s research

base.

(2) Many of the empirical studies are on a few sites in one or several states, lacking national representativeness. The analytical focus of an empirical study in this literature is more often on child care quality as a whole and the estimate for the B.A. effect is often reported as a byproduct. More than half of the empirical studies on the effect of teacher education used other measures instead of the B.A. dummy, precluding a threshold interpretation (Kagan, Kuaerz, & Tarrant, 2008; Tout et al., 2006; Whitebook, 2003). Simply knowing more education is better is not helpful in deciding which level of education to be set as a threshold or minimum requirement for the teachers. And among the limited studies that did involve nationwide or multi-state sites and that also used the B.A. dummy measure, the estimation of B.A. effect on child development outcomes is again often a byproduct because their focused research question is for the effect of CDA-A.A. (Burchinal et al., 2002) or the effect of child care quality (e.g., Howes et al., 2008). Studies using the well-known NICHD dataset only examined the effect of teacher education using the continuous measure of education (e.g., NICHD, 2002). Thus a new study based on a new dataset can enrich the literature by using the B.A. dummy measure for the analysis, making it the major research question, and analyzing it in detail.

(3) Most of the studies with a B.A. dummy measure only reported findings for the comparison between the B.A. teachers (B.A. and above) and non-B.A. teachers (A.A. and below). However, estimating the effect based on the comparison of B.A. versus A.A. is a more germane policy question. Similarly, the comparison between B.A. with specialized education in ECE and B.A. without specialized education in ECE is policy relevant, for which only Early et al. (2006) and Early et al. (2007) studied it. New B.A. studies should try to include in their analysis these two policy relevant comparisons.

(4) Fewer studies looked at B.A. impacts on child development outcomes than on the intermediate outcomes (teacher-child interactions, classroom quality, and program quality). This dissertation study can therefore add more information regarding the B.A. effect on child development outcomes.

(5) The existing research has been criticized for using narrow measures of child development outcomes towards the academic or early learning skills (e.g., early reading and math skills) as compared to social emotional development indicators (e.g., social competence). Also, the use of measures for learning approach dimension of child development outcomes (e.g., attention skills and eagerness to learn) is comparatively limited. Beyond these measures, other measures of child development could be explored, if data are available.

(6) While policy makers are concerned about both the effect of B.A. and the effect of B.A. with specialized education in ECE to guide threshold regulation for preschool teachers, the existing empirical base for the latter is less clear. First, findings from limited empirical studies are equivocal about the relationship between bachelor's degree and specialized training. Neither did they reach consensus on the relative importance of the two features. More studies are needed to test the relationship and their relative importance of B.A. and the major (or specialization) of the degree. Second, the measures for specialized educaion in ECE mostly rely on a dummy indicative of whether a teacher/caregiver has received training in ECE or not (not necessarily pre-service or obtained for the B.A. degree), or whether her major for B.A. is ECE among those whose highest degree is a bachelor's degree or above. More nuanced measures for the quality and content of the degree are needed for better understanding of the B.A. effect.

(7) Studies rarely mentioned theoretical or conceptual frameworks for the pathways of the B.A. effect, but the need to identify the pathway is acknowledged and addressed empirically in

some way. We don't lack studies aimed to identify the effect of B.A. on intermediate outcomes (e.g., teacher-child interactions or classroom quality). However, one important limitation in the empirical literature is that only a few studies like Early et al. (2006) and Early et al. (2007) looked at the B.A. effects on both child development outcomes and intermediate outcomes. It is important to estimate both types of outcomes together in one dataset, and report both the effect of B.A. on child development outcomes and the effect of B.A. on teacher-child interactions. This dissertation study will enrich the literature by estimating the effects for both types of outcomes using the same dataset, and by examining the links between B.A., teacher-child interactions, and child development outcomes.

(8) The B.A. studies on child development outcomes rarely examined effect differences in terms of preschool type, which can be done either by adding an interaction term of B.A. with preschool type in the estimation model or doing it separately for subsamples of children and teachers from different programs. Even when they did, the relevant studies didn't report their results on the differential impact by program type. There are only a few exceptions. For instance, Vu et al. (2008) added an interaction term between preschool type and B.A. and found that in some types of preschool (private non-profit programs and Head Start/general child care programs), having a B.A. does matter, whereas in school district and state preschool programs, having a B.A. does not make as much of a difference.[64] Because of this, for a new B.A. study, it is important to see if the B.A. effect differentials by type of preschool still exist for a new generation of children and teachers. Another important reason for examining differential effect

[64] This finding of a larger B.A. effect in non-public school settings is consistent to the empirical findings based on samples of children in a certain type of preschool settings. Just as stated before, the B.A. was found not to matter for preschool teacher quality and child development in state pre-K programs (Early et al., 2007; Howes et al., 2008) and Head Start programs (Currie & Neidell, 2007). Both the Early et al. (2007) and the Currie and Neidell study were cited by Fuller (2011) as one point supporting his conclusion that a bachelor's degree does not lift child development. Epstein (1999) also pointed out that teachers' formal education was positively related to program quality only in public schools, whereas in-service training was related to quality only in Head Start.

by preschool type lies in that different preschool programs have different mandates, standards, per child investments (and therefore different non-teacher resources), and so on (Epstein & Barnett, 2012; Kagan, 2009). In different programs, "specific features of such programs, such as length of day or curriculum content may interact with teacher background so as to impact program quality and child outcomes differently than in other types of early care and education services" (Whitebook, 2003, p.5). Other dimensions of difference include salary, work environment and support. All in all, to a certain degree, looking at differential effect by program type would help answer the policy question regarding "in what context does B.A. matter" (Barnett, 2011).

(9) None of the studies that used the B.A. dummy measure for teacher education have analyzed or reported the interaction effect of B.A. with children's family background (e.g., low SES). However, it is important to study variation of the B.A. effect by family SES, because children from low SES families may benefit more and this needs empirical test.

For point (8) and point (9), we need newer datasets with more diversified samples of children, teachers, and programs to test whether the effect of B.A. is smaller in Head Start and state preK settings than in other settings, and whether the effect of B.A. is larger for children from low SES families. Such type of analysis is also endorsed by Steven Barnett in National Research Council (2012)'s workshop report regarding the early childhood care and education workforce.

Chapter 3

Research Design

This chapter describes the research design for this dissertation. Section 3.1 concisely presents the two research questions. Linking these questions to the literature introduced in Chapter 2, a brief summary of the significance of this dissertation study is presented in Section 3.2. In Section 3.3, a conceptual framework for child development and an analytical framework for the two research questions are presented. Section 3.4 introduces data and sampling, followed by Section 3.5 which outlines the measures of child development outcomes, teacher–child interactions and teacher education. Sections 3.6 and 3.7 respectively explain the methods for estimating the B.A. effects and the two-step OLS connecting B.A., teacher–child interactions and children's development outcomes. Section 3.8 summarizes the research design presented in this chapter.

3.1 Research questions

To fill some of the knowledge gaps in the current literature, this dissertation study aims to answer two research questions. The first research question deals with final child development outcomes while the second one with the intermediate outcome of teacher–child interactions.

Research Question 1 (RQ1): For a representative sample of American children born in 2001 and observed to be in preschool at age 4, what is the effect of experiencing a preschool teacher with a bachelor's degree (B.A.) on children's development outcomes during the preschool year (age 4) and one year later at age 5? With regard to the B.A. treatment, two comparisons will be analyzed: (1) children whose teachers have a B.A. or higher and children whose teachers do not

have a B.A.; (2) children whose teachers have a B.A. or higher and children whose teachers hold an Associate's degree (A.A.).

Research Question 1 is broken down into four sub-questions:

Sub-question 1 (RQ1.1): What is the effect of having a B.A. teacher in preschool on children's development outcomes at age 4 (the preschool year) and one year later at age 5?

Sub-question 2 (RQ1.2): Does the impact of B.A. vary by preschool type? In this study, preschool types include Head Start, State Pre-K, partially public funded private preschools, and exclusively private preschools.

Sub-question 3 (RQ1.3): Is the impact of B.A. larger for children from low socioeconomic status (SES) families?

Sub-question 4 (RQ1.4): Does specialized education in ECE interplay with the B.A. effect? Specifically, does having a degree in ECE or the number of college courses in ECE increase the effect of B.A.?

RQ1.1 is the major sub-question of RQ1 and the focus of this dissertation study; RQ1.2, RQ1.3 and RQ1.4 are extended analyses of RQ1.1.

Research Question 2 (RQ2): During the preschool wave (age 4), what are the relationships between a B.A., teacher-child interactions and child development outcomes? In other words, what is the role of teacher-child interactions in mediating the effect of a B.A. on child development outcomes?

3.2 Significance of the study

This study has several implications. It is the first quasi-experimental study for the effect of a B.A. on child development. Even though it is not an experimental study, findings from two

quasi-experimental methods (PSM and IV) move understanding beyond that of prior work, because a quasi-experimental method takes a substantial step toward casual-inference analysis than simple regression (Fuller, 2011).[65] Also, this study improves estimation over previous studies through enhanced richness of controls, based on the detailed information available in the Early Childhood Longitudinal Survey, Birth cohort (ECLS-B) dataset. For example, the model of OLS with rich controls incorporates many confounders missing in other studies, such as prior cognitive and social development outcomes at age 2, maternal depression, number of books at home, the quality of parenting, length of preschool day, auspices (for profit or nonprofit), and home-preschool connection.

Second, it is the first to use a post-2000 nationally representative sample from the ECLS-B that allows for the diversity in children's backgrounds and center quality to conduct an effectiveness analysis for teachers' bachelor's degrees (B.A.), both on the intermediate outcome of teacher-child interactions and the final outcomes of child development. The effect of B.A. for a new generation of teachers on the child development outcomes of a new generation of children may differ from the previous generation and this generates new knowledge. Apart from the effect of a B.A., the national representativeness of the children in this study also means that this study's results for the magnitude and effect sizes of various predictors of child development outcomes in the early years can be generalized to children across the country.

Third, this study also contributes to the literature on the often-missed outcomes, such as eagerness to learn, persistence, independence, self-control and attention skills. These outcomes are important in predicting academic skills and school readiness, but were not analyzed or emphasized in previous studies about the B.A. effects.

[65] This is especially valuable as conducting an experimental study for the B.A. effect is costly (Barnett, 2011b).

Fourth, the two extended analyses on the heterogeneous effects by preschool type and by family SES would help policymakers decide on the degree of flexibility and select a socially efficient as well as equitable target when they set up the B.A. threshold policy. It attempts to answer for which type of preschool and for which group of children a B.A. teacher is important. Also, the other extended analysis regarding the importance of an ECE major for the degree would address policymaker's interest regarding an intriguing question: Does a bachelor's degree serve only as a screening tool for academic aptitude and persistence (Tout, Zaslow, & Berry, 2006), or does it actually elevate a preschool teacher candidate's knowledge, skills and practices in the preschool classroom and more importantly, affect child development outcomes?

Fifth, guided by a conceptual framework for teacher quality, this dissertation seeks to extend the research on the effect of B.A. by focusing on the decomposition of teacher quality into components other than teacher qualifications, and to meet the policy needs by partially exploring the sequential paths of the B.A. effect through an essential intermediate outcome (i.e., teacher-child interactions). To some degree, this study therefore enriches the empirical literature by continuing to explore the full linkage of the B.A. input, process and child development outcomes with the same dataset. It helps to identify important factors that contribute to teacher quality and child development outcomes that are essential for the future of children and society.

Altogether, findings from the empirical investigations in this dissertation will speak to the B.A. threshold debate. For policy use, it is important to distinguish between an association and a causal relationship; it is "more than just an academic exercise – it is crucial to policymakers" (Goldhaber, 2007, p.3). Although the three methods in this study are not experimental, they control for more potential confounding variables of B.A. and suffer less from bias than simple

methods used in some prior studies. They can be informative for the next policy steps regarding the qualification requirement of the ECE teaching workforce.

3.3 Conceptual frameworks

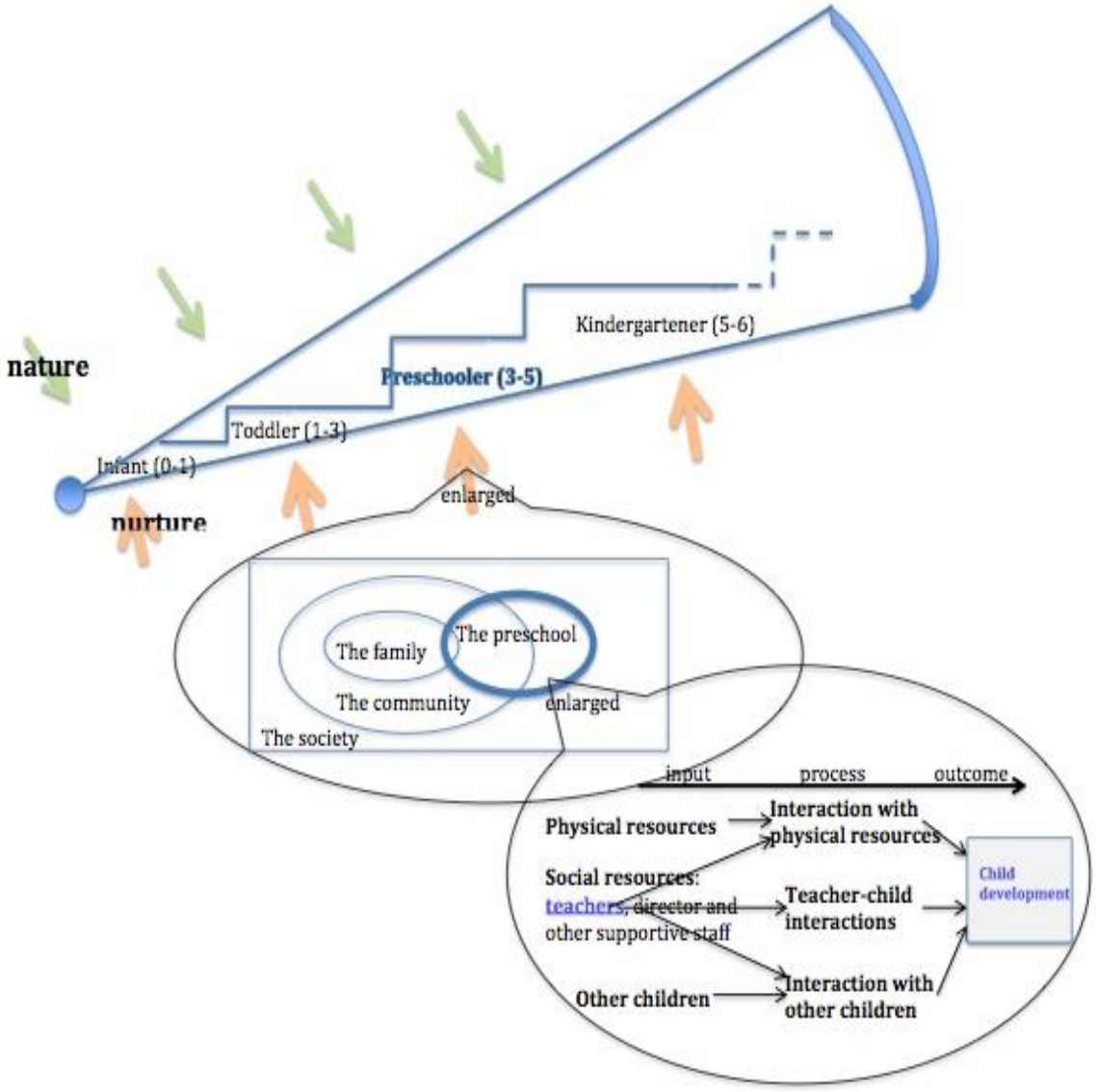
As mentioned in the literature review, the theory and conceptual framework is quite limited for this applied research question of the effect of a B.A. on developmental outcomes. To facilitate understanding of the B.A. effect on child development (RQ1), this dissertation situates the B.A. as a component of teacher quality, which was shown earlier in Figure 2-1 of Chapter 2. With that in mind, it is important to note that teacher quality can be taken as a contextual factor in the model for child development, as presented in Figure 3-1 of this chapter (Subsection 3.3.1). Drawing on these two frameworks, I also established a simplified analytical framework to guide my empirical efforts in linking the B.A. to an important teacher behavior dimension (i.e., teacher–child interactions) and to child development outcomes in RQ2. This simplified analytical framework is shown in Figure 3-2 of Subsection 3.3.2.

3.3.1 A conceptual framework linking teacher quality to child development

Applying the bioecological model of child development for preschool programs (Bronfenbrenner & Morris, 2006; Mashburn & Pianta, 2010; NICHD, 2002), a unified theory for child development (Sameroff, 2010), and NCES's conceptual model for the ECLS-B data collection, and the theory of education production function (Coleman, 1988; Cohn & Geske, 1990), I construct a synthesized conceptual framework for studying the relationship between teacher quality and child development. According to the upper section of this framework shown in Figure 3-1, child developmental outcomes are influenced by the genetic construct ("nature")

and contextual factors ("nurture") grouped in several social biological subsystems or environmental subsystems in society as a main system.

Figure 3-1 A conceptual framework linking teacher quality and child development



At this stage, preschool is an important subsystem for child development (Nores & Barnett, 2010). It has an input–process–outcome mechanism. In this preschool subsystem, teachers and physical resources contribute to the process of care and education. The widely used measures for

preschool program quality cover the quality of both types of resources. Preschool quality is often conceptualized in terms of structural and process domains (Currie, 2001), and teacher attributes or factors can be viewed as having both a structural domain and a process domain.

Similarly to K-12 education, the teacher is one of the most important variables that contribute to children's learning and development. As shown in the bottom enlarged graph for the preschool subsystem, the characteristics of a preschool teacher need to translate into a certain level of teacher–child interactions in the classroom so that the child can benefit from the teacher. Interactions define experience. According to the bioecological development model, everyday interactions between adults and children are proximal processes through which children develop (Bronfenbrenner & Morris, 2006); they serve as the primary mechanisms. Various teacher quality components have been listed in the conceptual model for teacher quality (Figure 2-1). By specifying each characteristic and their pathways to child development outcomes, we are actually exploring the educational production process or, in other words, the mechanism of the impact of a particular teacher characteristic. Also, another understanding of the child development model is the interplay between teacher resources with other non-teacher resources (e.g., physical resources, center director, staff, and other children). Take physical resources (such as classroom materials and interest areas) as an example, on the one hand, children's interactions with physical resources are largely dependent on the teacher's actions: As the only adults in the preschool classroom, teachers are central in terms of organizing activities using physical resources. On the other hand, a minimum level of physical resources may be required for the B.A. effect to take place. Accounting for this possibility, as an additional analysis after the main model without interactions, this dissertation also tests whether an interaction term of teacher education with other preschool resources contribute to child development outcomes.

In summary, a preschool teacher's B.A. attainment is a component of teacher quality. It affects child development outcomes by being part of the preschool environment.

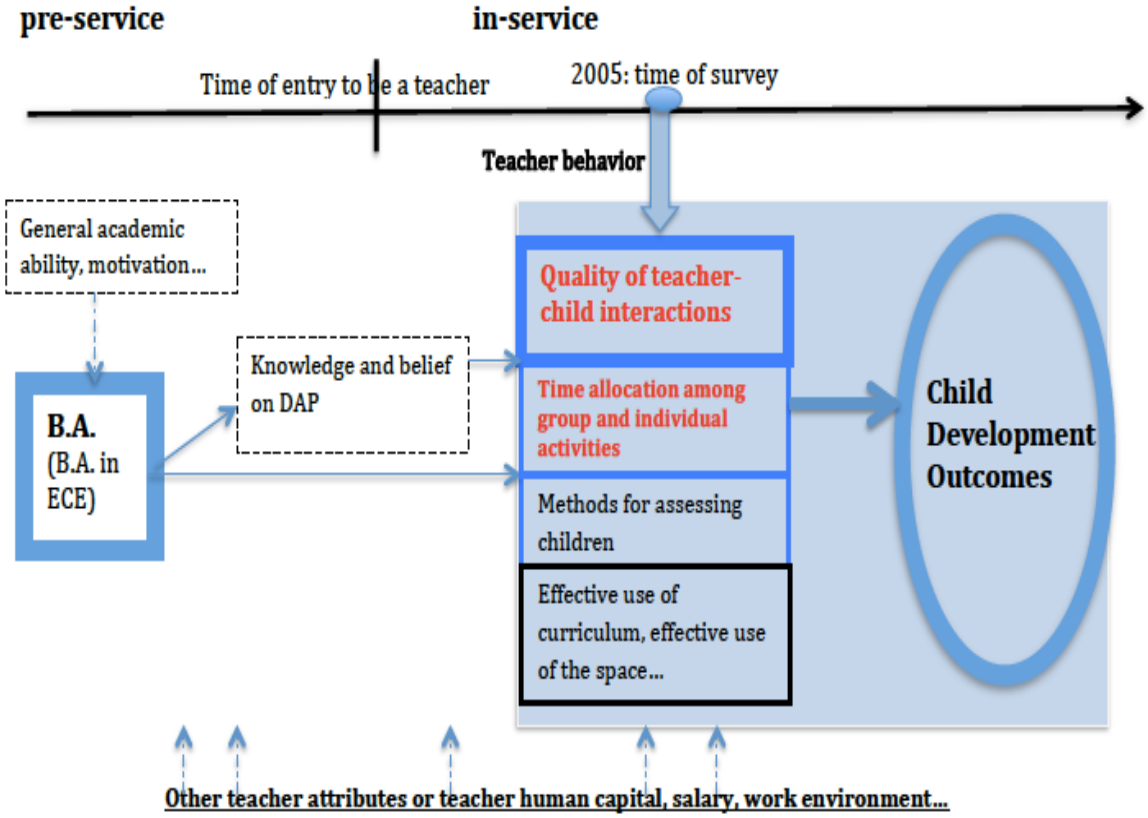
3.3.2 A simplified analytical diagram linking the B.A., teacher behavior and child development outcomes

Based on these understandings, I have constructed a simplified analytical diagram to guide both of the two research questions of this dissertation. As shown in Figure 3-2, it incorporates teacher behavior (i.e., teacher-child interactions and other dimensions of teacher behavior) into the pathways of the B.A. effect. It is a simplified version of Figure 2-1.

Figure 3-2 shows that, before being a preschool teacher, a person makes a decision about whether to get a bachelor's degree for the career of being a preschool teacher based on his/her ("her" will be used hereafter for brevity) own interest; she may also be affected by the regulation context (in some places a B.A. is an entry requirement). It is possible that a person who feels a calling to work with young children decides to get a B.A. in order to be a better teacher. A person may also attain a B.A. only because she has a higher academic ability and could use it to gain better job opportunities in general, even before she makes her decision to be a preschool teacher. The motivation and general academic ability of teachers are two factors behind the B.A. degree that may underline some of the B.A. effect on child development. They could be the confounders of B.A. in the estimation model for a B.A. effect. If the two factors are not controlled in the model, the B.A. effect estimate may be upwardly biased. Yet this issue is not problematic for the purpose of identifying a high-quality preschool teacher at the hiring stage, if B.A. is a valid signal for higher motivation and academic ability that might also contribute to teacher performance and child development. Note also, the person's family background may

affect her decision to pursue a B.A. It is possible that a teacher candidate from low SES background may underinvest her college education when faced with high cost and thus she is less likely to pursue a B.A. than teachers from better family background. Yet this dissertation is not able to empirically account for this factor due to the lack of sufficient information.

Figure 3-2 A simplified framework linking B.A., teacher behavior and child development



Note. The dashed arrows cannot be empirically tested in the ECLS-B dataset directly but may be inferred. The factors in the dashed boxes or black boxes are measured or tested directly in this dissertation study. DAP stands for developmentally appropriate practice.

After receiving formal education and getting a B.A. (especially a B.A. in ECE or a related field), the person then becomes a teacher in preschool. If the B.A. training improves teacher's knowledge and skills about how to interact with children, organize activities, use the curriculum effectively, and assess children properly, the teacher may exhibit a higher quality of behavior in the classroom. An additional channel for the B.A. effect could be that B.A. teachers are usually

more highly paid than non-B.A. teachers (Barnett, 2003b; Whitebook, Phillips, & Howes, 2014), which may affect their professional commitment or satisfaction with the profession and their emotional well-being, which can be further transmitted into positive teacher behavior and child development outcomes.

The first research question of this dissertation estimates the connection between the attainment of a bachelor's degree and child development, and this framework implies that the quality of teacher behavior, including teacher-child interactions, may be an important process variable that mediates the connection between the B.A. and various child development outcomes. In the second research question, this dissertation study will analyze one dimension of teacher behavior, namely, teacher-child interactions. Note also that, another two dimensions of teacher behavior, time allocation and assessment method, are presented in Figure 3-2 only to facilitate thinking and to guide extended analysis in the post-dissertation stage, they will not be examined in this dissertation study.

The framework, when linked to Figure 3-1's model for child development, implies that the B.A. effect may differ across different domains of child development. For example, the B.A. training may tend to help teacher candidates gain superior knowledge about academic skills and improve their own pedagogical strengths in delivering the knowledge, whereas teachers' emotional support to children might be more related to teachers' personality skills that have little to do with the B.A. training. If so, we may find that the B.A. has a larger impact of on children's cognitive and academic outcomes, than on social-emotional development. Findings for the two research questions of this dissertation study will speak to this possibility.

Overall, three conceptual figures guide this dissertation study. Figure 2-1 in Chapter 2 depicts the components of teacher quality and how they are related to child development. Figure

3-1 is the conceptualization of child development, introducing the system of genetic and environmental factors in which a preschool teacher is an input of child development. Figure 3-2 can be viewed as a simplified version of the other two figures. It highlights the two research questions of this dissertation study and shows how B.A. may be related to children's development through several potential pathways.

3.4 Data and sampling

Data for this study are from the Early Childhood Longitudinal Survey, Birth Cohort (ECLS-B), a longitudinal dataset that follows a sample of about 10,700 children born in 2001. Sampling was conducted mainly according to the birth certificate frame (or in some cases, replaced by the hospital sampling frame) by primary sampling unit (PSU).[66] This sample is designed to be representative of American children born in 2001, except that it oversampled Chinese children, other Asian and Pacific Islander children, American Indian and Alaska Native children, twins, and children born with low and very low birth weight. Data collection started when the children were 9-months-old, with follow-ups at age 2, age 4, and in kindergarten. Because children attended kindergarten at different ages, the follow-up year for kindergarten data collection is different for two subgroups of children: (a) for 75% of the children who attended kindergarten at age 5, it is 2006; (b) for the remaining 25% of the children who attended kindergarten later at age 6, it is 2007.

Table 3-1 shows the different stages of data collection for the same cohort of children. Parental interview, direct child assessment, and phone interviews with the child's early care and education provider (ECEP), were conducted in the preschool wave (the age 4 follow-up). Among them, child assessment was done for each wave. Apart from the outcomes measured in the

[66] There are about 100 of such PSUs in 46 states and D.C., either individual counties or groups of adjacent counties.

preschool wave, child development outcomes at age 2 (though in many cases the measures are different from measures used in the preschool and kindergarten waves) and in the kindergarten year (age 5 or 6) are also available. Furthermore, preschool teacher, as this dissertation study's focal subject, is one type of early care and education provider, i.e., the ECEP in center-based preschool programs. The ECEP instrument includes questionnaires for the focal child's caregiver. Very importantly, for a subsample of children, direct observations of 3.5 hours duration of the child care settings were also conducted. This subsample is called "preschool child care observation subsample" in the Data File User's Manual.

Table 3-1 Display of the ECLS-B sample by wave

Wave	Longitudinal follow-ups for children born in 2001	Sample size
1	9 months old: 2001–2002	about 10700 participated in this first time survey
2	Age 2: 2003	9800 or 91% completed the preschool parent interview
3	Age 4 (preschool): 2005	(1) 8900 ECEP cases, and among them, 8750 include child assessment information: (2) 5300 used in center-based care or Head Start, and about 4500 have information on teachers' education attainment
4	Age 5 (kindergarten 2006, 75%): 2006–2007	85% subsampling (a reduced sample) in order to reduce field costs; about 75% enter kindergarten in 2006
5	Age 6 (kindergarten 2007, 25%): 2007–2008	The remaining 25% from wave 4 enter kindergarten in 2007

Note. This data structures can be summarized as "one cohort with 5 waves of follow-ups". All sample size information in this table and in the rest of the dissertation was rounded to the nearest 50, as required by the data provider, i.e., Institute of Education Science (IES) of National Center for Education Statistics (NCES).

Due to sample attrition, about 8900 children participated throughout until the preschool wave of the survey. Since our main analysis is based on the preschool wave (2005) data,

sampling weights accounting for attrition and oversampling must be used to make the sample estimates representative of the nation's 4-year-olds in 2005. This dissertation applies weights at every step of data analysis.

Among the 8900 children followed up at age 4, around 5300 of them were in center-based programs: about 4150 in "center-based care" and 1150 in "Head Start" according to ECLS-B's terminology. Here center-based care refers to centers other than Head Start. Given that not all Head Start programs are center-based, I only count those center-based Head Start programs (defined by the reported information from the center director instead of parents),[67] and under this restriction, the total number of children in center-based programs/preschools number about 5200. This leads to a weighted percentage of about 57% (5200 out of 8900) enrolled in center-based programs/preschools. This ratio is quite close to the one presented online for the national statistics in NCES's Fast Fact table.[68]

For those 4-year-old preschoolers in preschools, about 4500 had information on preschool teachers' degree/educational attainment (B.A.), and therefore the missing rate for this key predictor is not high.[69] Note that the sample of preschools and the teachers are not necessarily representative of the preschool settings or the preschool teacher workforce in this country, because the teachers were identified by the children (one to one) rather than randomly selected from preschools directly. However, the preschool teachers in the dataset stand for the preschool

[67] The identification strategy of Head Start centers here is to use center directors' reports on the location of the Head Start program, specifically by excluding those located at home.
[68] See https://nces.ed.gov/fastfacts/display.asp?id=4
[69] Admittedly, in addition to non-response of this key variable of interest, the definition of some controls may lead to a reduction of effective sample size (e.g., 96% of the preschool wave parent interviewees are mothers or a female guardian), but this will not significantly reduce the operational sample size.

teachers of the representative sample of children enrolled in center-based programs,[70] as also pointed out by the "FAQs for data analysis" section on the ECLS-B dataset's website.[71]

This leads to an operational sample of around 4300 with non-missing sample weight and non-missing child development outcomes. Data show that 56% of the children with teacher education information experienced a teacher with a bachelor's degree (B.A. or higher), and the percentage of teachers without a B.A. (A.A. or lower) is thus 44% (17% with an A.A., 16% some college, 11% with high school diploma or lower). Therefore there is sufficient variation in teachers' education attainment for an estimation of the B.A. effect.

3.5 Measures for key variables

Child Development Outcomes. At the preschool wave, information for children's cognitive and socio-emotional outcomes came from direct assessment by trained assessors during the home visit period, as well as from outcomes reported by parent respondents, preschool teachers (or other types of ECEPs) and therefore were called "indirect assessment" as used in the data provider's language. Cognitive skills are measured exclusively through direct assessment, which included items used in the Early Childhood Longitudinal Study-Kindergarten Cohort's (ECLS-K) kindergarten wave, providing data on early reading (including language and literacy) and mathematics skills. Unique to the preschool wave, the direct assessment of children's knowledge of basic colors was also included. All assessments were developed for ECLS-B, including combinations of items from other established child assessment tools with strong reliability and validity (Najarian, Snow, Lennon, Kinsey, & Mulligan, 2010). Social-emotional development and related outcomes of the child were measured mostly indirectly through reports by parents

[70] This will not, however, invalidate the use of state-level policy variables or characteristics as instrumental variables for the study's focused treatment to the children, i.e., experiencing a teacher with a B.A. or higher.
[71] Source: http://nces.ed.gov/ecls/birthfaq.asp?faq=3

and teachers. The ECLS-B dataset includes social-emotional Likert scale items originated from the Preschool and Kindergarten Behavior Scales–Second Edition (PKBS-2).

Based on the above information, operationally, this dissertation uses eight child development outcome constructs for the children at age 4 and one year later (aged 5), in three domains of child development, as defined in the key term section of Chapter 1. The three domains are cognitive, social emotional and approaches to learning. Among them, children's early reading and math scores are in the cognitive domain, while measures for social competence, externalizing behavior problems and internalizing behavior problems are in the social-emotional development domain. Moreover, attention skills, persistence, independence, and eagerness to learn that were found to be empirically related to learning outcomes (Blair & Razza, 2007; Duncan & Magnuson, 2011), are included as a third category of child developmental outcomes defined in this study, which is often called "approaches to learning" (Kagan, Moore, & Bredekamp, 1998). The social-emotional domain and approaches to learning domain are often treated as non-cognitive in the economics literature for child care and early education (Gupta & Simonsen, 2010). The non-cognitive measures used in this study are all reported by parents; nonetheless, results for teacher-reported measures are available in the Appendices table of this dissertation. The teacher-reported measures are not shown in the main text for two reasons: (1) this is a study about teacher effect, and using teacher-reported child outcomes may increase the amount of selection bias (B.A. teachers may report child outcomes systematically differently than non-B.A. teachers); and (2) the missing rate for the teacher-reported measures are higher; there are fewer items for each construct available from the teachers than from the parents. Studies such as Abner et al. (2013) and Colwell et al. (2013) also used parent-reported social outcome measures.

The composition and source of the eight constructs at age 4 and age 5 are listed in Table 3-1. Age 4 outcomes are constructed based on information from the preschool wave. Age 5 outcomes are derived from the Kindergarten 2006 wave in a similar manner except that the assessment of color knowledge is not included.[72] Each of the eight social outcome measures is constructed either using exploratory factor analysis (EFA)[73] or simple average (for externalizing behavior), as is often used in other ECLS-B studies. For example, parent-reported social competence at age 4 is based on 7 items, such as "accepted by other children", "makes friends easily", "comforts other children". Factor analysis leads to "measurement of underlying or latent factor representing non-cognitive skills or distribution of them" (Almlund et al., 2013, p.49), and has been used in many other ECLS-B studies, such as Sabol and Pianta (2014), Keys et al. (2013), Abner et al. (2013), Lee et al. (2013), and Pearson (2013). Most measures are consistent with previous studies to facilitate cross-study comparisons. The outcome constructs are briefly explained below.

(1) *Early reading*. The early reading score is a summary score of 73 items on basic skills (English language skills, letter knowledge, letter-sound knowledge, print conventions, and world recognition) and vocabulary skills (receptive and expressive language). This was claimed to include "literacy and receptive language items" in the Data File User's Manual for the newest

[72] Also, two nuances are noted here: (1) At the kindergarten year, cognitive assessment was modified to include additional items (for 75% of the sampled children who attended kindergarten in 2006) and it was used again with minimal modification during the kindergarten 2007 wave (for the remaining 25% of children). (2) For the analysis of the teacher reported non-cognitive age 5 outcomes, there is an issue of "reporter difference" for children who attended kindergarten in different years (2006 or 2007). Because of that, the indirect measures are from kindergarten teachers for 75% of the children; yet these measures are from preschool teachers for the remaining 25% of children who were still enrolled in preschools at age 4. Even though the reporters are different, it will not be a big issue because the outcome scale is the same. Moreover, in the model to predict outcomes at age 5, this reporter difference will be accounted for by a dummy variable indicative of whether the child belonged to the 25%, i.e., whether the child attended kindergarten at age 5 (in 2006).

[73] This analysis was implemented through the "factor" command in STATA 12.

release of data; and "the previously released preschool scale scores for literacy and language were dropped from the data file" (p. xvii).

(2) *Story-telling skills or expressive language.*[74] It is an approximate measure indicative of the degree to which the child is able to construct a grammatically correct, coherent story. Two scores for two stories are available in ECLS-B, which originated from the Let's Tell Stories subscale from the Pre Language Assessment Scale (PreLAS 2000) (Duncan & DeAvila, 1998; Data File User's Manual, 2010, p.61). Each score has an ordered scale from 0–5, from low fluency to high fluency. The average of two scores is used as the measure for story-telling skills in this study. Another ECLS-B study, Lee et al. (2013), used the term "expressive language" for this measure.

(3) *Math*. The math score is a summary score of 45 items in three dimensions: number sense, properties, operations; geometry and spatial sense; as well as patterns, algebra and functions.

(4) *Knowledge of basic colors*. This direct assessment measure is unique to the preschool wave and has values from 0 to 10.

(5) *Parent-reported social competence.* This is a summary index obtained by combining a set of seven behavior items reported by parents, including "accepted by other children", "makes friends easily", "comforts other children", "shares", "invites other children to play", "tries to understand others", "volunteers to help others", and "uses words to describe feelings". EFA is used to construct this continuous index.

[74] Some studies obtained separate measures for children's literacy, expressive language and receptive language skills (e.g., Fram & Kim, 2012). Information for constructing such measures was included in the dataset at preschool wave release. However, I am using the full dataset released for the kindergarten 2007 wave, in which this new "early reading" score that compasses all language skills represents the best language outcome; and the separate scores are no longer available. At the same time, a story-telling scale with five values is available, which measures expressive language skills.Thus I have two outcome measures in total for language skills.

(6) *Parent-reported externalizing behavior*. This is a summary index derived by combining a number of seven items indicative of problem behaviors, including "over active", "has temper tantrums", "acts impulsively", "restless". "disrupts others", "annoys other children", and "physically aggressive". Sabol and Pianta (2014) used a similar scale for age 5 outcomes in their ECLS-B study.

(7) *Parent-reported internalizing behavior*. This averages two items with a progressive scale from 1 to 5 that were reported by parents, including worrying about things and being unhappy. This measure was also used in other ECLS-B studies (e.g., Han, Lee, & Waldfogel, 2012).

(8) *Parent-reported approaches to learning (ATL) skills.* As mentioned earlier, this domain is empirically highly related to learning outcomes, and here it was constructed by EFA on the basis of four items reported by parents, including "eagerness to learn", "attention", "independence", and "persistence". Each of the items has a progressive scale from 1 to 5. This construct is likely to be highly correlated with "cognitive self-regulation", an increasingly researched non-cognitive skill in the economics and educational psychology literature (Duncan & Magnuson, 2011, p.7). It is also likely to be closely related to "conscientiousness" and "agreeableness", two of the Big Five personality traits (Komarraju, Karau, Schmeck, & Avdic, 2011).[75] This measure of ATL is similar to what is used in prior studies. For example, ECLS-K studies by Li-Grining et al. (2010) defined ATL for grade K to 5, based on 6 items ("keeps belongings organized", "shows eagerness to learn new things", "works independently", "easily adapts to changes in routine", "persists in completing tasks", and "pays attention well"). Each of the items has a scale from 1 to 4. Also, my measure of ATL is consistent with other ATL studies

[75] In psychology, the Big Five personality traits are five broad dimensions of personality that are used to describe human personality. The five factors are openness, conscientiousness, extraversion, agreeableness, and neuroticism (OCEAN).

(Crosnoe, 2006; Duncan et al., 2007; Galindo & Fuller, 2010). Other ECLS-B studies such as Rispoli et al. (2013) and Iruka et al. (2014) also used the term “ATL”.

Table 3-2 Outcome constructs, index construction, and reliability coefficient

Child development domain	Eight constructs	Compositions of each construct	Important ECLS-B studies that used similar outcome measures at age 4	Reliability: Cronbach's alpha
Cognitive	(1) Early reading	an index provided by the dataset, based on 73 items: basic skills (language, letter, letter-sound, print conventions, and world recognition) + vocabulary skills	"early reading" for kindergarten readiness (Lee et al., 2013); early reading for K (Han et al., 2012), "literacy/receptive language" (Keys et al., 2013), "reading" for age 5 (Sabol & Pianta, 2014), "literacy and receptive language" (Fram & Kim, 2012); "early literacy" (Hawkinson, Griffen, Dong, & Maynard, 2013); "literacy" (Bassok, 2009)	NA
	(2) Story-telling skills/expressive language skills	grammatically correct, coherent story; averaged across 2 scores, each has a scale of 0–5	"expressive language" for K readiness (Lee et al., 2013), "expressive language" for K (Han et al., 2012), "expressive language" (Fram & Kim, 2012)	NA
	(3) Math	an index provided by the dataset, based on 45 items: number sense, properties, operations; geometry and spatial sense; patterns, algebra and functions	"math" for kindergarten readiness (Lee et al., 2013; Han et al., 2013); "math" (Keys et al., 2014; Fram & Kim, 2012), "early math" (Hawkinson et al., 2013), "mathematics" (Bassok, 2009)	NA
	(4) Color recognition	knowledge of basic color, unique to the preschool wave; 0-10	No ECLS-B studies	NA
Social emotional	(5) Social competence: parent reported	a summary index constructed by factor analysis (makes friends easily, comforts other children…, 7 items)	"social competence" (Colwell et al., 2013; Abner et al., 2013; Rispoli et al., 2013), "social skills" (Keys et al., 2013), "prosocial skills" for age 5 (Sabol & Pianta, 2014), "prosocial behavior" for kindergarten readiness (Han et al., 2012; Lee et al., 2013), "sociability" (Pearson, 2013)	0.7863
	(6) Externalizing behavior problem: parent reported	factor analysis (7 items: over active, has temper tantrums…)	Sabol and Pianta (2014) used the term "externalizing behavior" for age 5 outcome; others used "externalizing problem behaviors" (Keys et al., 2013), "emotional and behavioral regulation" (Colwell et al, 2013; Abner et al., 2013), or "conduct problem, hyperactivity" for kindergarten readiness (Lee et al., 2013), "conduct problems for K (Han et al., 2012), "externalizing negativity" (Pearson, 2013), "behavior problems" (Rispoli et al., 2013)	0.7098

	(7) Internalizing behavior problem: parent reported	average across 2 items (worries about things, and unhappy)	"internalizing problems" for kindergarten (Han et al., 2012), "intrinsic negativity" (Pearson, 2013)	NA
Approaches to learning (ATL)	(8) Approaches to learning skills: parent reported	factor analysis, including 4 items ("attention", "eagerness to learn", "independence", "persistence")	"ATL" (Iruka et al., 2014, ECRQ; Rispoli et al., 2013; Ligrining., 2013, ECLS-K), "attention and concentration" (Colwell et al., 2013; Abner et al., 2013), "attention" (Pearson, 2013), "hyperactivity/inattention" for K (Han et al., 2012), "eagerness to learn" (Fram & Kim, 2012)	0.6566

Note. The reliability index reported here is only for those three constructs that were obtained through factor analysis. The other established constructs are the original scale generated by the ECLS-B data provider and have been proven to be of high reliability.

The reliability coefficient for the outcome construct is also presented in Table 3-2. All of the three new constructs generated by factor analysis have values over 0.6. Early reading and math scores are provided by the ECLS-B crew; they are the two most reliable outcome measures given that they are derived from direct assessment, have sufficient variation and are very comprehensive. Further, in the final estimation models, all of the outcomes are standardized to have a mean of 0 and a standard deviation of 1 for ease of interpretation of the estimated coefficients.

Measures for the frequency and quality of teacher–child interactions. High-quality teacher–child interaction fosters cognitive, behavioral, and social-emotional skills, and has both interpersonal and instructional features (Howes et al., 2008; Kontos, Burchinal, Howes, Wisseh, & Galinsky, 2002; La Paro, Pianta, & Stuhlman, 2004). It also has two dimensions: frequency and quality. According to Layzer and Goodson (2006), teacher-caregiver behavior and interactions, including "how they interact with children and how they structure their activities, their emotional tone, and the content of their interactions with children", primarily define children's daily experience in the early care and education setting (p. 562). First, for the full sample of approximately 5000 children in center-based programs for this dissertation study, frequency of interaction or activity measures include a teacher' instructional activities and other type of activities: six in total, including reading books, telling stories, song-singing, playing games with the focal child,[76] and asking questions to the focal child about what was read to them. The frequency of each type of activity is a measure for the frequency of teacher–child interactions, not in a summary form but activity specific. Each type of activity is meant to support the development of a certain type of child outcome or more than one type of outcome (Hamre et al., 2013). Note that this measure of interaction frequency is for each focal child in

[76] "Games" here include board games, card games, and guessing games.

care, because the teacher was asked to report the frequency of interaction with the focal child rather than the whole group of children when she/he was interviewed over by telephone.[77] This is different from the "interaction and activity" component of the ECER-S measure of center quality for the subsample of care and education settings in direct observation; here interaction means for all children in the observed classroom, including the focal child and his/her classmates. Note also that time allocation among different activities, another perspective of a teacher's behavior, is not available in the dataset and will not be studied. Second, in the direct observation subsample of about 1100 children, the quality of the teacher's interactions with the children was assessed using the well-established Arnett Caregiver Interaction Scale (CIS, Arnett, 1989). The CIS is a summary score of 26 items aggregated to four scales: sensitive, harsh, detached, and permissive. Note that this particular teacher–child interaction measure only covers the "emotional support" domain of the CLASS measure; the full list includes "classroom climate", "teacher sensitivity" and "regard for student/ child perspective (CLASS implementation guidance, 2009). Another name for the CIS measure in some studies is closer teacher–child relationships (Howes et al., 2008; Peisner-Feinberg & Burchinal, 1997). The use of the CIS scale is consistent with many prior studies on teacher–child interactions (Arnett, 1989; Burchinal et al., 2002; De Kruif, McWilliam, Ridley, & Wakely, 2000), although recent studies tend to use the above mentioned CLASS measure for teacher–child interactions (e.g., Curby et al., 2009; Howes et al., 2008; Hamre et al., 2013). In theory and also verified in some empirical studies, teachers who interact with students in a warm, sensitive, and responsive style cultivate a sense of security that supports active engagement in the classroom and learning activities, which are helpful for children's

[77] Another potential measure of frequency is at the classroom level (not child level) and thus is not analyzed at this time. Teachers reported the frequency of specific activities, such as reading and language activities and math activities, rated from "never", "once a month or less", "two or three times a month", "once or twice a week", "four times a week", to "everyday", on a scale of 0–5. This can be noted for future studies for classroom level behavior.

language and cognitive development (Howes & Smith, 1995; NICHD, 1998, 2000; Domitrovich et al., 2008).

Measures for teacher education and the covariates. As mentioned in Chapter 1, the key predictor of interest for this dissertation is a B.A. dummy, i.e., whether the child is cared for and educated by a preschool teacher with a bachelor's degree (B.A.). This B.A. can be B.A. in any major, either related to ECE or not. My analysis involves two comparisons: the first between children whose teachers have a B.A. or higher and children whose teachers do not have a B.A, and a second comparison between children whose teachers have a B.A. or higher and children whose teachers hold an Associate's degree. A specialized college education in ECE is indicated by two measures: (1) whether the teacher has a degree in early childhood education or a related major; and (2) the number of college courses in ECE. Covariates in the analysis will be explained in Section 3.6.

Treatment of Missing Data. The missing rate is low for most of the covariates if the variables are viewed separately. In fact, 15 out of the 44 covariates have missing values. The missing rates are mostly under 1%, and some are around 5%–8%. If counting all the missing covariates together, 26% of the observations have missing information in one of the covariates. This overall missing rate is normal, and therefore this dissertation follows the economic standard of employing a dummy flag as the missing data treatment technique. Operationally, for each of the covariates that suffer from missing values, this technique replaces each of the missing values in that particular covariate with an impossible value and creates an indicator for values that are missing (Allison, 2002). Alternatively, results for the complete case are compared to the results based on the dummy flag. This serves as one of the robustness checks. The differences across the two sets of results are trivial, and this means the missing data problem is not severe. Note also,

multiple imputation (MI) is not used here for two reasons: (a) While it is feasible to run OLS models in STATA on a selected number of imputed datasets (often five) based on MI and to report combined estimates of the effects of interest (both the coefficient and standard error), there hasn't been a command available for the estimation of PSM and IV models on the set of imputed datasets. Because of this, it is so far not feasible to combine the standard errors of the estimates from different imputed datasets. (b) For some type of datasets, there is a chance of not achieving convergence in the MI estimation process.

3.6 Methods for Research Question 1: The effect of B.A. on child development outcomes

3.6.1 RQ 1.1: The effect of having a B.A. teacher on child development

For this research question for the B.A. effect, as indicated earlier, the major comparison will be between children whose teachers have a B.A. or higher, and children whose teachers do not have a B.A. degree or lower. Relatedly, A.A. is selected as a second comparison group to check if a bachelor's degree is marginally better than an Associate's degree

The usual model set-up in the literature. In prior studies, the OLS regression equation for estimating the impact of B.A. or other measures of teacher education on child development outcomes generally takes the following form: The child development outcome at a certain time after preschool enrollment is modeled as a product of child, parent, family, preschool and classroom characteristics, with teacher education being the predictor of interest.[78]

[78] Covariates for the B.A. dummy or other measures of teacher education are usually displayed in blocks.

$$Child\ development\ outcomes_i$$
$$= \beta_0 + \beta_1 BA_i(or\ other\ measusres\ of\ teacher\ education) + \boldsymbol{\gamma Child}$$
$$+ \boldsymbol{\delta Maternal} + \boldsymbol{\theta FAM} + \boldsymbol{\tau T_{others}} + \boldsymbol{\rho Classroom} + \boldsymbol{\varphi Center_{others}}$$
$$+ state\ or\ site\ dummies + \varepsilon_i$$

Identification challenges. As discussed in Chapter 2, such estimations may be biased due to several identification challenges. First, omitting prior outcomes that are usually highly predictive of current outcomes would lead to a biased estimate for the B.A. effect, as teachers with different degree qualifications may be assigned or selected to children with different initial outcomes, either because of family selection or teacher/preschool's decision. Or B.A. teachers may treat children with different starting developmental statuses differently than non-B.A. teachers (Pianta, 1999; O'Connor & McCartney, 2007).[79] The direction of bias could go either way. Most studies have not controlled for children's prior outcomes, in fact only five studies on child development outcomes did so: four B.A. studies (Early et al., 2006; Early et al., 2007; Henry et al., 2004; Howes et al., 2008) and one study that used a continuous measure of teacher education instead of a B.A. dummy measure (NICHD & Duncan, 2003). However, the control of prior outcomes is very important given the continuity of development. Omitting this variable would jeopardize our ability to determine whether child development outcomes are predicted by the quality of earlier care and family experience or the quality of the current preschool experience.

Second, selection bias emerges because of the sorting effect of prevalent parental choices in preschool arrangements given the disparity of center and teacher quality and their link to various family background factors (Phillips, Mekos, Scarr, McCartney, & Abbott–Shim, 2000; NICHD & Duncan, 2003). The sorting mechanism often leads to the selection of children into different

[79] For example, in O'Connor and McCartney (2007), teachers were found to attend less often to children with whom they had higher quality relationships.

preschool programs and/or classrooms with differing teacher quality.[80] Especially for child follow-up data, it is likely that we may be comparing a child in one center with a teacher with a bachelor's degree with another child at another center with a teacher who is not a bachelor's degree holder. Some family selection factors (e.g., income, motivation, parental expectations, cultural beliefs about child rearing) for preschool choice are not necessarily measured in the survey. These selection factors are usually associated with teacher's degree level. If omitting these variables that are also predictive of child development outcomes, the estimation of the B.A. effect will be biased. One way to deal with this problem is to control, in the OLS model, factors that affect the selection of preschools, for example, family and community SES, which this study will do given that we have information for family income and mother's education in our dataset. Still, some unobserved factors such as parental expectations and cultural beliefs can be the selecting factors that enable children to experience a B.A. teacher in preschool while at the same time directly predicting outcomes. This important source of bias can be alleviated by the quasi-experimental methods used in this dissertation study. For instance, one could introduce some exogenous variation for B.A. attainment by using the instrumental variable (IV) approach; this method will be explained in more detail.

Third, teacher's self-selection into B.A. Among the teacher candidates, those who chose to pursue a bachelor's degree or higher are more willing to be a teacher, for example, more likely to see caring and educating young children as a calling or as a passionate job (Kagan, Kaurez, & Tarrant, 2008; Goldhaber, 2007[81]). This motivation may also push them to provide better care and education to the children. In such cases, the effect of the B.A. may not solely be a

[80] Here non-random classroom assignment is not severe in our data because most of the sampled children attended different preschool centers so they will not be in different classrooms at the same center. However, the selection of children into different preschool centers with different teacher qualifications should be taken into account.
[81] This paper discussed methodological issues for estimating the effect of an M.A. in K-12 education.

consequence of knowledge and practice learned from the B.A. program; rather, it may only reflect the effect of the enthusiasm of the teacher toward the profession. Those who are enrolled in a B.A. program are also more likely to be able to pay for it, and the family SES of the teachers may in turn be related to their reservation wage and teaching/caring quality. It is also difficult to partial out the influence of other attributes of preschool teachers who earn a bachelor's degree, such as training, credentials, etc. (Fuller, 2011; NICHD & Duncan, 2003). These together may contaminate the independence of the B.A. effect on child development outcomes.

Fourth, there could be interdependence between B.A. and other features of preschools. The chance of getting a teacher with a B.A. may confound with other center quality indicators, such as group size, child-staff ratio, center director's leadership, and supportive environment. This is often called the problem of "all good things co-occur in child care" (Phillips, 1987, p.5).[82] The "iron triangle" of structural variables (child-adult ratio, group size, and teacher education/training) is the often-cited example of the "good things". If these center quality indicators are missing in the model, the OLS estimates may overstate the impact of a B.A. teacher. So the best way to deal with this problem is to hold other exogenous teacher characteristics and other elements of program structure constant. For example, child–adult ratio, group size, materials in the classroom, space of the classroom, and center director's characteristics, and program length (full-day or part-day) could be added in the model. Another example is to control for the peer context because peers' ability were found to matter for children's development (Henry & Rickman, 2007) and a study that used the ECLS-B found a much larger peer effect than teacher effect (Fram & Kim, 2012).[83]

[82] It is also possible that more highly educated teachers choose to work in higher quality settings (Hamre & Bridges, 2004).

[83] It is also possible that the coefficient of peer quality may have captured other features, such as a center's overall quality or some unobserved family characteristics through parental choice, or community characteristics.

OLS with rich controls. Following the experience from prior empirical studies, an OLS model controlling for prior outcomes and a rich variety of factors, OLS with rich controls, will be used first, where the key variable will be a B.A. dummy for the child's preschool teacher. HLM is not suitable here because this dataset belongs to the "child follow-up" data: Similar to the NICHD data, the sampled children in the ECLS-B dataset attended different preschools at age 4, so that the data is not clustered in preschools or preschool classrooms. The "OLS with controls" model has strength over a simple OLS models in that it controls more potentially confounding factors than in prior studies and as such can be viewed as a rigorous method (Magnuson, Ruhm, & Waldfogel, 2007; Han et al., 2012; Lee et al., 2013; Miguel, Gertler, & Levine, 2005). In the OLS model with rich controls, the B.A. dummy takes the value of 1 if the child has a teacher with a B.A. or higher, and 0 if otherwise (i.e., some college, A.A. and lower). The full model is as follows:

$$Child\ development\ outcomes_i = \beta_0 + \beta_1 BA_i + \beta_2 prior\ outcomes\ at\ age\ 2_i + \boldsymbol{\gamma C} + \boldsymbol{\delta F} + \boldsymbol{\theta HOME} + \boldsymbol{\tau T_{others}} + \boldsymbol{\rho Classroom_{others}} + \boldsymbol{\varphi Preschool_{others}} + \boldsymbol{H - P\ connection} + \boldsymbol{Region\ dummies} + \varepsilon_i$$

The dependent variable is child development outcomes (eight outcome constructs introduced earlier). The term *BA* indicates the B.A. dummy (1 if the teacher has a bachelor's degree or higher). *Prior outcomes at age 2* are controlled for. The measured prior outcomes at age 2 include: (1) the Bayley Scale for mental ability, a construct measuring problem solving and language acquisition skills (33 items from the Bayley Short Form-Research Edition, or BSF-R), to serve as the control for the model of cognitive outcomes at age 4 and age 5; and (2) social skills constructed by EFA based on items from the hotspot scores of Toddler Attachment Sort

(TAS)-45 that measure children's attachment to their parent, to serve as a control for the model of social-emotional outcomes and other outcomes in the third domain.[84] Important child characteristics are represented by *C* (age at time of assessment measured in months, gender, race, low birth weight or not, whether attended center-based care at age 2, hours per week in the current preschool, and for age 5 outcomes, a dummy indicating whether the child attended kindergarten in 2006). *F* stands for family characteristics, especially the family selection factors related to the children's chance of experiencing a B.A. teacher, such as mother's education, degree of maternal depression, family income (averaged across the first three waves of data) (Magnuson & Votruba-Drzal, 2009; Lee et al., 2013), and marital status. *HOME* indicates home environment, which contains three dimensions: (a) learning materials measured by number of children's books at home; (b) home language is not English; and (c) a summary index of the quality of parenting constructed by PCA (Principal Component Analysis) based on measures of parenting behavior, including number of times per week the parent spanks the child, time out, number of times the parent reads to the child, plays with the child and expresses affection to the child. T_{others} represents other teacher characteristics: gender, racial match to the child, experience, CDA (Child Development Associate), other certificate, in-service training, history with this focal child, and intensity of care and education (hours in care), the last of which was controlled for given that NICHD & Duncan (2003) found hours of care an important control.[85] Other classroom level characteristics include child–adult ratio (which accounts for the number of assistant teachers in the classroom), age range in the classroom (if larger than two or not), and materials in

[84] These prior outcome measures used at children's age 2 by the ECLS-B data collectors are not the same as the instruments/metrics used for age 4 and 5, perhaps mainly because of the concern that toddlers need differential assessments to preschoolers and kindergarteners (see the "Choosing score" PDF document provided by ECLS-B data website). Retrieved from: http://nces.ed.gov/ecls/pdf/birth/ChoosingScores.pdf

[85] Note that teacher salary is not controlled for here because it does not make sense to hold salary constant while estimating the effect of the B.A. in that teacher education does not vary much if salary is held constant (Barnett, 2011b). Nonetheless, I will test for robustness by including it in the model and compare the new results to the old from the model that excludes the teacher salary variable.

the classroom (number of books plus number of interest areas). *Preschool* indicates preschool attributes and non-teacher center quality factors: a) preschool type, including Head Start, state pre-K, partially publicly funded private preschools, and exclusively private preschools (without any public funding); b) auspices of the center (for profit vs. "non-profit"), in which the public schools are deemed to be non-profit; c) the full-day status of the preschool, constructed by using the variable for children's "per-week hours in preschool"; and d) whether the center has been licensed by any national, state or local organizations. *H-P connection* is a variable indicating the frequency of home–preschool connection, i.e., teacher-reported parental involvement in preschools (Arnold, Zeljo, Doctoroff, & Ortiz, 2008). *Regional dummies* are Northeast, Midwest, and West, using South as the reference group.

Apart from the above covariates of B.A., "teacher's earnings" is a relevant variable that is very informative for understanding the things associated with a bachelor's degree. As expected, B.A. teachers and non-B.A. teachers differ a lot in their annual earnings: $23 259 vs. $8 635. Although there is a need to study the "more distal measures such as wages and turnover" that have been implicated in a few studies for their predictability of child outcomes in the large studies (Bryant, Zaslow, & Burchinal, 2010, p. 49), one should not control for earnings in the model to estimate the B.A. effect because there will not be any variation in the B.A. dummy if holding earnings constant (Barnett, 2011). Therefore I will not add earnings as a covariate of B.A. in the main analysis.[86] Nonetheless, I will add the earnings variable in the estimation model as a strategy of robustness checks in Section 5.4 of Chapter 5.

Specific measures of the predictors are presented in Table 3-3 below.

[86] In this regard, one of the mechanisms for any B.A. effect may include better teacher practices induced by better compensation, to a certain degree (the other two are human capital enrichment and signaling, to be discussed further in Chapter s5, 6 and 7).

Table 3-3 Definition and measures of the key treatment variable and its covariates

Predictors		Measures
1. Variables of interest		
B.A.		1 if the preschool teacher has a B.A. or above; 0 otherwise
Specialized education in ECE for a college degree		2 measures: (1) having a degree in ECE for a degree (B.A. or A.A.); (2) number of college courses in ECE
2. Covariates of B.A.		
prior outcome at age 2		Barley mental score; social competence (TAS-45 HOTSPOT) (both are standardized)
Child characteristics	age at assessment (measured in month)	continuous: measured in months (Lee, Zhai, Brooks-Gunn, Han, & Waldfogel, 2013)
	gender: male	1 if male
	race/ethnicity	African American, Hispanic, Asian/Pacific Islander, Native American, Multi-racial, as compared with non-Hispanic White
	low birth weight (LBW)	1 if low birth weight
	hours per week in preschool	continuous (Lee et al., 2013)
	attended center-based programs at age 2	1 if yes
	attend kindergarten in 2006 (for age 5 outcomes only)	1 if attended kindergarten in 2006, as used in Votruba-Drzal et al. (2013)
Family characteristics	mother's age	continuous
	mother's education	B.A. or above (Keys et al., 2013)
	maternal depression	constructed using the Center for Epidemiological Studies–Depression Scale (CES-D; Radloff, 1977), averaging the total scores of 12 items that ranged from 0 to 36: Higher scores indicates higher levels of depression symptoms
	married to biological father	1 if married to biological father
	family income	continuous or categorical
	number of siblings	count
Home environment	number of children's books	count
	language spoken at home: non-English	1 if non-English
	quality of parenting	This is an index constructed by PCA, based on number of times per week the parent spanks the child, time out, number of times the parent reads to the child, plays with the child and expresses affection to the child. It is standardized.

Other teacher attributes	gender	1 if the teacher is male
	racial match to the child	the same race or not
	experience	measured in years and month; not accounting for raising own children
	in-service training in the last 12 months	1 if having experienced training (by definition ECE training) in the last 12 months (and have more than 1 years' experience)
	CDA	1 if having a CDA
	Other certificate	1 if having other certificates (e.g., state awarded certificate)
	history with this focal child	month
	intensity of care and education for the focal child	continuous: hours per week
Other classroom characteristics	group (class) size	continuous
	child–adult ratio	number of children divided by the number of adults; here adult includes teacher assistants and volunteers
	materials	• number of books; • number of interest areas: reading area, listening center, areas for playing with puzzles and blocks (Lego), art area, etc.
	peer's racial composition (for supplemental analysis)	Fram and Kim (2012), on racial composition
Preschool characteristics	preschool type	following Lee et al. (2013), 3 types of preschools are categorized: Head Start, state pre-K, other preschools Other preschools will serve as the reference group)
	licensed	1 if the center is licensed
	auspices	For profit vs. non-profit
	full-day status	1 if full-day (not currently having a good measure, but to be constructed from the child level indicating "hours per week in preschool")
Home-Preschool connection	teacher-reported parental involvement	ordered scale 1 to 4: From almost never to always
Region dummies	Northeast, Midwest, South and West	South serves as the reference region

Two quasi-experimental methods. Propensity score matching (PSM) and Instrumental Variable (IV) are the two quasi-experimental tools for this study.

(1) Propensity score matching

Propensity score matching is a quasi-experimental method that intentionally constructs the treatment group and control group from the observed sample, by "matching the groups determined most likely to be equal" (Goldhaber, 2007, p.11). The treatment effect can be obtained by comparing means across the matched pairs (Angrist & Pischke, 2009; Rosenbaum & Rubin, 1983). It is a more nonparametric way to control for confounding covariates and model the selection directly,[87] although it requires the conditional independence assumption (CIA), meaning that the treatment status should not be correlated to the potential outcomes in any unobserved way. This method has been used to estimate the effect of teacher qualifications (e.g., master's degree) in K-12 education (Goldhaber, 2007). For RQ1, the application of PSM can balance the possible confounders of the B.A. for children experiencing a B.A. teacher vs. children without experiencing a B.A. teacher.

The usual procedure for implementing PSM involves four iterative steps. The first is to select potential confounders and to use a probit or logit model to estimate a propensity score for each individual, namely, the probability of being in the treatment group (i.e., experiencing a B.A. teacher for this dissertation study), conditional on covariates. Such a propensity score is a linear combination of covariates for a single score; it balances treatment and control groups on a large number of covariates without losing a large number of observations. Its value should be bounded between zero and one (Hirano & Imbens, 2001). A common probit model specification is presented below:

[87] It relaxes the parametric assumption implied in the OLS model. Also, it produces more robust estimates given the similarity between the pairs (Conniffe, Gash, & O'Connell, 2000; Rubin & Thomas, 2000).

$$p_i = P(\mathrm{D}_i = 1|\mathrm{X}_i) = \alpha + \boldsymbol{\omega} X_i + u_i$$

where p_i is the propensity score of being treated, D_i is a dummy indicating the treatment status, X_i are confounding covariates, which include the prior outcomes of the child at 2 years old, age and gender of the child, maternal education, maternal depression, family income, number of books in the classroom, parenting, preschool type and its license status, and preschool director's experience, etc. A list of the confounders is presented in Section 5.2 of Chapter 5.

Second, during the estimation of the propensity score, following DuGoff et al.'s (2014) procedure, a sample weight for each of the child observations in the operational sample is added as an additional predictor, so that the model is adjusted to achieve national representativeness.[88]

The third step is matching: Each child with a B.A. teacher is matched with another child who has a similar propensity/likelihood of being treated (i.e., to have a B.A. teacher) but is not (i.e., does not have a B.A. teacher). Several matching algorithms can be chosen, for example, a nearest neighbor matching method matches a treated child with an untreated child with the closest propensity score. As the default option, one-to-one nearest neighbor matching with replacement[89] is used as the major algorithm for the matching ("psmatch2" in STATA). Alternative matching techniques such as within calipers (radius) matching, and Kernel matching (Epanechnikov as the default option), are also available in the "psmatch2" command in STATA, and are used as robustness checks. Specifically, one-to-one nearest neighbor matching with replacement matches a treated child with an untreated child that has the closest propensity score and an untreated child can be used multiple times as a control. Radius/caliper matching performs

[88] In the future, when techniques advance, sample weight could also be applied in the outcome model: the frequency of the matched child being used as a match (currently accounted for) multiplied by the sample weight.

[89] Matching "with replacement" is the default option in the STATA command "psmatch2". Allowing for replacement has the merit of reducing bias (Stuart, 2010), but standard error may need adjustment, which is already involved in "psmatch2". Another reason for selecting "with replacement" is that my sample contains more treated children than non-treated children. I also note that, it is possible that when matching with replacement, the treatment effect estimate will be based on just a small number of controls, and the number of times each control is matched should be monitored (Stuart, 2010).

radius matching within the specified radius given by the caliper (usually 0.1), to avoid the risk of bad matches. Kernel matching relies on the weighted average of all cases in the untreated group to estimate the counterfactual outcomes, for which the weight is the distance in propensity scores. The closest control is given the greatest weight (Heckman, Ichimura, & Todd, 1998). For the matching process, it is important to check whether there is enough "overlap" of the treated and untreated group (the "common support" requirement). Another test is to see if the matched groups are actually balanced through mean comparisons of important covariates. A treated child differs from the untreated child matched to him/her only in terms of his/her teacher's degree status (having a B.A. or not), otherwise they should, on average, be similar in observed characteristics (e.g., prior outcomes, gender, mother's education, family SES, home-learning environment, teacher's experience). This can be done by t-test, however a better way is to compare the pre-matching balance to the post-matching balance for each of the covariates (standardized mean difference between the matched sample and the non-matched sample), available from the "psbal2" command developed by causal method expert and statistician Jenifer Hill for supplemental use with the "psmath2" command in STATA (Gelman & Hill, 2006). If the matching quality is not satisfied, one could go back to the first step to modify the specifications of the propensity model by including higher-order or interaction terms of the current-used covariates or adding additional covariates in an iterative process, until both acceptable quality of balance and acceptable common support are reached.

The final step is to estimate the effect. After finishing the matching process, the average treatment effect on the treated (ATT) as the default estimand will be obtained by comparing the mean outcomes across matched groups. This is without additional covariates (i.e., beyond the ones in the probit model). In addition, to get the average treatment effect (ATE) of matching

analysis that is comparable to the OLS results of the full sample, inverse probability of treatment weighting (IPTW) that was first proposed by Rosenbaum (1987), will be conducted and reported.[90] In IPTW, the weight is the inverted propensity score of being in the group: p for the treated children and 1-p for the untreated children (here p is the propensity of having a B.A. teacher, which is generated in the propensity model). This operation resembles the use of survey sampling weights that are employed to weight survey samples so that they are representative of specific populations (Morgan & Todd, 2008).

(2) Instrumental Variable

The instrumental variable approach (IV) is a widely used quasi-experimental method for estimating causality when experimental data is not available. Though still new, some applications are found in the early childhood education literature (Magnusson, Ruhm, & Waldfogel, 2007; Crosby, Dowsett, Gennetian, & Huston, 2010). An instrument is obtained and used to disentangle the exogenous variation of an endogenous treatment variable, and the exogenous portion of the variation will be used to estimate the effect to be taken as causal. To be a valid instrument, it has to be correlated with the endogenous explanatory variables, conditional on the other covariates; but it cannot be correlated with the error term in the explanatory equation, that is, the instrument cannot predict outcomes directly. The only way for the instrument to affect outcomes is through the treatment variable. Under these assumptions, IV estimates may be accepted as causal estimates.

For this research question on the B.A. effect, the major instrumental variable is the percentage of B.A. holders in the ECE workforce in the state where the children live, in 2005. This IV is a supply factor that affects a teacher's degree status and therefore the chance of a child

[90] This is a necessity because "psmatch2" doesn't usually report the standard errors for ATE and ATU estimates, which is the case for my data. I also tried "attnd", but encountered the same problem.

experiencing a B.A. teacher. Such a factor is above the individual teacher and child level and is not likely to directly predict individual child outcomes.

A common mode for IV analysis is the two-stage least squares (2SLS) analysis. Applied to my research question, the set-up for a 2SLS is:

1st stage: $BA_i = \delta_0 + \delta_1 state\ level\ percentage\ of\ B.A.\ teachers + \boldsymbol{\eta X_i} + \varepsilon_i$

2nd stage:

$$Child\ development\ outcomes_i = \beta_0 + \beta_1 \widehat{BA}_i + \beta_2 prior\ outcomes\ at\ age\ 2_i + \boldsymbol{\gamma C} + \boldsymbol{\delta F} + \boldsymbol{\theta HOME} + \boldsymbol{\tau T_{others}} + \boldsymbol{\rho Classroom_{others}} + \boldsymbol{\varphi Preschool_{others}} + \boldsymbol{H - P\ connection} + \varepsilon_i$$

In the first stage, B.A. will be regressed on the instrument variable (i.e., the state-level percentage of B.A. teachers), controlling for a set of covariates, which will be the same as the covariates in the second stage (Angrist & Piscke, 2009). Then in the second stage, the predicted probability of experiencing a B.A. teacher will be replace the B.A. dummy in the OLS model and the coefficient of beta 1 will be the IV estimator, and the region dummies will not be used to allow for more state-level variation of the IV. The 2SLS can be estimated using the "ivreg" command in STATA.

Data for the IV, the percentage of B.A. teachers in each state, is obtained from the American Community Survey (ACS) data (2005) through the special tabulation service provided by the U.S. Census Bureau. I chose the 2005 data because preschool teachers for the children in our sample were interviewed in 2005. The ACS data has sufficient observations to produce state-level average estimates, as is implied by the description of the National Research Council (2012) and my comparison of it to Bassok et al. (2013) and the relative sample size of ACS versus CPS.

Also, it is intuitive to expect that the average educational attainment value for the larger workforce will affect the chances of a child having a B.A. teacher in preschool to a large enough degree.[91] This is confirmed by the *F* test value of the 1st stage correlation. A preview of results for this test in Chapter 5 shows that the *F* values range from 30 to 40 for this test for the eight child outcomes, higher than the critical value of 10 (Staiger & Stock, 1997). A Hausman test for exogeneity will also be conducted (Hausman, 1978).

To summarize, results for the B.A. effect will be compared between the three methods (OLS controlling for prior outcomes, PSM and IV). For all of them, weights for correcting the oversampling issue as mentioned earlier will be applied to produce national-level estimations for the children born in 2001.

3.6.2 RQ1.2: Does the effect of B.A. vary by preschool type (Head Start, State Pre-K, partially public-funded private preschool, and exclusively private preschool)?

In the ECLS-B dataset, children in Head Start centers can be identified by two variables from the parent interview: One variable indicates whether the child was in a Head Start program; the other variable indicates whether the Head Start program's location was in a public school, a private school, a college or university, a community center, or its own building. Defining state Pre-K is not easy for the ECLS-B dataset. Another ECLS-B study, Lee et al. (2013), defined children in state pre-K as those who attended a "prekindergarten",[92] and my dissertation study follows this definition. Preliminary analysis indicates that 21% of the center-attending children were enrolled in this definition of state pre-K. This is consistent with the figure of 20% in 2005

[91] Note that for the ACS data, we cannot distinguish between early care and education workers who work with infants and toddlers and those who work with preschoolers, or between preschool teachers and kindergarten teachers. However, this test removes the potential detrimental effects of this limitation on the validity of the IV.

[92] This is from a question in the parent interview (CC432), where a parent selects the program from the following options: (1) a day care center; (2) a nursery school; (3) a preschool; (4) a pre-kindergarten; (5) something else.

reported by NIEER's state preschool yearbook (Barnett et al., 2006). "Exclusively private preschool" programs are the private settings (non-public school prekindergarten)[93] that either did not receive any type of public sponsorship, such as Head Start, public school/board of education, state or local government[94] or claimed to be a for profit center. Partially public-funded private preschool programs can be defined as the remaining center-based preschool programs that cannot be identified as a Head Start center, a state pre-K, or an exclusively private preschool.[95]

For RQ1.2, the heterogeneous effects by preschool type are estimated by adding an interaction term between the B.A. dummy and the other three center-type dummies in the model, taking children in Head Start centers as a reference group. A separate subsample analysis is not conducted because the sample size shrinks a lot in the subsample.

3.6.3 RQ1.3: Is the impact of B.A. larger for children from low socioeconomic status (SES) backgrounds?

In the ECLS-B dataset, a socioeconomic status (SES) index of social standing that reflects the socioeconomic status of the household at the time of the preschool parent interview in 2005 is readily available. The components for this SES index include: father's/male guardian's education; mother's/female guardian's education; father's/male guardian's occupation; mother's/female guardian's occupation, and household income. Based on the value of the SES variable, I will divide the children into three groups: (1) those with the lowest 20 percent in the

[93] This means "private school pre-K, child care center, or preschool/nursery school" according to question CI002 in the ECEP interview in the ECLS-B dataset.

[94] This organization sponsorship question (CI023) has the following options: (1)Head Start; (2) social service organization or agency; (3) church or religious group; (4) public school/board of education; (5) private school, religious; (6) private school, non-religious; (7) college or university; (8) private company or individual; (9) non-government community organization; (10) state or local government.

[95] Alternatively, the "partially public-funded private preschools" are in the above-mentioned private settings but with somewhat public sources of sponsorship. Some private preschools may also receive children with subsidies (CI045b) and they will also be taken as "publicly funded".

value of SES are called the low-SES children; (2) the middle 60 percent; (3) the highest 20 percent. This definition follows Chernoff, Flanagan, McPhee, and Park's (2007) report.

To obtain the separate B.A. effect for low-SES children, I will add an interaction term between B.A. and a variable indicative of the child's low-SES status into the model for RQ1.1.

RQ1.4: Does having a degree in early childhood education (ECE) or a related field (or the number of college-level ECE courses) increase the effect of having a bachelor's degree?

Information about specialized education in ECE could be obtained from a question asked to the teachers ("Do you have any degree in ECE or a related field?"). This information is not sufficient to define B.A. in ECE, because a degree in ECE can be either a B.A. in ECE, or an A.A. in ECE. Another measure for specialized education in ECE is the number of college courses taken in ECE, including relevant classes taken to earn a degree or CDA qualification. These courses are not necessarily B.A. courses either. Although these two measures are not necessarily exclusively associated with the B.A., they are still helpful for our understanding of the interaction between the degree and ECE content training in formal education to see if a bachelor's degree has both a screening effect and a content effect (Early et al., 2006). As they stated:

> *"If teacher' education is important because it provides teachers with specific insights into child development and pedagogy, the content of the teacher's education, and not simply its length or level of education are associated with classroom quality and child outcomes" (p. 176).*

Two methods are employed to answer this question. First, an interaction term of the B.A. dummy with one of the two measures for specialized education in ECE is added into the models

for RQ1.1. This will be the main analysis for RQ1.4. Second, consistent with Early et al. (2006) and Early et al. (2007), the incremental effect of specialized education in ECE (based on the previously mentioned two measures) is estimated among the children whose preschool teachers are B.A. holders.

3.7 Methods for Research Question 2: Estimating B.A.'s effect in two steps

RQ2: What is the role of teacher–child interactions in mediating the effect of B.A. on child development outcomes?

A two-step OLS analysis will be conducted to see if the B.A. effect is mediated through the quality of teacher–child interactions, and to what extent.

Stage 1:

$$T-C\ interaction = \beta_0 + \beta_1 BA_i + \beta_2 prior\ child\ outcomes_i + \boldsymbol{\gamma C} + \boldsymbol{\delta PF} + \boldsymbol{\tau T_{others}} + \boldsymbol{\rho Classroom} + \boldsymbol{\varphi Center_{others}} + \varepsilon_i$$

where "T–C interaction" is a measure of the quality of teacher–child interaction.

Stage 2:

$$Child\ development\ outcomes = \beta_0 + \beta_1 T-C\ interaction + \beta_2 y_i + \boldsymbol{\gamma C} + \boldsymbol{\delta P} + \boldsymbol{\theta F} + \boldsymbol{\rho Classroom} + \boldsymbol{\varphi Center_{others}} + \varepsilon_i$$

In the first step regression, a similar equation to the one for RQ1 will be used by replacing the outcome variable with a measure of the quality of teacher-child interactions. Controls in the model predicting teacher-child interactions are the same as in the equation for child development outcomes. As robust checks, some family characteristics are removed to increase degree of freedom, for example, home learning environment, which may not affect the quality of teacher-

child interactions directly. In the second stage, child development outcomes will be regressed upon measures of teacher-child interactions, controlling for the same set of covariates such as classroom resources and center level characteristics. The two-step analyses will reveal whether or not teacher-child interaction is an important intermediate outcome linking B.A. to child development outcome.

3.8 Summary of the research design

As a concluding section of this chapter, a summary of the research design for RQ1 and RQ2 of this dissertation is provided in Table 3-4.

Table 3-4 Summary for the research design for RQ1 and RQ2

<table>
<tr><th colspan="3">Research Questions</th><th>Methods (all models will control for prior outcomes)</th></tr>
<tr><td rowspan="6">RQ1: The B.A. effect</td><td colspan="2" rowspan="2">RQ1.1: The estimation of the effect of B.A. using quasi-experimental methods</td><td>(1) OLS with rich controls</td></tr>
<tr><td>(2) PSM</td></tr>
<tr><td rowspan="3">Heterogeneous effect</td><td rowspan="2">RQ 1.2: Varying B.A. effect by type (Head Start; state pre-K; private)</td><td>(3) IV</td></tr>
<tr><td rowspan="2">interaction terms of preschool type and B.A.</td></tr>
<tr><td>RQ 1.3: Effect larger for low-SES children?</td><td>interaction terms of low SES and B.A.</td></tr>
<tr><td colspan="2" rowspan="2">RQ1.4: Does a major in ECE and the number of college courses in ECE interplay with the B.A. effect?</td><td>(1) Add interaction terms of B.A. and major into the models for RQ1.1</td></tr>
<tr><td>(2) Subsample analysis (Early et al., 2007)</td></tr>
<tr><td colspan="3" rowspan="2">RQ2: Relationships among B.A., teacher–child interactions and child development outcomes</td><td>(1) Step 1 OLS linking B.A. to T–C interactions: Same as above</td></tr>
<tr><td>(2) Step 2 OLS linking T–C interactions to child development outcomes: OLS</td></tr>
</table>

Chapter 4

Descriptive Statistics

This chapter depicts the important features of the child-teacher sample, highlighting the basic patterns and distributions of the child development outcomes and the teacher behavior measures for this study (i.e., the frequency and quality of teacher-child interactions). Descriptive statistics are also presented separately for the children with B.A. teachers and the children with non-B.A. teachers, so that the similarities and differences between the treated group and the untreated group are explicit. All the descriptive statistics are weighted by sample weight to produce national representative estimates for the children in center-based programs.

4.1 Features of the sample: Child, family, teacher, classroom and preschool characteristics

Table 4-1 shows the basic features of the center-attending children, with information on their family, their lead teachers, and their preschools. I name this sample the child-teacher sample.

Consider first the child characteristics. Some 51.5% of the children are boys. On average, a child spends about four days a week in preschools. Approximately twenty-two percent of them attended center-based programs at the age of two. About 33% of the children's mothers have a B.A. or above. The average family income level for this sample of children is eight, meaning that the average annual income (averaged across waves 1 to 3) for this sample of children is between 35001 and 40000 dollars. On average, there are about 66 books in a child's home at age 2 (wave 2) and age 4 (wave 3).

As expected, only a small proportion of the teachers are male, namely, around 1%. About 82% of teachers received some form of ECE training in the past year, in terms of "courses,

workshops or seminars" defined in the ECLS-B survey instrument. As regards teachers' educational levels, 56% of the teachers have a B.A. or above; and 31% hold a Child Development Associate (CDA) qualification.

In terms of preschool characteristics, about 22% of this sample of center-attending children attended Head Start centers, 15% state preKs, 25% partially publicly funded private preschools; and 38% in exclusively private preschools. Among the designated preschools for the children, 76% of them have been licensed.

Table 4-1 Descriptive statistics for the analytical sample: B.A. and the covariates

Variable		N (rounded)	Mean	Std. Dev.	Min	Max
Key predictor	Teacher has a B.A. or above ("the B.A.")	4300	0.559	0.497	0	1
Prior outcome at age 2	mental ability (Bayley)	4000	0.162	0.969	-2.984	4.21
	social skills (TAS-HOT SPOT)	4200	-0.062	0.967	-1.731	3.22
Child characteristics	age at assessment measured in month	4300	52.667	3.921	44	64.8
	boy	4300	0.515	0.500	0	1
	White-non Hispanic (reference group)	4300	0.581	0.493	0	1
	African American	4300	0.143	0.350	0	1
	Hispanic	4300	0.207	0.405	0	1
	Asian	4300	0.024	0.154	0	1
	Native American	4300	0.005	0.073	0	1
	multi-race	4300	0.036	0.185	0	1
	low birth weight	4300	0.012	0.110	0	1
	days per week in preschool	4300	4.321	1.018	1	6
	having attended a center before (at age 2)	4300	0.215	0.411	0	1
Family characteristics	mother's age (years)	4300	28.055	6.249	15	50
	mother's education is BA or plus	4300	0.331	0.471	0	1
	mother's degree of depression	4000	1.358	0.647	1	4
	mother was married at child birth	4300	0.705	0.456	0	1
	family income	4300	7.972	3.205	1	13
	number of siblings	4300	1.341	1.053	0	8

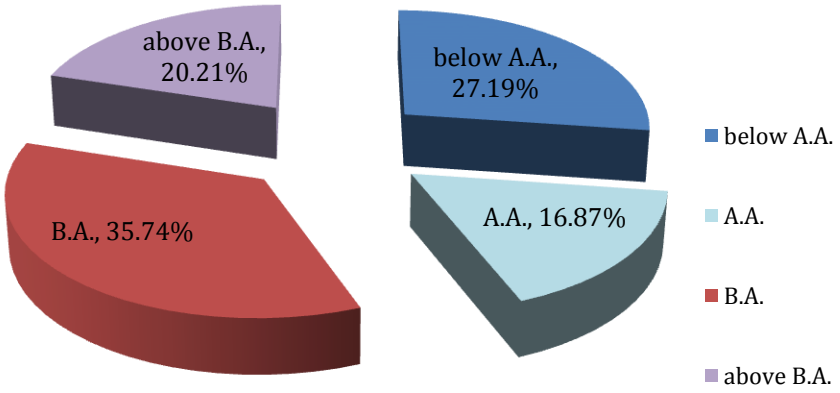
Home environment	number of books at home (wave 2 and 3)	4300	66.249	64.456	0	550
	home language is not English	4300	0.158	0.365	0	1
	the quality of parenting	4000	0.020	0.876	-25	1.465
Other teacher attributes	teacher's gender is male	4300	0.012	0.111	0	1
	teacher's race is different from the child	4300	0.312	0.463	0	1
	experience	4300	13.057	8.736	0	70
	in-service ECE training	4300	0.815	0.388	0	1
	CDA	4300	0.309	0.462	0	1
	other certificate	4300	0.313	0.464	0	1
	history with the child	4300	7.225	7.051	1	56
	hours in care (per week)	4300	21.229	12.977	1	87
Other classroom characteristics	group size	4300	14.196	5.439	1	50
	child-adult ratio	4250	6.676	3.466	0.125	40
	number of books in the classroom	4300	114.62	172.338	0	1000
	number of interest areas in the classroom	4300	8.710	1.689	0	10
	age range larger than 2	4200	0.297	0.457	0	1
Preschool characteristics	Head Start center	4300	0.215	0.411	0	1
	state preK	4300	0.150	0.357	0	1
	partially public funded private programs	4300	0.249	0.432	0	1
	exclusively private programs (reference group)	4300	0.386	0.487	0	1
	for profit	4300	0.246	0.431	0	1
	preschool was licensed	4300	0.761	0.426	0	1
	full day	4300	0.317	0.466	0	1
	home-preschool connection	4300	1.920	1.225	0	4
Region dummies	Northeast	4300	0.189	0.391	0	1
	Midwest	4300	0.208	0.406	0	1
	West	4300	0.210	0.407	0	1
	South (reference group)	4300	0.393	0.489	0	1

Note. All the descriptive statistics in this table were weighted.

4.2 Teacher education, specialized college education in ECE, in-service training, CDA, experience and other potential confounders

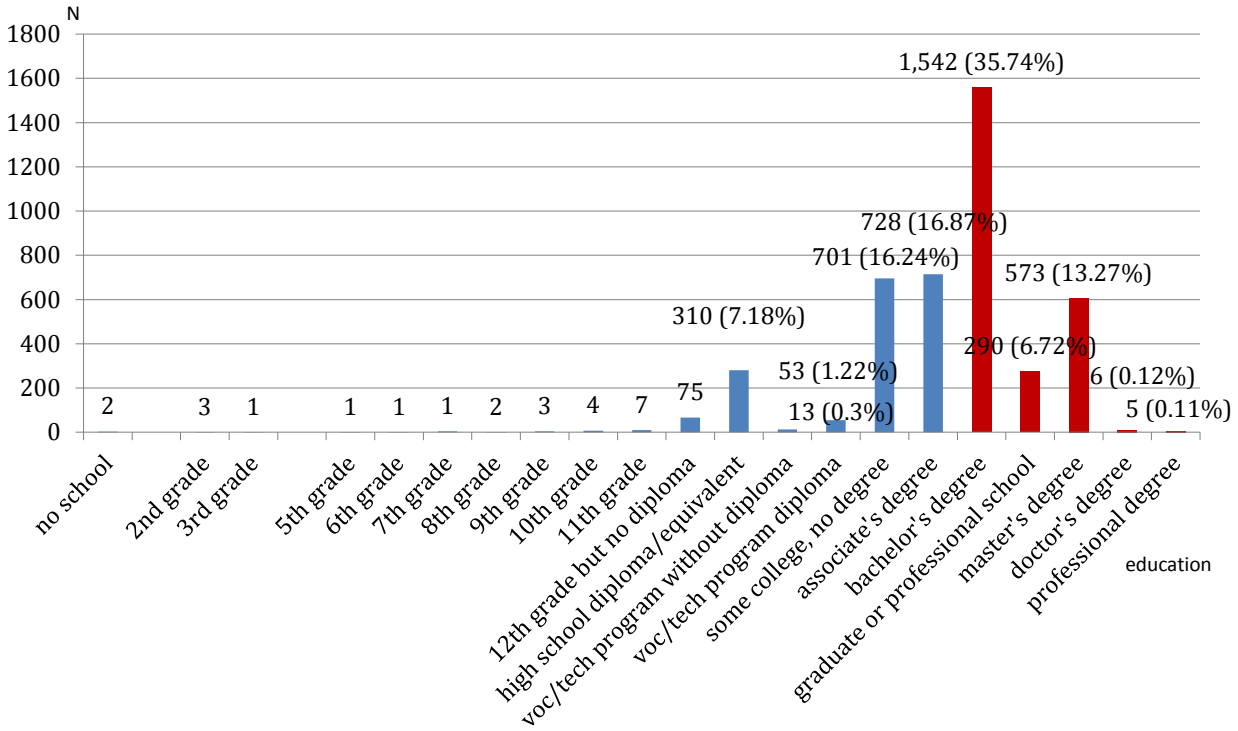
Teacher Education. As shown in Figure 4-1, weighted by sample weight, 55.96% of the children have a B.A. teacher (56.90% if not weighted). In the language of an experiment study, these are the children treated by the "B.A. teacher" treatment. The rest of the children (44.04%) either have an A.A. teacher or one with an education level lower than A.A. (such as some college, high school diploma, or even less). These are the children in the control group. A detailed distribution graph of teachers' level of education is presented in Figure 4-2.

Figure 4-1 Teacher education for the child-teacher sample



Note. In this figure, the dark red and purple segments denote the treated children, whereas the blue and light blue segments represent the untreated children.

Figure 4-2 The detailed distribution of teacher education for the child-teacher sample



Note. The total size of the child-teacher sample is about 4300 (rounded to the nearest 50). The overall percentage of children with B.A. teachers (including B.A. or above) is the sum of the displayed ratios of four degree levels (35.74%, 6.72%, 13.27%, 0.12%, and 0.11%), namely, 55.96%. Bars highlighted in dark red stand for treated children, namely, children whose teachers have a B.A. or an even higher level of degree.

Specialized college education in ECE and B.A. In relation to RQ 1.4, I also present the statistics for "specialized college education in ECE" and its correlation with B.A. As mentioned earlier in Chapter 3, this study relies on two measures for "specialized college education in ECE": (1) college major in ECE or a related field, and (2) the number of ECE courses in college. On average, 57.13% of the children's teachers reported having a degree majored in ECE or a related field. Broken down by degree levels, as shown in Figure 4-3 and the first column of Table 4-2, more than 75.50% of the B.A. teachers (including those above B.A.) have a degree majored in ECE, yet this college major may be associated with a prior Associate's degree (A.A.). Although information in this dataset prevents me from directly knowing exactly how many of the B.A. teachers with an ECE major obtained the specialized education in their pursuit of an A.A. instead of in the pursuit of a B.A., the proportion is likely to be small, as it can be roughly inferred that about 10% of college students get an A.A. before having a B.A. and the rest get the B.A. without getting an A.A. (Lichtenberger & Dietrich, 2013). [96] Also, among those teachers with some college or above, 63% reported having an ECE major for a degree, and the percentage is higher among those with B.A. or above (85.76%); and 80% of those with ECE majors have currently achieved a degree level of at least a B.A. [97] Thus the B.A. indicator pretty much captures the ECE degrees. A similar conclusion also applies to another measure for specialized college education in ECE, i.e., the number of college courses in ECE, as shown in the second column of Table 4-2. In fact, those with no college degrees also took some college courses in

[96] For example, according to this study, for the Illinois high school graduating class of 2003, some of them enroll in four-year colleges with the aspiration of B.A (N=21522). Some start in 2-year colleges with an aspiration to transfer to get a B.A. (N=2154), although a portion of them choose not to obtain the A.A. certificate during the process. Thus 10% is the highest possible percentage of teachers who hold an A.A. before earning a B.A. Also, as a note for future survey questionnaires should note that one can ask the teachers questions like "what is your major for the highest degree", as done so in the NICHD and other 6 datasets used in Early et al. (2007). In so doing, as regards those whose highest degree level is B.A, we can be clearer about whether the ECE major is associated with the current B.A. or with the A.A. obtained earlier.

[97] It makes sense to see some teachers who reported an ECE major for an education level of "some college"'; still, they compose a very small proportion of the sample, i.e., 5%.

ECE. Another interesting finding is that A.A. teachers in this sample on average took more ECE courses than B.A. teachers.

Figure 4-3 B.A. and college major in ECE

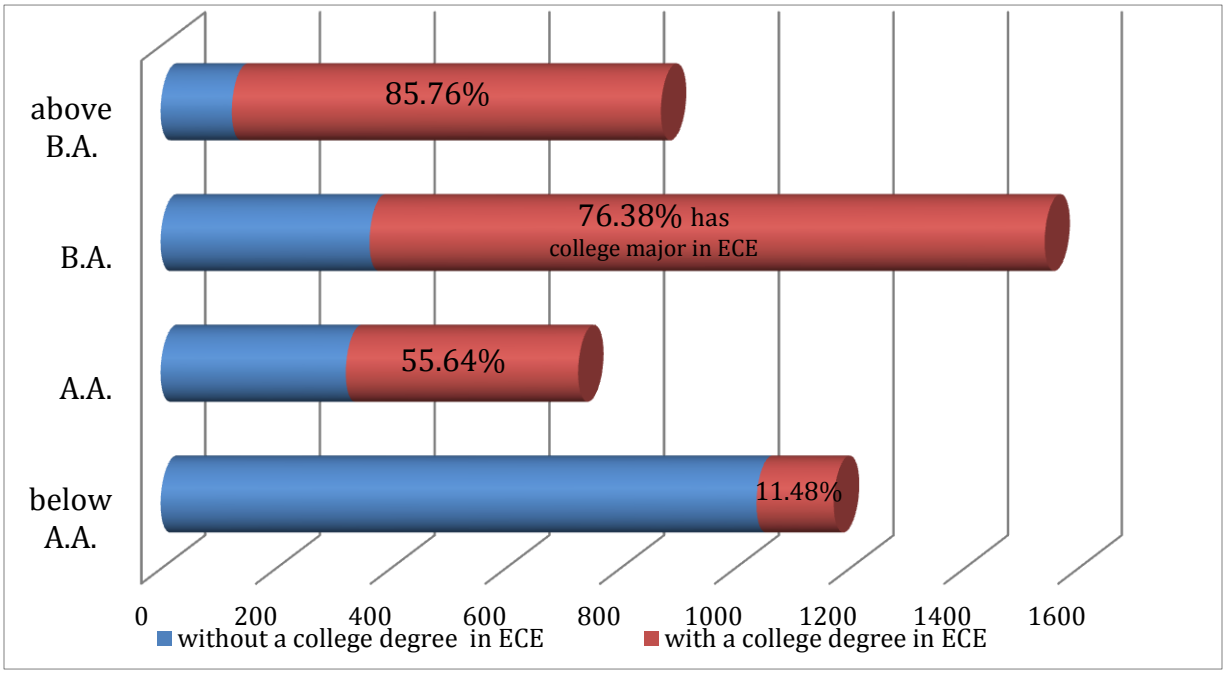


Table 4-2 Specialized college education in ECE, in-service training and CDA by education level

Education level of teachers		Specialized college education in ECE		Had in-service training in ECE (in the past 12 months)	CDA	Other certificate
		ECE major	# of college courses in ECE			
Below A.A.		11.48%	3.124	79.36%	34.96%	16.88%
	high school or below	2.10% [(a)]	0.251	78.03%	38.52%	10.68%
	some college	17.80%	5.061	80.25%	29.68%	21.05%
A.A.		55.64%	3.283	83.83%	56.48%	19.95%
B.A.		76.38%	1.554	79.47%	23.09%	34.52%
Above B.A.		85.76%	1.233	86.01%	17.70%	54.55%
Total		57.13%	2.208	81.50%	30.86%	31.31%

Note. (a) Some teachers who claimed to have an education level of high school or below also claimed to have a degree in ECE, which may be attributed to someone's treating CDA as a degree, a reporting error, or a particular scenario. I have retained the original value of the ECE major dummy variable in RQ 1.4 as a baseline. Alternatively, these observations are to be treated as not having a college major in ECE, as a sensitivity check. (b) All the statistics are weighted by sample weight.

In-service training and Child Development Associate (CDA). As presented in Table 4-2, 81.50% of the children's teachers have participated in early childhood training in the past 12 months (mostly likely in-service training). The ratio is quite consistent across different degree levels (79.36%, 83.83%, 79.47%, and 86.01%), indicating that the variation of in-service training is smaller, and that the correlation between in-service training and level of education is small. The latter implication reduces the potential concern about multicolinearity between these two observable attributes of teachers. As for the national credential for early childhood, CDA, those without a B.A. or higher tend to have a higher chance of holding a CDA. This makes sense given that CDA is a widely recognized credential and it often serves as an alternative teacher qualification standard in some states for state preK programs (e.g., in Connecticut, see Barnett, Carolan, Fitzgerald & Squires, 2012).

Experience. Regarding the relation between teacher's degree status and teacher's experience, I found that teachers with fewer than 10 years of experience have a very similar likelihood of having a B.A. as compared to teachers with 20 to 30 years of experience and those with more than 30 years of experience (55.99% vs. 54.27% vs. 58.86%). This pattern eases concern about the possibility that young teachers (and often in-experienced ones) are more likely to have a B.A. than older (often experienced) teachers. If the majority of young teachers have already obtained a B.A., it calls into question the use of the B.A. as a threshold for young teachers. This is not the case for the sample of this dissertation.

The iron triangle. Further, regarding the possibility of an "iron triangle" of "good things" going together (Phillips, 1987), often referred to the correlation among teacher education, in-service training, and group size (or child-adult ratio), I tested the Pearson correlation coefficients among them. Almost all the correlations are below 0.1, suggesting that the co-placement of in-

service training, group size and child-adult ratio in the model will not affect the estimate of the B.A. coefficient. An exception is the correlation between group size and child-adult ratio (0.571, also shown in Table 4-4), which is as expected. Yet, this correlation is not very high, and does not necessarily cause a multicolinearity problem for the estimation of the B.A. effect. This is also confirmed by the VIF (Variance Inflation Factor) test of multicolinearity after the OLS regression, to be introduced in the next chapter.

Balance check for the covariates by B.A. status. For better understanding of the full list of the observed potential confounders of B.A. (including family selection of B.A. teachers, correlation of B.A. with other teacher factors, etc.), the means and standard deviations of the covariates by B.A. status are presented in Table 4-3. The two subsamples have similar characteristics in terms of children's prior mental ability and social skills at age 2, percentage of boys, racial composition, maternal depression and marital status, percentage of children whose home languages are not English, teacher gender, racial match, experience and in-service training status, group size, child-adult ratio, and number of interest areas in the classroom. However, some observed factors exhibit positive selections, such as the age measured in months (children with B.A. teachers are on average 18 days older than children with non-B.A. teachers), mother's age, mother's B.A. status[98] (children with B.A. teachers are more likely to be children who have B.A. mothers than their counterparts), family income (children with B.A. teachers have higher family income than children with non-B.A. teachers), number of books at home, and the quality of parenting. There are also negative selection factors, such as CDA, history with the child (for children with B.A. teachers, their teachers had fewer history with them), and hours of care.

[98] Another interesting pattern related to this factor is that, for the operational sample, the percentage of teachers with a B.A. or above is higher than the percentage of B.A. mothers. This may be related to the ECE policy context in the U.S., where there are many federal/state programs targeting low-income children and therefore it increases their chance of getting a teacher with a high-level of degree.

Children with B.A. teachers are also less likely to be from Head Start centers, for-profit preschools, and full-day preschools. In order to derive the independent effect of B.A., these positive and negative selection factors should be controlled for in the OLS and IV models, and be matched well in the PSM model. This dissertation did so.

Table 4-3 Balance check for the covariates: By teachers' B.A. status

Covariates	Children with B.A. teachers		Children with non-B.A. teachers		Mean difference (t test)
	Mean	Std. Dev.	Mean	Std. Dev.	
mental ability at age 2 (Bayley)	0.140	1.015	0.190	0.907	-0.050
social skills at age 2 (TAS-45 HOTSPOT)	-0.058	0.971	-0.067	0.962	0.009
age at assessment measured in month	52.939	3.849	52.321	3.985	0.618***
Boy	0.530	0.499	0.497	0.500	0.032
African American	0.140	0.347	0.146	0.354	-0.007
Hispanic	0.203	0.402	0.211	0.408	-0.008
Asian	0.027	0.163	0.021	0.142	0.007**
Native American	0.005	0.068	0.006	0.080	-0.002
Multiple race	0.035	0.185	0.036	0.187	-0.001
low birth weight	0.013	0.114	0.011	0.105	0.002
days in preschool	4.289	1.024	4.362	1.008	-0.073*
attended center based programs before	0.191	0.393	0.245	0.430	-0.055***
mother's age	28.456	6.285	27.546	6.168	0.910***
mother has B.A. or plus	0.377	0.485	0.272	0.445	0.105***
mother's degree of depression	1.351	0.641	1.368	0.656	-0.017
mother is married	0.717	0.451	0.690	0.462	0.027
family income	8.129	3.212	7.773	3.185	0.356***
number of siblings	1.393	1.066	1.275	1.033	0.118***
number of books at home (average for wave 2 and wave 3)	69.592	66.617	62.004	61.359	7.589***
home language is not English	0.165	0.371	0.150	0.357	0.016
quality of parenting	0.095	0.805	-0.075	0.950	0.170***
teacher is male	0.016	0.125	0.008	0.089	0.008*
teacher has another race	0.316	0.465	0.307	0.461	0.009
experience	13.155	8.891	12.934	8.536	0.221
in-service training in ECE	0.818	0.386	0.811	0.392	0.008
CDA	0.211	0.408	0.432	0.495	-0.221***

other certificates	0.418	0.493	0.181	0.385	0.237***
history with the child(month)	6.804	6.298	7.759	7.875	-0.954***
hours of care for the child	20.371	12.594	22.318	13.372	-1.947***
group size	14.336	5.083	14.019	5.857	0.317
child-adult ratio	6.757	3.485	6.573	3.441	0.185
# of books in the classroom	129.874	191.249	95.237	142.500	34.637***
# of interest areas in the classroom	8.761	1.602	8.646	1.793	0.114
age range of the classroom larger than 2 years or more	0.279	0.449	0.321	0.467	-0.042**
Head Start center	0.181	0.385	0.259	0.438	-0.078***
state preK	0.233	0.423	0.044	0.206	0.189***
partially public	0.258	0.438	0.237	0.426	0.021
for profit	0.182	0.386	0.328	0.470	-0.146***
preschool licensed	0.708	0.455	0.829	0.377	-0.120***
full day	0.271	0.445	0.376	0.484	-0.105***
home-preschool connection	2.029	1.173	1.783	1.275	0.246***
Northest	0.243	0.429	0.120	0.325	0.122***
Midwest	0.226	0.418	0.185	0.389	0.041**
West	0.162	0.369	0.270	0.444	-0.108***

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.
*p<0.1, ** p<0.05, *** p<0.01

Finally, I present a correlation matrix for all the predictors in Table B1 of the Appendix B. This includes B.A. and all of the covariates. In Table 4-4 below, I only present the pairs of variables that have a correlation coefficient of 0.2 or larger, because the number of covariates is too large (44). As one could see, only 11 out of the 990 pairs of variables exhibit a correlation over 0.4; and only 53 pairs' correlation coefficients fall into the range of 0.2 to 0.4. All other pairs have a correlation under 0.2. These statistics imply that the multicolinearity issue is not severe in my estimation models, be it the OLS model or the IV model.

Table 4-4 Degree of correlations among the predictors

Range of the Pearson Correlation Coefficient	N of pairs	Pairs and the values
>0.4 (11 pairs)	1	mother is married & mother's age (0.455)
	1	mother's age & mother has B.A. or above (0.409)
	5	family income & mother's age (0.518); family income & mother has B.A. or above (0.546); family income & mother is married (0.538); family income & number of books at home (0.409); family income and Head Start center (-0.479)
	1	home language not English & Hispanic (0.592)
	1	group size & child-adult ratio (0.571)
	2	full day & days in preschool (0.443); full day and hours of care per week (0.515)
[0.2, 0.4] (53 pairs)	3	B.A. & CDA (-0242); B.A. & other certificates (0.244); B.A. & state preK (0.263)
	5	mental ability at age 2 & social skills at age 2 (0.311); mental ability at age 2 & mother has B.A. or above (0.224), mental ability at age 2 & family income (0.288); mental ability at age 2 & number of books at home (0.209); mental ability at age 2 & home language not English (-0.217)
	6	African American & Hispanic (-0.215); African American & days in preschool (0.237); African American & mother is married (-0. 362); African American & family income (0.347); African American & quality of parenting (-0.255); African American & hours of care (0.256)
	5	Hispanic & mother has B.A. or plus (-0.215); Hispanic & family income (-0.238); Hispanic & number of books at home (-0.239); Hispanic & teacher has another race (0.366); Hispanic & Head Start center (0.230)
	1	Asian & home language not English (0.241)
	1	multiple race & teacher has another race (0.289)
	3	days in preschool & family income (-0.293); days in preschool & number of books at home (-0.283); days in preschool & quality of parenting (-0.201);
	1	prior attendance of center &hours of care (0.201)
	3	mother's age & number of books at home (0.271); mother's age & quality of parenting (0.271); mother's age & Head Start center (0.265)
	4	mother has B.A. or above & mother is married (0.368); mother has B.A. or above & number of books at home (0.311); mother has B.A. or above & quality of parenting (0.275); mother has B.A. or above & Head Start center (-0.321)
	3	mother is married & number of books at home (0.288); mother is married & quality of parenting (0.217); mother is married & Head Start center (-0.259)
	1	family income & quality of parenting (0.325)
	3	number of books at home & home language not English (-0.267); number of books at home & quality of parenting (0.235); number of books at home & Head Start (-0.246)

	1	home language not English & teacher has another race (0.230)
	1	in-service ECE training & number of interest areas (0.242)
	1	CDA & Head Start (0. 265)
	2	number of interest areas & Head Start center (0.224); number of interest area & licensed (0.210)
	3	Head Start center & state preK (-0.217); Head Start center & partially publicly funded private preschool (-0. 293); Head Start center & for profit (-0.218)
	2	state preK & partially publicly funded private preschool (-0. 240); state preK & for profit (-0.242)
	1	partially publicly funded private preschool & for profit (-0.326)
	2	Northeast & Midwest (-0. 242); Northeast & West (-0.236)
	1	Midwest & West (-0.259)

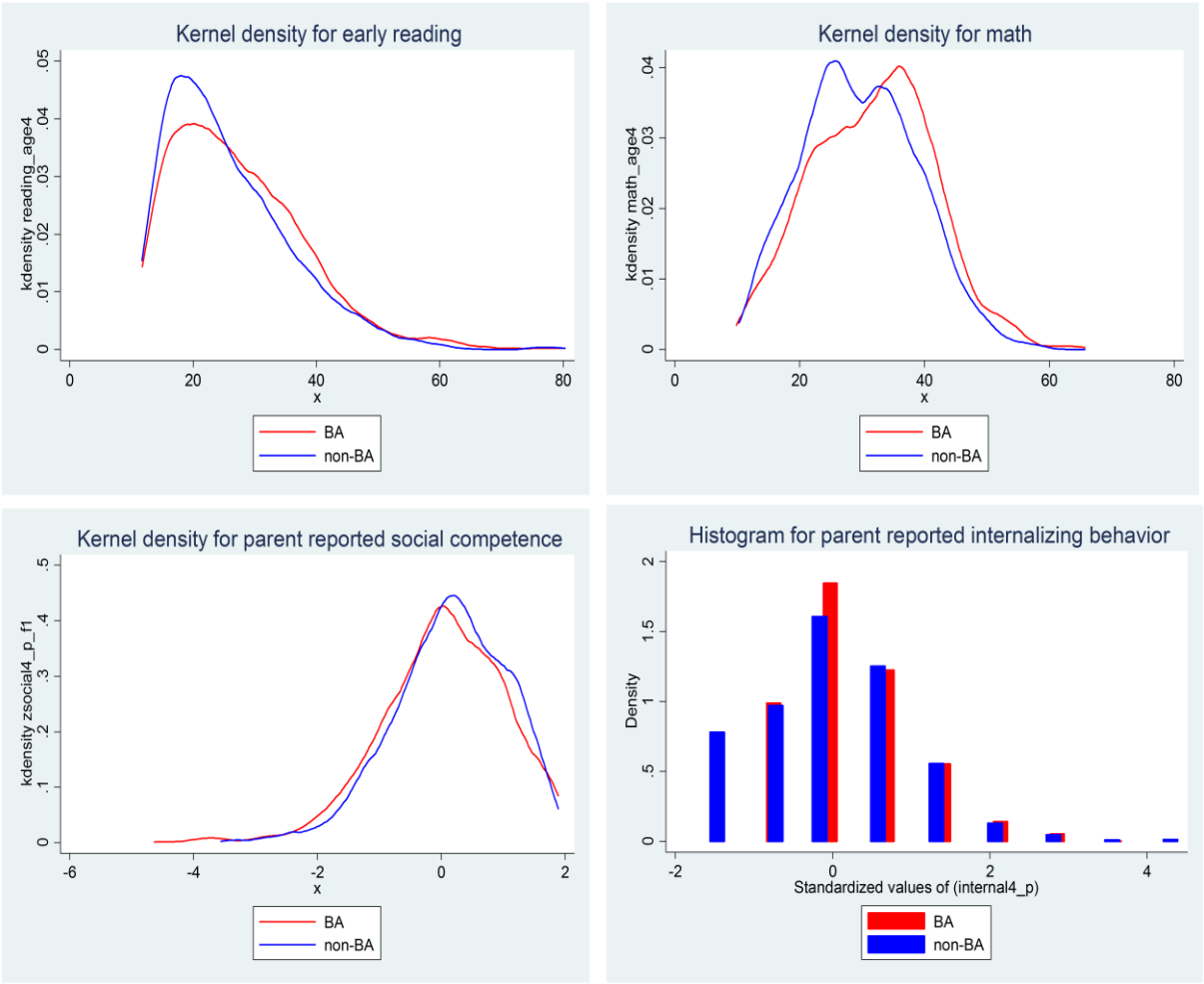
4.3 Child development outcomes and teachers' B.A. status

I present the distribution of each of the eight outcome constructs in Figure 4-4. Most of the constructs exhibit a certain type of normal distribution. For the full sample of children, the average score for early reading skills is 26.69 (as shown in Table 4-5), on a spectrum of 11.71 to 80.29, and the distribution is left skewed and left censored (the first graph in Figure 4-5). Somewhat similarly, the average math score for the whole sample of children is 30.84, on a spectrum from 9.87 to 65.74. In the second graph for math in Figure 4-5, one can observe a difference in the average math scores between children with B.A. teachers and children without B.A. teachers. This raw difference is still descriptive in that it did not account for any covariates, and a final conclusion needs to be drawn after application of the rigorous methods in Chapter 5.

Note also that the distribution of the parent-reported social competence score has a normal type and is somewhat right censored, as shown in Figure 4-4 (the third graph). A particularity of measure lies in the parent-reported internalizing behavior problem construct. It is the average of two ordered scales (one to five) and therefore this variable is also ordered with nine values. Accordingly, an ordered probit model is used as a check for robustness in Chapter 5 for the

parent-reported internalizing behavior outcome. Additionally, I include descriptive features of four teacher-reported non-cognitive outcomes in Table B2 of Appendix B.

Figure 4-4 Distribution of four of the eight outcome constructs at age 4 by B.A. status



Note. Only the last graph (histogram) was not weighted because weighting is not allowed for the command in STATA. However, the graph should be similar if weighted.

Table 4-5 Means and standard deviations of outcomes: Full sample and by B.A. teacher status

Outcomes	Full sample		B.A. teachers		Non-B.A. teachers		Mean differences (t test)
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	
Age 4 outcomes							
(1) Early reading	26.695	10.435	27.443	10.708	25.750	10.004	1.692***
(2) Story telling	2.489	0.977	2.485	0.983	2.493	0.970	-0.008
(3) Math	30.842	9.492	31.754	9.691	29.702	9.112	2.051***
(4) Color recognition	8.988	2.041	8.987	2.059	8.990	2.018	-0.003
(5) Social competence: parent reported	0.094	0.962	0.041	0.994	0.161	0.915	-0.120***
(6) Externalizing behavior problems: parent reported	-0.041	0.945	-0.055	0.954	-0.025	0.935	-0.030
(7) Internalizing behavior problems: parent reported	-0.020	0.993	-0.029	0.966	-0.009	1.026	-0.020
(8) Approaches to learning(ATL) skills: parent reported	0.060	0.966	0.043	0.993	0.081	0.930	-0.038
Age 5 outcomes							
(1) Early reading	41.091	14.557	41.909	14.739	40.084	14.272	1.825***
(2) Story telling	3.422	0.821	3.404	0.860	3.445	0.769	-0.041
(3) Math	30.842	9.492	31.754	9.691	29.702	9.112	2.051**
(4) Social competence: parent reported	0.057	0.977	0.031	0.998	0.089	0.950	-0.058
(5) Externalizing behavior problems: parent reported	0.044	0.978	0.004	0.972	0.091	0.985	-0.087
(6) Internalizing behavior problems (worry): parent reported	-0.042	0.968	-0.042	0.973	-0.043	0.963	0.001
(7) Approaches to learning(ATL): parent reported	0.007	0.997	-0.005	0.983	0.020	1.014	-0.025

Note. (a) All descriptive statistics in this table are weighted by sample weight. (b) 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. (c) The child development outcomes and measures for teacher-child interactions are presented in its original scale score, but during the estimation the standardized counterparts were used. (d) Color recognition is not available for age 5. *p<0.1, ** p<0.05, *** p<0.01

Table 4-5 also presents the raw differences of child outcomes at age 4 and age 5 for two groups of children: children with B.A. teachers and children with non-B.A. teachers. As for age 4, one can see raw differences in most of the outcomes: positive for early reading and math, and negative for parent reported social competence.

Also worth noting is that, correlations among child development outcomes at age 4 are mostly small, as presented in Table 4-6. This is one of the reasons for not using simultaneous equation modeling (SEM) for the estimation of the B.A. effect, in addition to the fact that I have already controlled for initial cognitive and social emotional ability at age 2 in the estimation model of each of the outcome construct.

Table 4-6 Correlation matrix for the outcomes at age 4

Outcomes	early reading	story-telling	Math	Color	social_p	external_p	internal_p	ATL_p
early reading	1.000							
story telling	0.290	1.000						
math	0.771	0.302	1.000					
color	0.373	0.224	0.451	1.000				
social_p	0.194	0.180	0.185	0.163	1.000			
external_p	-0.201	-0.074	-0.238	-0.153	-0.285	1.000		
internal_p	0.063	0.062	0.081	0.022	-0.104	0.252	1.000	
ATL_p	0.277	0.140	0.277	0.182	0.550	-0.473	-0.084	1.000

Note. Social_p, external_p, internal_p and ATL_p stand for parent-reported social competence, externalizing behavior problems, internalizing behavior problems and approaches to learning (ATL) skills, respectively.

4.4 Teacher-child interactions and teachers' B.A. status

What is the status of teacher-child interactions---frequency of activities and the quality of emotional support---for American preschoolers in 2005, as represented by this national sample? As regards the frequency measures of teacher-child interactions, for the full sample of children, teachers read to the focal child about seven times per week. Also, on average, the number is 8.5

for song-singing activities and 4.6 for games, and 3.7 for building something (likely to be related to math skills). These statistics are shown in the first panel of Table 4-7. These numbers are consistent with studies that found "children spent much of their time in language/literacy, social studies, and art, and less time in math and gross motor activities" for state funded preKs (Early et al., 2010, p.177). As regards the quality of teacher-child interactions (i.e., emotional support as measured by CIS) that is available for a subsample, the quality of teacher-child interactions ranges from 4 to 78, with a mean of 66. The lower right graph of Figure 4-5 also shows that the overall index of teacher-child interactions is right skewed, which confirms the findings in Colwell et al. (2013). So is the pattern for each of the component score of CIS: sensitivity, less harshness, less detachedness and less permissiveness. This picture of high quality emotional support is consistent with findings in some state preK studies (e.g., Pianta et al., 2005; Burchinal, Vandergrift, Pianta, & Mashburn, 2010).[99] However, compared with child care centers in the NICHD study, the center-attending children (born in 2001) in ECLS-B experienced higher interaction quality in 2005 than children (born in 1991) in the NICHD sample, where "observed positive caregiving" was "very uncharacteristic" for 6% of the children, "somewhat uncharacteristic" for 51%, "somewhat characteristic" for 32%, and only "highly characteristic" for 12% (NICHD, 2000, p.116). [100]

[99] This has to be noted with low level of instructional support in state preK classrooms, featured by whole-group directed activities, much time spent in management and routine activities, rote and lack conceptual focus, varied vocabulary, and interactive feedback (Pianta et al., 2005; LoCasale-Crouch et al., 2007). The level of instruction was also found to be low in another study regarding publicly funded preschool programs (Justice, Mashburn, Hamre, & Pianta, 2008) and in Head Start settings (Dickinson, Darrow, & Tinubu, 2008). For this perspective in the ECLS-B dataset, I am not going to elaborate here given the dissertation focus of teacher behavior is the frequency of activities and the emotional support measure of teacher-child interactions.

[100] Although NICHD data used the 4-point ratings of caregivers' sensitivity/responsivity, stimulation of cognitive development, intrusiveness (reflected), and detachment (reflected) to define positive caregiving (a subcomponent of the Observational Record of the Caregiving Environment, or ORCE), their findings on the teacher sensitivity status is still comparable to the CIS measure here for ECLS-B.

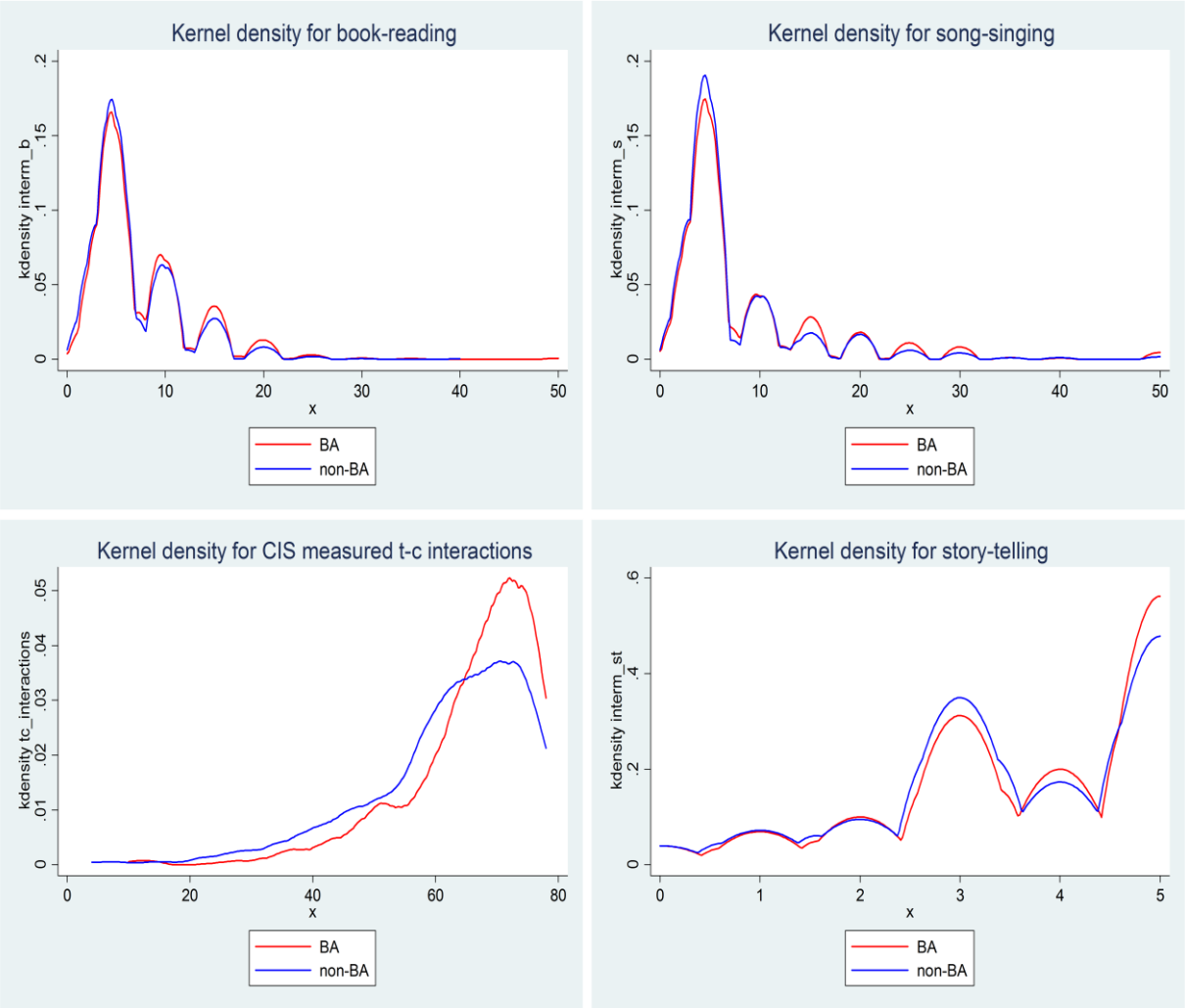
Note also that there are some differences across preschool types. For instance, on average, children in Head Start and state prekindergartens experience higher frequencies of all types of activities than those in private preschools (whether partially public or not), as indicated in Table 4-7. Comparatively, the partially publicly funded preschools show the highest CIS measured teacher-child interaction quality. Still, the quality of teacher-child interactions is high across all preschool types.

Table 4-7 Children's experience of teacher-child interactions: Total and by preschool type

Teacher-child interactions/Mean		All types	Head Start	state preK	partially publicly funded private preschools	exclusively private
Full sample: frequency of activities	book-reading (times per week read books to child)	7.459	8.117	8.656	6.296	6.516
	song-singing	8.450	9.535	10.026	6.931	6.889
	playing games or do puzzle with the child	4.587	5.221	4.924	4.240	4.256
	building something with the child	3.679	4.287	3.818	3.595	3.701
	tell stories to the class (1-5)[(a)]	3.753	3.885	3.724	3.634	3.636
	ask question about what is read to the children (1-5)	3.273	3.473	3.305	3.220	3.221
Subsample (around 1000), and restricted to the same teacher and same child in focus	CIS measure for the quality of interactions: total	66.064	64.855	64.304	65.630	62.785
	sensitivity score	22.989	22.088	22.042	22.842	21.289
	harshness score: less harsh	24.258	24.089	23.973	24.181	23.494
	detachment score: less detached	11.209	11.100	10.893	10.999	10.667
	permissive score: less permissive	7.608	7.578	7.396	7.609	7.335

Note. (a) 1-5 standards for "how often": once a month or less--> everyday; (b) The reported statistic are the means of the variables.

Figure 4-5 Distribution of the frequency and quality measures of teacher-child interactions



How do teacher-child interactions experienced by the children differ according to their teachers' B.A. status? As shown in Table 4-8, children with B.A. teachers experienced higher frequency of book-reading activities, song-singing activities and story-telling activities. [101] Moreover, for values before standardization, the CIS measure of the teacher-child interaction quality is higher for children with B.A. teachers, not only in the total score, but also in terms of

[101] In addition, another perspective of teacher behavior, which is related to teacher-child interactions, is teacher's time allocation among different activities (whole-child, group, individual) and the assessment tool for children's performance. Information regarding assessment tools is also available in the ECLS-B dataset, although not included in the second research question due to the author's choice of research scope.

subscores. In other words, descriptively, children with B.A. teachers are more sensitive, less harsh, less detached and less permissive. These patterns are confirmed in Figure 4-5.

Table 4-8 Children's experience of teacher-child interactions by teachers' B.A. status

Teacher-child interactions/Mean	Total		B.A. teachers		Non-B.A. teachers		Mean difference (t test)[(b)]
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	
Book reading (times per week the teacher reads books to child)	7.126	4.951	7.459	5.265	6.704	4.488	0.755***
Tell stories to the class (1-5)[(a)]	3.702	1.388	3.753	1.384	3.638	1.391	0.115**
Ask question about what is read to the children (the stories: 1-5)	3.288	0.783	3.273	0.786	3.306	0.779	-0.033
Song singing	7.938	7.482	8.450	8.051	7.289	6.638	1.162***
Playing games or do puzzle with the child	4.559	3.429	4.587	3.562	4.524	3.252	0.063
Building something with the child	3.818	2.941	3.679	2.808	3.994	3.094	-0.315**
CIS measure for the quality of interactions: total score	64.263	11.963	66.064	10.635	62.203	13.027	3.861***
Sensitivity score	21.974	6.431	22.989	5.930	20.813	6.781	2.175***
Harshness score: less harsh	23.899	3.886	24.258	3.496	23.488	4.254	0.771***
Detachment score: less detached	10.910	1.836	11.209	1.500	10.569	2.108	0.640***
Permissive score: less permissive	7.480	1.564	7.608	1.475	7.333	1.649	0.275**

Note. (a) 1-5 standards for "how often": once a month or less--> everyday; (b) For the t-test, in order to apply sample weight, we used regression: each of the variables was regressed upon a B.A. dummy as the only predictor, and the t statistic showed the statistical significance of the mean difference of that variable between children with B.A. teachers and children with non-B.A. teachers.
*p<0.1, ** p<0.05, *** p<0.01

4.5 Missing rates of the variables

Table 4-9 provides the missing rates of the dependent and explanatory variables. For the outcomes, all of the missing rates are below or around 5%. This is rather good for statistical

inference. We do not impute outcomes as this is generally a widely accepted rule in empirical research. For the predictors, the missing rate differs by variables. Most of them have a missing rate lower than 5%, except for mental ability at age 2 (7.76%), mother's degree of depression (8.02%), and quality of parenting (7.83%). As mentioned in the Research Design chapter, I use the dummy flag technique to deal with missing data in the covariates, and observations with missing data in the treatment (teachers' B.A. status) are not used.

Table 4-9 Missing rates for each of the outcomes and affected predictors

Type of variable	No.	Full size N (rounded): 4300		
		Variable	N (rounded)	missing rate
Outcome at age 4	1	Early reading	4150	3.96%
	2	Expressive language	4100	5.12%
	3	Math	4100	4.43%
	4	Color recognition	4150	3.64%
	5	Social competence	4250	1.32%
	6	Externalizing behavior	4150	4.40%
	7	Internalizing behavior	4300	0.39%
	8	Approaches to learning skills	4300	0.09%
Predictors	1	Mental ability at age 2	4000	7.76%
	2	Social skills at age 2	4200	3.13%
	3	Mother's age	4300	0.83%
	4	Mother's degree of depression	3950	8.02%
	5	Mother is married	4300	0.83%
	6	Number of books at home	4300	0.07%
	7	Quality of parenting	4000	7.83%
	8	Teacher has another race	4300	0.25%
	9	Experience	4300	0.12%
	10	History with the child (month)	4300	0.30%
	11	Hours of care per week	4300	0.30%
	12	Group size	4300	0.51%
	13	Child-adult ratio	4250	0.95%
	14	Number of books in the classroom	4300	0.09%
	15	Children's age range larger than 2	4200	2.46%

Chapter 5

Findings and Discussion I: The Effects of B.A. on Child Development Outcomes

This chapter presents the findings for the first research question, regarding the effect of B.A. on child development outcomes at age 4 and age 5. All of the results are based on the dummy flag approach for treating missing values. The findings for the concurrent/immediate outcomes at age 4 are the focus of the analysis and are presented in Sections 5.1 to 5.6; findings for age 5 are presented in Section 5.7. Specifically, for age 4 outcomes, Sections 5.1 to 5.3 present the results for RQ1.1, for OLS, PSM and IV estimation methods respectively. Section 5.4 is a summary of the findings across the three estimation methods, along with some robustness checks. Extended analyses pertaining to the heterogeneous effects of B.A. and B.A.'s interplay with specialized education in ECE (RQ1.2 to1.4) are presented in Section 5.5. Section 5.6 shows the full list of predictors of child outcomes other than teacher education. After Section 5.7's presentation of age 5 findings, Section 5.8 provides a summary of the findings in this chapter.

Since all of the eight outcome constructs have been standardized to have a mean of zero and a standard deviation of one, in all the results tables, the coefficient of the B.A. dummy itself reflects the difference in outcome in terms of standard deviation between the treated (children with B.A. teachers) and the untreated (children with non-B.A. teachers). It is comparable to the effect size definition in Kelley and Camilli's mate-analysis. In another perspective, the effect size for each of the predictor in the full list of child development predictors in the model of OLS with rich controls in Section 5.6, stands for the degree of change in outcome by standard deviation, given a standard deviation change in the predictor of interest.

5.1 OLS with rich controls

The first quantitative method for the estimation of the B.A. effect is OLS with rich controls, whereby each of the child outcomes is regressed on a set of predictors, including B.A. and its covariates. The so-called "rich controls" include children's mental ability and social skills at age 2, age at assessment measured in terms of months, boy dummy, racial dummies, low birth weight indicator, days per week in preschool, status of attendance in a center at age 2, mother's age, education, degree of depression, marital status at child birth, family income level, number of siblings, number of books at home, home language is other than English, quality of parenting, teacher's gender, racial match with the child, experience, in-service training, CDA, other certificate, history with the child (in month), hours in care per week, group size, child-adult ratio, number of books in the classroom, number of interest areas in the classroom, age range greater than two, preschool type, for profit status of the preschool, licensed or not, full-day status, the degree of home-preschool connection, and region dummies.

The OLS results are presented in Table 5-1. Two sets of the estimates of the B.A. effects for the major comparison (B.A or above vs. A.A. or below) for age 4 outcomes are presented. The model of OLS with rich controls is the focus, and the simple model of OLS without covariates is presented to indicate the direction of bias in the B.A. estimate due to a lack of covariates. Table 5-1 shows that the two sets of B.A. estimates are quite different.

Table 5-1 B.A. estimates from the OLS model: B.A. or above vs. A.A. or below

Outcome measures: age 4		N (rounded)	OLS without covariates	OLS with rich controls	
			B.A. estimates	B.A. estimates	*R* squared
(1)	Early reading (language + literacy)	4150	0.155*** (0.041)	0.079** (0.039)	0.360
(2)	Story-telling/expressive language	4100	-0.008 (0.042)	-0.040 (0.042)	0.197
(3)	Math	4100	0.206*** (0.040)	0.142*** (0.036)	0.412
(4)	Color recognition	4150	-0.002 (0.042)	-0.064 (0.041)	0.205
(5)	Social competence: parent reported	4250	-0.120*** (0.040)	-0.093** (0.042)	0.171
(6)	Externalizing behavior problem: parent reported	4150	-0.030 (0.040)	-0.020 (0.041)	0.199
(7)	Internalizing behavior[(a)] problem: parent reported	4300	-0.020 (0.042)	-0.100** (0.043)	0.092
(8)	ATL (approaches to learning): parent reported	4300	-0.038 (0.040)	-0.060 (0.041)	0.170

Note. (a) Internalizing behavior includes worries and the state of being unhappy. (b) All results were weighted by sample weight. (c) The sample sizes are rounded to the nearest 50, as required by the data provider. (d) The possibility of multicolinearity among covariates was tested and the VIF test indicated little evidence for multicolinearity.
*** p<0.01, ** p<0.05, *p<0.1; and standard errors are in parentheses.

For the model of OLS with rich controls, the model quality is quite good. First, my VIF test after the regressions indicates negligible multicolinearity. VIF stands for Variance Inflation Factor, available after a multiple regression. It indicates the magnitude of the inflation in the standard errors associated with a particular beta weight that is due to multicolinearity. A value of 10 is often seen as the maximum level of VIF (e.g., Hair, Black, Anderson, & Tatham, 1995). Among the predictors, only three variables have VIF values over two, and they are: family income (2.61), Hispanic (2.54), and Head Start center (2.46). All the other predictors have a VIF below two. In particular, the value is 1.61 for B.A. These test results remove the worries about multicolineairty among predictors for the OLS model and the IV model that also builds on OLS.

Second, the *R* squared feature is reasonable for most of the outcomes. As one can see from the last column of Table 5-1, models for most of the outcomes have a value of *R* squared larger than 0.1. The model for math skills has the largest explanatory power at 0.401. A finding worth noting is that the *R* squared value for cognitive outcome is larger than that for the non-cognitive outcomes. This may be related to the instability of non-cognitive skills across time points and the greater measurement difficulty regarding non-cognitive skills (Thiel & Thomsen, 2011; Borghans, Golsteyn, Heckman, & Humphries, 2011).

According to this focused OLS model with rich controls, the effect of B.A. differs by child development outcomes. Specifically, compared with children with non-B.A. teachers, children with B.A. teachers have higher early reading and math skills, and fewer parent-reported internalizing behavior problems, but they also have lower social skills (social competence). For the other four outcome constructs (story telling skills, color recognition, parent-reported externalizing behavior problems, and approaches to learning skills reported by parents), the B.A. estimates are not statistically significant.

Also worthy of mentioning is that, most of the statistically significant B.A. estimates in the OLS model with rich controls have smaller magnitude than the OLS without any controls, indicative of the importance of adding rich controls in the model to reduce upward bias. Otherwise, if we rely on the model of OLS without any controls, we could mistakenly attribute a larger effect size to the B.A. factor. The importance of adding rich controls is confirmed here.

5.2 Propensity score matching

As explained in Chapter 3, for PSM, I focus on the average treatment effect on the treated (ATT) generated by "psmatch2" by the algorithm of nearest-neighbor matching with replacement (with sample weight applied in the 1st stage propensity model), as the main results, for two reasons: (1) the common version of PSM (i.e. not regression adjusted) does not require the fitness of the outcome model, unlike IPTW (inverse probability treatment weighting) that is often regression adjusted and removes more imbalance between the treated and untreated group of the observational sample in some settings (Austin, 2011); and (2) the standard error for the average treatment effect (ATE) was not available in "psmatch2" for my dataset. Also, I report the IPTW's ATE estimates alongside, so that the PSM estimates are applicable for all of the matched children, including those who currently have a B. A. teacher and those who currently have a non-B.A. teacher. In so doing the PSM estimates can be compared with the OLS and IV estimates that are also based on information from the full sample. Additionally, given that computationally it is not possible to apply sample weight in PSM via STATA's "psmatch2" command,[102] the unweighted results for PSM are also available in Table B3 in the Appendices for some researchers' interest.

A brief overview of the matching and estimation process precedes the presentation of the findings. The process involves several iterative steps. First, Table 5-2 shows the set of potential confounders finally selected for the matching, ranging from the most important to the least important, and classified according to the descriptive statistics in Section 4.2 of Chapter 4 and the prior literature. The above-mentioned descriptive statistics, combined with the first-stage propensity model estimates shown in Table 5-3 of this chapter, together indicate the degree of importance of achieving balance for certain covariates.

[102] This command combines the first stage propensity model and the ATT estimates in one step.

Table 5-2 List of potential confounders for the matching

Type of confounders	List of confounders that are likely to be related to both the treatment and the outcomes
Most important to balance	Mental ability ate age 2 (Bayley scale), social skills age 2, age measured in month, racial dummies, attended center based programs before, mother's age, mother's education level is BA or higher, family income level, number of siblings, number of books at home, home language is not English, the quality of parenting, recent training, CDA, having other certificates, hours in care with the child, group size
Somewhat less important to balance[(a)]	Mother is married, degree of maternal depression, child-adult ratio, number of books in the preschool classroom, number of interest areas, preschool type dummies, preschool being for-profit, preschool was licensed
Least important to balance	Boy, low birth weight, days in preschool, teacher is male, , full-day preschool, region dummies

Note. (a) Some of them are correlated with B.A., and may have independent effects on child outcomes. (b) Two covariates that might be post-treatment variables are excluded from the list: history of the teacher with the child measured in months, and the connection between home and preschool.

Second, when part of DuGoff et al. (2014)'s two-step procedure for conducting propensity score matching with sample weight is applied, the sample weight is added as an additional predictor in the first stage probit model for generating the propensity score. Table 5-3 presents the model results for this first stage, showing the coefficient estimates for each of the potential confounders. Two square terms are added as two additional controls to obtain the model with best fitness according to the criterion for checking the balance provided by Jennifer Hill's STATA package "psbal2".

Table 5-3 Probit model estimates for the first stage of PSM: the propensity model

Potential confounders for matching		Coef. (s.e)
(1) Child characteristics	mental ability at age 2 (Bayley)	0.016 (0.026)
	social skills at age 2 (TAS-45 HOTSPOT)	0.019 (0.028)
	square term of the mental score at age 2	0.006 (0.015)
	square term of the social score at age 2	-0.021 (0.020)
	age measured in month	0.017*** (0.005)
	boy	0.051 (0.043)
	African American	-0.048 (0.080)
	Hispanic	0.068 (0.086)
	Asian	0.159 (0.114)
	Native American	-0.239 (0.154)
	Multiple race	0.199* (0.103)
	low birth weight	0.033 (0.081)
	days in preschool	0.012 (0.028)
	attended center based programs before	-0.128** (0.055)
(2) Family characteristics	mother's age	0.009** (0.004)
	mother has B.A. or above	0.218*** (0.056)
	mother’s degree of depression	0.024 (0.034)
	mother is married	-0.062 (0.062)
	family income	0.013 (0.011)
	number of siblings	0.031 (0.022)

	number of books at home (average for wave 2 and wave 3)	-0.000 (0.000)
	home language is not English	-0.067 (0.079)
Home process	quality of parenting	0.030 (0.024)
(3) Teacher characteristics	teacher is male	0.127 (0.184)
	teacher has another race	0.031 (0.060)
	experience	0.003 (0.002)
	in-service ECE training	-0.004 (0.057)
	CDA	-0.575*** (0.048)
	other certificates	0.662*** (0.048)
	hours of care for the child	-0.007*** (0.002)
(4) Non-teacher classroom characteristics	group size	0.012** (0.005)
	child-adult ratio	-0.001 (0.008)
	number of books in the classroom	0.001*** (0.000)
	number of interest areas in the classroom	0.039*** (0.014)
	age range of the classroom: 2 years or more	-0.012 (0.048)
(5) Preschool characteristics	Head Start center	0.129 (0.080)
	state preK	1.104*** (0.092)
	partially public	0.146** (0.067)
	for profit	-0.178*** (0.066)
	preschool licensed	-0.192*** (0.056)
	full day	-0.136** (0.055)

(6) Region dummies (South as the reference group)	Northeast	0.356***
		(0.069)
	Midwest	0.096
		(0.061)
	West	-0.290***
		(0.062)
Sample weight		0.000
		(0.000)
Constant		-1.669***
		(0.363)
N		4300

Note. Estimates for the dummy flags are omitted due to space limit.
*** p<0.01, ** p<0.05, * p<0.1; and standard errors are in parentheses.

Third, using the nearest-neighbor matching with replacement algorithm, each of the treated children is matched to an untreated child with the closest propensity score. Figure 5-1 shows the common support and overlap graphs of the data, and the conditions are satisfied. All of the treated observations are on support, although some of the untreated are not used, namely "unmatched". Additionally, results for the above mentioned balance table generated by "psbal2" for each of the confounders are presented in Table 5-4. After matching, most of the confounders have a standardized mean difference smaller than 0.10, and all of the standardized mean differences are below 0.15. In addition, the pre and post balance histograms suggested by Austin (2011) for certain important potential confounders (i.e., age in months, family income, quality of parenting) are presented in Figure 5-2. These balance graphs look better in the post-matching scenario than in the pre-matching scenario.

Figure 5-1 Common support and overlap

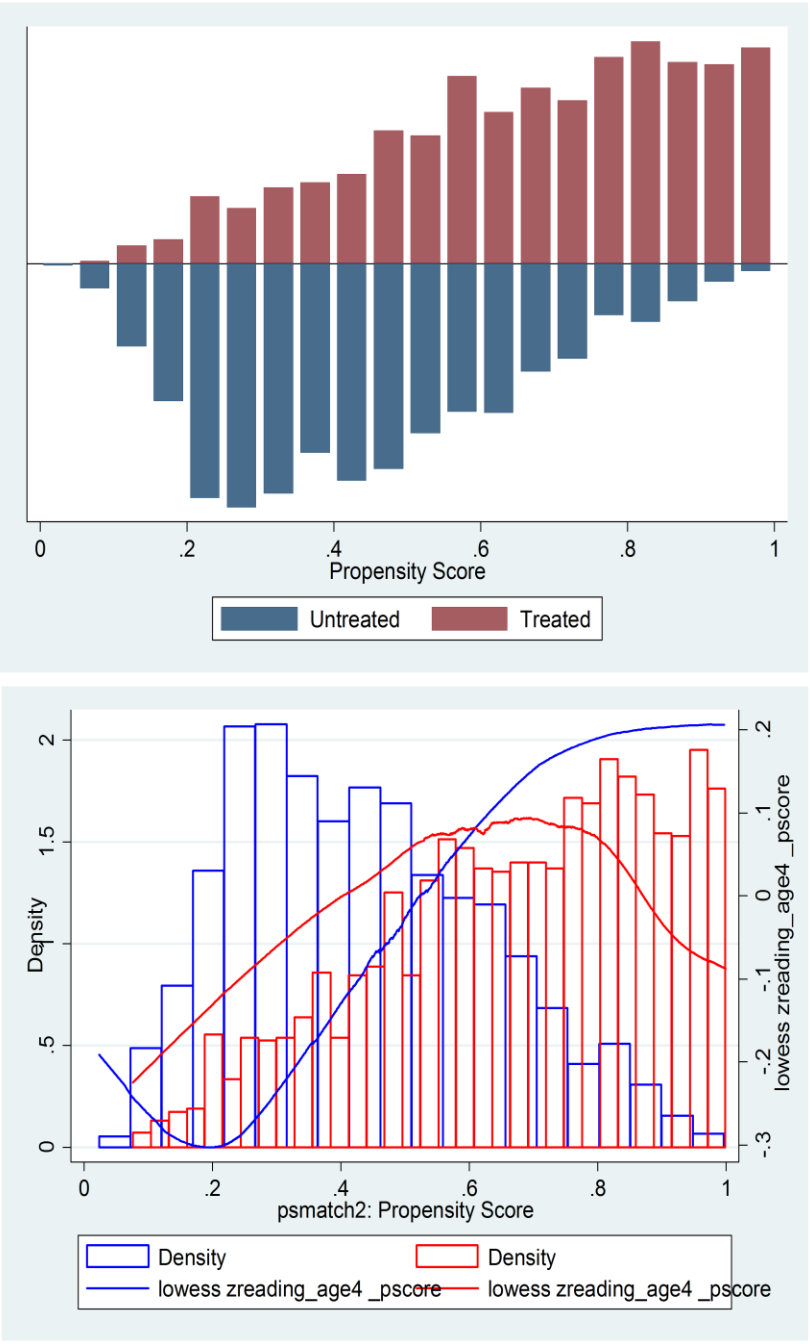


Table 5-4 Balance table for PSM using psbal2

Variable	Sample	Mean		SD		Standardized mean difference	Ratio of SDs
		Treated	Control	Treated	Control		
Mental ability at age 2 (Bayley)	Unmatched	782.096	687.023	2685.50	2530.09	0.035	1.06
	Matched	782.096	816.213	2685.50	2739.56	-0.013	0.98
the square term of the above	Unmatched	782.999	687.799	2685.24	2529.88	0.035	1.06
	Matched	782.999	817.304	2685.24	2739.24	-0.013	0.98
Social skills age at 2 (TAS-45 HOTSPOT)	Unmatched	269.230	360.079	1618.93	1863.70	-0.056	0.87
	Matched	269.230	192.361	1618.93	1374.17	0.047	1.18
the square term of the above	Unmatched	270.197	361.029	1618.77	1863.52	-0.056	0.87
	Matched	270.197	193.290	1618.77	1374.04	0.048	1.18
Age of assessment measured in month	Unmatched	53.298	52.816	3.87	4.05	0.125	0.95
	Matched	53.298	53.254	3.87	3.79	0.011	1.02
Boy	Unmatched	0.525	0.502	0.50	0.50	0.046	1
	Matched	0.525	0.470	0.50	0.50	0.11	1
African American	Unmatched	0.142	0.172	0.35	0.38	-0.086	0.93
	Matched	0.142	0.171	0.35	0.38	-0.084	0.93
Hispanic	Unmatched	0.151	0.161	0.36	0.37	-0.029	0.97
	Matched	0.151	0.124	0.36	0.33	0.075	1.09
Asian	Unmatched	0.119	0.096	0.32	0.30	0.07	1.1
	Matched	0.119	0.134	0.32	0.34	-0.045	0.95
Native American	Unmatched	0.024	0.041	0.15	0.20	-0.112	0.77
	Matched	0.024	0.015	0.15	0.12	0.059	1.26
Multiple race	Unmatched	0.090	0.076	0.29	0.27	0.048	1.08
	Matched	0.090	0.064	0.29	0.25	0.091	1.17
Low birth weight	Unmatched	0.098	0.095	0.30	0.29	0.011	1.01
	Matched	0.098	0.091	0.30	0.29	0.023	1.03
Days in preschool	Unmatched	4.326	4.397	1.00	0.98	-0.07	1.01
	Matched	4.326	4.350	1.00	1.01	-0.024	0.99
Prior attendance to centers age at age 2	Unmatched	0.180	0.239	0.38	0.43	-0.155	0.9
	Matched	0.180	0.180	0.38	0.38	0	1
Mother's age	Unmatched	88.500	149.344	769.14	1094.53	-0.079	0.7
	Matched	88.500	80.230	769.14	712.65	0.011	1.08
Mother has a B.A. or above	Unmatched	0.400	0.324	0.49	0.47	0.156	1.05
	Matched	0.400	0.429	0.49	0.50	-0.058	0.99
Mother's degree of depression	Unmatched	727.770	843.370	2595.80	2777.42	-0.045	0.93
	Matched	727.770	659.431	2595.80	2480.67	0.026	1.05
Mother is married	Unmatched	60.561	122.559	771.28	1097.49	-0.08	0.7
	Matched	60.561	52.005	771.28	714.65	0.011	1.08
Family income	Unmatched	8.276	7.798	3.18	3.21	0.151	0.99
	Matched	8.276	8.474	3.18	2.99	-0.062	1.06

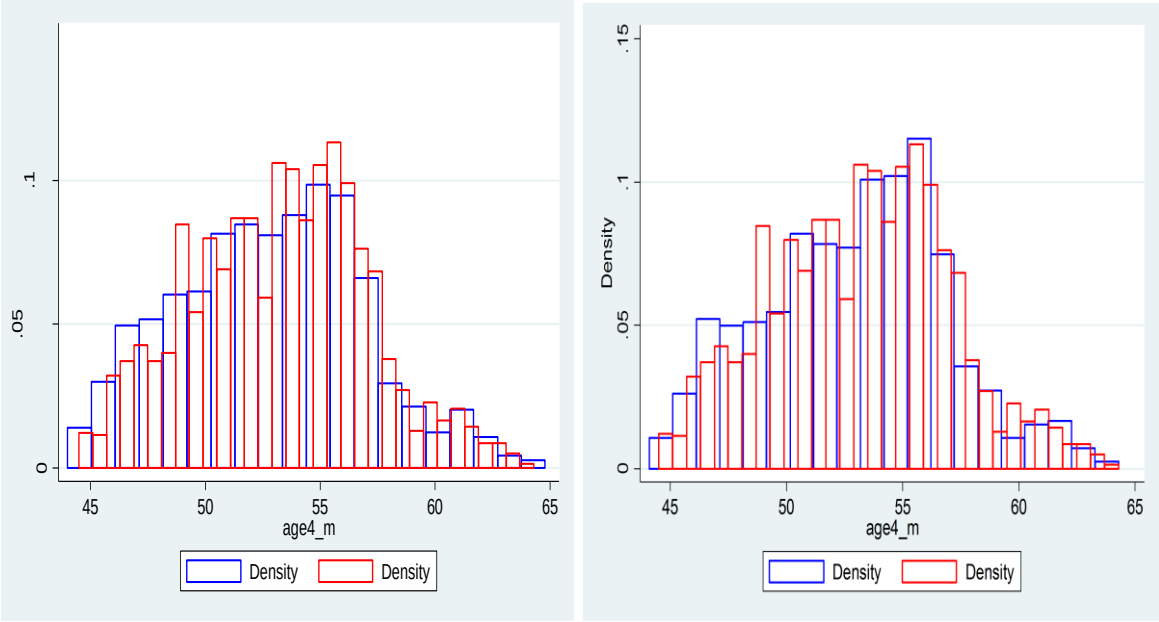
Number of siblings	Unmatched	1.435 1.332	1.09 1.05	0.094	1.03
	Matched	1.435 1.406	1.09 0.97	0.027	1.12
Number of books at home (wave 2 and 3)	Unmatched	75.250 60.749	297.21 61.09	0.049	4.87
	Matched	75.250 68.581	297.21 62.28	0.022	4.77
Home language not English	Unmatched	0.181 0.168	0.38 0.37	0.034	1.03
	Matched	0.181 0.183	0.38 0.39	-0.006	1
The quality of parenting	Unmatched	807.749 764.453	2725.37 2658.00	0.016	1.03
	Matched	807.749 782.077	2725.37 2686.54	0.009	1.01
Teacher is male	Unmatched	0.015 0.013	0.12 0.11	0.01	1.04
	Matched	0.015 0.012	0.12 0.11	0.021	1.1
Teacher has another race	Unmatched	21.797 33.654	461.84 575.73	-0.026	0.8
	Matched	21.797 81.592	461.84 897.93	-0.129	0.51
Experience	Unmatched	22.044 29.666	292.03 407.01	-0.026	0.72
	Matched	22.044 26.470	292.03 357.65	-0.015	0.82
ECE training in the past 12 months	Unmatched	0.819 0.811	0.39 0.39	0.021	0.98
	Matched	0.819 0.806	0.39 0.40	0.034	0.97
CDA	Unmatched	0.205 0.411	0.40 0.49	-0.511	0.82
	Matched	0.205 0.194	0.40 0.40	0.025	1.02
Other certificates	Unmatched	0.414 0.187	0.49 0.39	0.46	1.26
	Matched	0.414 0.415	0.49 0.49	-0.002	1
Hours in care for the child	Unmatched	63.371 39.740	651.23 406.73	0.036	1.6
	Matched	63.371 90.601	651.23 822.82	-0.042	0.79
Group size	Unmatched	57.133 80.273	651.54 811.77	-0.036	0.8
	Matched	57.133 48.471	651.54 583.25	0.013	1.12
Child-adult ratio	Unmatched	70.657 133.835	797.72 1121.19	-0.079	0.71
	Matched	70.657 151.832	797.72 1196.54	-0.102	0.67
Number of books in the classroom	Unmatched	134.137 110.717	280.37 427.33	0.084	0.66
	Matched	134.137 128.001	280.37 185.50	0.022	1.51
Number of interest areas in the classroom	Unmatched	8.766 8.557	1.55 1.94	0.135	0.8
	Matched	8.766 8.926	1.55 1.48	-0.103	1.05
Age range of children in the classroom larger than 2	Unmatched	175.497 343.804	1312.28 1821.69	-0.128	0.72
	Matched	175.497 184.032	1312.28 1343.84	-0.007	0.98
Head Start center	Unmatched	0.171 0.259	0.38 0.44	-0.233	0.86
	Matched	0.171 0.178	0.38 0.38	-0.017	0.99
State preK (prekindergartens)	Unmatched	0.233 0.040	0.42 0.20	0.457	2.16
	Matched	0.233 0.235	0.42 0.42	-0.004	1
Partially publicly funded private preschools	Unmatched	0.260 0.245	0.44 0.43	0.034	1.02
	Matched	0.260 0.241	0.44 0.43	0.043	1.02
For profit	Unmatched	0.194 0.319	0.40 0.47	-0.318	0.85
	Matched	0.194 0.195	0.40 0.40	-0.003	1

Preschool was licensed	Unmatched	0.713 0.821	0.45 0.38	-0.238	1.18
	Matched	0.713 0.732	0.45 0.44	-0.041	1.02
Full day preschool	Unmatched	0.293 0.384	0.46 0.49	-0.199	0.94
	Matched	0.293 0.317	0.46 0.47	-0.052	0.98
Northeast	Unmatched	0.213 0.100	0.41 0.30	0.276	1.36
	Matched	0.213 0.242	0.41 0.43	-0.071	0.96
Midwest	Unmatched	0.235 0.213	0.42 0.41	0.052	1.04
	Matched	0.235 0.204	0.42 0.40	0.072	1.05
West	Unmatched	0.188 0.287	0.39 0.45	-0.254	0.86
	Matched	0.188 0.188	0.39 0.39	0	1

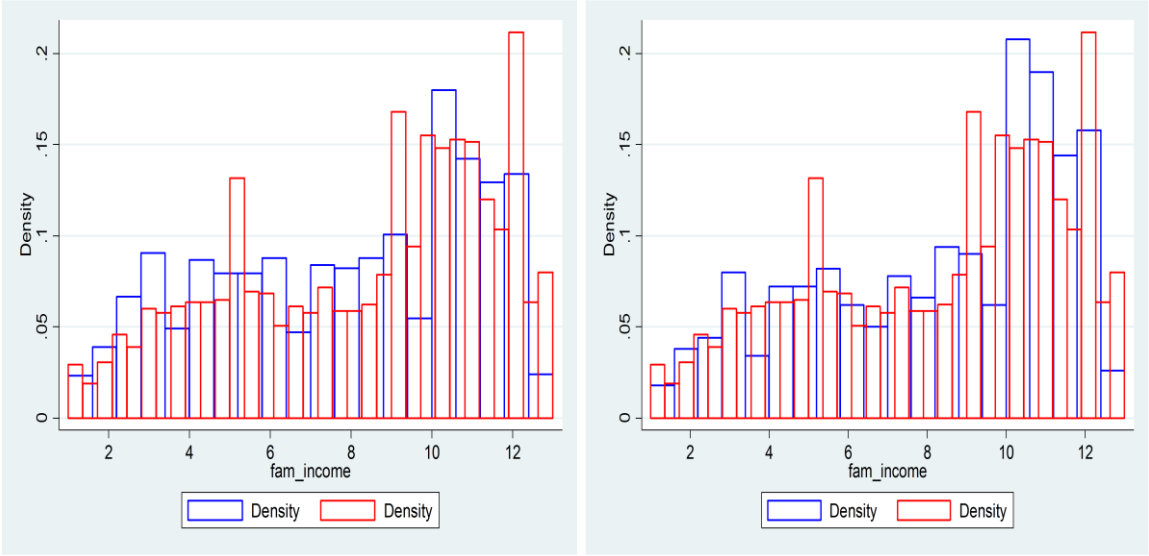
Note. The "unmatched" controls are those untreated observations that are not used as controls in the matching. The balance statistics for the dummy flags are omitted due to space limit, and the balances are pretty good for most of them.

Figure 5-2 Pre and post balance graph on certain important potential confounders: Pre (left) vs. Post (right) and Treated (red) vs. Untreated (blue)

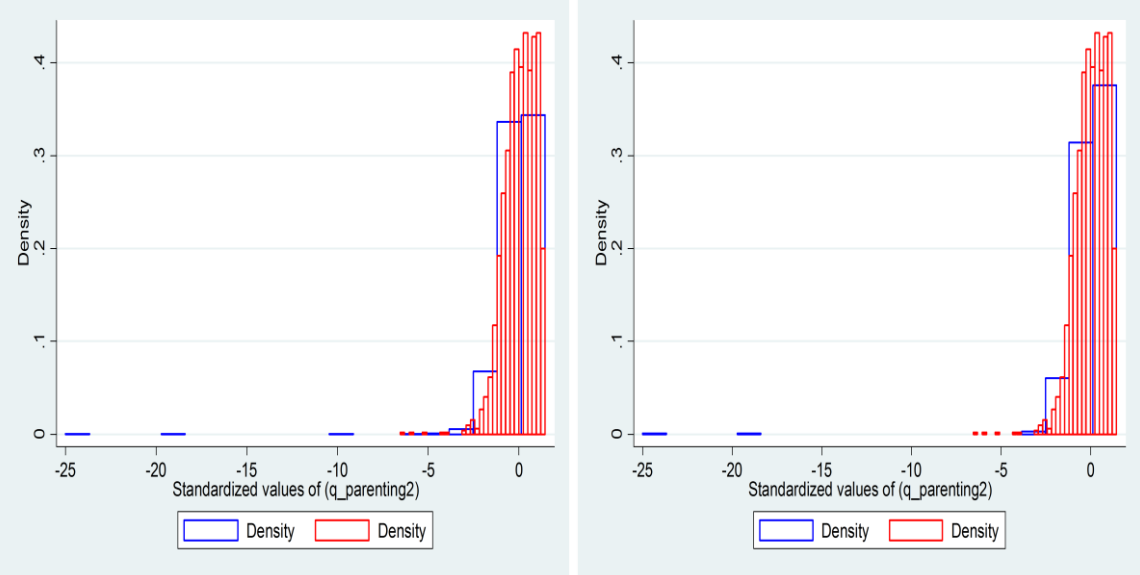
A. age of assessment measured in months



B. family income



C. quality of parenting



Fourth, B.A. estimates in terms of average treatment effect on the treated (ATT) from PSM are reported in the third column of Table 5-5. The PSM estimated B.A. effect is only statistically significantly positive for math and negative for social competence. The significance of the B.A. effects on math skills and social competence are consistent with the OLS model estimates. Apart from this similarity the PSM estimates are different from the OLS estimates in two ways: (1) the effect of B.A. on early reading skills is not statistically significant in PSM; and (2) the effect on internalizing behavior problems is not statistically significant in PSM. Note also that results are not straightly comparable because the PSM estimates pertain to the average effect for the treated children (ATT), whereas the OLS estimates pertains to the average effect for the whole sample of children (ATE), no matter whether their teachers have a B.A. or not.

I also present the IPTW estimates for the average treatment effect for both the treated and the untreated observations in the last column of Table 5-5. Both the PSM ATT and IPTW results were derived using "psmatch2" in STATA, and the estimation processes have accounted for sample weight according to DuGoff et al. (2014): for PSM ATT, only in the first stage propensity model; for IPTW, also in the second stage effect estimation. Another command in STATA, "attnd," was also tried, but it experienced the same problem as "psmtach2" of not being able to report the standard error for ATE; and that's why IPTW is a necessity for the calculation of ATE. This IPTW ATE has better comparability with OLS estimates because the OLS estimates are ATE by nature too. According to the ATE estimates, only one outcome (math skills) is statistically significantly and positively affected by preschool teachers' B.A. status. Therefore, among all the outcomes, B.A.'s effect on math skills is the most consistent: across PSM and OLS and across two alternative estimands for matching analysis (PSM's ATT and IPTW's ATE).

Table 5-5 B.A. estimates for PSM (nearest-neighbor matching with replacement): ATT for PSM and ATE using IPTW after the PSM

Outcome measures: age 4		N (rounded)	PSM (not regression adjusted): ATT	IPTW (without covariates): ATE
(1)	Early reading	4150	0.066 (0.069)	0.045 (0.058)
(2)	Story telling/Expressive language	4100	0.046 (0.070)	0.029 (0.062)
(3)	Math	4100	0.122* (0.069)	0.131** (0.052)
(4)	Color recognition	4150	0.016 (0.067)	-0.033 (0.056)
(5)	Social competence: parent-reported	4250	-0.190** (0.067)	-0.079 (0.057)
(6)	Externalizing behavior problem: parent-reported	4150	0.097 (0.069)	0.052 (0.060)
(7)	Internalizing behavior (worries and unhappy): parent-reported	4300	0.073 (0.072)	-0.031 (0.048)
(8)	ATL (approaches to learning): parent-reported	4300	0.030 (0.068)	-0.045 (0.037)

*** p<0.01, ** p<0.05, * p<0.1; and standard errors are in parentheses.

Additionally, sensitivity checks for PSM by matching methods are shown in Table 5-6. First, if we look at the with/without replacement choice in the nearest-neighbor matching method, the positive effects on math skills are there for both scenarios. For the early reading outcome, however, only the "without replacement" scenario shows a positive effect of B.A. Also, the negative B.A. effect on social competence disappears if we use the "without replacement" option for the nearest-neighbor matching. Note that nearest-neighbor matching without replacement is more suitable when there are many untreated observations, and this is not the case for my sample. Second, the B.A. estimates are not very robust across matching methods, and across the three alternative matching algorithms. Estimates from Kernel matching and radius matching point to the negative effect of B.A. on social competence, whereas nearest-neighbor

matching indicates significance for math and social competence. Still, the insignificances of B.A. effect on story-telling (expressive language), color recognition, and internalizing behavior are robust across all scenarios.

Overall, the positive effect on math skills is consistent for "replacement/without replacement," and the negative effect of B.A. on social competence is consistent across three matching methods. For expressive language, color recognition, and internalizing behavior, the insignificances are robust across all scenarios.

Table 5-6 Sensitivity check for PSM ATT:
Without replacement and alternative matching method

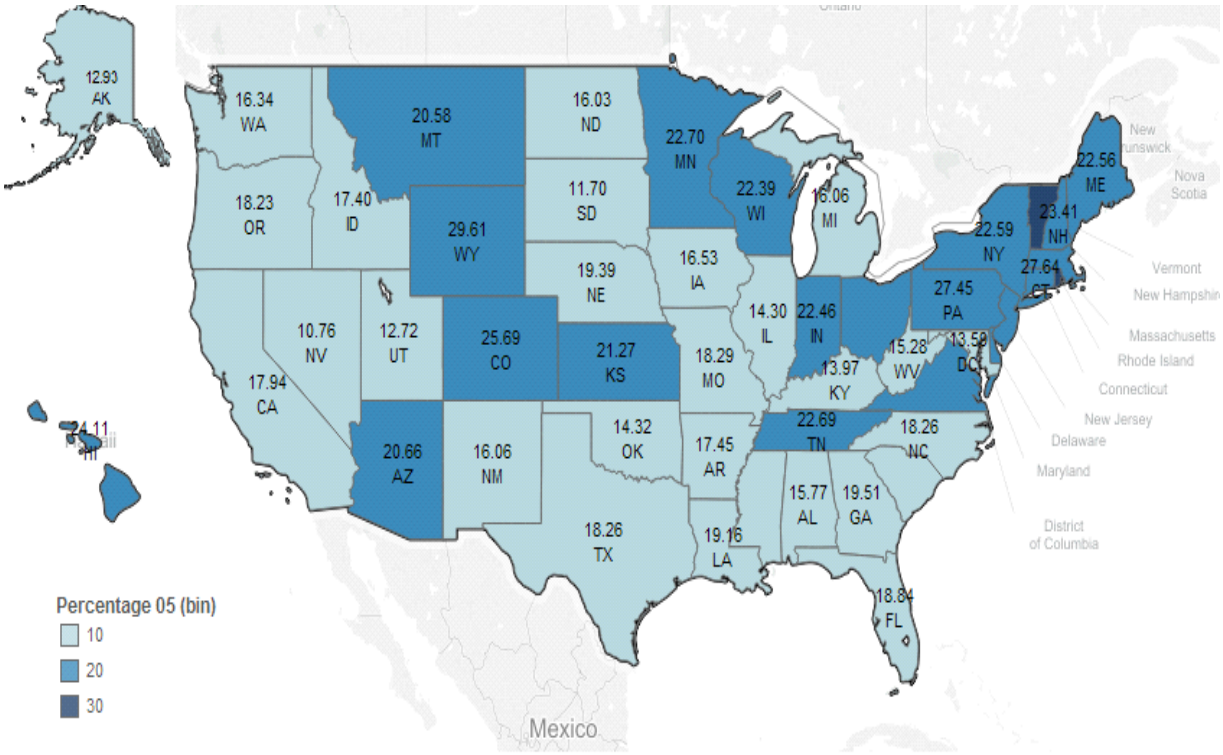
Outcome measures		N (rounded)	Nearest Neighbor Matching		Kernel matching (Epanechnikov)	Radius/Caliper matching
			with replacement	without replacement		
(1)	Early reading	4150	0.066 (0.069)	0.155*** (0.033)	-0.015 (0.050)	-0.043 (0.044)
(2)	Story telling/expressive language	4100	0.046 (0.070)	0.032 (0.033)	-0.007 (0.051)	-0.039 (0.045)
(3)	Math	4100	0.122* (0.069)	0.176*** (0.033)	0.028 (0.050)	0.006 (0.045)
(4)	Color recognition	4150	0.016 (0.067)	0.046 (0.033)	0.032 (0.052)	0.006 (0.046)
(5)	Social competence: parent reported	4250	-0.190** (0.067)	-0.051 (0.032)	-0.118** (0.049)	-0.150*** (0.043)
(6)	Externalizing behavior problem: parent reported	4150	0.097 (0.069)	-0.062* (0.033)	0.109 (0.052)	0.062 (0.046)
(7)	Internalizing behaviour problem: parent reported	4300	0.073 (0.072)	-0.012 (0.033)	0.047 (0.053)	0.025 (0.046)
(8)	ATL (approaches to learning): parent reported	4300	0.030 (0.068)	0.010 (0.032)	-0.048 (0.050)	-0.062* (0.034)

Note. The radius used in the Radius/Caliper matching is 0.1, which is standard choice.
*** p<0.01, ** p<0.05, * p<0.1; and standard errors are in parentheses.

5.3 Instrumental variable

As mentioned in Chapter 3, the major instrumental variable (IV) used for the IV model in this dissertation is the state-level percentage of early childhood care and education workers with a B.A. or above. There is variation in the distribution of this state level IV, i.e., state-level percentage of B.A. workers in the early childhood care and education workforce in 2005 based on data from American Community Survey, as shown in the map in Figure 5-3. For example, Nevada has the lowest percentage (10.76%). South Dakota has a low percentage too, at 11.70%. New York's percentage is 22.59%. Comparatively, Massachusetts's B.A. percentage is higher (29.06%), and Rhode Island has the highest among all states (35.23%). This IV, which was derived from outside data, is also highly correlated with the state level average calculated from the ECLS-B dataset. More importantly, as seen in Table 5-8, for each of the outcome constructs, the first stage of the 2SLS estimation of the IV method returns a F-value of 30 to 40, surpassing the threshold of 10 recommended in Staiger and Stock (1997). This implies that this IV is not a weak IV. The full estimates for the first stage for the IV model of the "early reading outcome" are presented in Table B4 of the Appendices.

Figure 5-3 Distribution of the IV across state:
Percentage of B.A. holders in the early childhood care and education workforce



Note. (a) The sub-maps of Alaska and Hawaii are embedded into the main map, and the size does not represent their comparative size with other states. (b) The unit for the values is percentage points (%). (c) The darker the area, the higher the value of the percentage. (d) This map was generated by Tableau Desktop.

As also explained in the Research Design chapter, this IV is likely to satisfy the "independence" assumption in Angrist and Pischke (2009), or in other words, the "ignorability" assumption), because families of young children are not likely to select preschools across states. At the same time, the "exclusion restriction" assumption is also likely to be satisfied given that my control of individual level, classroom level and preschool level characteristics helps rule out the influence of other contemporary state-level policies/factors that are also likely to affect child development. The reason is that the other contemporary state-level policies/factors (apart from teacher qualification policy that is related to the education levels of the workforce) have to go through the lower level characteristics actually to affect child outcomes. Holding these micro-

level factors constant in the outcome model reduces the possibility that the IV's supposed exogenous effect is contaminated by these contemporary polices. One example is the state's standard on group size that often comes with the state standard on teachers' degrees. This standard takes effect only by affecting the actual group size in the child's classroom, and this classroom level variable has been controlled for in both stages of the IV method.

The B.A. estimates from the IV method are presented in Table 5-7. In the IV models, statistically significant B.A. effect was only found for reducing externalizing behavior. This effect applies to the "compliers" only, namely, to those children whose chance of meeting a B.A. teacher was affected by their state residence (and the state stock of highly educated early childhood teachers), according to the "local treatment effect on the treated" (LATE) framework (Angrist & Pischke, 2009; Angrist, Imbens, & Rubin, 1996; Imbens & Angrist, 1994).[103] Still, these "compliers" are "unseen" and cannot be specifically identified because we cannot observe the counterfactuals (Murnane & Willett, 2011, p.280). In other words, we do not know which subset of children among the full sample of children the effect can be applied to. We only know the effect is estimated through the comparison of children across states with different probabilities of having a B.A. teacher. This may imply that when a child is moved to another state where the educational levels of the early childhood workforce is higher, he/she may benefit from having a B.A. teacher in terms of reducing his/her externalizing behavior problems. Note also that, because of this locality, the IV estimates of the B.A. effects are not directly comparable with the OLS estimates.

[103] As also stated in Murnane and Willett (2011), the IV estimator only capitalizes "on variation in the question predictor that is sensitive to variation in the instrument" (p. 280).

Table 5-7 B.A. estimates for the IV model

Outcome measures: age 4		N	Coefficient for B.A. (s.e.)	F statistic	Endogeneity test: p value	Reduced form estimates: for the IV
(1)	Early reading	4150	0.310 (0.313)	36.837	0.541	0.004 (0.004)
(2)	Story telling/expressive language	4100	0.457 (0.327)	35.784	0.116	0.006 (0.004)
(3)	Math	4100	0.452 (0.321)	37.086	0.095	0.004 (0.004)
(4)	Color recognition	4150	0.264 (0.284)	36.163	0.190	0.006 (0.004)
(5)	Social competence: parent reported	4250	0.529 (0.335)	35.818	0.050	0.007* (0.004)
(6)	Externalizing: behavior parent reported	4150	-0.729** (0.297)	38.402	0.013	-0.011** (0.004)
(7)	Internalizing behavior: parent reported	4300	0.183 (0.340)	40.647	0.409	0.00005 (0.005)
(8)	ATL (approaches to learning): parent reported	4300	0.003 (0.307)	40.670	0.831	0.00005 (0.004)

*** p<0.01**, p<0.05, *p<0.1; and standard errors are in parentheses.

Note that when we look at the reduced form estimates for the instrumental variable, the effect of B.A. on reducing externalizing behavior is there, consistent with the IV estimate. Very interestingly, a 10% significant effect was also found for social competence, and it is positive. This, when compared with the OLS finding of a negative B.A. effect on social competence, may indicate some degree of endogeneity of the B.A. treatment. Additionally, according to the p value of the endogeneity test shown in the second to the last column of Table 5-8, the IV results are mostly similar to the OLS results, except for social competence and externalizing behavior. These two exceptions are the cases where the B.A. effect was "thought" by this test to be endogenous, although this endogeneity test has its own limitations.[104]

[104] In statistics, it is very difficult to test endogeneity. This endogeneity test, available in STATA after "ivregress2" and similar to the Hausman test, aims to provide some inferences about the degree of the endogeneity problem by comparing result differences across the IV and OLS models, but these are incomplete. Models that pass the test may still suffer from endogeneity problems.

In addition, two other IVs have been considered in this study: (1) an indicator, based on data from the State of Preschool Yearbook, of whether the state has adopted a B.A. standard for its state prekindergarten programs; (2) the number of B.A.-granting higher education institutions for ECE in each of the states. The results for the latter are not presented here because these two instrumental variables were weak according to their first-stage estimates. Descriptive statistics and the F values of the first stages of these two IVs are provided in Table B5 of the Appendices. This information may be helpful to future researchers for selection of instrumental variables.

5.4 Summary of the main findings and robustness checks

Table 5-8 puts together the findings from all three estimation methods. This allows for the direct comparison of the B.A. estimates across the three methods. According to the model of OLS with rich controls, children with B.A. teachers are known to exhibit better early reading and math skills than children with non-B.A. teachers and fewer parent-reported internalizing behavior problems. For example, holding other factors constant, on average, a child with a B.A. teacher is 0.142 standard deviation higher in math scores than a child with a non-B.A. teacher. No effects are found for story-telling skills, color recognition, parent-reported externalizing behavior problems or approaches to learning skills. Yet the children with B.A. teachers are found to exhibit lower parent-rated social competence than children with non-B.A. teachers.

Table 5-8 Main results for the B.A. estimates in three methods (OLS, PSM, and IV): B.A. (B.A. or above) vs. non-B.A. (A.A. or below)

Outcome measures		N (rounded)	OLS	PSM	IV
			Coefficient for B.A. (s.e.)(a)	ATT, with replacement	Coefficient for B.A. (s.e)
(1)	Early reading	4150	0.079** (0.039)	0.066 (0.069)	0.310 (0.313)
(2)	Story-telling/ expressive language	4100	-0.040 (0.042)	0.046 (0.070)	0.457 (0.327)
(3)	Math	4100	0.142*** (0.036)	0.122* (0.069)	0.452 (0.321)
(4)	Color recognition	4150	-0.064 (0.041)	0.016 (0.067)	0.264 (0.284)
(5)	Social competence: parent-reported	4250	-0.093** (0.042)	-0.190** (0.067)	0.529 (0.335)
(6)	Externalizing: behavior parent-reported	4150	-0.020 (0.041)	0.097 (0.069)	-0.729** (0.297)
(7)	Internalizing behavior: parent-reported	4300	-0.100** (0.043)	0.073 (0.072)	0.183 (0.340)
(8)	Approaches to learning: parent-reported	4300	-0.060 (0.041)	0.030 (0.068)	0.003 (0.307)

Note. (a) 44 covariates are controlled for in the model of OLS with rich controls. (b) All results were weighted by sample weight. (c) Sample sizes for PSM differ from the OLS in that some treated observations are off support, and that the estimates are for the treatment on the treated (ATT), regression adjusted. (d) The coefficient of B.A. here is actually the effect size: from having a non-B.A. teacher to having a B.A. teacher. (e)
*** p<0.01, ** p<0.05, *p<0.1; and the standard errors are in parentheses.

Comparatively, the PSM and IV estimates tend to be less statistically significant than the OLS estimates: most of the effects disappear in the PSM models (except for math and social competence) and in the IV models (an effect for reducing externalizing behavior was found instead). Magnitude wise for math skills and social competence, PSM estimates of B.A. are somewhat smaller than the OLS estimates, possibly because of upward bias in the OLS that did not ensure the group of children with non-B.A. teachers was a similar group to the group of children with B.A. teachers. Still, some of the differences may be attributable to the heterogeneous effects because some of the children with non-B.A. teachers were not used in PSM. Note, however, such differences in magnitude between the OLS and PSM estimates for

early reading and math are small, whereas the magnitudes and signs are less consistent for the other outcomes. Also, the IV estimates of B.A. are many times larger in absolute magnitude than the OLS estimate (e.g., early reading, expressive language, math, color recognition, social competence, externalizing behavior, and internalizing behavior), except in one case (approaches to learning skills). Notably, the standard errors of the IV model are also much larger than those in the OLS model. This is typical and consistent with many other IV studies (e.g., Angrist & Krueger, 1991; Magnuson, Ruhm, & Waldfogel, 2007; Auger, Farkas, Burchinal, Duncan, & Vandell, 2014), given that the actual treatment variation in the IV estimation comes from the impact of the instrumental variable on the treatment.

Overall, five out of the eight outcome constructs at age 4 were found to be affected by preschool teachers' B.A. status in at least one of the three models: positive effect on early reading and on math, and negative effect on social competence, externalizing behavior problems, and internalizing behavior problems.[105] The differences across the three models are attributed to either bias or heterogeneity, given that each has its own strengths and weaknesses.

A comparison of the three methods is summarized in Table 5-9. Specifically, OLS with rich controls controlled for a rich set of confounders of B.A. so that it can remove the selection bias that are possibly generated by observed differences in these characteristics, but it relies on parametric assumptions and cannot deal with the unobserved confounders. PSM has advantages over OLS in relaxing the parametric assumption and in deriving more robust estimates (Conniffe, Gash, & O Connell, 2000; Rubin & Thomas, 2000), yet it estimates the average treatment effect on the treated only, not for the average. At the same time, the IV method can do a better job than OLS with rich controls and PSM in removing some endogeneity of the treatment, especially the

[105] When the approaches to learning skills is broken down into items (i.e., attention", "eagerness to learn", "independence", and "persistence), no statistically significant B.A. effect was found for each of the items.

bias from unobserved characteristics, but the IV estimate is only local to a subset of children in the sample (i.e., the "compliers").

Table 5-9 Relative strengths and weaknesses of the three quantitative methods

Strengths and weaknesses	OLS with rich controls	PSM	IV
Strength	Rich controls account for the observed differences between the two groups of children-teachers	Relax the parametric assumption; more robust estimates given the similarity between the pairs (Conniffe, Gash, & O'Connell, 2000; Rubin & Thomas, 2000)	Can reduce bias from the unobserved sources, though the degree of bias reduction depend on the exogeneity of the IV
Type of bias reduced (the four identification challenges listed in Section 3.6.1 of Chapter 3, p. 83-85)	(1), part of (2), part of (4)	(1), part of (2), part of (4)	The observed part of (1) (2), part of (4)+ Unobserved part of (2) &(3)
Weakness	Cannot deal with the unobserved confounders; rely on parametric assumptions	Can only address selection on observables; treatment effect on the treated (ATT) only, such estimates are not exactly desirable given that they are not for the average population and cannot be directly comparable with OLS unless using IPTW	Local effect: for the compliers only (LATE)

Therefore, as cautious researchers, we need to make conclusions based on all three models together, considering the relative strengths and weaknesses of each. The best way to draw conclusions is to summarize the findings for all three methods: (1) For the whole picture, our findings indicate that B.A. matters for certain types of child development outcomes in certain methods, particularly so for early reading and math. (2) The evidence is not very strong, given the inconsistency of findings across methods, the negative effect of B.A. on social competence

and the small effect size even for the mostly affected child outcome (i.e., math) using Cohen's criterion.[106]

The rest of this section presents the findings of some additional analyses.

First, I separate the existing treatment (B.A., including above B.A.) into "above B.A." and "B.A. only." The comparison group is still "A.A. or below". I find that the existing treatment effects are mostly attributable to the "B.A. only" treatment rather than the "above B.A." treatment (M.A., Ph.D., professional degrees, etc.). The only exception is for the negative effect on social competence, which is driven mainly by the "above B.A." treatment. Results are not shown here but are available in Table B6 of the Appendices.

Next, Table 5-10 is presented to address a relevant policy, that is, what's the effect of a B.A. teacher as compared with an A.A. teacher on children's development outcomes. For this question, the sample is restricted to children with either B.A. or A.A. teachers. Answers to this question help policy-makers to decide which educational level to set as a threshold for teacher qualifications. The table shows that children with B.A. teachers have higher math skills and fewer internalizing behavior problems than children with A.A. teachers, but there is a negative B.A. effect on color recognition. Comparing these findings with the comparison between at least B.A. vs. below B.A. analysis (OLS in Table 5-8), we see that both found a positive effect on math and reduction effect on internalizing behavior. The difference is that the positive effect on early reading is no longer statistically significant for this new comparison; neither is the effect on social competence statistically significant in the new comparison. These findings mean that B.A. teachers still have some advantages over A.A. teachers in terms of giving children better math

[106] It is also small when compared to the effect size of some prior empirical studies (Kelley & Camilli, 2007).

skills and reducing internalizing behavior, but the comparative advantage does not pertain to early reading.

Table 5-10 OLS estimates for the B.A. effects:
B.A. (not including those above B.A.) compared with A.A.

Outcomes: age 4		N (rounded)	OLS coefficient (s.e.)
(1)	Early reading	2200	0.065 (0.054)
(2)	Story telling/Expressive language	2150	-0.035 (0.061)
(3)	Math	2150	0.113** (0.049)
(4)	Color recognition	2200	-0.130** (0.058)
(5)	Social competence: parent reported	2250	-0.063 (0.061)
(6)	Externalizing: behavior parent reported	2200	-0.054 (0.058)
(7)	Internalizing behavior (worries and unhappy): parent reported	2250	-0.126** (0.061)
(8)	ATL (approaches to learning): parent reported	2250	-0.020 (0.058)

*** p<0.01, ** p<0.05, *p<0.1; and standard errors are in parentheses.

Finally, I conduct robustness analyses on the main findings in Table 5-8, focusing on the model of OLS with rich controls. These analyses consist of seven checks on internal validity and one check on external validity.

Let us consider the robustness checks on internal validity first.

(1) Using an alternative missing data treatment technique: complete cases instead of dummy flag. Findings from the complete case scenario are fairly similar to the dummy flag case. The two exceptions are that, for the complete cases technique, the effect on social competence was not

statistically significant in the OLS models, and an additional significant effect on externalizing behavior was found.

(2) Adding additional covariates in the model, including major in ECE for a degree, teacher age, earnings, and peer context for a sub-sample. The inclusion of major in ECE for a degree addresses the concern that specialized college education may affect both the B.A. treatment and the child development outcomes. The inclusion of the teacher's age when a teacher's experience was controlled in the model accounts for the possibility that age has an independent effect on child outcomes other than experience (for example, similarity to mother's age) and that the younger generation of teachers may have more chance of a B.A. education. The inclusion of teacher's earnings is to account for the differences between the earnings of B.A. teachers and non-B.A. teachers.[107] The inclusion of peer effect addresses the concern that children may behave or develop differently in different peer contexts (racial composition) through several mechanisms (language development and social identity, problem-solving and conflict resolution) (Fram & Kim, 2012).[108] As presented in Table 5-11, the new B.A. estimates are close to those presented in Table 5-8. A slight difference lies in the absolute values of the coefficient of B.A., which tend to be larger when these new variables are added in the model.

[107]As also mentioned in the Chapter 3, I did not control earnings in the baseline model because holding earnings constant is likely to reduce or remove the variation of teacher education, as pointed out by Barnett (2011). Still, in practice or technically, this does not prevent inclusion of the earnings variable in the model. Note that in my dataset, some of the variation is still there because this is a national sample that follows children from different states, so that even when it holds earnings constant, across-state variation in B.A. status remains, given that the cost of living differs by state.

[108] Luckily the racial composition for each child's center is available in ECLS-B for a subgroup of the children in the center director's self-answered questionnaire (SAQ), and it was used in the analysis here.

Table 5-11 New estimates of the B.A. effects in OLS after adding additional covariates: For the comparison with Table 5-8

Outcome at age 4	Full sample			subsample	
	OLS in Table 5-8	adding ECE major and teacher's age	further add earnings	before	add racial composition
early reading	0.079** (0.039)	0.096** (0.043)	0.094** (0.044)	0.116 (0.084)	0.116 (0.084)
expressive language/story telling	-0.040 (0.042)	-0.019 (0.048)	-0.003 (0.049)	-0.024 (0.085)	-0.019 (0.085)
math	0.142*** (0.036)	0.175*** (0.040)	0.184*** (0.041)	0.149** (0.075)	0.147* (0.077)
color recognition	-0.064 (0.041)	-0.065 (0.048)	-0.062 (0.048)	-0.105 (0.092)	-0.113 (0.091)
social competence	-0.093** (0.042)	-0.113** (0.048)	-0.087* (0.049)	0.058 (0.090)	0.056 (0.091)
externalizing behavior	-0.020 (0.041)	-0.031 (0.046)	-0.040 (0.047)	0.003 (0.090)	-0.008 (0.089)
internalizing behavior (worries and unhappy)	-0.100** (0.043)	-0.134*** (0.049)	-0.137*** (0.049)	-0.027 (0.087)	-0.020 (0.087)
ATL (approaches to learning)	-0.060 (0.041)	-0.059 (0.047)	-0.040 (0.047)	-0.044 (0.088)	-0.053 (0.089)

*** p<0.01, ** p<0.05, *p<0.1; and standard errors are in parentheses.

(3) Removing potential outliers for some variables: for the age variable, exclude children older than 60 months at the time of assessment (there were 204 of them); for the experience variable, exclude children with teachers that claimed to have more than 60 years of experience or those whose teachers' experience was reported to be greater than age; and for group size, exclude group size lower than five. The new OLS results are consistent with the results in Table 5-8, except that the B.A. effect for early reading is not statistically significant at the 10 percent level.

(4) Replacing age 2's social outcome measure with alternative measures: use parent-reported social skills instead of the TAS-45 scale. New results are similar to the previous ones.

(5) Replacing region dummies with state dummies in the OLS model. Most of the coefficients for the state dummies are not statistically significant in all of the outcome models, and the new set of B.A. estimates doesn't change much.

(6) Combining teacher reported non-cognitive outcome measures with parent-reported counterparts. Results indicate that the B.A. estimates for non-cognitive outcomes are similar to those in the original case.

(7) Using ordered probit model as replacement of OLS for one of the outcomes that has an ordered scale, i.e., parent-reported internalizing behavior problems. The new estimate signifies similar signs of the B.A. effect on reducing internalizing behavior.

Next, an additional sensitivity check that is somewhat related to external validity is also conducted. When the sample is restricted to children with non-missing outcome information for all of the eight outcome constructs, instead of allowing for different sample sizes for different outcomes (as in the baseline in Table 5-8), the key findings remain almost the same.[109]

In short, the eight additional analyses reported above indicate that the findings for the OLS model in Table 5-8 are robust. A summary table on the eight robustness analyses is available in Table B7 of the Appendices.[110]

5.5 Other important predictors of child development outcomes

Table 5-12 presents the estimates of other predictors in the child development models, but only for four of the outcomes because of the space limitation. Many of the outcomes are significantly affected by B.A. in the model of OLS with rich controls. Note that these estimates

[109] This way it allows for the maximization of information for each of the outcome variables.
[110] Note also that, since it is related to robust check (6), the B.A. effects for the teacher reported non-cognitive outcomes are available in Table B8 of Appendix B.

are for the covariates, and the main point of adding these covariates is "holding other factors constant" for the key treatment of interest (B.A.).

Table 5-12 OLS Estimates for all predictors of child development outcomes at age 4

Variables	Early reading	Math	Social competence: parent reported	Internalizing behavior: parent reported
B.A. (or above)	0.079** (0.039)	0.142*** (0.036)	-0.094** (0.042)	-0.100** (0.043)
mental ability at age 2 (Bayley)	0.223*** (0.021)	0.240*** (0.020)	0.153*** (0.025)	0.031 (0.024)
social skills at age 2 (TAS-45 HOTSPOT)	0.023 (0.017)	0.038** (0.017)	0.044** (0.020)	0.006 (0.022)
age measured in month	0.069*** (0.005)	0.084*** (0.004)	0.017*** (0.005)	0.015*** (0.005)
boy	-0.101*** (0.034)	-0.042 (0.032)	-0.243*** (0.037)	0.014 (0.041)
African American	0.096* (0.055)	0.062 (0.056)	0.173*** (0.067)	-0.357*** (0.073)
Hispanic	-0.131* (0.069)	-0.159*** (0.060)	0.005 (0.070)	-0.064 (0.086)
Asian	0.373*** (0.083)	0.292*** (0.073)	-0.163* (0.092)	-0.275*** (0.098)
Native American	-0.192*** (0.071)	-0.298*** (0.098)	-0.078 (0.109)	0.154 (0.119)
multiple race	0.028 (0.074)	-0.052 (0.077)	0.008 (0.104)	-0.060 (0.101)
low birth weight	-0.049 (0.042)	-0.178*** (0.048)	-0.061 (0.065)	0.080 (0.060)
days per week in preschool	0.028 (0.022)	0.027 (0.021)	-0.022 (0.025)	-0.002 (0.025)
attended center based programs before	-0.005 (0.047)	0.050 (0.042)	0.027 (0.049)	-0.053 (0.049)
mother's age	0.008** (0.003)	0.010*** (0.003)	-0.007* (0.004)	-0.013*** (0.004)
mother has B.A. or plus	0.225*** (0.047)	0.165*** (0.041)	-0.096** (0.047)	0.280*** (0.051)

mother is depressed	-0.033 (0.025)	-0.043* (0.026)	-0.127*** (0.035)	0.201*** (0.036)
mother is married	0.052 (0.049)	0.045 (0.045)	-0.007 (0.053)	0.102* (0.059)
family income	0.055*** (0.008)	0.055*** (0.008)	0.009 (0.009)	0.002 (0.010)
number of siblings	-0.134*** (0.015)	-0.097*** (0.016)	-0.102*** (0.020)	-0.026 (0.020)
number of books at home (average for wave 2 and wave 3)	0.0008** (0.000)	0.0007** (0.000)	0.001*** (0.000)	-0.0001 (0.000)
home language is not English	-0.093 (0.062)	0.077 (0.059)	0.031 (0.069)	0.290*** (0.081)
quality of parenting	0.076*** (0.020)	0.075*** (0.020)	0.173*** (0.026)	-0.036 (0.025)
teacher is male	-0.172 (0.160)	-0.265 (0.161)	-0.137 (0.184)	0.021 (0.164)
teacher has another race	0.055 (0.045)	0.034 (0.043)	0.023 (0.052)	0.005 (0.059)
experience	-0.003* (0.002)	0.0006 (0.002)	-0.003 (0.002)	-0.002 (0.002)
In service ECE training	-0.035 (0.050)	0.008 (0.043)	9.44E-05 (0.050)	0.038 (0.054)
CDA	0.094** (0.044)	0.054 (0.036)	-0.017 (0.043)	-0.082* (0.046)
other certificates	0.027 (0.038)	-0.054 (0.036)	-0.124*** (0.043)	-0.012 (0.046)
history with child(month)	-0.003 (0.002)	0.003 (0.002)	0.001 (0.003)	-0.003 (0.003)
hours of care for the child	-0.0003 (0.002)	-0.002 (0.002)	0.002 (0.002)	0.001 (0.002)
group size	-0.0004 (0.004)	-0.003 (0.004)	0.0004 (0.004)	0.008* (0.005)
child-adult ratio	0.013* (0.007)	0.013** (0.006)	0.012* (0.007)	-0.013* (0.007)
number of books in the classroom	0.0002** (0.000)	5.45E-05 (0.000)	0.0002* (0.000)	0.0002* (0.000)
number of interest areas in classroom	-0.018* (0.011)	-0.009 (0.010)	-0.003 (0.012)	-0.008 (0.013)

age range of the classroom larger than 2 years or more	-0.003 (0.039)	0.004 (0.036)	-0.050 (0.044)	0.069 (0.047)
Head Start center	-0.054 (0.063)	-0.143** (0.060)	-0.051 (0.077)	-0.109 (0.077)
state preK	-0.089 (0.065)	-0.132** (0.061)	-0.184** (0.077)	0.011 (0.080)
partially public	-0.044 (0.056)	-0.106** (0.050)	-0.037 (0.060)	-0.122** (0.062)
for profit	-0.015 (0.053)	-0.038 (0.049)	-0.035 (0.060)	-0.069 (0.063)
preschool licensed	0.074* (0.043)	0.097** (0.040)	-0.044 (0.050)	-0.018 (0.053)
full day	-0.001 (0.044)	0.016 (0.042)	0.0006 (0.049)	-0.057 (0.055)
home-preschool connection	0.021 (0.014)	0.024* (0.013)	0.056*** (0.016)	-0.003 (0.017)
Northeast	0.038 (0.056)	0.0423 (0.049)	0.033 (0.056)	0.118* (0.062)
Midwest	-0.035 (0.045)	-0.013 (0.044)	-0.235*** (0.055)	0.248*** (0.057)
West	-0.035 (0.047)	0.028 (0.045)	-0.087 (0.056)	0.088 (0.059)
constant	-4.342*** (0.297)	-5.353*** (0.257)	-0.162 (0.308)	-0.713** (0.322)
N (rounded)	4150	4100	4250	4300
R-squared	0.360	0.412	0.171	0.092

Note. Estimates for the dummy flags are omitted in the table due to space limit, but available upon request. Effect sizes for the other four outcome constructs are also omitted given the space limit.

Several findings can be highlighted. First, prior mental ability at age 2 and age differences in month are significant factors for most of the outcomes, and prior social skills are important for math skills and social competence. These findings confirm the nature of continuity of young children's development. Second, there have already been some racial/ethnic differences even at such a young age. For instance, compared to Whites, Asian children exhibit higher scores in early reading and math, but lower scores in parent reported social competence. Also, Hispanic and Native American children exhibit lower early reading and math skills. African American shows somewhat higher early reading scores and social competence. Third, low birth weight has a detrimental association only with early math skills. Fourth, many family predictors have been found to be important. For instance, mother's education exhibits a positive association with early reading and math, and with the reduction of internalizing behavior, but it is negatively associated with social competence. Degree of maternal depression is also an important predictor for many outcomes, with an expected sign. Moreover, family income matters for early reading and math; and the quality of parenting is statistically significantly associated with three of the four child outcomes presented here (except for internalizing behavior problems). Fifth, having a teacher with a Child Development Associate (CDA) qualification is associated with higher early reading skills. Sixth, the number of books in the classroom shows the expected positive sign.[111] A somewhat puzzling finding is the unexpected positive sign for child-adult ratio, given that child outcomes are as expected and demonstrably better the fewer children there are per adult (Vandell & Wolfe, 2000; NICHD). However, the effect is not statically significant. This insignificant finding is consistent with Love et al. (1992) and Scarr et al. (1994). Similarly, the insignificant

[111] According to Howes et al. (2008), examples of appropriate materials in preschool classroom include manipulatives, books, blocks, and dramatic play props. This emphasis on providing appropriate materials for children to use in exploring and mastering concepts is historically heavily influenced by Piagetian notions of learning, which is further reflected in the DAP indicators and more generally in the field of early childhood education.

effect of the full-day status is different from the findings in studies like Elicker and Mathur (1997), and Meisels and Atkins-Burnett (2006). Seventh, preschool features matter for some of the outcomes. For example, children from Head Start centers, state preKs and partially publicly funded preschools exhibit lower math skills than those from exclusively private preschools. In comparison with the exclusively private preschools, state prekindergarten children also have lower social competence, whereas children in partially publicly funded private preschools are reported by parents to have fewer internalizing behavior problems. Also important is that children in licensed preschools tend to have higher early reading and math sores. Eighth, the frequency of home-school connection is associated with children's higher math and social competence. This is consistent with other research (Christopher & Blackman-Jones, 2006). Ninth, more statistically significant predictors are found for the math outcome, and the *R* squared of its model is highest among the models of all outcomes; in other words, math skills seem to be more malleable than other child development outcomes measured in this study.

Table 5-13 presents the effect size of various predictors of child development outcomes, based on the OLS model. For math outcome,[112] the most important predictors in terms of effect size are age measured in months (0.344), mental ability at age 2 (0.240), family income (0.185), number of sibling (-0.107), mother with a B.A. or above (0.084), the quality of parenting (0.075), and teachers with a B.A. or above (0.074). Compared to other predictors, B.A.'s effect size is above the medium level, and it is larger than any other teacher level, classroom level or preschool level predictors. For early reading, the relative ranking of B.A.'s effect size is similar. The exception is that CDA, another teacher feature, shows a positive effect size similar to B.A.'s.

[112] The reason for the focus on math is that statistically significant B.A. effect was found for math more often than other child development outcomes. Also, there are more statistically significant predictors in the math model than in the model for other outcomes.

Table 5-13 Effect size for predictors of child development outcomes at age 4 in the OLS model

Variables	early reading	math	social competence: parent reported	internalizing behavior: parent reported
B.A. (or above)	0.041**	0.074***	-0.048**	-0.050**
mental ability at age 2 (Bayley)	[0.223]***	[0.240]***	[0.153]***	[0.031]
social skills at age 2 (TAS-45 HOTSPOT)	[0.023]	[0.038]**	[0.044]**	[0.006]
age measured in month	0.280***	0.344***	0.067***	0.059***
boy	-0.053***	-0.022	-0.126***	0.007
African American	0.035*	0.023	0.063***	-0.126***
Hispanic	-0.054*	-0.066***	0.002	-0.026
Asian	0.060***	0.048***	-0.026*	-0.042***
Native American	-0.015***	-0.023***	-0.006	0.011
multiple race	0.005	-0.010	0.002	-0.011
low birth weight	-0.005	-0.020***	-0.007	0.009
days per week in preschool	0.030	0.029	-0.024	-0.002
attended center-based programs before	-0.002	0.022	0.011	-0.022
mother's age	[0.051]***	[0.064]***	[-0.045]*	[-0.084]***
mother has B.A. or plus	0.111***	0.082***	-0.047**	0.133***
mother is depressed	[-0.022]	[-0.029]*	[-0.086]***	[0.135]***
mother is married	[0.024]	[0.020]	[-0.003]	[0.046]*
family income	0.182***	0.185***	0.029	0.006
number of siblings	-0.147***	-0.107***	-0.112***	-0.027
number of books at home (average for wave 2 and wave 3)	[0.050]**	[0.041]**	[0.079]***	[-0.007]
home language is not English	-0.034	0.029	0.012	0.107***
quality of parenting	[0.076]***	[0.075]***	[0.173]***	[-0.036]
teacher is male	-0.020	-0.031	-0.016	0.002
teacher has another race	[0.028]	[0.017]	[0.011]	[0.003]
experience	[-0.029]*	[0.005]	[-0.027]	[-0.016]
In service ECE training	-0.014	0.003	0.000	0.015
CDA	0.045**	0.026	-0.008	-0.038*
other certificates	0.013	-0.026	-0.060***	-0.006
history with the child(month)	[-0.022]	[0.024]	[0.008]	[-0.025]
hours of care for the child	[-0.004]	[-0.031]	[0.032]	[0.013]
group size	[-0.002]	[-0.018]	[0.002]	[0.046]*

child-adult ratio	[0.044]	[0.046]**	[0.042]*	[-0.046]*
number of books in the classroom	[0.032]**	[0.021]	[0.031]*	[0.030]*
number of interest areas in the classroom	-0.032*	0.009	-0.006	-0.014
age range of the classroom larger than 2 years or more	[-0.001]	[-0.016]	[-0.023]	[0.032]
Head Start center	-0.023	-0.061**	-0.022	-0.045
statepreK	-0.033	-0.049**	-0.068**	0.004
partially public	-0.020	-0.049**	-0.017	-0.053**
for profit	-0.007	-0.017	-0.016	-0.030
preschool licensed	0.033*	0.043**	-0.019	-0.008
full day	-0.001	0.008	0.000	-0.027
home-preschool connection	0.027	0.031*	0.071***	-0.004
Northeast	0.016	0.017	0.013	0.047*
Midwest	-0.015	-0.006	-0.099***	0.101***
West	-0.015	0.012	-0.037	0.036

Note. The effect size is traditionally defined, referring to the degree of change of standard deviation in outcome given a one standard deviation change in the predictor. However, the effect sizes presented in brackets are the effect size recalculated manually for the predictors with missing values (e.g., mental ability at age 2), by using their raw standard deviation (not including the standard deviation for the transformed variable with imputed values), so that the effect sizes of these variables with missing values are comparable to the effect sizes of other predictors without missing values. Also, as mentioned in the previous text (the first page of this chapter), when thinking about effect size for a binary predictor like B.A., a better interpretation of definition is the change in outcome's standard deviation given a 0 to 1 change in the predictor, rather than for a one standard deviation change in the predictor.

5.6 Extended analysis: Heterogeneous effects of B.A. and the contribution of specialized education in ECE on the B.A. effects

This section presents findings for RQ1.2-1.4. For those four child development outcomes affected by B.A. in the model of OLS with rich controls (i.e., early reading, math, social competence and internalizing behavior problems), RQ1.2 asks whether the B.A. effect differs by preschool type; RQ1.3 examines whether the B.A. effect is larger for children from low socioeconomic status (SES) families. RQ1.4 asks if specialized college education in ECE increases B.A.'s effect on child development outcomes. Results are shown in Table 5-14.

First, for RQ1.2 and RQ1.3, according to Table 5-14, addition of interaction terms to the model of OLS with rich controls shows little evidence of differential effects by preschool type, and the B.A. effects are no larger for children from families of low SES. The only exception is that children with B.A. teachers were not shown to have fewer internalizing behavior problems than children with non-B.A. teachers in Head Start center and state preKs; in other words, the effect only exists for the private settings. Other child or family related heterogeneity factors also considered include gender, children's initial mental ability at age 2 (as done so in NICHD & Duncan, 2003[113]), non-B.A. mothers, and poverty. However, I did not find statistically significant heterogeneous effects except for gender. There is a statistically significant interaction between being a boy and having a B.A. teacher on the child's math skills. Compared with a girl, a typical boy in this sample benefits more from having a preschool teacher with a bachelor's degree for his math skills. Similar inference can be given to the negative effect on social competence in that boys may have been the major recipients of the negative effects of B.A. on social competence, although the interaction term is not statistically significant. Differences in the B.A. effect by race were not found. Results for these additional heterogeneity tests are available in Table B9 in the Appendices.

Second, for RQ1.4, specialized college education in ECE, as measured by whether the major of college degree is in ECE or a related field and the amount of college coursework in ECE (not including a related field), doesn't show a statistically significant interaction effect with the B.A. effect. This can be seen in the last column of Table 5-14. As a sensitivity check, I further broaden the number of college course to include ECE related field such as special education or psychology; however, conclusions regarding specialized training's role for B.A. are almost unchanged. Apart from adding interaction terms, another way of deriving heterogeneous effects

[113] "Children with low and normal early cognitive scores" in their Table 4.

for specialized college education in ECE is to restrict the sample to children with B.A. teachers only, and to directly compare children whose teacher has a B.A. and a college major in ECE with children whose teacher has a B.A. without a college major in ECE. Similar procedure was done in Early et al. (2007). However, in this case, among the group of children whose teachers hold a B.A., ECE major does not exhibit positive and statistically significant teacher effects on the child development outcomes in my study.

Also, for the readers' interest, I classify the children's teachers into eight categories: (1) B.A. with a degree in ECE; (2) B.A. without a degree in ECE; (3) A.A. in ECE; (4) A.A. in non-ECE; (5) some college in ECE; (6) some college in non-ECE; (7) high school with CDA or other certificate; (8) high school without anything. This is in some way similar to Howes et al.'s (1992) analysis. No effect was found for this new analysis.

In addition, I tested the effect of an interaction of B.A. with the number of books in the classroom and an interaction of B.A. with whether the length of the preschool day is full-day or not, inspired by the model for child development (Figure 3-1 of Chapter 3) as well as by Whitebook and Ryan (2011). However, no statistically significant interaction effects were found.

Table 5-14 Heterogeneous effects of B.A. and B.A.'s interaction with specialized college education in ECE: Results for RQ1.2-1.4

Outcome: age 4	N (rounded)	RQ1.2				RQ1.3		RQ1.4			
		preschool type (control group: exclusively private)				low SES		specialized education in ECE			
								measure 1		measure 2	
		B.A.	Head Start*B.A.	State preK*B.A.	partially public*B.A.	B.A.	low SES*B.A.	B.A.	ECE major*B.A.	B.A.	B.A.*# of courses in ECE
Early reading	4100	0.086 (0.060)	-0.099 (0.090)	0.0158 (0.149)	0.055 (0.091)	0.092** (0.044)	-0.063 (0.068)	0.1056 (0.062)	-0.018 (0.078)	0.086** (0.043)	0.0003 (0.0165)
Math	4100	0.186*** (0.053)	-0.088 (0.089)	0.073 (0.145)	-0.113 (0.081)	0.150*** (0.039)	-0.038 (0.076)	0.172*** (0.056)	0.007 (0.072)	0.164*** (0.040)	0.009 (0.013)
Social competence: parent reported	4250	-0.077 (0.060)	0.083 (0.104)	-0.054 (0.155)	-0.126 (0.092)	-0.104** (0.045)	0.049 (0.098)	-0.099 (0.064)	-0.032 (0.089)	-0.130*** (0.047)	0.024 (0.018)
Internalizing behavior: parent reported	4300	-0.213*** (0.063)	0.221** (0.109)	0.359** (0.172)	0.134 (0.100)	-0.131*** (0.048)	0.152 (0.100)	-0.159** (0.066)	0.058 (0.090)	-0.105** (0.049)	-0.006 (0.020)

*** p<0.01, ** p<0.05, *p<0.1; and standard errors are in parentheses.

5.7 Longer-term effect of B.A.: Age 5 findings

This section presents the effects of B.A. on age 5 child development outcomes, shown in Table 5-15. The methods are the same as those used for age 4 outcomes, with the only difference being that an additional covariate is added to the age 5 models. This additional variable is a dummy indicative of whether the child attended kindergarten in wave 4 (2006), given that 75% attended kindergarten whereas the rest of them did not. The age 5 results are not encouraging: B.A.'s once statistically significant effects on early reading and math at age 4 disappear at this time; fewer outcomes have been affected by the B.A. in a statistically significant way. Exceptions are: (1) in both the PSM and IV models, preschool teachers' B.A. status reduces children's externalizing behavior at age 5; (2) B.A. was also found to improve children's approaches to learning skills at age 5 in PSM.

Several interesting findings are worth noting in comparing the results of age 4 with the results for age 5: (1) in the models of OLS with rich controls, all the previously statistically significant effects for age 4 disappear for age 5; (2) in the PSM models for age 5, the effects on the reduction of children's externalizing behavior and learning skills are new: they are not there for the same outcomes at age 4; and (3) in the IV models, the B.A. effect for reducing externalizing behavior at age 5 is statistically significant, fairly consistent with the age 4 finding. However, preschool teachers' B.A. status negatively affects children's early reading skills at age 5 whereas the effect is not statistically significant in the IV model for age 4.

Table 5-15 Age 5 B.A. estimates in three models (OLS, PSM, and IV):
B.A. (B.A. or above) vs. non-B.A. (A.A. or below)

Age 5 outcomes	N (rounded)	OLS with rich controls	PSM (ATT)	IV		
		coefficient for B.A.	coefficient for B.A.	coefficient for B.A.	Endogeneity test: p value	F value: 1st stage
(1)early reading	3400	0.063 (0.042)	0.010 (0.079)	-0.848** (0.335)	NA	39.305
(2) expressive language	3300	-0.041 (0.048)	-0.024 (0.081)	0.098 (0.323)	NA	39.190
(3) math	3400	0.064 (0.039)	0.002 (0.078)	-0.473 (0.300)	NA	39.557
(4)social competence: parent reported	3400	-0.037 (0.053)	-0.115 (0.080)	0.226 (0.345)	0.453	37.755
(5)externalizing behavior problems	3350	0.035 (0.049)	-0.168* (0.089)	-0.907** (0.360)	0.006	37.581
(6)internalizing behavior(only "worry")	3450	-0.009 (0.051)	0.106 (0.079)	0.050 (0.352)	0.888	39.129
(7)ATL (approaches to learning)	3450	-0.007 (0.050)	0.175* (0.094)	0.107 (0.341)	0.101	38.392

Note. (a) The "color recognition" outcome is not available for age 5. (b) Covariates for the age 5 models are almost the same as in the age 4 models, except the addition of the dummy indicator of whether the child attended Kindergarten at age 5 (wave 4). (c) "NA" means that the Endogeneity test was not feasible in that context and was not reported by STATA.

*** p<0.01 ** p<0.05 *p<0.1

To summarize, the findings for age-4 and age-5 children indicate that the effect of having a preschool teacher with a bachelor's degree on early reading and math fades out over time. It makes sense because "the effects of particular teachers on students' skills are likely to become muted over subsequent years, as the students experience other teachers" (Murnane & Willett, 2011, p. 342). This statement regarding K-12 education applies to the preK and K grades in this study. In contrast, the somewhat robust effects of preschool teachers' B.A. status on reducing parent-reported externalizing behavior problems at age 5 is positive evidence for the longer-term B.A. effect.

5.8 Summary and discussion

According to this chapter's findings in relation to the first research question of the dissertation, having a B.A. teacher in preschool exhibits some statistically significant effect on some measures of children's cognitive and non-cognitive skills at age 4 in one of the three estimation methods (OLS with rich controls, PSM and IV), although the findings are not very consistent across models. Specifically, in the OLS model, four of the eight comprehensive outcome constructs are affected by teachers' B.A. status. Children with B.A. teachers (B.A. or above) exhibit higher early reading and math skills than children with non-B.A. teachers (A.A. or below) and fewer parent-reported internalizing behavior problems. No effects are found for story-telling skills, color recognition, parent-reported externalizing behavior problems or approaches to learning skills. Yet the children with B.A. teachers are found to have lower parent-rated social competence. In the PSM model, B.A. shows positive effects on math but negative effects on social competence. In the IV models, a statistically significant B.A. effect is found only for reducing parent-reported externalizing behavior; no significant B.A. effect is present for

the other seven outcomes. This difference may be attributed to either bias or heterogeneity, given that IV estimates remove some endogeneity of the treatment but is only local to compliers. When the results of the three estimation methods are taken together, one may conclude that the empirical evidence of a significant effect of B.A. on cognitive and non-cognitive child development outcomes is rather weak.

Moreover, in the comparison between children with B.A. teachers and children with A.A. teachers, there are fewer statistically significant effects on child development outcomes; B.A. is found to increase math skills and to reduce internalizing behavior only. In addition, there is little evidence of heterogeneous effects for those outcomes; that is, the B.A. effect does not vary by preschool type, by children's family SES, or by ECE major. Most of the B.A. effects at age 4 disappear at age 5 when 75% of the children attend kindergartens, or in other words, enroll in schools.

Putting these findings in the context of prior studies helps researchers and policy-makers understand the findings of this dissertation study. Also helpful is the discussion about the potential reasons for the statistically significant effects in certain models.

Let us first take a closer look at RQ 1.1 and outcomes during the preschool year (age 4), and an outcome-by-outcome summative statement of the findings: having a B.A. teacher is found to produce some positive effects on children's math skills (OLS and PSM), and to some degree on early reading skills (OLS), reduction of internalizing behavior problems (PSM) or externalizing behavior problems (IV); yet having a B.A. teacher is also associated with lower social competence according to the OLS and PSM models. For the ease of comparison, I have classified several important prior studies according to the type of data used ("child follow-up" or

"site/preschool based"), as presented in Table 5-16. Among them, the NICHD ECCRN (National Institute of Child Health and Human Development Early Child Care Research Network) data[114] and the Early Head Start (EHS) follow-up data are especially comparable for two reasons: (1) these two datasets are also of the child follow-up type of data; and (2) the percentages of B.A. teachers in these two samples are very close to ECLS-B (56% for ECLS-B, about 58% for NICHD, and 55% for the EHS follow-up). Another NICHD study, NICHD and Duncan (2003), was also selected because of its methodological rigor despite its focus on the overall effect of "child care quality" and its use of "years of education" rather than the B.A. dummy as the teacher education measure.[115]

Three major points can be made regarding the comparison summary in Table 5-16. First, compared with studies based on the other two child follow-up datasets (i.e., NICHD and Early Head Start follow-up) that showed no effects on receptive language, pre-reading and math skills,[116] my findings deliver somewhat stronger evidence of statistically significant effects in at least one or two of the three estimation methods, as shown in the first part of Table 5-16. Second, as regards the comparison of my findings with the most recent and important B.A. study based on seven large datasets, i.e., Early et al. (2007), two different observations can be made. On the one hand, my B.A. estimates are consistently significant in at least two models for two types of outcomes (math skills and social competence). This is somewhat different from the general conclusion of insignificant B.A. effects on academic skills (receptive vocabulary, pre-reading

[114] This is a longitudinal follow-up of 1300 children, starting in 1991 at one month of age, with follow-ups at four phases (birth through three years of age, 54 months through first grade, second through sixth grades, and seventh through ninth grades).

[115] Specifically, in this study, three methods are used, including "level model" (multiple regression models of 54-month child outcomes, with prior outcome as a control), "simple change model" (using the differences between 24-54 child outcomes as the outcome), and "residualized change model" (allowing for early inputs and parent and child characteristics in the change equation). Compared to NICHD and Duncan (2003), my model of OLS with rich controls is like the level model in NICHD and Duncan (2003); yet I cannot apply the change models because they require common measures of child outcomes in different waves.

[116] Early et al.'s (2007) reanalysis of NICHD even found negative effect on math.

skills and early math skills) in Early et al. (2007),[117] where no obvious pattern was found.[118] On the other hand, my finding of the overall "insignificance" of B.A. effects for early reading in PSM and IV and for expressive language skills (story-telling skills) across three methods is consistent with Early et al. (2007)'s overall finding for pre-reading scores. Third, compared with Kelley and Camilli's (2007) meta-analysis of 32 studies that returned an effect size of B.A. (relative to high school, some college, or A.A.) of about 0.14-0.50 for cognitive outcomes and 0.03-0.17 for social outcomes,[119] my ECLS-B findings on the B.A. effects are smaller in magnitude and less consistent. Apart from differences in data type and research design, this might be related to the trend whereby more teachers now have a B.A. than in the older studies.

[117] Note that the B.A. dummy is one of the two measures used in this study for the effect of degree; the other measure is "the highest degree" with four categories.

[118] In more detail, the picture is that "none of the seven studies found an association between the highest degree and receptive language skills, and only a few studies reported associations with reading or math" (p. 572).

[119] Overall, if we also count in the final set of outcomes the intermediate outcomes (e.g., teacher belief, interactions, and instruction activities), in addition to the final child development outcomes, the effect size is about 0.16, according to their multilevel analysis (Table 7).

Table 5-16 Comparison of the main findings of RQ1.1 with important prior studies of the B.A. effect for concurrent outcomes: By type of data

Important datasets/studies	Measures of teacher education	Sample	Findings	Implications for my study: similarity or difference
1.Child-follow up data:				
NICHD (National Institute of Child Health and Human Development) in Early et al. (2007)	B.A. vs. non-B.A: 58% have B.A. (its Table 2)	639 children (from 639 classrooms)at the 10 locations in 9 states[120]who were observed to be in center based care (all types) the year before kindergarten (1995-1996)	Using value-added specification of OLS, no association was found between whether the teacher has a B.A. (or teacher's highest degree) and children's receptive language and pre-reading skills; children whose teacher has a B.A. scored lower on early math scores (the Woodcock-Johnson Applied Problems test) than children whose teacher does not have a bachelor's degree (d=-0.18) (its Table 6)	The NICHD findings in Early et al. (2007) are different from findings from my OLS with rich controls, where I found positive effects on early reading and math. This may indicate that the new generation of nationwide preschoolers in 2005 (about 10 years later than the NICHD preschoolers) benefit cognitively from having a B.A. teacher.
NICHD in NICHD and Duncan (2003), also reexamined by Kelley and Camilli (2007)	Years of education: 13.35, 13.34, 13.5, 13.8, and 15.1 respectively for 1 month, 6 months, 15 months, 24 months, 36 months and 54 months	Same as above, with outcomes measured at 15, 24, 36, and 54 months, for children born in 1991about 1000-1100 children in the operational samples	With three statistical methods, found positive correlation between caregivers' years of education and 54-month cognitive outcomes (math, reading, phonological knowledge); also true for the other two quality indicators in the study (staff/child ratio and group size). This finding is relatively consistent in the change models but not in the level models. The correlation coefficient is 0.12 in Kelley and Camilli's (2007) analysis.	

[120] These sites are: Little Rock, Arkansas; Irvine, California; Lawrence, Kansas; Boston, Massachusetts; Philadelphia, Pennsylvania; Pittsburgh, Pennsylvania; Charlottesville, Virginia; Morganton, North Carolina; Seattle, Washington; and Madison, Wisconsin.

Early Head Start (EHS) follow-up in Early et al. (2007)	B.A. vs. non-B.A.: 55% have B.A. (its Table 2)	887 children who had been in EHS Evaluation Study as infants or toddlers but later observed in the preschool year (2001-2003, three cohorts)	With the same method described for the NICHD dataset in Early et al. (2007), no association between receptive language, pre-reading, and early math skills and whether the teacher had a bachelor's degree (or teacher's highest degree)	
2.Site (preschool/classroom) based data:				
NCEDL data that combines the Multi-State Study of Pre-Kindergarten and Study of State-Wide Early Education Programs (SWEEP), in Early et al. (2007)	B.A. vs. non-B.A.: 71%	2966 children in 721 state preK classrooms from 11 participating states	Children whose teacher had a bachelor's degree or higher had slightly higher pre-reading scores (d=0.09) and math scores (d=0.07); no statistical difference for receptive language	Small effect size or not statistically significant for cognitive outcomes, similar to mine to some degree
NCEDL's Multi-State Study of Pre-Kindergarten and SWEEP in Howes et al. (2008)	B.A. and early childhood certification vs. others[121]: on average, 60.2% held both a B.A. and a specialized early childhood certification; by dataset, 48 % for Multi-State Study and 62% for SWEEP	2800 children randomly selected from approximately 700 randomly selected state-funded pre-K classrooms in 11 states (4 per classroom)	Using HLM, not related to cognitive skills (pre-academic achievement); only marginally significant for social competence (its Table 5)	
Data for the Observational Study of Early Childhood Programs (OSECP) by Layzer et al. (1993), reexamined by Kelley and Camilli(2007) for the B.A. effect	B.A. vs. high school diploma: "55% of lead teaches with a college degrees or higher" percentage of B.A. should be lower	Teachers for 4-year-olds in 119 classrooms from 5 sites in 1989: both lead and assistant teachers lead teacher as the focus	By multiple regression, the effect size is 0.29 for cognitive outcomes (e.g., task engagement) and 0.00 for social (e.g., use of higher level social strategies with other children) when compared to high school diploma	Fairly large effect size for the cognitive skills, very impressive, but not so comparable with my data structure given its comparatively narrow sites

[121] Note also that 11.1% of all teachers possessed a B.A. but without early childhood specialized training; however, they do not seem to be in the treatment grou according to the authors' description in the article.

Small sample study by Vandell and Powers (1983), reexamined by Kelley and Camilli (2007) for the B.A. effect	B.A. vs. high school diploma	55 children (3- to 4-year-olds) from six different day care centers	Through ANOVA, effect size is 0.18 for social outcomes (positive & negative behavior with adults) when B.A. is compared with high school diploma	Reduction of externalizing behavior and internalizing behavior problems

Note. These studies were selected by the author from the 12 studies using the B.A. dummy measure listed in the Literature Review chapter (Chapter 2); note that an NICHD study (NICHD & Duncan, 2003) that uses the years of education measure is also presented here given the importance of that dataset and its similarity to ECLS-B.

Additionally, there are two probable explanations for the B.A. effects identified in some of the models. The first is the "substance" explanation, related to the quality of the teacher preparation programs that grant the bachelor's degree. Preschool teachers who received formal education in a B.A. program may build up more human capital than the non-B.A. teachers, in terms of both general knowledge and specific knowledge in early childhood education/child development. Because of B.A. training, B.A. teachers may be more likely to know how to implement developmentally appropriate practice (DAP) in promoting children's learning so that the math skills and earning reading skills of their 4-year-olds in care are better, and they may be more sensitive to children's emotional needs so that some of the internalizing behavior problems and externalizing behavior problems are reduced. The second explanation is that a B.A. teacher may have higher academic ability or learning skills than a non-B.A. teacher, so that he/she can better apply the most recent research and training into the early childhood practice. In the hiring stage and from the perspective of preschool directors, this is like the signaling effect of a college degree on workplace productivity in the eyes of an employer. In this regard, what a B.A. signifies can also encompass positive motivation and commitment to the early childhood field, which is quite difficult to control for even in the three rigorous quantitative methods of this dissertation.

As regards the unexpected negative effects on social competence, there are two possible reasons for this finding. [122] One is that the B.A. teachers may spend less time on social activities (and more time on learning activities) than non-B.A. teachers, which is to some degree confirmed in Chapter 6. A preview of the findings in chapter 6 indicates that B.A. is associated

[122] It is unexpected because some people think specialized knowledge and ability of teachers matter, and because some studies did find positive effects (see Kelley and Camilli's Table 4). Also, Howes et al. (2008), who used a comparison of B.A. and certification with the others, reported marginally significant effects for social competence, even when there was no effect for pre-academic achievement.

with higher frequency of reading and story-telling activities, and there is no difference in the frequency of playing games or building something between B.A. and non-B.A. teachers. Another is that parents of children with B.A. teachers may be stricter when rating their child's social skills than parents of children whose teachers have lower education, and therefore empirically the data show that B.A. teachers are related to lower parent-reported social skills. Note also that the negative effects of teachers' B.A. on social competence are mostly applicable to the boys and for the children whose teachers have above B.A. level of education. This supplemental finding provides some hints of the second reason. Still, the negative effect on social competence is puzzling because findings in Chapter 6 also indicate that B.A. teachers in this child-teacher sample tend to be more sensitive, less harsh, less detached, and less permissive than non-B.A. teachers, and in terms of the higher quality of teacher-child interactions. The B.A. effect on social skills is likely to be positive. Given that neither the NICHD nor the EHS follow-up studies researched the B.A. effect on social emotional outcomes, more research on social competence could be performed on such data in the future.

Note also that the findings in Chapter 6 for RQ2 may help to explain the link between preschool teacher's B.A. status and children's development outcomes, and why there is no effect or no strong effect for certain types of outcomes.

For RQ 1.2 and 1.3, my study provides some answers to the heterogeneous effects by preschool type and by SES, which were not presented or examined in previous studies. The overall answer to this question is negative. The B.A. effect doesn't differ much by preschool type except for Head Start's much smaller effect on internalizing behaviour problems. This is based on a child follow-up dataset of ECLS-B. Such findings of not much differential B.A. effects on cognitive outcomes by preschool type are different from Early et al. (2007). According to

specific breakdown analysis of the Early et al. study by Bryant et al. (2010, p.49), B.A. modestly predicted better cognitive outcomes in community child care, but was not related to either observed quality or cognitive child outcomes in publicly funded Head Start or prekindergarten programs. For RQ 1.4, my finding of no additional effect by specialized college education in ECE from ECLS-B is consistent with Early et al. (2007), and with Kelley and Camilli (2007), who compared the child outcomes for teachers with and without ECE training separately for B.A.s and non-B.A.s.[123] To a large extent, these findings of no statistically significant heterogeneous effects by preschool type or by children's socio-economic status are unexpected. The reason could be the limited sample size,[124] which means that the sample size may not be sufficient for the detection of heterogeneous effects. This may also be related to family's response to teacher input, with one possibility being that the families of children with B.A. teachers may have made supplemental investment to their children.[125]

For RQ1.4, the scant evidence on ECE major/courses is, however, expected, for two reasons. First, most of the B.A. teachers of the children in this sample have a college major in ECE or a related field, and therefore an ECE major may not exhibit any difference in teacher performance. Second, Early et al. (2007) also found that the effect of ECE major is mostly negligible among those teachers whose highest degree is a B.A. Mainly because of the first reason, an insignificant interaction between B.A. and major in ECE in the ECLS-B sample does not mean that specialized college education (in the pre-service sense) does not matter. The effect may not have

[123] In their study, the estimates for those with ECE training have a somewhat wider range than those without ECE training, but the difference seems not statistically significant.
[124] Although the ECLS-B dataset is a national sample, the sample size of about 4300 children in center-based care across the country is relatively small. National studies in K-12 education often have larger sample sizes.
[125] For this dataset, I didn't find statistically significant difference in the quality of parenting by teacher's B.A. status, but future studies are recommended to test this possibility. Studies like Gelber and Isen (2011) and Todd and Wolpin (2003) are good references for this.

been identified because of limited variation in the “specialized college education” indicator in this dataset.

Chapter 6

Findings and Discussion II: B.A.'s Effects in Two Steps

As shown in the analytical framework presented in Figure 3-3 of Chapter 3, the first potential pathway by which a higher level of teacher education may influence children's developmental outcomes is through positive teacher behaviors with the child. Accordingly, this chapter presents the findings for the dissertation's second research question (RQ2), which explores the relationships between whether a teacher has a B.A., the frequency and quality of teacher-child interactions, and child development outcomes at age 4. I examine these relationships using a two-step OLS regression. The first step examines whether having a B.A. affects teacher-child interactions, based on the measures of teacher-child interactions that are available in the ECLS-B dataset. The second step examines whether teacher-child interactions contribute to child development outcomes. The two steps together provide information on whether the teacher-child interactions mediate the effects of teachers' level of education on the relevant child development outcomes.

Section 6.1 discusses the findings for the first step of the OLS regression—the effect of teachers' level of education on the frequency of classroom activities in several domains and the quality of teacher-child interactions. Section 6.2 discusses the findings for the second step of the OLS regression. Section 6.3 connects the findings for the two steps, organizing the results by child outcome domain and by the significance of the B.A. effects on specific outcomes identified in Chapter 5. Section 6.4 summarizes the findings in this chapter and discusses some of their implications.

6.1 Findings for the first step linking B.A. and teacher-child interactions

Table 6-1 presents the findings for the first step of the OLS regression, showing the effect of whether a teacher has a B.A. on two types of process variables that are available in the ECLS-B dataset. One is the frequency of teacher-reported activities by type, and the other is the quality of teacher-child interactions as measured by the Caregiver Interaction Scale (CIS). The first measure can be viewed as the frequency of teacher-child interactions in specific activities. The second measure defines quality of teacher-child interactions mainly in terms of emotional support (Arnett, 1989).

Using an OLS regression with rich controls, I find that children who have a teacher with a B.A. rather than a teacher without a B.A. experience a greater number of book-reading activities, story-telling activities, and song-singing activities in a week. No differences were found in the frequency with which teachers asked questions about the reading, played games, or built something with children. In addition, the average frequency of the above-mentioned activities is also positively predicted by the B.A. status of children's teachers. Note that the sample sizes for this analysis for these frequency measures of teacher-child interactions are around 4300, and their *R* Squared values range from 0.063-0.252. The same covariates as in the model of OLS with rich controls in Chapter 5 are controlled for here, including children' prior mental ability and social skills at age 2, family income, teacher's gender, certificate, in-service training, etc. As robust checks, some variables are excluded because they are not expected to be directly related to teacher-child interactions, for example, number of books at home; but the estimates and their statistical significance didn't change much.

Table 6-1 Effects of B.A. on the frequency and quality of teacher-child interactions

Measures of teacher-child interactions: Frequency and quality		Coefficient	Effect size
Full sample	Frequency of book-reading (how many times per week teacher read books to child)	0.737*** (0.199)	0.074
	Frequency of teacher telling stories to the class (1-5: from once a month or less to everyday)	0.117* (0.062)	0.042
	Frequency of teacher asking questions about what was read (i.e., the stories) to the children (1-5)	-0.003 (0.036)	-0.002
	Frequency of singing songs with children (how many times per week)	0.836** (0.358)	0.055
	Frequency of playing games or doing puzzles with the child (how many times per week)	0.209 (0.143)	0.030
	Frequency of building something with the child (how many times per week)	-0.119 (0.118)	-0.020
	Average frequency of the above mentioned activities (how many times per week)	0.292*** (0.108)	0.057
Subsample	CIS measure for the quality of interactions: Total score	0.287*** (0.079)	0.146
	Sensitivity subscore	0.335*** (0.086)	0.167
	Harshness subscore: Less harsh	0.157*** (0.078)	0.082
	Detachment subscore: Less detached	0.227*** (0.079)	0.118
	Permissiveness subscore: Less permissive	0.153** (0.075)	0.079

Note. The refined model of OLS removes covariates such as group size and the number of interest areas in the classroom. The subsample was further restricted to the same teacher and same child in focus. Sample size this table has been rounded to the nearest 50 according the requirement by IES (Institute of Education Science) of National Center for Education Statistics (NCES).

*** p<0.01, ** p<0.05, * p<0.1; and standard errors are in parentheses.

In addition, for the subsample of about 1000 children for whom the CIS measure of the quality of teacher-child interactions is available, having a teacher with a B.A. is also linked with a better quality of teacher-child interactions. This is true not only as measured by the total CIS scores but also as measured by the four subscores: Teachers with a B.A. tend to have interactions that are more sensitive, less harsh, less detached and less permissive. Sample size for this analysis is about 1050 for the overall CIS quality measure, with an *R* Squared of 0.229. This model controls the same covariates pointed out in the previous paragraph. Information regarding sample size and effect size of the models of CIS's sub-scores is also shown in Table 6-1.

These findings are consistent with findings from prior empirical studies. First, as mentioned in the literature review in Chapter 2, the impact of teachers having a B.A. on teacher-child interactions is mostly positive. Ten out of the 14 studies on the effects of a B.A. on the quality of teacher-child interactions reported positive findings. The studies used a variety of measures of teacher-child interactions, mostly focusing on the quality of interactions, in particular the dimensions of emotional support or the teacher sensitivity. Among the eight studies that used the CIS measure (Arnett, 1989; Blau, 1997; Burchinal, Cryer, Clifford, & Howes, 2002; Howes, Whitebook, & Phillips, 1992; Howes, 1997; Layzer et al., 1993; Phillipsen, Burchinal, Howes, & Cryer, 1997; Whitebook M., Howes, & Phillips, 1989), most found positive effects. For example, in a seminal study, Arnett (1989) found that the Bermuda College training program, a program specific to early childhood education at the college level, was associated with positive teacher-child interactions as defined by CIS. According to Kelley and Camilli's (2007) calculation of the effect size, the effect size of Arnett (1989) is 1.18 for the comparison between teachers with a B.A. and those with a high school diploma, and 1.46 for the comparison between teachers with a B.A. and those with some college. Using the CIS measure of teacher sensitivity, Howes et al.

(1992) concluded that both formal education and very high levels of specialized training prepare teachers to be effective in the classroom, but that formal education was a better predictor of teacher performance than specialized training. My findings on the effects of teachers' level of education on positive teacher-child interactions are consistent with the above findings.

Fewer studies have used measures related to the frequency of activities performed in teacher-child interactions. However, studies that have used activity-related measures of teacher behavior have mostly shown that higher levels of teacher education have positive effects (Barnett, Lamy, & Frede, 2001;[126] Howes, James, & Ritchie, 2003; McMullen, 2003). For example, Howes et al. (2003) looked at three measures of "effective teaching" through classroom observations: responsive involvement, adult engagement, and learning activities (creative and language arts activities). Among the three, responsive involvement and learning activities are related to the activity measures in this dissertation. Howes et al. measured responsive involvement by the proportion of time the children were within 3 feet of the teacher and the teacher was responsively engaged. The indicator they used for learning activities was the percentage of the observation period that the teachers and children were doing creative activities or language arts activities. Although these measures are not the same as the frequency measures used in my study, they are relevant. According to Kelley and Camilli (2007)'s reanalysis of Howes et al. (2003), the effect size of teachers having a B.A. found in the Howes et al. study is 0.37 for instructional activities (B.A. compared with an associate degree) and 0.30 for interaction. My first-step OLS findings regarding the effects of teachers' level of education on reading and story-telling activities are consistent with the findings from these studies. As mentioned in Chapter 5, these effects are statistically significant on learning activities but not on the more social activities, such as playing

[126] Barnett et al. (2001) did not adopt the B.A. dummy measure of teacher education; instead, they estimated the correlation between level of education and instructional activities.

games; this may partially explain why the B.A. has a negative effect on social competence in the OLS and PSM models.

Overall, the first-step OLS estimation adds one piece of positive evidence regarding the effects of a B.A. on the quality of teacher-child interactions and the frequency of some learning activities. These findings send a positive message that B.A. education affects teacher's professional practice in the preschool classroom. More importantly, this evidence is based on a nationally representative sample of center-attending children.

6.2 Findings for the second step linking teacher-child interactions and child outcomes

The second step of the OLS regression examines the relationship between teacher-child interactions and child development outcomes. As shown in Table 6-2, most of the teacher activity measures do not exhibit significant impacts on child development outcomes, with a few exceptions. The majority of the frequency measures for specific activities are not associated with children's cognitive skills, social-emotional development, or approaches to learning skills; but the signs are mostly positive. Notably, the frequency of story-telling is positively associated with math skills (at the 10% level) and parent-reported social competence. Frequency of playing games is positively associated (at the 10% level) with children's early reading scores. The games activity includes guessing games. If teachers who have a B.A. play more guessing games with children than do teachers without a B.A., this activity might be mediating the positive effect on early reading. But it is not statistically significant here. Finally, frequency of building is negatively associated (at the 10% level) with parent-reported approaches to learning skills, of which attention skills is a component.

Table 6-2 also shows that the quality of teacher-child interactions, as measured by CIS score, is associated with a reduction in externalizing and internalizing behavior problems, whereas the effects are not statistically significant on other child development outcomes. Specifically, as shown in the last four rows, teachers' being less harsh and being less permissive are negatively associated with parent-reported externalizing behavior problems. Also, parents reported fewer internalizing behavior problems for children whose teachers were more sensitive and less harsh.

Table 6-2 Effects of teacher-child interactions on child development outcomes

Teacher behavior measures		Early reading	Expressive language	Math	Color recognition	Social competence	Externalizing behavior	Internalizing behavior	ATL
Full sample	book-reading	0.003 (0.004)	0.002 (0.004)	0.002 (0.003)	0.003 (0.004)	-0.0005 (0.004)	0.005 (0.004)	0.0004 (0.005)	-0.005 (0.004)
	Story-telling to the class	0.008 (0.012)	0.012 (0.014)	0.030* (0.012)	0.002 (0.015)	0.044*** (0.015)	-0.015 (0.013)	0.003 (0.015)	0.018 (0.014)
	asking questions (about reading)	0.026 (0.022)	0.006 (0.025)	0.014 (0.021)	0.034 (0.025)	-0.004 (0.024)	0.002 (0.024)	-0.005 (0.027)	0.010 (0.024)
	singing songs	0.0009 (0.002)	-0.0003 (0.003)	-0.003 (0.002)	-0.0007 (0.003)	0.003 (0.002)	0.002 (0.003)	-0.002 (0.003)	0.0002 (0.002)
	playing games or doing puzzle with the child	0.011* (0.006)	-0.002 (0.007)	0.002 (0.005)	0.008 (0.006)	0.005 (0.006)	0.004 (0.006)	-0.008 (0.006)	-0.001 (0.007)
	building something with the child	0.001 (0.007)	-0.006 (0.006)	0.0006 (0.006)	-0.0004 (0.007)	-0.010 (0.008)	0.009 (0.007)	-0.008 (0.007)	-0.013* (0.007)
Subsample, and restricted to the same teacher and same child in focus	CIS quality of interactions: total score	0.002 (0.033)	0.025 (0.038)	0.005 (0.034)	-0.021 (0.043)	-0.002 (0.036)	-0.060 (0.042)	-0.103** (0.046)	-0.037 (0.040)
	sensitivity subscore	0.009 (0.032)	0.030 (0.039)	-0.003 (0.034)	-0.027 (0.039)	0.018 (0.037)	-0.016 (0.042)	-0.074* (0.043)	-0.016 (0.040)
	harshness subscore: less harsh	-0.018 (0.033)	0.010 (0.034)	0.016 (0.031)	-0.012 (0.048)	-0.029 (0.032)	-0.134*** (0.039)	-0.128*** (0.048)	-0.034 (0.039)
	detachment subscore: less detached	0.007 (0.035)	-0.004 (0.041)	-0.004 (0.038)	-0.011 (0.042)	0.040 (0.035)	0.017 (0.042)	-0.060 (0.048)	-0.064* (0.037)
	permissive subscore: less permissive	0.015 (0.032)	0.037 (0.037)	0.011 (0.031)	-0.004 (0.047)	-0.064* (0.035)	-0.070* (0.042)	-0.067 (0.048)	-0.052 (0.039)

Note. ATL stands for approaches to learning. The non-cognitive outcomes are parent reported.
*** p<0.01, ** p<0.05, * p<0.1; and standard errors are in parentheses.

Let's compare the findings in Table 6-2 with those in the literature. The findings reported above, which show few significant associations, are a bit surprising because previous studies have found that teacher-child interactions predict children's cognitive and social development. The comparison involves two parts (i.e., the frequency measure and the quality measure of teacher-child interactions), given that the relationships between interactions and child outcomes are domain-specific (Hamre, Hatfield, Pianta, & Jamil, 2013; Bryant, Zaslow, & Burchinal, 2010).[127]

First, consider prior studies on the frequency of classroom activities. For instructional activities, some studies have found that teachers' instructional interactions are particularly valuable for fostering children's academic learning in language, literacy, and math skills (Mashburn et al., 2008). When broken down into domain-specific activities, shared reading has been the focus of most of the studies on instructional activities, most of which have found positive associations between shared reading and children's academic skills (Cabell, Justice, Vukelich, Buell, & Han, 2008; Justice, Kaderavek, Fan, Sofka, & Hunt, 2009; Zucker, Cabell, Justice, Pentimonti, & Kaderavek, 2013). In the current study, I classify book-reading as a language activity, but recall in Table 6-1 that book-reading is not significantly related to any of the eight child development outcome constructs.

Similarly, prior studies have found that teacher-student verbal interactions that include rich and varied vocabulary, back-and-forth exchanges between the teacher and students, and decontextualized and cognitively challenging talk are associated with children's cognitive and social competence in preschool (Kontos & Wilcox-Herzog, 1997). These types of interactions are also related to children's later language and emergent literacy abilities (Beals, DeTemple, &

[127] As pointed out by Bryant, Zaslow and Burchinal (2010), "it would be unexpected for any single measure to be the best predictor of school readiness, because we have so many different desired outcomes for children" (p. 53).

Dickinson, 1994; (NICHD, 2000)). In my study, story-telling and asking questions about reading are examples of such activities. Interestingly, though, I only find that story-telling has an effect on math skills (at the 10% level); but not on early reading skills or other outcomes. The frequency of asking questions doesn't show any statistically significant effects.

Previous studies on math-related activities and curriculum show the importance of the intensity of math activities (Brendefur, Strother, Thiede, Lane, & Surges-Prokop, 2013; Klein, Starkey, Clements, Sarama, & Iyer, 2008). However, the two most relevant activities in my study (i.e., playing games and building something) are not necessarily math-related activities. They are thus not strictly comparable to these earlier studies.

In addition, studies indicate that children who are more engaged in playing activities show stronger academic gains (e.g., Ginsburg, Lee, & Boyd, 2008; Wolfgang, Stannard, & Jones, 2003) and stronger social outcomes (see the review by Fisher, Hirsh-Pasek, Golinkoff, Singer, & Berk, 2011; Singer, 2006). In my study, then, it is a bit surprising that, among such types of activities, only playing games or puzzles showed a small positive effect, and that effect was on early reading outcomes rather than social outcomes.

Overall, it is difficult to explain why the frequency of specific teacher behavior is not significantly associated with outcomes in related domains of child development. It may be that there is not adequate information for each type of activity beyond the frequency (e.g., time involved in each of the activities during a day or a week), and therefore it is not possible to define other measures like the proportion of time spent in each type of activity. Therefore, the current measure, which relies on the frequency of the limited types of activities, is likely to suffer from significant measurement error, which may make it more difficult to detect significant effects on child development outcomes. In addition, several important activity types are not

measured in the study, such as emergent literacy and math activities. These processes have recently been shown to be important in empirical studies (e.g., Justice, Mashburn, Hamre, & Pianta, 2008; Brendefur et al., 2013). Further, "activity" is not equal to "experience." The measured frequency of an activity may not apply directly to the child, depending on the responsiveness of the child to each type of interaction. Particularly for the ECLS-B dataset, the child's individual engagement level in the activities is not captured in the current measure of activities, and therefore the higher intensity of such activities may not necessarily translate into better child development outcomes.

Next, consider studies on the quality of child-teacher interactions. The current study's finding that CIS-measured quality reduces internalizing behavior is consistent with existing theories[128] and some prior empirical studies (e.g., Howes, Phillips, & Whitebook, 1992). For example, based on a composite measure including CIS scores and three other observational measures of quality, Peisner-Feinberg et al. (2001) found small contemporaneous associations with child attention, problem behaviors, and sociability, with effects sizes of 0.03 to 0.11. Other studies using different scales for teacher sensitivity have shown similar effects on children's cognitive and social-emotional outcomes (for a meta-analysis, see Burchinal, Kainz, & Cai, 2011). One example is the Early Childhood Environment Rating Scale-Revised (ECERS-R)'s Vygotskian-based measure of caregiver sensitivity, which was found to have a modest-to-moderate relationship to language and social skills (Vandell, 2004). In theory, such effects are larger for social-emotional skills, although more empirical studies are needed to support this point.

[128] In child development theory, teacher sensitivity is important for children's social emotional well-being (Arnett, 1989; Cole, Cole, Lightfoot, & Lightfoot, 2005).

Given this context for the current study, the CIS measure of teacher sensitivity might have been expected to predict social competence and externalizing behavior; however, no significant associations were found. There are several possible explanations for these null findings. The first is that the CIS measure is narrow, in the sense that it does not include the educational perspective of teachers' behavior beyond emotional support. The Observational Record Caregiving Environment (ORCE) scale of teacher sensitivity used in the NICHD dataset does not capture the educational dimensions of classroom interactions (e.g., specific learning-focused exchanges, the curriculum, or available learning materials) (NICHD & Duncan, 2003). This same limitation exists for the CIS measure. Comparatively, the Classroom Assessment Scoring System (CLASS) is a relatively more comprehensive measure of teacher-child interactions.[129] Another possible factor contributing to the null effects pertains to the measurement of CIS scores for the ECLS-B dataset. According to Colwell et al. (2013), the total score of the quality of teacher-child interactions as measured by CIS is right-skewed in ECLS-B: Most of the preschool teachers were rated high in terms of the quality of teacher-child interactions (see also in Chapter 4 of this dissertation). This clustering of teachers' CIS scores of the dataset may make it difficult for researchers to find significant effects of CIS-measured quality on child outcomes. Indeed, researchers have only found small effects of CIS scores on children's cognitive and social-emotional outcomes, especially for preschoolers in centers (e.g., Colwell, Gordon, Fujimoto, Kaestner, & Korenman, 2013).

[129] As pointed out by Burchinal (2010, p. 6), "CLASS added ratings of the quality of instruction and classroom management to ratings of caregiver responsiveness and sensitivity." The "quality of instruction and classroom management" is different from the frequency of activities. The effects of such additional dimensions have also been empirically demonstrated (Burchinal et al., 2008; Howes et al., 2008; Mashburn et al., 2008; Moiduddin, Aikens, Tarullo, West, & Xue, 2012).

6.3 Connecting the two steps and implications for the B.A. effect estimates in Chapter 5

In this section, I bring together the findings for the two steps of the OLS regression presented in this chapter and ascertain whether these findings are consistent with those B.A. estimates presented in Chapter 5. The goal is to clarify the findings on the relationship between teachers' level of education and children's development outcomes.

The findings from the model of OLS with rich controls in Table 5-8 of Chapter 5 show that children whose teachers have a B.A. exhibit higher early reading skills, better math skills, and fewer internalizing behavior problems, but lower social competence as reported by parents. No effects were found for other outcomes. To shed light on the mechanisms behind these effects, I examine one by one the connections between teachers' B.A. status, teacher-child interactions, and child development outcomes.

In Table 6-3, I categorize the eight child development outcome constructs used in this dissertation into three categories based on the main findings from Chapter 5. The first category contains four child development outcomes: early reading, math, social competence, and internalizing behavior problems. These are the outcome constructs for which a statistically significant B.A. effect was found in the OLS model with rich controls; math and social competence are associated with B.A. effects in the PSM model as well. The second category includes externalizing behavior, for which a B.A. effect was found in the IV model. The rest of the outcomes (i.e., story-telling/expressive language, color recognition and approaches to learning) are in the third category.

Table 6-3 Connection between the two steps by outcome construct and implications for RQ1 in Chapter 5

Type of outcomes	Outcome construct	RQ1 findings from Chapter 5	Step 1: picking the activity/interaction relevant for step 2	Step 2	Connected?
Outcomes with significant B.A. effects in the OLS model (math and social competence also have B.A. effects in the PSM model)	early reading	0.079** in OLS	B.A. does not predict the frequency of playing games	0.011* play games---> early reading	NO, could have other unidentified channels
	math	0.142*** in OLS and 0.122* in PSM	0.117* B.A.---> story-telling	0.030* story-telling---> math	YES/NO: somewhat, but not very reasonable
	social competence	-0.093** in OLS and -0.190** in PSM	0.117* B.A.---> story-telling	0.044** storytelling--->social competence	YES/NO: but the two-step results do not explain the negative B.A. effect in RQ1; could have other unidentified channels→Not much
	internalizing behavior	-0.100** in OLS	0.287*** B.A.---> CIS teacher-child interactions	-0.103** t-c interactions--->internalizing behavior	YES
Outcomes with significant B.A. effects in the IV model	externalizing behavior	-0.729** in IV	+ B.A.---> CIS teacher-child interactions (including less harshness and less permissiveness)	-0.134** less harsh ---> externalizing behavior -0.070* less permissive ---> externalizing behavior	YES, somewhat: may explain why there is an B.A. effect in the IV model

Other outcomes	Story-telling /expressive language	NS	-	No activity/interaction is predicting story-telling skills	NO
	color recognition	NS	-	No activity/interaction is predicting color recognition	NO
	ATL (approaches to learning)	NS	B.A. does not predict the frequency of building something	-0.013* building something---> ATL	NO

Note. Coefficient estimates and significant levels were presented. NS = not statistically significant.
*** p<0.01, ** p<0.05, * p<0.1; and standard errors are in parentheses.

Looking at outcomes in the first category, it is interesting to note that playing games is positively associated with early reading skills, but that teachers having a B.A. does not predict the frequency of playing games. This means that the two steps of the OLS regression for RQ2 are not well enough connected to explain the positive effect of teachers' level of education on early reading outcomes. This effect could occur through other unidentified channels, which could be a topic for future research.

For math, the two steps are connected in some way. Teachers who have a B.A. engage in story-telling activities with the focal child in this sample more frequently, and this increased frequency is also positively associated with math skills. Still, one would expect the use of math activities, if measured, may be a more important mediator of the connection between teachers' level of education and children's math skills than story-telling activities.

For social competence, teachers' B.A. status has a positive effect on the CIS-measured emotional support index, and the higher interaction quality predicts higher social competence. However, this finding is contradictory to the negative effects of the B.A. on children's social competence found in Chapter 5. There could be other types of process variables that are negatively affected by teachers' level education, so that the overall B.A. effect on social competence is negative. These other process variables cannot be identified in the ECLS-B dataset given the current information, but they may be explored in future studies.

For internalizing behavior problems, the two steps are well connected by the CIS-measured quality of interaction. In step 1, having a B.A. predicts the quality of interactions with a comparatively large coefficient (0.287) and effect size (0.146); in step two, the higher quality of interactions reduces internalizing behavior problems with an effect size of 0.103.

The second category includes outcomes for which significant effects were found in the IV model. Externalizing behavior problems is the only outcome in this category; the IV model in Chapter 5 indicates that teachers' B.A. status reduces children's externalizing behavior. The second step in Section 6.2 shows that teachers' being less harsh and less permissive predicts the reduction of externalizing behavior. This can be linked to the first step in Section 6.1, which shows that teachers who have a B.A. tend to be less harsh and less permissive. The two steps appear to be connected and are supportive of the findings in Chapter 5. The two subcomponents of interaction quality (lower levels of harshness and permissiveness) could be mediating the effect of teachers' level of education on externalizing behavior shown in the IV model.

The third category that includes those outcomes for which there are no positive significant predictors in the second step OLS regression—story-telling (expressive language), color recognition, and ATL. Likewise, there are no B.A. effects on these three outcomes shown in Chapter 5. To some degree, the statistically insignificant second step may partially explain why there are no B.A. effects for these three outcomes in Chapter 5. However, the available information is still preliminary for this conclusion because there may be other unexplored mediators that are related to B.A. but not connected to the child outcomes in this category.

Notably, data on quality of teacher-child interactions is only available for a subsample of children, so there is a concern about the smaller sample size. The smaller sample size could potentially result in different estimates of the effects of teacher education on child outcomes for the subsample than are seen in the full sample analysis. To check this issue, I conduct a separate set of estimates for the B.A. effects on child development outcomes for the subsample. As shown in the results in Table B10 in Appendix B, the effects on child development outcomes are almost

the same as in the full sample. This removes the concern about small sample size for the CIS-measured quality of interactions.

In short, the two-step analysis provides some explanations for the findings in Chapter 5 by providing information on the mechanisms by which teachers' level of education may be affecting child outcomes. This is especially true for the OLS finding that having a teacher with a B.A. tends to reduce internalizing behavior problems; the two-step analysis reveals that the high quality of teacher-child interactions among teachers with a B.A. is at least a connector or intermediate factor. However, for other outcomes, such as early reading and math skills, this study has not identified a likely channel for the effects. The weak findings on the effects of teacher-child interactions on child development outcomes may be partly due to the inadequate measures of teacher behavior, and there may also be important intermediate factors not identified by this study (e.g., other unidentified but essential behaviors of teachers).

6.4 Summary and Discussion

Teachers structure the daily preschool schedule in different ways and spend varying amounts of time engaging children and interacting with children in particular settings and learning activities, according to several large-scale studies (Chien et al., 2010; Early et al., 2010; Fuligni, Howes, Huang, Hong, & Lara-Cinisomo, 2012; Winton & Buysse, 2005). Linking the variations in teacher behavior (see Section 4.4 of Chapter 4) to levels of teacher education in the ECLS-B dataset, findings from this chapter's first-step analysis indicate that some of the differences in teacher-child interactions are related to teachers' level of education. Specifically, having a teacher with a B.A. results in better teacher-child interactions in several ways: more book-reading, story-telling, and song-singing activities. Moreover, having a teacher with a B.A. is

associated with better quality of teacher-child interactions; teachers with a B.A. tend to be more sensitive, more responsive, less harsh, and less permissive than teachers without a B.A. However, the second step of the analysis reveals that only a few activity measures are associated with certain outcomes (e.g., playing games on early reading scores), and that the quality of teacher-child interactions only exhibits an impact on reducing internalizing behavior problems and to some degree on reducing externalizing behavior problems.

Connecting the two steps of the analysis for each of the eight outcome constructs, this chapter provides some information about the possible pathways or channels by which having a teacher with a B.A. may impact children's internalizing behaviour and externalizing behaivor, as shown in Chapter 5. Among the B.A.-affected outcomes, of particular interest is the mediating effect of the quality of teacher-child interactions, as measured by CIS score, in connecting the B.A. and the reduction of internalizing behavior and externalizing behavior problems. The two-step analysis also sheds some light on the reasons for the null effects on story-telling/expressive language, color recognition and approaches to learning skills. This analysis may be helpful in outlining a framework for future research on the mechanisms by which teachers' "structural" characteristics (such as B.A., certificate, and in-service training) affect children's development outcomes. In addition, regarding the effects of teachers having a B.A. on outcomes such as early reading and math, somewhat but not a very strong reasonable channel has been identified; this may also be a worthwhile avenue for future research.

It is important to note that these analyses are supplemental and preliminary, given that the measures of teacher-child interactions in this study may not adequately capture some quality aspects. There is still work to be done in terms of measuring the essential dimensions of "process quality" or "practice" (Pianta et al., 2005, p.145; Peisner-Feinberg & Yazejian, 2010). The need

to obtain accurate measurements remains an issue in early childhood education research. As pointed out by Layzer and Goodson (2006), "the ways in which researchers currently measure early care environments are flawed and… the conclusions from existing studies of the relationship between these measures and outcomes for children are often incorrect or overstated" (p. 558).

In this regard, two important observations are worth noting. First, before the introduction of CLASS, the existing research regarding teacher-child interactions often used "narrow measures of what teachers do" (Goldhaber, 2007, p.146; Hamre et al., 2013). This dissertation relies only on the CIS measure of teacher-child interaction quality. While the CIS-measured quality of interactions is valuable as it captures teacher's emotional support, other aspects of teacher behavior, such as instructional support and classroom management, are also important (Burchinal, 2010; Layzer & Goodson, 2006). Because there are not enough measures of teachers' instructional support in the ECLS-B dataset (at least for the full sample), it is unclear whether the variation in the quality of instructional support for learning is sufficient for detecting its effect on child outcomes; it is possible that the average level of instructional support was quite low, as indicated for the pre-K classroom in Pianta et al. (2005). In this regard, there is a need for the use of multiple measures together (as in Early et al., 2006, where two observational systems, ECERS-Rand and CLASS, are used). In addition, the insignificant associations between teacher-child interactions and other non-cognitive skills (i.e., social competence and ATL) may be partially due to the fact that the CIS measure of teacher-child interactions used in the ECLS-B dataset has a skewed distribution, toward the high values (Colwell et al., 2013). This type of distribution, which means a small variance, lowers the chance of detecting statistically significant effects of the quality of teacher-child interactions on child development outcomes.

Second, the effect of teachers' education may depend on other process aspects of the classroom, such as "curriculum implementation, roles of peers and assistant caregivers in the setting, and engagement of the child" (Burchinal, 2010, p.7). Many of these, however, are not available in the dataset. In particular, the engagement of the child may or may not be related to teacher characteristics.

Chapter 7

Conclusion and Implications

This chapter summarizes the dissertation's findings and implications. Section 7.1 reviews and discusses the main findings of the study. Section 7.2 and Section 7.3 consider the implications of the findings for research and policy. Section 7.4 assesses the limitations of this study and highlights areas for future research.

7.1 Summary and discussion

Increasingly, children in the United States spend much of their time in early child care and education (ECCE) environments, especially in preschools (Snyder, Dillow, & Hoffman, 2008; Belfield, 2012). Children's experience in such environments depends largely on the quality of their teachers. What defines a highly qualified preschool teacher is a complex question (Kagan et al., 2008), and there has been a heated policy debate over whether to make a bachelor's degree (B.A.) a minimum education requirement for preschool teachers in publicly funded programs. One partial source of this controversy is the mixed and non-causal research base on the effects of a B.A. on teacher behavior and child development outcomes (Kelley & Camilli, 2007; Early et al., 2007). No prior experimental or quasi-experimental studies have been conducted on this topic (W. Steven Barnett, 2011b). In order for policy makers to make sound decisions about what level of education should be required for preschool teachers, it is critical that they have rigorous evidence to support those decisions (What Works Clearinghouse, 2013; Goldhaber, 2007).

In this context, this dissertation makes a valuable contribution to the literature on preschool teacher credentials: It estimates the effects of having a teacher with a B.A. on children's

developmental outcomes at age 4 and age 5 for a representative sample of American children born in 2001 who attended center-based programs in 2005, using three rigorous quantitative methods (OLS with rich controls, PSM, and IV; the latter two are both quasi-experimental estimation methods). Additionally, it asks three sub-questions: (a) Do the effects of a B.A. differ across types of preschools, including Head Start programs, state pre-kindergartens, partially publicly funded preschool programs, and exclusively private child care centers? (b) Are the effects larger for children from low socioeconomic status (SES) families? and (c) Does specialized education in early childhood education (ECE) interact with the effects of a B.A.? Further, the dissertation explores the relationships among whether a teacher has a B.A., teacher-child interactions, and child development outcomes during the preschool wave (at age 4).

This dissertation yields several important findings.

First, the estimates for the main research question (RQ1) show that having a teacher with a B.A. in preschool has statistically significant effects on some measures of children's cognitive and non-cognitive skills at ages 4 and 5 in at least one of the three estimation methods; but the evidence is not very strong for its inconsistency across the three methods.

(1) In the OLS model for outcomes at age 4, children whose teachers have a B.A. or higher are found to exhibit higher early reading skills, higher math skills, and fewer parent-reported internalizing behavior problems compared with children whose teachers have an A.A. or below. Children who have a teacher with a B.A. exhibit math skills that are 0.14 standard deviations higher than those of children whose teachers have lower levels of education. This effect is the largest seen among the eight outcome constructs examined in this study (early reading, story-telling/expressive language, math skills, color recognition, social competence, externalizing behavior, internalizing behavior, and approaches to learning). Whether a teacher has a B.A. has

no significant effects on children's story-telling skills, color recognition, parent-reported externalizing behavior problems, or approaches to learning skills; and children whose teachers have a B.A. have lower parent-rated social competence. Note also that when statistically significant, the B.A. effects are larger in magnitude than most of other teacher variables such as in-service ECE training, Child Development Associate (CDA) degree, and experience. Compared with the effect sizes of other predictors of child development outcomes, the B.A.'s effect on the four significant outcomes is above the median level, and it is larger than many other teacher-level, classroom-level, and preschool-level predictors.

(2)The PSM and IV models indicate fewer significant effects than the OLS model. In the PSM model, having a B.A. is found to have statistically significant effects on two outcomes, with positive effects on math skills and negative effects on social competence; no statistically significant effects are found for the other six outcomes.

(3) In the IV model, a statistically significant effect for the B.A. is found only for reducing parent-reported externalizing behavior.

(4) Comparatively, the PSM and IV estimates also tend to differ in magnitude from the OLS estimates. The PSM estimates of the effect of teachers having a B.A. on children's math skills and social competence are mostly smaller than the OLS estimates. The IV estimates, although mostly statically insignificant, are larger in absolute magnitude than the OLS estimates for most of the outcomes; and the IV estimates also have larger standard errors than that of the OLS models, which are typical in IV studies. As discussed in Chapter 5, each of the three estimation methods has its strengths and weaknesses in terms of reduction of bias and maintenance of sample representativeness.

(5) Overall, there are some positive effects of having a teacher with a B.A. identified on some of the child development outcomes at age 4 (e.g., some robust evidence of positive B.A. effects on early reading and math); still, the evidence is not very strong, given the inconsistency of the findings across models and the negative effect of the B.A. on parent-reported social competence in the OLS and PSM models.

(6) Further, when I compare teachers who have just a B.A. (as opposed to a B.A. or higher) with teachers who have just an A.A., the B.A. is found to have fewer statistically significant effects in the OLS model. Significant effects are found for two outcomes only: Having a teacher with a B.A. increases math skills and reduces internalizing behavior.

(7) One year later, at age 5, when 75% of the children in the sample attended kindergarten, most of the above-mentioned effects fade out, except for the reductions in children's externalizing behavior in the PSM and IV models and the improvements in children's approaches to learning skills in the PSM model.

(8) Subgroup analyses show that there is not much heterogeneity in the B.A. effects by preschool type or by children's SES status, when looking at children's age 4 outcomes. For example, different settings do not systematically foster different levels of effects for teacher education, except that the B.A.'s reduction effect on internalizing behavior is driven mostly by the outcomes for children in the exclusively private setting and that the negative effect of B.A. on parent-reported social competence applies mainly to the boys.

(9) In addition, analysis based on available measures shows little evidence of interactions between specialized college education in ECE measured by this study and the effects of a B.A. on child development. In other words, having a degree in ECE or a related field does not increase the B.A. effect found for early reading, math, social competence, and internalizing behavior.

Likewise, the number of college courses teachers took in ECE does not increase the B.A. effects either. This finding should be noted with the fact that the more than 75% of the children in the sample whose teacher have at least a B.A. level of education hold a degree in early childhood education.

Second, in the supplemental research question (RQ2), of this dissertation, I aimed to provide some information about *why* there may be associations between teachers' level of education and children's developmental outcomes, on the basis of available intermediate measures in the ECLS-B dataset. It finds that whether a teacher has a B.A. has direct effects on the frequency of several classroom activities and on the quality of teacher-child interactions. However, the two steps of the analysis—which attempt to connect teachers' level of education with teacher-child interactions, and then to connect those teacher-child interactions with child development outcomes—only provide slight evidence for the mediating role of teacher-child interactions.

(1) The first step of the OLS analysis shows that having a B.A. is associated with higher frequencies of book-reading, story-telling, and song-singing activities, as well as more sensitive, less harsh, less detached, and less permissive teacher-child interactions.

(2) The second step, however, indicates only slight evidence of teacher-child interactions' contributions to child development outcomes. The majority of the frequency measures for specific classroom activities are not associated with children's cognitive skills, social-emotional development, or approaches to learning skills, with a few exceptions. The quality of teacher-child interactions, as measured by Caregiver Interaction Scale, has statistically significant effects on reducing externalizing and internalizing behavior problems but not on other outcomes. Note also that these findings should be noted with limitations in the measures of teacher behavior.

(3) Linking the two steps, I can conclude that there is evidence of some connection between teachers' level of education, teacher-child interactions, and certain child development outcomes—particularly for the B.A.'s effect on reducing internalizing behavior problems—but for most outcomes, including early reading and math, the relationship is not clear.

These findings, based on the ECLS-B sample, apply to a newer generation of children and their teachers, and have wider generalizability than those reported in the famous National Institute of Child Health and Human Development (NICHD) studies (NICHD, 2002; NICHD & Duncan, 2003; Early et al., 2007). First, the ECLS-B cohort of children is 10 years younger than the NICHD cohort. Second, while the NICHD studies also utilize a child-follow-up dataset like the ECLS-B and include a large and diverse sample of children and their families, the children in the NICHD dataset were selected at only 10 locations across the United States; thus, the sample is not nationally representative.

Compared with some of the NICHD studies like NICHD and Duncan (2003), this dissertation finds weaker evidence of positive effects of teacher education on child development outcomes; but the evidence in this dissertation is stronger than what is found in Early et al. (2007) for the NICHD sample. On the one hand, in NICHD and Duncan (2003), teacher education levels were found to be correlated with children's cognitive and academic outcomes at 54 months, but this study did not use the B.A. dummy measure for teacher education. On the other hand, B.A.'s effect on math found in two methods of this dissertation is stronger than Early et al. (2007)'s findings of insignificant B.A. effects on prereading and negative effect on math outcomes for their reanalysis of the NICHD sample. Since Early et al. (2007) directly compares children's academic skills between B.A. teachers and non-B.A. teachers, it is more comparable to my study. It is also a very important study in this empirical literature. In addition, it should be noted that the

frequent findings of insignificant and inconsistent B.A. effects in the current study are also concluded in Early et al. (2007) for seven large studies of center-based programs.

The positive effects of a B.A. on some outcomes found in some of the three analytical methods used in this dissertation study may be driven by the specific human capital teachers accumulated in the B.A. training program, teachers' general academic ability (general human capital), or their motivation for teaching young children. Although the OLS, PSM, and IV methods are capable of removing some of the potential bias due to the omission of observed factors (all three methods) and unobserved factors (IV only), the possibility that the B.A. has a signaling effect for general ability, motivation, beliefs, or other unobserved teacher characteristics still exists to some degree. In addition, the oftentimes insignificant findings may be related to the low quality of many of the B.A. programs for teachers in the sample; in other words, it could be that teachers with a B.A. do not learn more than teachers without a B.A., though further research is needed to explore this possibility.

This study has scholarly significance for several reasons. First, it is the first quasi-experimental study of the effects of having a teacher with a B.A. on child development. Even though it is not an experimental study, findings from the two quasi-experimental methods move beyond prior work because they take a significant methodological step toward establishing causality (Fuller, 2011).[130] Additionally, based on detailed information available in the ECLS-B dataset, the OLS model controls for many more variables than previous studies, and it removes a certain amount of biases that are associated with observed factors. Second, this dissertation is the first study on this topic to use a nationally representative sample that allows for diversity in children's background and center quality, and it uses more recent data than other similar studies.

[130] This is especially valuable because conducting an experimental study on the effects of whether a teacher has a B.A. is costly (Barnett, 2011b).

This study evaluates the effects of a B.A. for a new generation of teachers on the child development outcomes for a new generation of children born after the year 2000, adding new empirical knowledge to the field. Third, this study contributes to the literature on outcomes that are often overlooked, such as eagerness to learn, independence, persistence, and attention skills.

7.2 Implications for research

The findings in this dissertation have important implications for the literature on program quality and teacher quality in early childhood equation. First, this study contributes to the current knowledge base on the importance of teachers' level of education, showing some positive evidence for a statement like "children whose teachers have a B.A. have better outcomes than children whose teachers have lower levels of education". In fact, the B.A. was found to matter for certain types of outcomes using certain methods, especially math. However, the findings are not consistent across models, and the overall evidence that having a teacher with a B.A. improves child outcomes is not very strong.

This dissertation's findings also demonstrate the importance of adding sufficient controls to the regression models used to estimate the effects of teachers' level of education. The difference between the results in the OLS model without any controls and the OLS model with rich controls (see Table 5-1 in Chapter 5) provides evidence for this implication. The effects of a B.A. will be overestimated if the analytical model does not control for enough potential confounding factors.

In addition, the current study highlights the need for caution in interpreting findings that are based on a single analytical method. As detailed in Chapter 6, the main results of this dissertation differ across OLS, PSM, and IV models; relying on only one model would have resulted in different conclusions. It is better to draw conclusions based on several methods together, when

each has its own strengths and weaknesses and it is difficult to assign the relative degree of bias versus heterogeneity for each of the three methods.

Furthermore, guided by a set of conceptual frameworks pertaining teacher quality and child development, and a rigorous analysis of potential confounding variables for the relationship between teacher education and child outcomes, this study provides some preliminary information about the mechanisms for the B.A. effect. According to Figure 3-2 of Chapter 3, there are three channels by which a B.A. might influence child development outcomes. The first is via human capital—that is, the knowledge stock that affects teacher behavior directly. This possibility is partially supported by the findings of this dissertation's second research question (as described in Chapter 6), especially regarding the effects of a B.A. on the reduction of children's internalizing behavior. Second, a B.A. may signal a higher level of teacher ability in some regard because the design of this study cannot fully rule out this possibility, which comes from unobserved factors. For example, teachers with a B.A. may have higher learning abilities than teachers without a B.A., and may be more likely to keep pace with the most recent research on best practices in early childhood education. Education level may also help to predict how much an early childhood teacher will respond to a quality-improvement approach (Jackson et al., 2007). If the higher ability of teachers with a B.A. is a strong channel for the effects on child development outcomes, it will need to translate into essential process variables, such as teacher-child interactions, to affect child outcomes. Additional research is needed to measure teacher ability and explore the importance of this channel. Third, having a B.A. may indicate better professional commitment and may be associated with better compensation, further motivating teacher performance. The evidence from this study suggests that this earnings channel may be excluded; in a robustness check in Chapter 5 (see Table 5-8), adding the earning variable did not reduce the

effects of a B.A. found in some of the models. However, given that teachers with a B.A. did enjoy better compensation, and because further analysis is beyond the scope of this dissertation, more rigorous studies need to be conducted in the future to determine whether increased professional commitment is a key channel for the effects of a B.A. on child outcomes.

Finally, the findings in this study imply some areas for further research. This dissertation finds only weak evidence of a relationship between preschool teachers' B.A. degrees and children's developmental outcomes; more research is still needed to identify the main factors that contribute to teacher behavior and child development outcomes. Research on the effects of a B.A. is part of a complex research question regarding teacher quality, and the B.A. should be considered as part of a larger system of factors that predict quality (Early et al., 2007). This larger system includes both teacher- and non-teacher-related factors. Delivering high-quality and stimulating preschool education is challenging and requires a great deal of skills. A B.A. may be one indicator that a teacher has such skills, but it is likely not the only indicator. Other teacher quality components, such as specific types of in-service training, teachers' non-cognitive skills, professional commitment, and teacher knowledge, should be studied. In addition, program factors such as work environment and "the quality of administrative leadership" may contribute to the quality of teacher's experience and children's experience together (Peisner-Feinberg & Yazejian, 2010, p. 37). As Peisner-Feinberg and Yazejian (2010) note, "quality is a complex construct with many interrelated aspects, and efforts to understand it require approaches that reflect those nuances" (p. 39).

7.3 Policy implications and recommendations

Apart from their implications for research, the findings from this dissertation also have important policy implications, including implications for the B.A. debate and for policies that aim to raise early learning and comprehensive school readiness. Below, I discuss the policy implications of my findings on the effects of whether teachers have a B.A., other teacher factors, other components of preschool settings, and other policy-relevant predictors of child development outcomes. Implication from the B.A. effect is presented in Sub-section 7.3.1 and implications from the non-B.A. factors are in Sub-section 7.3.2.

7.3.1 Implications from the effects of B.A.

In the context of the increased national attention to preschool teacher quality, including the Early Learning Challenge Fund and President Obama's preschool plan, the questions asked in this dissertation are relevant to the public policy discussion regarding teacher quality. According to Early et al. (2007), "identifying and supporting high-quality teachers is important," and "teachers who provide instruction that leads to positive child outcomes are high-quality teachers" (p. 575). Considering that teacher education is easier to measure than more abstract notions of teacher quality, whether having a B.A. matters for preschool teachers and children is an important empirical question. However, it is still a hard-to-answer question because there are many confounding factors that prevent researchers from providing an independent estimate of the effects of teacher education on child outcomes.

Accordingly, the B.A. findings from this dissertation, both on child development outcomes and teacher-child interactions, speak to the B.A. threshold debate. They provide some evidence of positive B.A. effects and add some weight to the pro side of the B.A. debate. The positive evidence first lies in that having a B.A. teacher produces some positive effects on children's

development outcomes, particularly in terms of early reading and math skills; the positive evidence also lies in that B.A. teachers is linked to better interactions with the children in the classroom.

The effect of B.A on child development outcomes. Early learning and child development serves as an important goal of early childhood policy. Whether the teacher improves children's learning is one component of teacher rating for early childhood teachers in some states that have incorporated early childhood teachers in their state teacher evaluation systems (Connors-Tadros & Horowitz, 2014). The findings for the first research question of this dissertation give some evidence for B.A.'s effect on child development. Specifically, having a teacher with a B.A. has shown positive effects on early reading, math skills, the reduction of internalizing behaviour problems and the reduction of externalizing behaviour problems. Particularly speaking, B.A's effect on math is quite robust across the OLS and PSM models; and the effect in the IV model, although imprecisely estimated with a large standard error, is not negative. Therefore the positive signs are consistent across the three methods. This picture of positive B.A. effect on math and early reading is stronger than that of another child follow-up study in the NICHD sample estimated in Early et al. (2007).

The effect of B.A. on teacher-child interactions. In addition to student learning and growth, teacher's professional practice has a weight of at least 50% in some states' teacher evaluation system for early childhood teachers, particularly high for New Jersey and Delaware (Connors-Tadros & Horowitz, 2014). In this background of accountability, this dissertation shows positive effects of B.A. on the frequency and quality of teacher-child interactions in the classroom. Teachers with a B.A. exhibit higher frequency of book-reading, storytelling and song-singing activities with the focal child. Also, compared to teachers with lower level of education, they

also produce a higher quality of teacher-child interactions: being more sensitive, less harsh, less detached and less permissive. These findings are consistent with one of arguments for the B.A. threshold policy. As introduced in Chapter 1, supporters of the B.A. debate believe that B.A., especially B.A. in early childhood education, contributes to better knowledge and skills of teachers, and in turn leads to richer cognitive stimulation and more sensitive interactions with the children in the classroom (Barnett, 2011b). B.A. teachers are also believed to better connect the newest research to practice (Goffin & Washington, 2007). Although most of the teacher behaviour measures used in this dissertation fail to be linked to child development outcomes, the findings of B.A.'s effect on professional practice is a good signal for two reasons: (a) This means that B.A. programs did make a difference in their teachers' classroom practice, which itself is a good message; and it further implies the possibility of changing teacher behavior by changing the content and quality of a B.A. program. (b) Findings in the second step OLS for the second research question implies a mostly positive association between teacher behavior and child development outcomes, even though they are not statistically significant. When better measures of teacher behavior are identified, B.A. programs can aim to achieve this type of identified high quality professional practice. [131]

Some reservations. At the same time, policy makers should note that the above-mentioned positive findings are not very strong, for a few reasons: (1) the B.A. effects on child development outcomes differ across analytical methods; (2) B.A. is shown to have a negative effect on parent-reported social competence at age 4; (3) the effect sizes are not large when compared to prior studies or assessed using the Cohen's criteria for judging the magnitude of an effect size; (4) the positive effects of B.A. on early reading, math and reduction of internalizing behaviour problems

[131] While observations of teacher practice and attributions to teachers of growth in student learning are challenging and have to come from reliable measures, the teacher behavior measure in this study may not have been well measured for the data to reveal statistically significant effect on child development outcomes.

at age 4 disappear at age 5; (5) compared to A.A. teachers, B.A. teachers are better only in math and the reduction of internalizing behaviour problems.

Overall, these findings at least contribute some positive evidence to the research base (especially for math) before entering into any value judgments or cost considerations. At the same time, policy makers should note that the positive findings are not very strong. The B.A. can sometimes be a policy parameter for elevating teacher quality, especially when information on classroom level teacher performance is limited and when the "identification" of a high-quality teacher takes place at the hiring stage (Temin, 2003; Staiger & Rockoff, 2010), but other teacher quality factors, such as professional commitment and teachers' non-cognitive skills, though difficult to measure, should still be considered.

Relevant factors for decision-making. Policy makers may want to examine a number of relevant factors together when making decisions about whether to set up a B.A. threshold as a teacher qualification. The first is the empirical evidence on the effects of having a teacher with a B.A. This study finds some positive effects of the B.A. on child development outcomes.

Second, when looking at the benefits of having a teacher with a B.A., policy makers may wish to focus on the whole-child approach to school readiness, defined to include a wide range of development domains, including academic development/early learning, social-emotional development, and physical development. This is a significant merit of preschool education, as compared with other stages of education (see chapters 15–20 in Zigler, Gilliam, & Barnett, 2011). The fact that B.A. effects are seen on some child development domains but not on others may be due to the inadequate quality or the narrow focus of ECE training in B.A. programs. If B.A. training programs do not pay much attention to the promotion of children's social-

emotional well-being but instead put a lot of emphasis on children's academic skills, researchers are unlikely to find a positive B.A. effect on children's social competence.

The third factor for policy makers to consider is the complementary service for implementing the B.A. threshold. Better compensation should accompany the B.A. requirement for preschool teachers, since the benefits may have an incentive effect, attracting teachers with a B.A. to join the profession and motivating teachers without a B.A. to get one. This is likely to result in significant additional costs to states who have to offer higher pay for B.A. teachers and to teachers as they incur more costs to obtain higher levels of education (Fuller, 2011). Cost may also include the crowding effect of the B.A. input on other inputs such as child-teacher ratio (Bassok, 2013). Future studies regarding the relative cost of implementing the B.A. threshold versus other highly relevant teacher quality strategies (e.g., professional development for non-B.A. teachers) should be conducted.

Fourth, policy makers can explore alternative components of teacher quality and compare both the costs and benefits of using the B.A. threshold with that of using other strategies for determining teacher quality. Although perhaps necessary, a B.A. alone may not be sufficient to ensure teacher quality (Bowman, 2011); an additional important focus is "caring quality" (Fuller, 2011), or the quality of teacher-child interactions in the classroom. Process quality, such as teacher-child interactions and other professional practice in the classroom measurable by several well-established quality scales (e.g., CIS and ECERS-R), could be used more widely in teacher quality ratings. Because the link between these measures and child outcomes found in this dissertation is not very strong (as shown in Chapter 6), policy makers should consider the use of alternative measures, such as CLASS; new measures also need to be developed. There is a strong need for a broader rating system for teacher quality because, as suggested by Early et al. (2007),

"teachers' education must be considered as part of a system of factors that contribute to teacher quality, which in turn is related to classroom quality and children's gains" (p.577). The conceptual framework of teacher quality presented in Figure 2-1 of Chapter 2 of this dissertation may be helpful for this type thinking.

Fifth, a comparison can be made between teacher quality strategies and non-teacher-related strategies (e.g., group size, physical environment, administrative leadership, etc.). An undue emphasis on teacher qualifications may obscure other aspects of children's preschool experience that would benefit from policy interventions. Regarding the need for a broader perspective on quality improvement strategies, and citing the conclusion from Early et al. (2007), one takeaway from this dissertation is that the inconsistent findings on the benefits of a B.A. for child outcomes may serve as "a springboard that moves research and policy regarding the role of teachers' education and, more broadly, teacher quality to a new level that is increasingly multifaceted and nuanced" (p. 577).

To summarize, the statistically significant positive B.A. effect on math skills seen in the OLS model with rich controls and PSM, plus the positive effect of B.A. on some classroom activities and the quality of teacher-child interactions, add some weight in favor of implementing a B.A. threshold. Still, given that the evidence is not very strong, policy makers should be careful in using this evidence and should consider other teacher factors together with teachers' level of education when making decisions about teacher quality. The new century has witnessed a continuous growth of public programs in the United States, including Head Start programs and state pre-Ks; and these implications for policy makers, based on the effect estimates of teachers'

B.A. status for the generation of children born in the new century using the latest large-scale early childhood dataset, enters timely for the B.A. debate.

7.3.2 Implications from the findings of non-B.A. factors

Implications from other teacher factors. There are several findings from this dissertation regarding other teacher factors, apart from whether a teacher has a B.A., that have policy relevance. First, specific pre-service training in ECE or a related field does not increase the B.A.'s comparative advantage in the relevant outcomes, such as math skills. If specialized education in ECE is measured well, this may imply that the bachelor's degree serves more as a screening tool for academic aptitude and persistence than as a route to specialized teaching knowledge (Tout, Zaslow, & Berry, 2006). Second, as shown in Section 5.7 of Chapter 5, in-service training does not have a statistically significant effect on any of the outcomes in this study, possibly because the measure is a bit coarse, and also because most of the teachers reported that they experienced ECE training within the past 12 months. Third, CDA has benefits for early reading and the reduction of internalizing behavior (see Table 5-13 in Chapter 5). Fourth, the available measures of teacher-child interactions in this dissertation do not show much of an effect on child outcomes, except on the reduction of externalizing behavior. Fifth, concerns that a higher degree of racial mismatch would follow the establishment of a B.A. threshold can be relieved a bit, because the matching dummy (i.e., whether the teacher has a different race than the child) is not statistically significant in the models used in this dissertation for the national sample of center-attending children.

Considering teachers' B.A. status alongside other teacher factors, it may be noted that each type of teacher quality ingredient has its strengths and weaknesses in serving as a policy

parameter for elevating teacher quality. For example, teacher education is easily measurable and has high malleability (is easy to be changed), and there is some evidence of its effect on child development, but establishing a B.A. threshold has high implementation costs. Comparatively, teachers' non-cognitive skills are difficult to measure and have moderate malleability, with unknown but likely high empirical effects and moderate to high costs. Keeping in mind the findings from this dissertation as well as the strengths and weaknesses of each type of teacher quality ingredient as a policy parameter, policy makers can decide which teacher factors (i.e., the teacher characteristics and behavior in Figure 2-1) to give more weight to in efforts to promote teacher quality. Similar ideas are also mentioned in Peisner-Feinberg and Yazejian (2010) and would benefit from additional research.

Implications from other components of the preschool setting that affect the impact of teacher quality on child outcomes. As Fuller et al. (2010) recommended, policy makers should support policies that address multiple components of teacher quality and the factors that might influence quality, such as mentoring, professional development support, monitoring and supervision, and accreditation of teacher preparation programs. The part on mentoring and professional development support is also consistent with Pianta (2011) and Kagan and Gomez (2011)'s suggestions in the *Pre-K Debates* book. Also, the current study indicates that "licensing" has some positive effects on early reading and math skills based on the OLS model with rich controls (see Section 5.5 of Chapter 5).

Implications from other policy-relevant predictors of child development outcomes at age 4. Additional predictors of child development outcomes at age 4 that are statistically significant and with a large effect size include prior child development status at age 2 (mental ability in particular), month difference in age at assessment, low birth weight, race/ethnicity, mother's

highest level of education, family income, quality of parenting, and frequency of home–preschool connection. Many of these are important factors for the consideration of policy interventions. Further, compared with other child outcomes, the model for math skills seem to be more malleable, meaning that policy efforts on the potential factors at child, family, teacher, classroom and preschool level are more likely to affect the growth trajectory of children's math skills. The findings for all of these policy relevant predictors described in this paragraph will be summarized and published in a policy brief in the post-dissertation stage.

7.4 Limitations and future research

Despite its scholarly significance, this dissertation study has several limitations. First, though much more comprehensive than prior studies of preschool teachers' level of education, the child development outcomes examined in the current study are short-term outcomes measured at just two time points (age 4 and age 5). Future studies with longer term follow-up data are needed to examine the effects of having a teacher with a B.A. on longer term child development outcomes.

Second, the B.A. measure used in this study is lacking in information regarding B.A. program quality, like the other B.A. studies in the literature (Peisner-Feinberg & Yazejian, 2010). Future data collection and research efforts related to teacher education should focus on the content and therefore the quality of B.A. teacher preparation programs. One such effort is the Early Childhood Higher Education Inventory conducted by the Center for the Study of Child Care Employment at the University of California, Berkeley (Whitebook et al., 2014). Admittedly, also, the ECLS-B dataset doesn't contain information on teacher's family background such as the SES of teachers. This prevents this dissertation to identify what type of

B.A. teachers the children are getting except his or her age, gender and race. Future studies are recommended in this regard.

Third, the estimates of the heterogeneous effects of teachers having a B.A. and the interplay of the B.A. effect with specialized college education in ECE are likely to be correlational estimates. Depending on the validity of the instrument for each subgroup, the models may or may not yield causal estimates. In the case of a non-causal estimate, some potential bias may not be removed. For example, the linkage between teacher-child interactions and children's social-emotional skills may indicate a backward pathway from outcome to interactions, because teachers may respond less sensitively to children who are less social. Because the equations control for prior child outcomes at age 2, this type of bias should be alleviated to some degree. Still, as in many other studies that have controlled for prior outcomes (e.g., Early et al., 2007), the prior outcome measures in this dissertation study are not the same as the outcome measures examined at ages 4 and 5, and this reduces the ability of the prior outcome measures to alleviate selection bias.

Fourth, some definitions regarding center types are limited by the data. For example, it is difficult to define "state pre-K" by the ECLS-B dataset, as the term was not used in the survey instrument. Instead, this study has used "prekindergarten" as a proxy for "state pre-K." Additionally, the definition of a full-day center is not ideal, because it is based on a child-level variable called "time of attendance." Future data collection could be improved by incorporating into the survey instrument a direct question about the preschool's full-day or part-day status.

Fifth, for the research question on specialized education in ECE, the measure of "specialized education in ECE" is not exclusively confined to the bachelor's degree. For example, those who reported having a degree in ECE or a related field may either have a

bachelor's degree in ECE or an associate's degree in ECE if they obtained an A.A. before earning a B.A. Fortunately, as explained in Chapter 3, this issue should not substantially alter the results, since it is inferred that only a small portion of the B.A. holders in the sample also obtained an A.A. This issue arises from a limitation in the survey design. Future data collection would benefit from a revision to the question regarding the ECE major, such that teachers are asked more explicitly whether the major for their bachelor's degree is ECE or a related field.

Sixth, there is a limitation in the study's generalizability due to the fact that data on outcomes is only available for children who have teacher information. A comparison of children with teacher information (84.31% of the children in the original sample) and children without teacher information (15.69%) shows that the two samples have different sample means. This limitation is rooted in the dataset itself, and it is not a big issue because most of the center-attending children have teacher information. Future data collection work could be improved in terms of achieving higher teacher response rate.

With these limitations in mind, the findings of this study suggest several directions for future research. First, more studies are needed to define what it means to be a high-quality preschool teacher. As Peisner-Feinberg and Yazejian (2010) point out, the quality of children's experiences in early childhood programs is "often far from optimal for promoting children's learning" (p.40). Policy makers and practitioners need to know the best and most effective ways for improving teacher quality. To this end, future research should work toward identifying teacher factors and strategies that promote desirable child development outcomes. Teacher factors of interest may include the quality of teachers' preparation and training, teacher personality, professional commitment, and teacher-child interactions. As for B.A.-related factors, information on the quality or content of teacher preparation programs that is not available in the existing datasets

should be gathered in the future. Professional development models should also be researched in detail (Early et al., 2007; Fuller et al., 2010). Perhaps more importantly, there is a great need for additional research on heretofore unobserved teacher characteristics. As pointed out by Early et al. (2007), "more fine-grained research is needed to address which aspects of teachers' attitudes, knowledge, and behaviors are affected by participation in higher education and in-service training" (p. 574). In a study on teacher value-added in K-12 education, teachers have been found to account for approximately 8.5 percent of the variation in students' 10th-grade achievement. However, less than 5 percent of the variation in teacher quality is explained by quantifiable characteristics, such as degrees and experience levels (Goldhader, Brewer, & Anderson, 1999). Similar findings may exist in early childhood education, but this needs empirical testing. For factors such as teacher-child interactions, researchers may want to investigate the level or threshold in quality related to better child outcomes, since such findings would be applicable in public policy and regulations (Tout, Zaslow, Halle, & Forry, 2009).

Second, there is a need for both highly focused and more comprehensive studies on preschool quality—that is, studies that examine specific inputs of child outcomes and studies related to multiple ingredients and common factors at the system level. On one hand, the efforts to disaggregate the aspects of quality should be enhanced, because "by teasing apart components of quality, researchers will be able to identify aspects of care that are most predictive of outcomes; help to guide the design of interventions to improve quality, including professional development initiatives, and help to refine quality rating and improvement systems by better specifying the most critical aspects of quality" (Peisner-Feinberg & Yazejian, 1999, p. 39). In order to identify the essential features that can be regulated, measured, and monitored, factors in those traditionally studied in the program quality literature (such as B.A. and certificate) should

be explored further ; other often missed factors like a well-designed certificate, intensity of professional development and teacher's non-cognitive skills should be examined, On the other hand, it would also be beneficial to conduct more research on the effect of multiple components of teacher quality together and the factors that might influence quality (e.g., mentoring, monitoring and supervision, and accreditation of teacher preparation programs), because the impact of teacher quality on classroom quality and child outcomes is likely to be influenced by many other components of the early care and education system.

Third, there is a need for stronger data-collection efforts to advance research in the above areas. For instance, measurement for traditionally difficult-to-measure teacher quality factors may be enhanced. This applies to classroom observations of teacher behavior, including teacher-child interactions. As Hamre et al. (2012) state, "one final factor promoting interest in interventions targeting improvements in teacher-child interactions is the inclusion of measures of teacher-child interactions in monitoring and quality improvement policies" (p. 91). CLASS is an example of one such effort: It is a scale developed by Hamre et al. (2007), which built and validated a theoretical model of classroom effects in over 4,000 early childhood and care settings. Several states plan to include the CLASS or other measures of teacher-child interactions as one component of their Quality Rating and Improvement Systems (QRISs) or other improvement efforts (Tout & Maxwell, 2010), which indicates that some progress is being made in this area.

REFERENCES

Abner, K. S., Gordon, R. a, Kaestner, R., & Korenman, S. (2013). Does child care quality mediate associations between type of care and development? *Journal of Marriage and the Family*, *75*(5), 1203–1217. doi:10.1111/jomf.12055

ACF (Administration for Children and Families). (2003). *Head Start FACES (2000): A Whole-Child perspective on program performance. Fourth progress report*. Head Start Information and Publication Center. Retrieved from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED478791&ERICExtSearch_SearchType_0=no&accno=ED478791

Ackerman, D. J., & Barnett, W. S. (2005). *Prepared for kindergarten: what does "readiness" mean?*. National Instititue for Early Education Research.

Almlund, M., Duckworth, A. L., Heckman, J. J., & Kautz, T. D. (2011). Personality psychology and economics. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education (Volume 4)* (pp. 1–179). doi:10.1016/B978-0-444-53444-6.00001-8

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*(434), 444–455.

Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics*, *106*(4), 979–1014.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Arnett, J. (1989). Caregivers in day-care centers: Does training matter? *Journal of Applied Developmental Psychology*, *10*(4), 541–552. doi:10.1016/0193-3973(89)90026-9

Arnold, D. H., Zeljo, A., Doctoroff, G. L., & Ortiz, C. (2008). Parent involvement in Preschool: predictors and the relation of involvement to pre-literacy. *School Psychology Review*, *37*(1), 74–90. Retrieved from http://eric.ed.gov/?id=EJ817289

Auger, A., Farkas, G., Burchinal, M. R., Duncan, G. J., & Vandell, D. L. (2014). Preschool center care quality effects on academic achievement : An Instrumental Variables analysis, *50*(12), 2559–2571.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, *46*(3), 399–424.

Barnett, W. S. (2004). *Better teachers, better preschools: Student achievement linked to teacher qualifications. NIEER Preschool Policy Matters* (Vol. 2). ERIC.

Barnett, W. S. (2008). *Preschool education and its lasting effects: Research and policy implications*.

Barnett, W. S. (2010). Universal and targeted approaches to preschool education in the United States. *International Journal of Child Care and Education Policy*, *4*(1), 1–12.

Barnett, W. S. (2011a). Effectiveness of early educational intervention. *Science*, *333*(6045), 975–978. doi:10.1126/science.1204534

Barnett, W. S. (2011b). Minimum requirements for preschool teacher educational qualifications. In & W. S. B. E. Zigler, W. S. Gilliam (Ed.), *The Pre-K Debates: Current controversies and issues* (pp. 48–54). Brookes Publishing Company.

Barnett, W. S. (2011). *Preschool education as an education reform: Issues of effectiveness and access* (pp. 1–14). Retrieved from http://nieer.org/publications/latest-research/preschool-education-educational-reform-issues-effectiveness-and-access

Barnett, W. S., Epstein, D. J., Friedman, A. H., Sansanelli, R. A., & Hustendt, J. T. (2009). *The state of preschool 2009*. National Institute for Early Education Research.

Barnett, W. S., & Hustedt, J. T. (2011). *Improving public financing for early learning programs*. Retrieved from http://nieer.org/resources/policybriefs/24.pdf

Barnett, W. S., Lamy, C., & Frede, E. (2001). *CEER Supplementary Technical Report: Preschool quality in Abbott Districts, 2000-2001*.

Bassok, D. (2009). *Three essays on early education policy*. Standford University.

Bassok, D. (2013). Raising teacher education levels in Head Start: Exploring programmatic changes between 1999 and 2011. *Early Childhood Research Quarterly*, *28*(4), 831-842.

Bassok, D., Fitzpatrick, M., Loeb, S., & Paglayan, A. S. (2013). The early childhood care and education workforce in the United States: Understanding changes from 1990 through 2010. *Education Finance and Policy*, *8*(4), 581–601. Retrieved from http://cepa.stanford.edu/sites/default/files/AEFP_ECCE Workforce.pdf

Beals, D. E., DeTemple, J. M., & Dickinson, D. K. (1994). Talking and listening that support early literacy development of children from low-income families. In *Bridges to literacy: Children, families, and schools* (pp. 19–40). Cambridge, MA: Blackwell.

Belfield, C. (2012). Early Childhood care and education: Enrollment patterns and expenditures. In R. Pianta, W. S. Barnett, L. M. Justice, & S. M. Sheridan (Eds.), *Handbook of Early Childhood Education*. Guilford Publication.

Berk, L. E. (1985). Relationship of caregiver education to child-oriented attitudes, job satisfaction, and behaviors toward children. *Child Care Quarterly*, *14*(2), 103–129. doi:10.1007/BF01113405

Bishop-Josef, S. J., & Zigler, E. (2011). The cognitive/academic emphasis versus the whole child approach: The 50-year debate. In E. Zigler, W. S. Gilliam, & W. S. Barnett (Eds.), *The Pre-K Debates: Current controversies and issues* (pp. 83–88). Brookes Publishing Company.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*(2), 647–663.

Blau, D. M. (1997). The production of quality in child care centers. *Journal of Human Resources*, *32*(2), 354–387.

Blau, D. M. (1999). The effect of child care characteristics on child development. *Journal of Human Resources*, *34*(4), 786–822.

Blau, D. M. (2000). The production of quality in child-care centers : another Look. *Applied Developmental Science*, *4*(3), 136–148.

Bogard, K., Traylor, F., & Takanishi, R. (2008). Teacher education and PK outcomes: Are we asking the right questions? *Early Childhood Research Quarterly*, *23*(1), 1–6.

Borghans, L., Golsteyn, B. H. H., Heckman, J., & Humphries, J. E. (2011). Identification problems in personality psychology. *Personality and Individual Differences*, *51*(3), 315–320.

Bowman, B. (2011). Bachelor's degrees are necessary but not sufficent: Preparing teachers to teach young children. In *The Pre-K Debates: Current controversies and issues* (pp. 54–57). Brookes Publishing Company.

Bredekamp, S., & Copple, C. (1997). *Developmentally appropriate practice in early childhood programs (revised edition)*. Washington D.C.: NAEYC.

Brendefur, J., Strother, S., Thiede, K., Lane, C., & Surges-Prokop, M. J. (2013). A professional development program to improve math skills among preschool children in head start. *Early Childhood Education Journal*, *41*(3), 187–195.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In *Handbook of child psychology* (pp. 793–827). Wiley Online Library. doi:10.1002/9780470147658.chpsy0114

Brooks-Gunn, J., Gross, R. T., Kraemer, H. C., Spiker, D., & Shapiro, S. (1992). Enhancing the cognitive outcomes of low birth weight, premature infants: for whom is the intervention most effective? *Pediatrics*, *89*(6), 1209–1215.

Bryant, D. M., Burchinal, M., Lau, L. B., & Sparling, J. J. (1994). Family and classroom correlates of head start children's developmental outcomes. *Early Childhood Research Quarterly*, *9*(3-4), 289–309. doi:10.1016/0885-2006(94)90011-6

Bryant, D. M., Zaslow, M. J., & Burchinal, M. R. (2010). Issues in measuring program quality. In *The quest for quality: Promising innovations for early childhood programs* (pp. 50–52). Baltimore, MD: Brookes.

Bueno, M., Darling-hammond, L., & Gonzales, D. (2010). *A matter of degrees: Preparing teachers for the pre-K classroom*.

Burchinal, M., Kainz, K., & Cai, Y. (2011). How well do our measures of quality predict child outcomes? A meta-analysis and coordinated analysis of data from large-scale studies of early childhood settings. In *Quality measurement in early childhood settings* (pp. 11–31). Baltimore, MD: Brookes.

Burchinal, M. R. (2010). *Differentiating among measures of quality: Key characteristics and their coverage in existing measures*. Retrieved from http://www.acf.hhs.gov/sites/default/files/opre/differ_measures.pdf

Burchinal, M. R., Cryer, D., Clifford, R. M., & Howes, C. (2002). Caregiver training and classroom quality in child care centers. *Applied Developmental Science*, *6*(1), 2–11. doi:10.1207/S1532480XADS0601_01

Burchinal, M. R., Howes, C., & Kontos, S. (2002). Structural predictors of child care quality in child care homes. *Early Childhood Research Quarterly*, *17*(1), 87–105. doi:10.1016/S0885-2006(02)00132-1

Burchinal, M. R., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher–child interactions and instruction. *Applied Developmental Science*, *12*(3), 140–153. doi:10.1080/10888690802199418

Burchinal, M. R., Hyson, M., & Zaslow, M. (2011). Competencies and credentials for early childhood educators: what do we know and hwat do we need to know. In *The Pre-K Debates: Current controversies and issues* (pp. 73–77). Brookes Publishing Company.

Burchinal, M. R., Roberts, J. E., Riggins, J., Zeisel, S. A., Neebe, E., & Bryant, D. (2000). Relating Quality of Center-Based Child Care to Early Cognitive and Language Development Longitudinally. *Child Development*, *71*(2), 339–357. doi:10.1111/1467-8624.00149

Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, *25*(2), 166–176. doi:10.1016/j.ecresq.2009.10.004

Buysse, V., Wesley, P. W., Bryant, D., & Gardner, D. (1998). Quality of early childhood programs in inclusive and noninclusive settings. *Exceptional Children*, *65*(3), 301–14. Retrieved from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=EJ593058&ERICExtSearch_SearchType_0=no&accno=EJ593058

Cabell, S. Q., Justice, L. M., Vukelich, C., Buell, M. J., & Han, M. (2008). *Strategic and intentional shared storybook reading: Achieving excellence in preschool literacy instruction* (pp. 198–220). Guilford Press New York, NY.

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

Carneiro, P., Crawford, C., & Goodman, A. (2007). *The impact of early cognitive and non-cognitive skills on later outcomes*. Centre for Economics of Education. Retrieved from http://cee.lse.ac.uk/ceedps/ceedp92.pdf

Cassidy, D. J., & Lawrence, J. M. (2000). Teachers' beliefs: The "whys" behind the "how tos" in child care classrooms. *Journal of Research in Childhood Education*, *14*(2), 193–204.

Chernoff, J. J., Flanagan, K. D., McPhee, C., & Park, J. (2007). *Preschool: First findings from the preschool follow-up of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B). First look. NCES 2008-025.* ERIC. Retrieved from http://nces.ed.gov/pubs2008/2008025.pdf

Chien, N. C., Howes, C., Burchinal, M., Pianta, R. C., Ritchie, S., Bryant, D. M., … Barbarin, O. A. (2010). Children's classroom engagement and school readiness gains in prekindergarten. *Child Development*, *81*(5), 1534–1549.

Child Care Aware of American. (2013). *We can do better: Chid Care Aware of America's ranking of state child care center regulations and oversight 2013 update*. Retrieved from http://www.naccrra.org/sites/default/files/default_site_pages/2013/wcdb_2013_final_april_11_0.pdf

Choi, J. Y. (2011). *Teacher qualifications and preschoolers' math outcomes: Mediation by teacher-child relationships?*. *ETD Collection for Purdue University*. Retrieved from http://docs.lib.purdue.edu/dissertations/AAI1510057

Christopher, H., & Blackman-Jones Ramona. (2006). Parent invovlemnt in preschool. In Z. E., W. Gillaim, & J. S. (Eds.), *A vision for universal preschool education* (pp. 149–168). Cambridge University Press.

Clarke-Stewart, K. A., Vandell, D. L., Burchinal, M., O'Brien, M., & McCartney, K. (2002). Do regulable features of child-care homes affect children's development? *Early Childhood Research Quarterly*, *17*(1), 52–86. doi:10.1016/S0885-2006(02)00133-3

Cohn, E., & Geske, T. (1990). *The economics of education*. Cambridge, Mass: Balhinger.

Cole, M., Cole, S. R., Lightfoot, S., & Lightfoot, C. (2005). *The development of children* (5th ed.). New York: Macmillan. Retrieved from http://catalog.bankstreet.edu/cgi-bin/koha/opac-detail.pl?biblionumber=69733

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, *94*, 95–120.

Colker, L. (2008). Twelve characteristics of effective early childhood teachers. *Young Children*, *63*(2), 68–73.

Colwell, N., Gordon, R. a, Fujimoto, K., Kaestner, R., & Korenman, S. (2013). New evidence on the validity of the Arnett Caregiver Interaction Scale: Results from the Early Childhood Longitudinal Study-Birth Cohort. *Early Childhood Research Quarterly*, *28*(2), 218–233. doi:10.1016/j.ecresq.2012.12.004

Committee on Family and Work Policies (2003). *Working families and growing kids: Caring for children and adolescents.* Washington, DC: The National Academies Press.

Commodari, E. (2013). Preschool teacher attachment, school readiness and risk of learning difficulties. *Early Childhood Research Quarterly*, *28*(1), 123–133. doi:10.1016/j.ecresq.2012.03.004

Conniffe, D., Gash, V., & O Connell, P. J. (2000). Evaluating state programmes:" natural experiments" and propensity scores. *Economic and Social Review*, *31*(4), 283–308.

Connors-Tadros, L., & Horowitz, M. (2014). How Are Early Childhood Teachers Faring in State Teacher Evaluation Systems? CEELO Policy Report. Retrieved from http://ceelo.org/wp-content/uploads/2014/03/CEELO_policy_report_ece_teachereval_march_2014.pdf

Crosby, D. A., Dowsett, C. J., Gennetian, L. A., & Huston, A. C. (2010). A tale of two methods: Comparing regression and instrumental variables estimates of the effects of preschool child care type on the subsequent externalizing behavior of children in low-income families. *Developmental Psychology*, *46*(5), 1030.

Crosnoe, R. (2006). Health and the education of children from racial/ethnic minority and immigrant families. *Journal of Health and Social Behavior*, *47*(1), 77–93.

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. In S. M. and L. W. Eric A. Hanushek (Ed.), *Handbook of the Economics of Education Volume 1* (Vol. 1, pp. 697–812). Elsevier. doi:10.1016/S1574-0692(06)01012-9

Curby, T. W., Rimm-Kaufman, S. E., & Ponitz, C. C. (2009). Teacher–child interactions and children's achievement trajectories across kindergarten and first grade. *Journal of Educational Psychology*, *101*(4), 912.

Currie, J. (2001). Early childhood education programs. *The Journal of Economic Perspectives*, *15*(2), 213–238. Retrieved from http://www.princeton.edu/~jcurrie/publications/Early_childhood_intervention.pdf

Currie, J., & Neidell, M. (2007). Getting inside the "Black Box" of Head Start quality: What matters and what doesn't. *Economics of Education Review*, *26*(1), 83–99. doi:10.1016/j.econedurev.2005.03.004

Data File User's Manual. (2010). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Kindergarten 2006 and 2007*.

Datta Gupta, N., & Simonsen, M. (2010). Non-cognitive child outcomes and universal high quality child care. *Journal of Public Economics*, *94*(1-2), 30–43. doi:10.1016/j.jpubeco.2009.10.001

De Kruif, R. E. ., McWilliam, R. ., Ridley, S. M., & Wakely, M. B. (2000). Classification of teachers' interaction behaviors in early childhood classrooms. *Early Childhood Research Quarterly*, *15*(2), 247–268. doi:10.1016/S0885-2006(00)00051-X

Dickinson, D. K., Darrow, C. L., & Tinubu, T. A. (2008). Patterns of teacher–child conversations in head start classrooms: Implications for an empirically grounded approach to professional development. *Early Education and Development*, *19*(3), 396–429.

DuGoff, E. H., Schuler, M., & Stuart, E. A. (2014). Generalizing observational study results: Applying propensity score methods to complex surveys. *Health Services Research*, *49*(1), 284–303.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., … Brooks-Gunn, J. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428.

Duncan, G. J., & Magnuson, K. (2011). The nature and impact of early achievement skills, attention skills, and behavior problems. *Whither Opportunity*, 47–69.

Duncan, G. J., & Magnuson, K. A. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps? *The Future of Children*, *15*(1), 35–54.

Duncan, S. E., & DeAvila, E. (1998). Preschool Language Assessment Survey 2000 Examiner's Manual: English Forms C and D. Monterey, CA: CTB/McGraw-Hill.

Dunn, L. (1993). Proximal and distal features of day care quality and children's development. *Early Childhood Research Quarterly*, *8*(2), 167–192. doi:10.1016/S0885-2006(05)80089-4

Early, D. M., Iruka, I. U., Ritchie, S., Barbarin, O. A., Winn, D.-M. C., Crawford, G. M., … Howes, C. (2010). How do pre-kindergarteners spend their time? Gender, ethnicity, and income as predictors of experiences in pre-kindergarten classrooms. *Early Childhood Research Quarterly*, *25*(2), 177–193.

Early, D. M., ZA, K. L., Alva, S., Bender, R. H., Bryant, D., Cai, K., … Zill, N. (2007). Teachers" education, classroom quality , and young children"s academic skills: results from seven studies of preschool programs. *Child Development*, *78*(2), 558–580. doi:10.1111/j.1467-8624.2007.01014.x

Early, D. M., Maxwell, K. L., Clifford, R. M., Pianta, R. C., Ritchie, S., Howes, C., … Barbarin, O. (2008). Teacher education and child outcomes: A reply to the commentary. *Early Childhood Research Quarterly*, *23*(1), 7–9. doi:10.1016/j.ecresq.2007.08.003

Elicker, J., & Mathur, S. (1997). What do they do all day? Comprehensive evaluation of a full-day kindergarten. *Early Childhood Research Quarterly*, *12*(4), 459–480.

Epstein, A. S. (1999). Pathways to quality in Head Start, public school, and private nonprofit early childhood programs. *Journal of Research in Childhood Education*, *13*(2), 101–119. doi:10.1080/02568549909594732

Epstein, D. J., & Barnett, W. S. (2012). Early education in the United States: Programs and access. In R. C. Pianta, W. S. Barnett, L. M. Justice, & S. M. Sheridan (Eds.), *Handbook of Early Childhood Education* (pp. 3–21). Guilford Publication.

Fisher, K., Hirsh-Pasek, K., Golinkoff, R. M., Singer, D. G., & Berk, L. (2011). Playing around in school: Implications for learning and educational policy. *The Oxford Handbook of the Development of Play*, 341–362.

Fram, M. S., & Kim, J. (2012). Segregated from the start: Peer context in center-based child care. *Children & Schools*, *34*(2), 71–82. doi:10.1093/cs/cds011

Fukkink, R. G., & Lont, A. (2007). Does training matter? A meta-analysis and review of caregiver training studies. *Early Childhood Research Quarterly*, *22*(3), 294–311. Retrieved from http://spinusa.net/Does Training Matter_2007_Fukkink.pdf

Fuligni, A. S., Howes, C., Huang, Y., Hong, S. S., & Lara-Cinisomo, S. (2012). Activity settings and daily routines in preschool classrooms: Diverse experiences in early learning settings for low-income children. *Early Childhood Research Quarterly*, *27*(2), 198–209.

Fuller, B. (2011). College credentials and caring: How teacher training could lift young children. In E. Zigler, W. S. Gilliam, & W. S. Barnett (Eds.), *The Pre-K Debates: Current controversies and issues* (p. 57). Brookes Publishing Company.

Fuller, B., Gasko, J. W., Anguiano, R., & Berkeley, U. C. (2010). *Lifting pre-K quality: Caring and effective teachers*.

Fuller, B., Livas, A., & Bridges, M. (2006). *How to expand and improve preschool in California : Ideals, evidence and policy options*.

Galindo, C., & Fuller, B. (2010). The social competence of Latino kindergartners and growth in mathematical understanding. *Developmental Psychology*, *46*(3), 579.

Gelber, A. M., & Isen, A. (2011). Children's schooling and parents' investment in children: Evidence from the Head Start Impact Study. NBER workign paper.

Gelman, A. (2005). Analysis of variance—why it is more important than ever. *The Annals of Statistics*, *33*(1), 1–53.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gerde, H. K., & Powell, D. R. (2009). Teacher education, book-reading practices, and children's language growth across one year of Head Start. *Early Education & Development*, *20*(2), 211–237. doi:10.1080/10409280802595417

Ginsburg, H. P., Lee, J. S., & Boyd, J. S. (2008). Mathematics education for young children: What it is and how to promote it. *Social Policy Report*, *22*(1).

Goelman, H. (1988). The relationship between structure and process variables in home and day care settings on children's language development. In A. R. Pence (Ed.), *Ecological research with children and families: From concepts to methodology* (pp. 16–34). New York, NY: Teachers College Press.

Goffin, S. G., & Washington, V. (2007). *Ready or not: Leadership choices in early care and education* (pp. 24–27). New York: Teachers College Press.

Goldhaber, D. (2007). *The importance of methodology in teasing out the effects of school resources on student achievement* (pp. 0–27).

Goldhader, D. D., Brewer, D. J., & Anderson, D. J. (1999). A three-way error components analysis of educational productivity. *Education Economics*, *7*(3), 199–208.

Gong, X. (2013). *The relationship between preschool teachers' B.A. attainment and child development outcomes: A review of the literature*. Unpublished manuscript. Teachers College, Columbia University.

Gormley, W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology*, *41*(6), 872.

Hair, J. F., Black, B., Anderson, R., & Tatham, R. (1995). *Multivariate data analysis: Text and readings*. MacMillan, International Edition.

Hamre, B., Hatfield, B., Pianta, R., & Jamil, F. (2013). Evidence for General and Domain-Specific Elements of Teacher-Child Interactions: Associations With Preschool Children's Development. *Child Development*, *85*(3), 1–18. doi:10.1111/cdev.12184

Hamre, B. K., & Bridges, M. (2004). *Early care and education staff preparation, quality, and child development: A review of the literature*. *University of California at Berkeley*.

Hamre, B. K., & Pianta, R. C. (2007). Learning opportunities in preschool and early elementary classrooms. In R. C. Pianta, M. J. Cox, & K. L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 49–84). Baltimore: Paul H Brookes Publishing.

Hamre, B. K., Pianta, R. C., Burchinal, M., Field, S., LoCasale-Crouch, J., Downer, J. T., … Scott-Little, C. (2012). A course on effective teacher-child interactions: Effects on teacher beliefs, knowledge, and bbserved practice. *American Educational Research Journal*, *49*(1), 88–123. doi:10.3102/0002831211434596

Hamre, B. K., Pianta, R. C., Mashburn, A. J., & Downer, J. T. (2007). Building a science of classrooms: Application of the CLASS framework in over 4,000 US early childhood and elementary classrooms. *Downloaded on March*, *27*, 2013.

Han, W.-J., Lee, R., & Waldfogel, J. (2012). School readiness among children of immigrants in the US: Evidence from a large national birth cohort study. *Children and Youth Services Review*, *34*(4), 771–782. doi:10.1016/j.childyouth.2012.01.001

Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early childhood environment rating scale*. Teachers College Press.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the Econometric Society*, 1251–1271.

Hawkinson, L. E., Griffen, A. S., Dong, N., & Maynard, R. A. (2013). The relationship between child care subsidies and children's cognitive development. *Early Childhood Research Quarterly*, *28*(2), 388–404.

Heckman, J. J. (2011). The economics of inequality. *American Educator*. Retrieved from http://www.aft.org/pdfs/americaneducator/spring2011/Heckman.pdf

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

Heckman, J. J., & Kautz, T. (2013). *Fostering and measuring skills: Interventions that improve chracter and cognition*.

Helburn, S. W. (1995). *Cost, quality and child outcomes in child care centers:Technical report, public report, and executive summary*. Cost, Quality, and Child Outcomes Study, Economics Department, University of Colorado at Denver. Retrieved from

http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED386297&ERICExtSearch_SearchType_0=no&accno=ED386297

Henry, G., Ponder, B., Rickman, D., Mashburn, A., Henderson, L., & Gordon, C. (2004). *An evaluation of the implementation of Georgia's pre-k program: Report of the findings from the Georgia Early Childhood Study (2002-03). Atlanta, GA: Georgia State University, Andrew Young School of Policy Studies*.

Henry, G., & Rickman, D. (2007). Do peers influene children's development in preschool? *Economics of Education Review*, *26*(1), 100-112.

Hirano, K., & Imbens, G. W. (2001). Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. *Health Services and Outcomes Research Methodology*, *2*(3-4), 259–278.

Honig, A. S., & Hirallal, A. (1998). Which counts more for excellence in childcare staff—years in service, education level or ECE coursework? *Early Child Development and Care*, *145*(1), 31–46.

Honig, A. S., & Lally, J. R. (1973). Assessing teacher behaviors with infants in day care.

Honig, A. S., & Lally, J. R. (1975). How good is your infant program? Use an observational method to find out. In *Child and Youth Care Forum* (Vol. 4, pp. 194–207). Springer.

Howes, C. (1997). Children's experiences in center-based child care as a function of teacher background and adult:child ratio. *Merrill-Palmer Quarterly*, *43*(3), 404–425. Retrieved from http://psycnet.apa.org/index.cfm?fa=search.displayrecord&uid=1997-07143-004

Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-Kindergarten programs. *Early Childhood Research Quarterly*, *23*(1), 27–50. doi:10.1016/j.ecresq.2007.05.002

Howes, C., James, J., & Ritchie, S. (2003). Pathways to effective teaching. *Early Childhood Research Quarterly*, *18*(1), 104–120. doi:10.1016/S0885-2006(03)00008-5

Howes, C., & Stewart, P. (1987). Child's play with adults, toys, and peers: An examination of family and child-care influences. *Developmental Psychology*, *23*(3), 423–430.

Howes, C., Whitebook, M., & Phillips, D. (1992). Teacher characteristics and effective teaching in child care : findings from the National Child Care Staffing Study. *Child & Youth Care Forum Youth Care Forum*, *21*(6), 399–414. doi:10.1007/BF00757371

Imbens, G., & Angrist, J. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*(2), 467–475.

Iruka, I. U., Gardner-Neblett, N., Matthews, J. S., & Winn, D.-M. C. (2014). Preschool to kindergarten transition patterns for African American boys. *Early Childhood Research Quarterly*, *29*(2), 106–117.

Jackson, R. H. (2007). *National evaluation of Early Reading First final report*. DIANE Publishing.

Justice, L. M., Kaderavek, J. N., Fan, X., Sofka, A., & Hunt, A. (2009). Accelerating preschoolers' early literacy development through classroom-based teacher–child storybook reading and explicit print referencing. *Language, Speech, and Hearing Services in Schools*, *40*(1), 67–85.

Justice, L. M., Mashburn, A. J., Hamre, B. K., & Pianta, R. C. (2008). Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Childhood Research Quarterly*, *23*(1), 51–68.

Kagan, S. L. (2009). *American early childhood education: Preventing or perpetuating inequity?*. *Research Review*.

Kagan, S. L., & Gomez, R. E. G. (2011). B.A. plus: Reconcilining reality and reach. In E. Zigler, W. S. Gilliam, & W. S. Barnett (Eds.), *The Pre-K Debates: Current controversies and issues* (pp. 68–73). Brookes Publishing Company.

Kagan, S. L., Kauerz, K., & Tarrant, K. (2008). *The early care and education teaching workforce at the fulcrum: An agenda for reform*. Teachers College Press.

Kagan, S. L., Moore, E., & Bredekamp, S. (1998). *Reconsidering children's early development and learning toward common views and vocabulary: National Education Goals Panel*. DIANE Publishing.

Kamerman, S. B., & Gatenio-Gabel, S. (2007). Early childhood education and care in the United States. *International Journal of Child Care and Education Policy*.

Kelley, P. J., & Camilli, G. (2007). *The impact of teacher education on outcomes in center-based early childhood education programs: A meta-analysis*. Retrieved from http://nieer.org/resources/research/TeacherEd.pdf

Keys, T. D., Farkas, G., Burchinal, M. R., Duncan, G. J., Vandell, D. L., Li, W., … Howes, C. (2013). Preschool center quality and school readiness: Quality effects and variation by demographic and child characteristics. *Child Development*, *84*(4), 1171–1190.

Kim, S., Chang, M., & Kim, H. (2011). Does teacher educational training help the early math skills of English language learners in Head Start? *Children and Youth Services Review*, *33*(5), 732–740. doi:10.1016/j.childyouth.2010.11.019

Klein, A., Starkey, P., Clements, D., Sarama, J., & Iyer, R. (2008). Effects of a pre-kindergarten mathematics intervention: A randomized experiment. *Journal of Research on Educational Effectiveness*, *1*(3), 155–178.

Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The big five personality traits, learning styles, and academic achievement. *Personality and Individual Differences*, *51*(4), 472–477. doi:10.1016/j.paid.2011.04.019

Kontos, S., Burchinal, M., Howes, C., Wisseh, S., & Galinsky, E. (2002). An eco-behavioral approach to examining the contextual effects of early childhood classrooms. *Early Childhood Research Quarterly*, *17*(2), 239–258.

Kontos, S., & Wilcox-Herzog, A. (1997). Influences on children's competence in early childhood classrooms. *Early Childhood Research Quarterly*, *12*(3), 247–262.

La Paro, K. M., Pianta, R. C., & Stuhlman, M. (2004). The classroom assessment scoring system: Findings from the prekindergarten year. *The Elementary School Journal*, 409–426.

Lamb, M. E. (1998). Nonparental child care: Context, quality, correlates and consequences. In *Handbook of child psychology,Vol. 4* (pp. 77–133). John Wiley & Sons Inc.

Layzer, J. I., & Goodson, B. D. (2006). The "quality" of early care and education settings: Definitional and measurement issues. *Evaluation Review*, *30*(5), 556–576.

Layzer, J. I., & Others. (1993). *Observational study of early childhood programs. Final report. Volume I: Life in preschool.* Retrieved from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED366468&ERICExtSearch_SearchType_0=no&accno=ED366468

Lee, R., Zhai, F., Brooks-Gunn, J., Han, W.-J., & Waldfogel, J. (2013). Head Start participation and school readiness: Evidence from the Early Childhood Longitudinal Study–Birth Cohort. *Developmental Psychology*.

Lee, V. E., & Burkam, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school.* ERIC.

Levin, H. M. (2013). The utility and need for incorporating noncognitive skills into large-scale educational assessments. In *The role of international large-scale assessments: Perspectives from technology, economy, and educational research* (pp. 67–86). Springer.

Lichtenberger, E. J., & Dietrich, C. (2013). The community college penalty and bachelor's degree completion: Fact or fiction? Policy Research: IERC 2013-1.*Illinois Education Research Council*.

Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreño, C., & Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology*, *46*(5), 1062.

LoCasale-Crouch, J., Konold, T., Pianta, R., Howes, C., Burchinal, M., Bryant, D., … Barbarin, O. (2007). Observed classroom quality profiles in state-funded pre-kindergarten programs and associations with teacher, program, and classroom characteristics. *Early Childhood Research Quarterly*, *22*(1), 3–17. doi:10.1016/j.ecresq.2006.05.001

Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance? *Economics of Education Review*, *26*(1), 33–51. doi:10.1016/j.econedurev.2005.09.008

Magnuson, K. A., & Votruba-Drzal, E. (2009). Enduring influences of child poverty. In *Changing poverty, changing policies* (pp. 153–179). NY: New York: Russell Sage Foundation.

Mashburn, A. J., & Pianta, R. (2010). Opportunity in early education: Improving teacher–child interactions and child outcomes. In A. J. Reynolds, A. J. Rolnick, & M. M. Englund (Eds.),

Childhood programs and practices in the first decade of life: A human capital integration (pp. 243–265). Cambridge University Press.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. a, Bryant, D., … Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, *79*(3), 732–49. doi:10.1111/j.1467-8624.2008.01154.x

Maxwell, K. L., Field, C.C. & Clifford, R. M. (2006). Quality and qualifications: Links between professional development and quality in early care and education settings. in M. Zaslow & I. Martinez-Beck (Eds.), *Critical Issues in Early Childhood Professional Development* (pp.21-48). Brookes.

McLanaham, S., Haskins, R., & Paxson, C. (2005). *School readiness: Closing racial and ethnic gaps*. Brookings Institution Press. Retrieved from http://futureofchildren.org/futureofchildren/publications/docs/15_01_FullJournal.pdf

McMullen, M. B. (2003). Acquiring and supporting developmentally appropriate beliefs and practices in early care and education professionals. In *Society for Research in Child Development Biennial Meeting, Tampa, Florida. Vartuli, S.(1999). How early childhood teacher beliefs vary across grade level. Early Childhood Research Quarterly* (Vol. 14, pp. 489–514).

McMullen, M. B., & Alat, K. (2002). Education matters in the nurturing of the beliefs of preschool caregivers and teachers. *Early Childhood Research & Practice*, *4*(2), 1–17.

McWilliam, R. A., Scarborough, A. A., Bagby, J. H., & Sweeney, A. L. (1998). *Teaching styles rating scale*. Chapel Hill, NC: Frank Porter Graham Child Development Center, University of North Carolina at Chapel Hill.

Meisels, S. J., & Atkins- Burnett, S. (2006). Evaluating early childhood assessments: A differential analysis. *Blackwell Handbook of Early Childhood Development*, 532–549.

Melhuish, E. C., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., Phan, M., & Malin, A. (2008). Preschool influences on mathematics achievement. *Science*, *321*(5893), 1161–1162.

Miguel, E., Gertler, P., & Levine, D. I. (2005). Does social capital promote industrialization? Evidence from a rapid industrializer. *Review of Economics and Statistics*, *87*(4), 754–762.

Mitchell, A. (2001). *Prekindergarten programs in the states: Trends and issues*. ERIC Clearinghouse.

Moiduddin, E., Aikens, N., Tarullo, L., West, J., & Xue, Y. (2012). Child Outcomes and Classroom Quality in FACES 2009. OPRE Report 2012-37a. *Administration for Children & Families*.

Morgan, S. L., & Todd, J. J. (2008). A diagnostic routine for the detection of consequential heterogeneity of causal effects. *Sociological Methodology*, *38*(1), 231–281.

Murnane, R., & Willett, J. (2011). *Methods matter: Improving causal inference in educational and social science research*. Oxford University Press.

Myers, J. L., Well, A. D., & Lorch, R. F. (2010). *Research design and statistical analysis* (3rd ed.). New York, NY: Routledge.

NAEYC. (2007). *Educational qualifications of program administrators and teaching staff: Building better futures for children and the profession.* Retrieved from http://www.naeyc.org/files/yc/file/200703/BTJProfDev.pdf

NAEYC. (2009). *Standards for early childhood professional preparation*. ERIC. Retrieved from http://www.naeyc.org/files/naeyc/file/positions/ProfPrepStandards09.pdf

NAEYC. (2011). *Early childhood education professional development: Training and technical assistance glossary*.

Najarian, M., Snow, K., Lennon, J., Kinsey, S., & Mulligan, G. (2010). Early childhood longitudinal study, birth cohort (ECLS-B). *Preschool-Kindergarten 2007 Psychometric Report*, 2009–2010.

National Research Council. (2012a). *The early childhood care and education workforce: Challenges and opportunities- a workshop report*. Washington, DC: The National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=13238

National Research Council. (2012b). *The Early Childhood Care and Education Workforce: Challenges and Opportunities: A Workshop Report*. The National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=13238

NICHD, & Duncan, G. J. (2003). Modeling the impacts of child care quality on children's preschool cognitive development. *Child Development*, *74*(5), 1454–1475. doi:10.1111/1467-8624.00617

NICHD Early Child Care Research Network. (1999). Child Outcomes When Child Care Center Classes Meet Recommended Standards for Quality. *American Journal of Public Health*, *89*(7), 1072–1077. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1508829/

NICHD Early Child Care Research Network. (2000). Characteristics and quality of child care for toddlers and preschoolers. *Applied Developmental Science*, *4*(3), 116. doi:10.1207/S1532480XADS0403_2

NICHD Early Child Care Research Network. (2002). Child-care structure -> process -> outcome: Direct and indirect effects of child-care quality on young children's development. *Psychological Science*, *13*(3), 199–206. doi:10.1111/1467-9280.00438

NICHD Early Child Care Research Network. (2003). Does quality of child care affect child outcomes at age 4 ½? *Developmental Psychology*, *39*(3), 451–469.

Nores, M., & Barnett, W. S. (2010). Benefits of early childhood interventions across the world: (Under) Investing in the very young. *Economics of Education Review*, *29*(2), 271–282. doi:10.1016/j.econedurev.2009.09.001

O'Connor, E., & McCartney, K. (2007). Examining teacher–child relationships and achievement as part of an ecological model of development. *American Educational Research Journal*, *44*(2), 340–369.

Pearson, K. M. (2013). *Attachment and Self Regulation in Preschool Age Children*. University of Rhode Island.

Peisner-Feinberg, E. S., & Burchinal, M. R. (1997). Relations between preschool children's child-care experiences and concurrent development: The Cost, Quality, and Outcomes Study. *Merrill-Palmer Quarterly*, *43*(3), 451–477.

Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Yazejian, N., Culkin, M. L., Zelazo, J., & Howes, C. (1999). *Children of the Cost, Quality, and Outcomes Study go to school*. Chapel Hill, NC.

Peisner-Feinberg, E. S., & Yazejian, N. (2010). Research on program quality: The evidence base. In V. Buysse & W. P. Wesley (Eds.), *The quest for quality: Promising innovations for early childhood programs* (pp. 21–45). Baltimore, MD: Brookes.

Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., & Yazejian, N. (2001). The relation of preschool child- care quality to children's cognitive and social developmental trajectories through second grade. *Child Development*, *72*(5), 1534–1553.

Phillips, D. (1987). *Quality in child care: What does research tell us?* (Vol. 1). National Association for the Education of Young Children Washington, DC.

Phillips, D., Mekos, D., Scarr, S., McCartney, K., & Abbott–Shim, M. (2000). Within and beyond the classroom door: assessing quality in child care centers. *Early Childhood Research Quarterly*, *15*(4), 475–496. doi:10.1016/S0885-2006(01)00077-1

Phillipsen, L. C., Burchinal, M. R., Howes, C., & Cryer, D. (1997). The prediction of process quality from structural features of child care. *Early Childhood Research Quarterly*, *12*(3), 281–303. doi:10.1016/S0885-2006(97)90004-1

Pianta, R. C. (2011). A degree is not enough: Teachers need stronger and more individualized professional development supports to be effective in the classroom. In E. Zigler, W. S. Gilliam, & W. S. Barnett (Eds.), *The Pre-K Debates: Current controversies and issues* (pp. 64–68). Brookes Publishing Company.

Pianta R. C., LaParo, M. K., & Hamre B. K. (2008). *Classroom Assessment Scoring System (CLASS)*. Baltimore, MD: Brookes.

Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, *9*(3), 144–159. doi:10.1207/s1532480xads0903_2

Rispoli, K. M., McGoey, K. E., Koziol, N. A., & Schreiber, J. B. (2013). The relation of parenting, child temperament, and attachment security in early childhood to social competence at school entry. *Journal of School Psychology*, *51*(5), 643–658.

Rohacek, M., Adams, G. C., & Kisker, E. E. (2010). *Understanding quality in context: Child care centers, communities, markets, and public policy*. Urban Institute Washington, DC.

Rosenbaum, P. R. (1987). Sensitivity analysis for certain permutation inferences in matched observational studies. *Biometrika, 74*(1), 13–26.

Rubin, D. B., & Thomas, N. (2000). Combining propensity score matching with additional adjustments for prognostic covariates. *Journal of the American Statistical Association, 95*(450), 573–585.

Sabol, T. J., & Pianta, R. C. (2014). Patterns of school readiness forecast achievement and socioemotional development at the end of elementary school. *Child Development, 83*(1), 282–99. doi:10.1111/j.1467-8624.2011.01678.x

Sameroff, A. (2010). A unified theory of development: a dialectic integration of nature and nurture. *Child Development, 81*(1), 6–22. doi:10.1111/j.1467-8624.2009.01378.x

Scarr, S., Eisenberg, M., & Deater-Deckard, K. (1994). Measurement of quality in child care centers. *Early Childhood Research Quarterly, 9*(2), 131–151. doi:10.1016/0885-2006(94)90002-7

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neuroscience & Biobehavioral Reviews, 30*(6), 855–863.

Snider, M. H., & Fu, V. R. (1990). The effects of specialized education and job experience on early childhood teachers' knowledge of developmentally appropriate practice. *Early Childhood Research Quarterly, 5*(1), 69–78. doi:10.1016/0885-2006(90)90007-N

Snyder, T. D., Dillow, S. A., & Hoffman, C. M. (2008). Digest of Education Statistics, 2007. NCES 2008-022. *National Center for Education Statistics*.

Staiger, D. O., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica, 65*(3), 557–586.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics, 25*(1), 1.

Temin, P. (2003). Low pay, low quality. *Education Next*, 8–13. Retrieved from http://educationnext.org/low-pay-low-quality/

Thiel, H., & Thomsen, S. L. (2011). *Noncognitive skills in economics: Models, measurement, and empirical evidence*. ZEW Discussion Papers. Retrieved from http://www.econstor.eu/bitstream/10419/54966/1/684245213.pdf

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal, 113*(485), 3-33.

Tout, K., & Maxwell, K. L. (2010). Quality rating and improvement systems: Achieving the promise for programs, parents, children, and early childhood systems. In P. W. Wesley & V. Buysse (Eds.), *The quest for quality: Promising innovations for early childhood programs* (pp. 91–112). Brookes.

Tout, K., Zaslow, M., & Berry, D. (2006). Quality and qualifications: Links between professional development and quality in early care and education settings. in M. Zaslow & I. Martinez-Beck (Eds.), *Critical Issues in Early Childhood Professional Development* (pp.77–110). Brookes.

Tout, K., Zaslow, M., Halle, T., & Forry, N. (2009). Issues for the next decade of quality rating and improvement systems. *Washington, DC: Child Trends*.

Vandell, D. (2004). Early child care: The known and the unknown. *Merrill-Palmer Quarterly*, *50*(3), 387–414.

Vandell, D. L., Henderson, V. K., & Wilson, K. S. (1988). A longitudinal study of children with day-care experiences of varying quality. *Child Development*, *59*(5), 1286–1292.

Vandell, D. L., & Powers, C. P. (1983). Day care quality and children's free play activities. *American Journal of Orthopsychiatry*, *53*(3), 493–500. doi:10.1111/j.1939-0025.1983.tb03393.x

Vandell, D., & Wolfe, B. (2000). *Child care quality: Does it matter and does it need to be improved?* (Vol. 78). Institute for Research on Poverty Madison, WI.

Vu, J. A., Jeon, H.-J., & Howes, C. (2008). Formal education, credential, or both: Early childhood program classroom practices. *Early Education & Development*, *19*(3), 479–504. doi:10.1080/10409280802065379

Walters, C. (2014). *Inputs in the production of early childhood human capital: Evidence from Head Start*.

What Works Clearinghouse. (2013). *What Works Clearinghouse Procedures and Standards Handbook (Version 3.0)*. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_draft_standards_handbook.pdf

Whitebook, M. (2003). Early education quality: Higher teacher qualifications for better learning environments. A review of the literature. Retrieved from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED481219&ERICExtSearch_SearchType_0=no&accno=ED481219

Whitebook, M., Gomby, D., Bellm, D., Sakai, L., & Kipnis, F. (2009). *Preparing teachers of young children: The current state of knowledge, and a blueprint for the future*. Retrieved from http://www.wcstonefnd.org/wp-content/uploads/2012/05/teacher_prep_2.pdf

Whitebook M., Howes, C., & Phillips, D. A. (1989). *Who cares? Child Care teachers and sthe quality of care in America.* Retrieved from http://www.eric.ed.gov/PDFS/ED323031.pdf

Whitebook, M., & Ryan, S. (2011). *Degree in context: Asking the right questions about preparing skilled and effective teachers of young children*. New Brunswick, NJ.

Whitebook, M., & Ryan, S. (2011). *Degrees in Context: Asking the Right Questions about Preparing Skilled and Effective Teachers of Young Children. Preschool Policy Brief. Issue 22. National Institute for Early Education Research*. ERIC. Retrieved from http://nieer.org/resources/policybriefs/23.pdf

Whitebook, M., Sakai, L., Gerber, E., & Howes, C. (2001). *Then & now: Changes in child care staffing, 1994-2000. Technical report.* ERIC. Retrieved from

http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED452984&ERICExtSearch_SearchType_0=no&accno=ED452984

Winton, P., & Buysse, V. (2005). How is the pre-k day spent? NCEDL prekindergarten study. *Early Developments*, *9*(1), 22–28.

Wolfgang, C., Stannard, L., & Jones, I. (2003). Advanced constructional play with LEGOs among preschoolers as a predictor of later school achievement in mathematics. *Early Child Development and Care*, *173*(5), 467–475.

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., … Zaslow, M. J. (2013). Investing in our future: The evidence base on preschool education. Retrieved from http://fcd-us.org/sites/default/files/Evidence Base on Preschool Education FINAL.pdf

Zigler, E., Gilliam, W. S., & Barnett, W. S. (2011). *The Pre-K debates: Current controversies and issues.* Batimore, MD: Brookes Publishing Company.

Zucker, T. A., Cabell, S. Q., Justice, L. M., Pentimonti, J. M., & Kaderavek, J. N. (2013). The role of frequent, interactive prekindergarten shared reading in the longitudinal development of language and literacy skills. *Developmental Psychology*, *49*(8), 1425.

APPENDICES

Appendix A. The list of the most relevant empirical studies reviewed in Chapter 2

Part 1: The B.A. studies for child development outcomes
(**12** studies with the B.A. dummy measures are listed here; the rest12 studies are not presented here given page limitations, but available upon request)

(1) Vandell, D. L., & Powers, C. P. (1983). Day care quality and children's free play activities. *American Journal of Orthopsychiatry*, *53*(3), 493–500. doi:10.1111/j.1939-0025.1983.tb03393.x

(2) Vandell, D. L., Henderson, V. K., & Wilson, K. S. (1988). A longitudinal study of children with day-care experiences of varying quality. *Child Development*, *59*(5), 1286–1292.

(3) Layzer, J. I., & Others. (1993). *Observational study of early childhood programs. Final report. Volume I: Life in preschool.*

(4) Helburn, S. W. (1995). *Cost, quality and child outcomes in child care centers:Technical report, public report, and executive summary*. Cost, Quality, and Child Outcomes Study, Economics Department, Campus Box 159, P.O. Box 173364, University of Colorado at Denver, Denver, CO.

(5) Burchinal, M. R., Roberts, J. E., Riggins, J., Zeisel, S. A., Neebe, E., & Bryant, D. (2000). Relating Quality of Center-Based Child Care to Early Cognitive and Language Development Longitudinally. *Child Development*, *71*(2), 339–357. doi:10.1111/1467-8624.00149

(6) Burchinal, M. R., Cryer, D., Clifford, R. M., & Howes, C. (2002). Caregiver training and classroom quality in child care centers. *Applied Developmental Science*, *6*(1), 2–11. doi:10.1207/S1532480XADS0601_01

(7) Henry, G., Ponder, B., Rickman, D., Mashburn, A., Henderson, L., & Gordon, C. (2004). *An evaluation of the implementation of Georgia's pre-k program: Report of the findings from the Georgia Early Childhood Study (2002-03)*. *Atlanta, GA: Georgia State University, Andrew Young School of Policy Studies*.

(8) Early, D. M., Bryant, D. M., Pianta, R. C., Clifford, R. M., Burchinal, M. R., Ritchie, S., … Barbarin, O. (2006). Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly*, *21*(2), 174–195. doi:10.1016/j.ecresq.2006.04.004

(9) Early, D. M., Maxwell, K. L., Alva, S., Bender, R. H., Bryant, D., Cai, K., … Zill, N. (2007). Teachers' education, classroom quality , and young children's academic skills: results from seven studies of preschool programs. *Child Development*, *78*(2), 558–580. doi:10.1111/j.1467-8624.2007.01014.x

(10) Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. a, Bryant, D., … Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child development*, *79*(3), 732–49. doi:10.1111/j.1467-8624.2008.01154.x

(11) Choi, J. Y. (2011). *Teacher qualifications and preschoolers' math outcomes: Mediation by teacher-child relationships? ETD Collection for Purdue University*. Retrieved from http://docs.lib.purdue.edu/dissertations/AAI1510057

(12) Kim, S., Chang, M., & Kim, H. (2011). Does teacher educational training help the early math skills of English language learners in Head Start? *Children and Youth Services Review*, *33*(5), 732–740. doi:10.1016/j.childyouth.2010.11.019

Part 2: The B.A. studies for teacher-child interactions
(**14** studies that used the B.A. dummy measures are listed here; the other studies are not listed)

(1) Arnett, J. (1989). Caregivers in day-care centers: Does training matter? *Journal of Applied Developmental Psychology*, *10*(4), 541–552. doi:10.1016/0193-3973(89)90026-9

(2) Whitebook et al. (1989). *Who cares? Child Care teachers and the quality of care in America.* Retrieved from http://www.eric.ed.gov/PDFS/ED323031.pdf

(3) Howes, C., Whitebook, M., & Phillips, D. (1992). Teacher characteristics and effective teaching in child care : findings from the National Child Care Staffing Study. *Child & Youth Care Forum Youth Care Forum*, *21*(6), 399–414. doi:10.1007/BF00757371

(4) Layzer, J. I., & Others. (1993). *Observational study of early childhood programs. Final report. Volume I: Life in preschool*

(5) Howes, C. (1997). Children's experiences in center-based child care as a function of teacher background and adult: child ratio. *Merrill-Palmer Quarterly*, *43*(3), 404–425. Retrieved from http://psycnet.apa.org/index.cfm?fa=search.displayrecord&uid=1997-07143-004

(6) Phillipsen, L. C., Burchinal, M. R., Howes, C., & Cryer, D. (1997). The prediction of process quality from structural features of child care. *Early Childhood Research Quarterly*, *12*(3), 281–303. doi:10.1016/S0885-2006(97)90004-1

(7) Blau, D. M. (1997). The production of quality in child care centers. *Journal of Human Resources*, *32*(2), 354–387.

(8) Honig, A. S., & Hirallal, A. (1998). Which counts more for excellence in childcare staff: Years in service, education level or ECE coursework? *Early Child Development and Care*, *145*(1), 31–46. doi:10.1080/0300443981450103

(9) De Kruif, R. E. ., McWilliam, R. ., Ridley, S. M., & Wakely, M. B. (2000). Classification of teachers' interaction behaviors in early childhood classrooms. *Early Childhood Research Quarterly*, *15*(2), 247–268. doi:10.1016/S0885-2006(00)00051-X

(10) Burchinal, M. R., Cryer, D., Clifford, R. M., & Howes, C. (2002). Caregiver training and classroom quality in child care centers. *Applied Developmental Science*, *6*(1), 2–11. doi:10.1207/S1532480XADS0601_01

(11) Howes, C., James, J., & Ritchie, S. (2003). Pathways to effective teaching. *Early Childhood Research Quarterly*, *18*(1), 104–120. doi:10.1016/S0885-2006(03)00008-5

(12) Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, *9*(3), 144–159. doi:10.1207/s1532480xads0903_2

(13) Early, D. M., Bryant, D. M., Pianta, R. C., Clifford, R. M., Burchinal, M. R., Ritchie, S., … Barbarin, O. (2006). Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly*, *21*(2), 174–195. doi:10.1016/j.ecresq.2006.04.004

(14) Vu, J. A., Jeon, H.-J., & Howes, C. (2008). Formal education, credential, or both: Early childhood program classroom practices. *Early Education & Development*, *19*(3), 479–504. doi:10.1080/10409280802065379

Appendix B. Additional tables for the dissertation

Table B1 Correlation matrix for all the predictors in the model of OLS with rich controls

Panel 1:

Variables	B.A. or above	zmental_age2	zsocial_age2	age4_m	boy	AfricanA	Hispanic	Asian	NativeA	Mrace	low_bweight	days_inp
B.A. or above	1											
zmental_age2	-0.0422	1										
zsocial_age2	0.0019	-0.3112	1									
age4_m	0.0937	-0.0178	0.013	1								
boy	0.0378	-0.1741	0.0579	0.0201	1							
AfricanA	-0.031	-0.1284	-0.021	-0.0141	-0.0147	1						
Hispanic	0.0001	-0.1946	0.0412	0.0849	0.0125	-0.2148	1					
Asian	0.0281	-0.0047	0.0255	0.0139	0.0062	-0.0622	-0.077	1				
NativeA	-0.0112	-0.0233	0.0199	0.0063	0.0025	-0.0264	-0.0327	-0.0095	1			
Mrace	0.0042	0.0035	0.0153	0.0065	0.0137	-0.0782	-0.0968	-0.028	-0.0119	1		
low_bweight	0.0063	-0.1007	0.0342	0.0008	-0.0061	0.05	-0.002	-0.0087	0.001	0.0115	1	
days_inp	-0.0382	-0.1318	0.0309	0.0985	0.0104	0.2365	0.1659	0.0163	0.0152	-0.0003	0.0216	1
prior2_center	-0.0589	0.1133	-0.0189	-0.0576	0.0362	0.0823	-0.0987	-0.0114	-0.0016	0.0172	-0.0034	0.158
m_age	0.0737	0.1063	-0.0482	-0.086	-0.0179	-0.1576	-0.1422	0.0541	-0.0235	-0.0007	-0.0004	-0.1925
m_BAplus	0.1138	0.2236	-0.0665	-0.0739	-0.0271	-0.1681	-0.2139	0.1179	-0.0291	0.0099	-0.0138	-0.1352
m_depressed	0.0051	-0.1176	0.0249	0.0175	0.0285	0.0686	0.0061	0.006	0.0084	0.0327	0.0218	0.0502
m_married	0.0085	0.1772	-0.0362	-0.0816	-0.0321	-0.3623	-0.1475	0.0842	-0.0338	-0.032	-0.0238	-0.2415
fam_income	0.0476	0.288	-0.0723	-0.0755	-0.005	-0.3468	-0.2384	0.0866	-0.0373	-0.0072	-0.019	-0.2934
n_sibling	0.0501	-0.0788	0.0118	-0.0125	-0.0039	0.0112	-0.0164	-0.0504	0.0147	-0.0052	0.0013	-0.118
n23_books	0.0574	0.2087	-0.047	-0.0452	-0.0245	-0.2354	-0.239	-0.045	-0.0193	0.0111	-0.0167	-0.2834
hl_nonEng	0.0178	-0.2165	0.0719	0.0832	0.028	-0.1481	0.5916	0.2409	-0.0205	-0.0651	-0.0162	0.1763

Variables	BA_plus	zmental_age2	zsocial_age2	age4_m	boy	AfricanA	Hispanic	Asian	NativeA	Mrace	low_bweight	days_inp
zq_parenting	0.0859	0.1582	0.0072	-0.0075	-0.036	-0.2551	-0.0182	0.0253	-0.0213	0.0004	-0.0268	-0.2006
t_male	0.0075	0.0343	-0.0024	0.02	0.0114	0.0032	0.0338	0.0071	0.0139	-0.0081	0.01	0.0295
t_another_race	0.0294	-0.125	0.037	0.0528	0.0296	0.1173	0.3663	0.1837	0.0976	0.2889	0.0168	0.1695
experience	0.0152	0.0133	-0.006	0.0331	-0.0058	0.0013	0.0103	0.0104	0.0085	0.0389	0.0121	-0.0616
training	-0.0221	-0.0423	0.0003	0.0031	-0.0022	0.0851	0.012	-0.0126	0.0034	0.0318	0.0077	0.0987
CDA	-0.2417	-0.0397	0.0217	-0.0062	-0.0066	0.0538	0.0839	-0.0006	0.0121	0.0022	-0.0081	0.1177
other_cert	0.2435	-0.0484	0.0142	0.046	0.0258	0.0392	0.0466	0.0227	-0.0007	-0.0054	0.0202	0.1142
history_m	-0.0433	-0.0057	0.0255	0.1029	-0.0174	0.0128	0.0216	0.0062	0.0003	0.0224	-0.0028	0.0795
hours_care	-0.0744	-0.0611	0.0075	0.045	-0.0205	0.2554	0.0227	0.0114	0.0128	0.0159	0.0113	0.5459
group_size	0.0292	0.0343	-0.0179	0.0221	-0.0293	-0.0478	0.0712	0.017	0.0116	0.0041	-0.0184	0.0753
ca_ratio	0.036	0.0197	-0.0327	-0.0073	-0.0467	-0.0383	0.0313	-0.0083	-0.0011	-0.0288	-0.0226	-0.0251
c_n_books	0.0994	0.0447	-0.0063	0.0048	-0.044	-0.0118	-0.0302	-0.0127	-0.003	-0.0044	-0.0049	-0.0409
c_n_interest	0.0121	-0.0517	0.0141	-0.0056	0.0198	0.1103	0.0811	-0.0349	-0.0103	0.0105	0.001	0.1771
c_ager2plus	-0.0375	-0.0419	0.0048	0.0211	0.0181	-0.0019	0.0087	0.0153	0.0258	0.0822	0.0091	0.0349
HS_center	-0.1014	-0.1785	0.0264	0.0017	0.0289	0.1943	0.2301	-0.0571	0.0359	-0.0131	0.013	0.2068
state_preK	0.2626	-0.1167	0.0369	0.1064	-0.01	0.0517	0.0399	-0.0201	0.0185	0.0064	0.0134	0.1591
partially_private	0.0315	0.0544	0.0276	0.0247	-0.0124	-0.0995	-0.0464	0.042	-0.0111	-0.0234	-0.0035	-0.1132
for_profit	-0.1562	0.1129	-0.0157	-0.0562	0.0027	-0.0348	-0.1007	0.0245	-0.0164	0.0408	-0.0061	-0.005
p_licensed	-0.1566	0.0036	0.0234	-0.0143	-0.0101	0.0874	0.0463	0.0271	-0.0216	0.027	-0.0169	0.1123
full_day	-0.1056	0.0166	-0.0261	0.0036	-0.0287	0.2483	-0.0058	0.0099	-0.002	0.0015	0.0019	0.4427
hp_connection	0.0993	0.0495	-0.0077	0.0676	0.0162	0.0115	0.0038	0.0204	-0.0019	0.0178	-0.0028	0
Northest	0.1507	0.0554	0.0924	0.0079	0.0281	-0.0152	0.0161	0.0154	-0.0197	-0.0093	-0.008	0.0022
Midwest	0.0419	0.0219	-0.0582	-0.0458	-0.0208	-0.0684	-0.1544	-0.0226	-0.0086	0.0206	-0.0025	-0.2124
West	-0.1257	0.022	-0.023	0.0334	-0.0081	-0.1483	0.1961	0.0704	0.0162	0.0386	-0.0148	-0.1301

Panel 2:

Variables	prior2_center	m_age	m_BAplus	m_depression	m_married	fam_income	n_sibling	n23_books	hl_nonEnglish	zq_parenting	t_male
prior2_center	1										
m_age	0.0537	1									
m_BAplus	0.1104	0.4086	1								
m_depressed	-0.0024	-0.0571	-0.1288	1							
m_married	-0.0275	0.4549	0.3681	-0.0867	1						
fam_income	0.0797	0.5183	0.5464	-0.2094	0.5379	1					
n_sibling	-0.0999	0.1813	-0.0079	0.0393	0.1168	-0.0365	1				
n23_books	-0.0001	0.2708	0.3108	-0.0984	0.2882	0.4085	0.1252	1			
hl_nonEng	-0.095	-0.0477	-0.1108	-0.0145	-0.0353	-0.1856	-0.0149	-0.2665	1		
zq_parenting	-0.0103	0.2702	0.2746	-0.1188	0.2165	0.3249	-0.0372	0.2354	0.0168	1	
t_male	0.0089	-0.0067	-0.0067	0.0124	-0.0264	-0.0264	0.0209	-0.0257	0.022	0.0005	1
t_another_race	0.0141	-0.0529	-0.0926	0.0132	-0.1368	-0.1285	-0.0421	-0.1639	0.2304	-0.0237	0.0115
experience	-0.0709	-0.0089	0.0196	0.0684	-0.0106	-0.0212	0.0204	0.0192	0.0256	0.0081	-0.0024
training	0.0019	-0.0409	-0.0399	0.0558	-0.103	-0.0861	0.0267	-0.0134	0.0355	-0.0534	0.0056
CDA	-0.0373	-0.0965	-0.142	0.0668	-0.1116	-0.1558	0.0044	-0.0946	0.1027	-0.0819	0.0003
other_cert	-0.0047	0.0058	-0.0241	0.0039	-0.0383	-0.0129	0.0234	-0.0314	0.046	0.0152	0.0082
history_m	0.1077	0.0132	0.0315	0.0036	-0.0033	-0.0065	-0.0114	-0.0192	0.0488	0.0202	0.0338
hours_care	0.2068	-0.1367	-0.0999	0.0352	-0.1775	-0.1865	-0.1458	-0.1729	0.0537	-0.175	0.0182
group_size	-0.0383	-0.0299	-0.0105	0.0339	0.0236	-0.0251	-0.0041	-0.0049	0.069	0.0179	0.024
ca_ratio	-0.0585	-0.0289	-0.0322	0.0006	0.0467	0.003	0.02	0.0029	0.0366	-0.0075	0.0336
c_n_books	-0.0144	-0.0021	-0.0202	-0.0444	0.0019	-0.0136	-0.0004	0.0206	-0.0357	0.0287	-0.0025
c_n_interest	-0.0096	-0.1003	-0.0866	0.0628	-0.1566	-0.1591	-0.0137	-0.0984	0.0698	-0.0488	0.0596
c_ager2plus	0.0027	-0.0828	-0.058	0.0575	-0.086	-0.1162	-0.0267	-0.0569	0.0056	-0.0029	0.0156

Variables	prior2_center	m_age	m_BAplus	m_depression	m_married	fam_income	n_sibling	n23_books	hl_nonEnglish	zq_parenting	t_male
HS_center	-0.0853	-0.2645	-0.3211	0.1093	-0.2586	-0.4785	0.0803	-0.2463	0.1883	-0.1799	-0.0174
state_preK	-0.0454	-0.0846	-0.1164	0.0252	-0.0904	-0.1413	0.0466	-0.074	0.0371	-0.0477	0.0392
partially_private	-0.0311	0.0898	0.1056	-0.0228	0.1006	0.15	-0.035	0.1246	-0.0313	0.0764	0.0209
for_profit	0.1598	0.0578	0.095	-0.1	0.0485	0.1904	-0.0902	0.0289	-0.0725	0.0465	-0.0192
p_licensed	0.0515	-0.0361	-0.0317	-0.004	-0.0686	-0.0423	-0.0568	-0.0503	0.0716	-0.0325	-0.0033
full_day	0.1983	-0.1045	-0.0508	-0.0197	-0.1788	-0.0944	-0.1386	-0.1497	-0.0171	-0.1286	0.0035
hp_connection	0.0102	0.0954	0.1507	-0.0095	0.0821	0.1105	0.0038	0.1064	0.0448	0.1442	0.0062
Northest	-0.0275	0.081	0.0978	-0.0056	0.0032	0.0781	-0.0156	0.0083	0.0635	0.1313	0.0553
Midwest	-0.0103	0.0098	0.0341	-0.003	0.0428	0.0594	0.0361	0.1266	-0.1205	0.0291	-0.0213
West	-0.0976	0.0747	-0.0512	0.0179	0.0292	-0.0025	0.0709	-0.0368	0.1434	0.0956	0.0268

Panel 3:

Variables	t_another_race	experience	training	CDA	other_cert	history_m	hours_care	group_size	ca_ratio	c_n_books	c_n_interest	c_ager2plus
t_another_race	1											
experience	0.0015	1										
training	0.0294	0.0625	1									
CDA	-0.006	0.1598	0.1054	1								
other_cert	0.0831	-0.011	0.0248	-0.0343	1							
history_m	0.027	0.0426	-0.0116	0.0518	0.0302	1						
hours_care	0.1388	-0.074	0.0619	0.0754	0.0643	0.0679	1					
group_size	0.0187	0.1072	0.0633	0.0693	-0.0004	-0.018	0.0751	1				
ca_ratio	-0.0378	0.0236	-0.0414	0.0077	-0.028	-0.0742	-0.0406	0.5922	1			
c_n_books	-0.0194	0.0742	-0.0124	-0.0588	0.0032	0.0508	-0.0298	0.0089	0.0077	1		
c_n_interest	0.0618	0.1153	0.2417	0.17	0.0622	-0.0031	0.1418	0.1535	-0.0301	0.0422	1	
c_ager2plus	0.0589	0.0488	-0.0055	0.089	0.0185	0.1738	0.0041	0.0304	-0.0384	-0.019	0.0101	1
HS_center	0.0781	0.0289	0.1038	0.2645	0.0294	-0.0464	0.1004	0.0854	0.0081	-0.0664	0.2239	0.1371
state_preK	0.0442	-0.0447	0.0322	-0.1187	0.0967	-0.0458	0.0405	0.0632	0.1206	0.0828	0.0738	-0.0645
partially_private	-0.0733	0.0753	-0.0106	-0.0677	-0.0096	0.0599	-0.0617	-0.0718	-0.0825	0.0278	-0.1537	-0.0018
for_profit	0.0334	-0.0594	-0.0658	-0.0138	-0.0356	0.059	0.092	-0.0191	-0.0251	-0.0435	-0.0318	0.019
p_licensed	0.0699	-0.0017	0.1464	0.1524	0.0016	0.0233	0.1508	0.0293	-0.0455	-0.0489	0.2104	0.0201
full_day	0.1054	-0.0774	0.0394	0.023	0.0508	0.0835	0.515	0.0113	-0.0664	-0.0487	0.0668	-0.023
hp_connection	-0.0232	0.0023	0.0454	0.0151	0.038	0.1733	-0.0061	0.0516	-0.0095	0.0202	0.0901	0.0043
Northest	0.0088	0.0128	0.025	-0.0646	0.1009	0.025	0.0059	-0.0472	-0.1	-0.0188	0.08	0.0188
Midwest	-0.108	-0.0141	-0.0089	-0.083	-0.0586	-0.0299	-0.1636	-0.0744	-0.0413	0.0672	-0.038	0.0536
West	0.0621	0.0853	-0.0456	0.0946	-0.0374	0.0516	-0.1782	0.0961	0.0356	0.0348	-0.0529	0.0989

Panel 4:

	HS_center	state_preK	partial_private	for_profit	p_licensed	full_day	hp_connection	Northest	Midwest	West
HS_center	1									
state_preK	-0.2168	1								
partially_private	-0.2929	-0.2399	1							
for_profit	-0.2179	-0.2416	-0.3264	1						
p_licensed	0.1489	-0.1617	-0.063	0.1724	1					
full_day	-0.0048	0.0046	-0.0601	0.1783	0.1655	1				
hp_connection	0.0369	-0.0735	0.0466	-0.1043	-0.0323	-0.0209	1			
Northest	-0.0654	0.0112	0.0751	0.0047	0.0334	0.0106	0.0515	1		
Midwest	-0.0522	-0.0735	0.0276	-0.0102	-0.0336	-0.1032	0.05	-0.2422	1	
West	0.0555	-0.0835	0.064	-0.0126	0.0591	-0.122	-0.0207	-0.2356	-0.2592	1

Table B2 Descriptive features of teacher-reported non-cognitive outcomes

Outcomes	B.A. teachers		Non-B.A. teachers		Mean differences (t test)
	Mean	Std. Dev.	Mean	Std. Dev.	
Age 4					
Social competence: teacher reported	0.021	0.989	0.088	0.957	-0.067*
Externalizing behavior problems: teacher reported	-0.012	0.971	-0.001	1.017	-0.011
Internalizing behavior problems: teacher reported	0.033	1.009	-0.059	0.969	0.092**
Approaches to learning(ATL) skills: teacher reported	0.004	0.993	0.041	0.946	-0.037
Age 5					
Social competence: parent reported	0.044	0.978	0.091	0.985	-0.048
Externalizing behavior problems: teacher reported	0.007	0.997	0.020	1.014	-0.013
Internalizing behavior problems: teacher reported	-0.015	1.013	-0.052	0.973	0.036
ATL: teacher reported	0.019	0.992	0.045	0.985	-0.026

*** p<0.01, ** p<0.05, * p<0.1, and standard errors are in parentheses.

Table B3 The unweighted B.A. estimates for PSM

Outcome measures: age 4		N (rounded)	PSM (not regression adjusted): ATT	IPTW(regression without covariates): ATE
(1)	Early reading	4150	0.036 (0.027)	0.022 (0.049)
(2)	Expressive language	4100	0.033 (0.071)	0.184 (0.163)
(3)	Math	4100	0.062 (0.068)	0.029 (0.061)
(4)	Color recognition	4150	0.008 (0.064)	0.109 (0.104)
(5)	Social competence: parent-reported	4250	-0.030* (0.086)	0.025 (0.085)
(6)	Externalizing behavior problem: parent-reported	4150	0.101 (0.071)	0.260 (0.171)
(7)	Internalizing behavior (worries and unhappy): parent-reported	4300	0.104 (0.072)	0.032 (0.052)
(8)	ATL (approaches to learning): parent-reported	4300	0.030 (0.067)	0.011 (0.064)

Note. This table is made to be compared with results in Table 5-5 of Chapter 5. The difference reflects the need to add sample weight in all models.

*** p<0.01, ** p<0.05, * p<0.1, and standard errors are in parentheses.

Table B4 Full estimates for the first stage for the IV model of the early reading outcome

Variables		Coef. (s.e.)
Outcome variable for the 1st stage IV: having a B.A. teacher		
The IV	percentage of B.A. degree holders in the early childhood workforce	0.014***(0.002)
Child	mental score at age two (Bayley)	-0.012 (0.011)
	social score at age two (TAS HOTSPOT)	-0.011 (0.010)
	age at assessment measured in month	0.007***(0.002)
	boy	0.028 (0.019)
	African American	0.008 (0.033)
	Hispanic	-0.017 (0.036)
	Asian	0.018 (0.044)
	Native American	-0.050 (0.053)
	Multiple race	0.024 (0.046)
	low birth weight	-0.010 (0.026)
	days in preschool	-0.007 (0.012)
Family	attended center before at age 2	-0.030 (0.024)
	mother's age	0.002 (0.002)
	mother has B.A. or above	0.124***(0.025)
	mother's degree of depression	0.008 (0.015)
	mother is married	-0.05**(0.025)
	family income	0.002 (0.005)
	number of siblings	0.010 (0.009)
	number of books at home	0.0001 (0.0001)
	home language is not English	0.017 (0.036)
Teacher	the quality of parenting	0.012(0.011)
	teacher is male	0.077(0.075)
	teacher has another race	0.014 (0.026)
	experience	0.001 (0.001)
	training	0.006(0.027)
	CDA	-0.177***(0.022)
	other certificates	0.214***(0.020)
	history with the child (month)	-0.004**(0.001)
Classroom	hours of care per week	-0.0004(0.0009)
	group size	0.002(0.002)
	child adult ratio	0.002 (0.003)
	number of books in the classroom	0.0002***(0.00005)
	number of interest areas in the classroom	0.005 (0.006)
	classroom children's age range larger than 2	-0.017 (0.021)

Preschool	Head Start center	0.039 (0.036)
	state preK	0.312***(0.033)
	partially publicly funded	0.036 (0.030)
	for profit	-0.066** (0.030)
	the preschool is licensed	-0.076***(0.023)
	full day	-0.063 (0.024)
	home-preschool connection	0.028***(0.008)
N (rounded)		4150
R squared		0.2413

*** p<0.01, ** p<0.05, * p<0.1, and standard errors are in parentheses.

Table B5 Descriptive statistics and F values of the first stages of the two additional IVs

State	State code	Alternative IV1	Alternative IV2
		Asked for a B.A. threshold for the state PreK or not? (2005) [a]	Number of B.A. granting higher education institutes for early childhood education (2005) [b]
Alabama	1	1	31
Alaska	2	0	0
Arizona	4	0	5
Arkansas	5	0	16
California	6	0	30
Colorado	8	0	4
Connecticut	9	0	2
Delaware	10	0	1
District of Columbia	11	0	7
Florida	12	0	15
Georgia	13	0	20
Hawaii	15	0	5
Idaho	16	0	4
Illinois	17	1	48
Indiana	18	0	11
Iowa	19	0	15
Kansas	20	1	4
Kentucky	21	1	8
Louisiana	22	1	14
Maine	23	1	1
Maryland	24	1	5
Massachusetts	25	0	40
Michigan	26	1	18
Minnesota	27	0	11
Mississippi	28	0	2
Missouri	29	1	21
Montana	30	0	2
Nebraska	31	1	4
Nevada	32	1	1
New Hampshire	33	0	7
New Jersey	34	1	2
New Mexico	35	0	3
New York	36	1	47

North Carolina	37	1	7
North Dakota	38	0	6
Ohio	39	0	51
Oklahoma	40	1	23
Oregon	41	0	10
Pennsylvania	42	0	48
Rhode Island	44	0	4
South Carolina	45	1	37
South Dakota	46	0	1
Tennessee	47	1	16
Texas	48	1	28
Utah	49	0	8
Vermont	50	1	7
Virginia	51	0	1
Washington	53	0	7
West Virginia	54	0	1
Wisconsin	55	1	12
Wyoming	56	0	0
Average value of the IV		0.3921569	13.15686
Std. Dev.		0.4930895	14.2385
First stage F value		7.859	0.258

Note. (a) Data for Alternative IV1 came from the National Institute for the Early Education Research (NIEER)'s State Preschool Yearbook, 2006, which contains the statistics for the prior year 2005. For those states that didn't have a state preK program in 2005, a value of 0 for the IV is assigned. (b) Data for Alternative IV2 was obtained from the survey data files for "completions" of the Integrated Postsecondary Education Data System (IPEDS). Based on the Classification of Instructional Programs (CIP) code 2000 Classification, I define B.A.-granting institutes as those higher education institutes that offer a bachelor's degree in the following major categories: 13.1015 Education/Teaching Individuals-Early Childhood Special Education programs; 13.1210 Early childhood Education and Teaching; 13.1288 Pre-Elementary/Early Childhood/Kindergarten Teacher Education. As long as an institute had non-zero students who completed a bachelor's degree in one of these majors, this institute is viewed as a B.A.-granting institute for early childhood education. The number of such institutes in the state is IV2.

Table B6 Findings from the model that separates the B.A. treatment into two: Above B.A. and B.A. only

Outcome measures: age 4		N (rounded)	OLS with rich controls		
			Coefficient for "above B.A."	Coefficient for "B.A. only"	*R* squared
(1)	Early reading	4150	0.053 (0.055)	0.089** (0.041)	0.3603
(2)	Expressive language/ story-telling	4100	-0.089 (0.057)	-0.022 (0.045)	0.1973
(3)	Math	4100	0.103** (0.049)	0.156*** (0.038)	0.4127
(4)	Color recognition	4150	-0.119** (0.057)	-0.043 (0.044)	0.206
(5)	Social competence: parent-reported	4250	-0.173*** (0.058)	-0.062 (0.045)	0.1722
(6)	Externalizing: behavior parent-reported	4150	-0.005 (0.056)	-0.026 (0.043)	0.1993
(7)	Internalizing behavior: parent-reported	4300	-0.090 (0.061)	-0.104** (0.046)	0.0916
(8)	Approaches to learning: parent-reported	4300	-0.104* (0.057)	-0.043 (0.044)	0.1709

*** p<0.01, ** p<0.05, * p<0.1, and standard errors are in parentheses.

Table B7 The eight robustness analyses for the B.A. effects in Section 5.4 of Chapter 5

Outcome at age 4	Baseline	Robust checks									
		(1)	(2)				(3)	(4)	(5)	(6)	(7)
	Full sample	Full sample	Full sample		Subsample		Full sample	Full sample	Full sample	Full sample	Full sample
	Dummy flag	Complete cases	Add ECE major, teacher's age	Further add earnings	Subsample B.A. effect	Add racial composition for the sub sample	Remove possible outliers in 3 variables	Alternative measure for social skills at age 2	Region dummies replaced by state dummies	Combine teacher & parent reported non-cognitive measures	Use ordered probit for internal behavior instead of OLS
Early reading	0.079** (0.039)	0.076* (0.043)	0.096** (0.043)	0.094** (0.044)	0.116 (0.084)	0.116 (0.084)	0.067 (0.041)	0.078** (0.039)	0.079** (0.039)	-	-
Expressive language	-0.040 (0.042)	-0.034 (0.051)	-0.019 (0.048)	-0.003 (0.049)	-0.024 (0.085)	-0.019 (0.085)	-0.042 (0.045)	-0.039 (0.042)	-0.031 (0.043)	-	-
Math	0.142*** (0.036)	0.144*** (0.040)	0.175*** (0.040)	0.184*** (0.041)	0.149** (0.075)	0.147* (0.077)	0.142*** (0.037)	0.140*** (0.036)	0.142*** (0.036)	-	-
Color	-0.064 (0.041)	-0.043 (0.047)	-0.065 (0.048)	-0.062 (0.048)	-0.105 (0.092)	-0.113 (0.091)	-0.067 (0.044)	-0.066 (0.041)	-0.070* (0.041)	-	-
Social competence	-0.093** (0.042)	-0.072 (0.048)	-0.113** (0.048)	-0.087* (0.049)	0.058 (0.090)	0.056 (0.091)	-0.090** (0.044)	-0.093** (0.042)	-0.101** (0.042)	-0.063* (0.034)	-
Externalizing behavior	-0.020 (0.041)	-0.08* (0.044)	-0.031 (0.046)	-0.040 (0.047)	0.003 (0.090)	-0.008 (0.089)	-0.025 (0.043)	-0.021 (0.041)	-0.021 (0.041)	-0.014 (0.035)	-
Internalizing behavior	-0.100** (0.043)	-0.137*** (0.049)	-0.134*** (0.049)	-0.137*** (0.049)	-0.027 (0.087)	-0.020 (0.087)	-0.100** (0.045)	-0.099** (0.043)	-0.102** (0.044)	-0.019 (0.032)	-0.104** (0.048)
ATL	-0.060 (0.041)	-0.035 (0.047)	-0.059 (0.047)	-0.040 (0.047)	-0.044 (0.088)	-0.053 (0.089)	-0.047 (0.043)	-0.059 (0.041)	-0.062 (0.041)	-0.038 (0.032)	-

Note. The last column shows the coefficient for the B.A. dummy in the ordered probit model; the size is not directly comparable but the sign is.

Table B8 Findings of RQ1 for teacher reported non-cognitive outcomes

Teacher reported non-cognitive outcome measures	N (rounded)	OLS without any controls	OLS with rich controls	PSM	IV
Social competence	4270	-0.067* (0.041)	-0.034 (0.045)	-0.113* (0.067)	-0.190 (0.340)
Externalizing behavior problems	4270	-0.011 (0.042)	-0.012 (0.045)	0.099 (0.073)	-0.646** (0.322)
Internalizing behavior problems	4257	0.092** (0.042)	0.064 (0.046)	0.213** (0.073)	0.003 (0.339)
ATL (approaches to learning) skills	4297	-0.037 (0.040)	-0.015 (0.042)	-0.071 (0.071)	-0.032 (0.297)

*** p<0.01, ** p<.05, * p<0.1, and standard errors are in parentheses.

Table B9 Estimates for other heterogeneity factors

Part A: B.A.'s interactions with gender and mental ability at age 2

Outcome measures: age 4	Heterogeneity factors					
	Gender: boy		Race: Hispanic		Mental ability at age 2: lower than average	
	B.A.	B.A.*boy	B.A.	B.A.* Hispanic	B.A.	B.A.*lower-mental ability at age 2
Early reading	0.028 (0.053)	0.100 (0.067)	0.100** (0.043)	-0.108 (0.086)	0.087* (0.051)	-0.026 (0.064)
Math	0.075 (0.048)	0.132** (0.062)	0.150*** (0.038)	-0.042 (0.083)	0.154*** (0.044)	-0.037 (0.064)
Social competence	-0.053 (0.055)	-0.079 (0.074)	-0.089** (0.045)	-0.021 (0.099)	-0.084* (0.049)	-0.024 (0.078)
Internalizing behavior	-0.092 (0.060)	-0.016 (0.080)	-0.076 (0.047)	-0.115 (0.110)	-0.150*** (0.053)	0.135 (0.084)

Part B: B.A.'s interactions with "has a non-B.A. mother" and "in a poverty family"

Outcome measures: age 4	Heterogeneity factors			
	Non-B.A. mother		In a poverty family: below the poverty line (21.46%)	
	B.A.	B.A.*non-B.A. mom	B.A.	B.A.*below poverty line
Early reading	0.092 (0.070)	-0.020 (0.079)	-0.022 (0.116)	0.056 (0.067)
Math	0.132** (0.057)	0.014 (0.069)	0.045 (0.137)	0.055 (0.074)
Social competence	-0.054 (0.062)	-0.060 (0.078)	-0.178 (0.185)	0.046 (0.099)
Internalizing behaviour	-0.143** (0.065)	0.065 (0.082)	-0.126 (0.185)	0.014 (0.100)

Table B10 B.A. estimates in the first-step OLS for the subsample with information on teacher-child interactions

Outcome measures: age 4	OLS with rich controls		
	Sub-sample		Full sample
	N (rounded)	Coef. (s.e.)	Coef. (s.e.)
Early reading	1000	0 .131* (0.076)	0.079** (0.039)
Story telling/expressive language	1000	-0.049 (0.083)	-0.040 (0.042)
Math	1000	0.137* (0.073)	0.142*** (0.036)
Color recognition	1050	-0.130 (0.083)	-0.064 (0.041)
Social competence	1050	0.023 (0.085)	-0.093** (0.042)
Externalizing behavior : parent reported	1000	-0.016 (0.089)	-0.020 (0.041)
Internalizing behavior : parent reported	1050	-0.198** (0.086)	-0.100** (0.043)
Approaches to learning (ATL)	1050	0.041 (0.080)	-0.060 (0.041)

Note. This table was made to be compared with the B.A. estimates in Table 5-8 of Chapter 5 and to provide information for a statement on page 192 of Chapter 6 regarding the two-steps connections.