EFFECTS OF FAMILY AND SCHOOL INSTITUTIONS ON CHILD OUTCOMES FROM EARLY CHILDHOOD THROUGH ELEMENTARY SCHOOL

by
Stephanie D'Souza

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
October 2019

ABSTRACT

Children from different family backgrounds receive unequal levels of investment in their development, prompting concern about social inequality trends. Complete explanations of inequality in child outcomes require serious treatment of both family and school institutions. This dissertation follows a three-paper format unified by a common theme concerning the joint effects of the home learning environment and school institutions on outcomes in early childhood and elementary school. Data come from the Early Childhood Longitudinal Study Program.

The first paper focuses on the role of parents in tandem with preschool programs for disadvantaged children using a birth-cohort dataset. The results indicate the home environment and preschool programs are associated with kindergarten readiness when both are specified in the estimation model. The moderation analysis suggests the effect of parents' emotional support is stronger for Head Start children than parent-care children. This paper assesses the psychometric properties of home environment measures based on rich, modern instruments from child psychology over traditional self-reports. The interaction findings between home environment and preschool are thus more nuanced and convincing.

The second paper examines the elementary school stage, calling attention to the need for continuous intervention in school and family institutions for disadvantaged children, using the kindergarten-cohort dataset. The home environment and school context were associated with outcomes and evidence of an interaction was found. This paper systematically examined available information regarding school environment, constructing scales with substantive clarity and high measurement reliability. The growth curve framework allows a thorough answer to the central question regarding family and school institution interactions with school-specific random effects in shaping the growth curves.

Applying a data-integration method, the third paper combines the birth-cohort dataset collected at an earlier time with a recent kindergarten-cohort dataset to answer the same questions from the first paper. School-based cohort surveys often cover one educational stage. Yet prior exposure to favorable or harmful environments may have long-term impacts. The exploration of one data combination method in this paper moves the research forward both methodologically and substantively. The results reiterate and update the importance of including both institutions to estimate kindergarten readiness for a more recent cohort.

Dissertation Advisor: Lingxin Hao

Second Reader: Julia Burdick-Will

Committee Members: Andrew Cherlin, Joyce Epstein, and Jeffrey Grigg

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Lingxin Hao, for the tremendous learning opportunities she provided during graduate school and for the investment, support, and guidance she offered throughout the dissertation writing process. I am grateful to my second reader, Dr. Julia Burdick-Will for her insightful comments and support. I would also like to thank my committee members Dr. Andrew Cherlin, Dr. Joyce Epstein, and Dr. Jeffrey Grigg for their helpful insights. I would like to thank Dr. Stephen Morgan for his wonderful classes and feedback and Dr. Faith Connolly for valuable research opportunities. I would also like to acknowledge Sociology Department staff and graduate student friends and peers for creating a sense of community. Finally, a heartfelt thank you to my family, friends, and book club for their endless love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Paper 1 The Joint Effects of the Home Environment and Early Childhood Programs on Kindergarten Readiness: Evidence from ECLS-B

1.1 Introduction

In early childhood, wide skill gaps emerge for children from different family backgrounds. By kindergarten entry, children from very low-income families are behind in reading and math skills by about one standard deviation than children from very high-income families (Reardon 2016). Entering kindergarten at such a disadvantage has troubling implications for later stage success. This concern is amplified within a context of rising levels of inequality in family incomes (Duncan & Murnane, 2016).

The principal sources of instruction and socialization during early childhood consist of the home environment and preschool programs. Enriching home environments featuring stimulating materials and warm and responsive parenting are essential to healthy cognitive development (Bradley et al., 1989; Connell & Prinz, 2002; Cristofaro & Tamis-LeMonda, 2012). Families facing financial constraints, however, are less able to provide enriching home environments for their children because of high levels of stress and insufficient financial resources to provide stimulating learning materials and experiences. The consequences of growing up in a household without such stimulation and support may be severe and long lasting. Public preschool programs emerged in the 1960s to compensate for poor quality home environments and ameliorate skill gaps for children from different family backgrounds. Indeed, children who attend high quality preschool programs outperform their counterparts in kindergarten readiness measures (Lee et al., 2014; Magnuson, Ruhm, & Waldfogel, 2007).

The early childhood literature, by and large, treats these sources of investment in isolation or emphasizes the effects of just one source. While important insights have emerged from these studies, they yield partial explanations for inequalities in child outcomes. A complete understanding of child outcomes requires consideration of the full set of learning opportunities available to children, also referred to as a child's instructional regime (Raudenbush & Eschmann, 2015). A critical question that remains unresolved concerns how these sources of investment operate in tandem to affect child outcomes.

Some studies have made headway in considering the joint effects of the home environment and preschool programs on child outcomes (e.g. Melhuish et al., 2008; Miller et al., 2014; Parcel & Dufur, 2001). With the exception of Miller et al. (2014), these studies have not used nationally-representative samples, instead, focusing on effects for children attending a specific preschool program, limiting the generalizability of their findings. Moreover, some of these studies tend to rely on a singular dimension of the home environment— cognitive stimulation, and do not consider the effects of emotional support.

Using nationally-representative data from the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B), this study examines the joint effects of the home environment and preschool programs on kindergarten readiness and is guided by the following research questions: 1) are higher quality home environments and public early education programs associated with improvements in kindergarten readiness? 2) Do public early education programs moderate the effect of the home environment on kindergarten readiness? Using information about preschool program participation and the home environment collected during the preschool wave of data collection, I consider differences in the effects of the home environment for children in different types of care. Conditional on preschool programs, stronger home environment effects for children from lower-income backgrounds, if confirmed, would undergird the importance of public programs in reducing inequality in early childhood outcomes. A primary contribution of this study is building upon the existing body of literature by using expansive measurement of the home environment based on trained observer reports to explain inequalities that emerge in early childhood.

1.1 Literature Review

A long line of research underscores the critical influences of high quality preschool programs and home environments on healthy cognitive and social development in early childhood. The small but growing set of studies that incorporate both sources of investment in meaningful ways provide additional insight into the nature of their effects and form the foundation of this study.

1.1.1 Early Childhood Program Effects

Early childhood programs serve as an important source of instruction and socialization. The best evidence of early childhood program impacts comes from the High/Scope Perry Preschool Project and the Carolina Abecedarian Project— high quality, intensive programs several decades old that served a set of highly disadvantaged children. Employing experimental study designs, the programs demonstrated impressive benefits for participants, particularly in the long-term (Schweinhart et al., 2005). In recent decades, early childhood programs have expanded dramatically and the majority of U.S. children now enroll in an early childhood program before entering kindergarten (Kena et al., 2016). Public preschool options today include Head Start and prekindergarten described in detail below.

1.2.1.1 Head Start

Head Start, a federally-funded means-tested preschool program targeting children in poverty, was originally founded in 1965 as part of President Johnson's War on Poverty. The program was intended to compensate for skills gaps that emerged in early childhood for children from different family backgrounds. Head Start uses a comprehensive, model that provides a variety of services and supports to children and their families related to health, nutrition, medical and dental care, and parenting education. Experimental evaluations of Head Start indicate that the program reduces a substantial portion of the black-white test gap (Feller, Grindal, Miratrix, & Page, 2016).

Head Start's model is also intended to have direct impacts on the home environment, engaging parents in a variety of ways. Its parent training programs covers various parenting techniques including discipline strategies, play that promotes cognitive development, and providing praise and encouragement. Positive results were observed with regard to parent-child interactions and child outcomes (Webster-Stratton & Hammond, 1998). Additionally, parental involvement is another way through which parents are exposed to the Head Start program. One study found that 60 percent of parents volunteered at their child's Head Start program at least once (Castro, Bryant, Peisner-feinberg, & Skinner, 2004).

1.2.1.2 Prekindergarten Programs

State-funded prekindergarten programs, operated by state and local education agencies, emphasize academic preparation for kindergarten, and some prekindergarten programs are closely aligned to local public school systems (Barnett et al., 2017). Prekindergarten programs do not have the same means-tested eligibility requirements as Head Start and, therefore, serve a set of more economically diverse children. As such, there is a lot of variability in the quality, funding, and curriculum of these programs. This heterogeneity makes reaching a consensus on impacts difficult to achieve (Gormley, Gayer, Phillips, & Dawson, 2005; Gormley, Phillips, & Anderson, 2018). Nonetheless, recent program evaluation findings are generally consistent with findings from seminal studies on preschool effects— high quality preschool programs have positive impacts on children. A review of state and district evaluations document gains in kindergarten readiness levels for children in contemporary prekindergarten programs (Philips & Dodge, 2017). Moreover, there is evidence to suggest these programs demonstrate greater benefits for children from lower-income families (Currie, 2001).

The prekindergarten model does not have the same direct links to parents embedded in the Head Start model. As such, the means by which the home environment may be affected would be less direct. Such changes may occur, for example, through parents' exposure to a more structure academic setting, interactions with preschool instructors, or through changes in child competencies or affinities.

1.1.2 Home Learning Environment Effects

A large body of literature has identified features of the early home learning environment that promote positive child outcomes. The home environment generally reflects quality of parent-child interactions, availability of stimulating materials in the home, and the physical environment (Bradley et al., 1989; Crane, 1996). Cognitive stimulation in the home receives the most attention in the literature as it plays a key role in the cognitive and language development of young children and is relatively straightforward to measure. Cognitive stimulation typically reflects age-appropriate, instructional parent-child interactions, high levels of language-use, book reading, and availability of stimulating games and materials.

Relative to cognitive stimulation, emotional support is a less commonly studied dimension of the home environment. Emotional support encourages a sense of safety for the child, effecting an environment in which children are able to learn and explore. Emotional support encompasses nurturing behaviors include expressions of physical affection toward the child, praise, and limited use of negative statements (Barnard 1997). Responsivity, another dimension of emotional support, refers to the parents' ability to address the needs of the child promptly (Bornstein & Tamis-lemonda, 1989).

Quality of the home environment has been shown to be correlated with socioeconomic status. Lower-income families face greater difficulty in providing high quality home environments because financial constraints limit parents' ability to provide stimulating materials in the home and high levels of stress strain parents' ability to provide emotional support. Critically, several studies have found that, although home environment and socioeconomic status are correlated, the home environment can be improved independently of financial resources (Bradley 1989; Melhuish 2008; Rodriguez and Tamis-Lemonda 2011). For example, a decrease in the use of physical discipline and an increase in the provision of cognitive stimulation have been observed for children in Head Start (John M. Love et al., 2005; Magnuson & Waldfogel, 2005). Moreover, effectiveness of home visits has been shown to increase with higher frequency of visits (Nievar, Van Egeren, & Pollard, 2010). Efforts to improve the quality of parental practices and behaviors in the home, therefore, offer a meaningful and practical approach to reducing inequality in early childhood outcomes.

1.1.3 Joint Effects of the Home Environment and Educational Institutions

A small but growing literature considers the effects of both the home environment and preschool programs on child outcomes and provides evidence that both sources of development are critical to promoting healthy outcomes (Anders et al., 2012; Bradley, Mckelvey, & Whiteside-Mansell, 2011; McLaughlin et al., 2007; Melhuish et al., 2008; Miller et al., 2014; Network, 2003; Philips & Dodge, 2017; Sammons et al., 2009). A recent study using ECLS-B data found reciprocal relationships in which the home environment prior to preschool enrollment was associated with preschool enrollment and preschool enrollment was, in turn, associated with the subsequent home environment (Ansari & Crosnoe, 2015).

Findings concerning the nature of the interaction between the two institutions, however, vary (Parcel, Dufur, & Zito, 2010). A few studies found evidence of a boosting relationship between the home environment and preschool programs (Anders et al., 2012; Padilla & Ryan, 2018; Pinto, Pessanha, & Aguiar, 2013). Other studies offer findings consistent with the compensatory aims of public preschool education (Bradley et al., 2011; Miller et al., 2014; Sammons et al., 2009). Miller et al. (2014), using data from the Head Start Impact Study, found largest gains in math skills for Head Start children whose home environments prior to enrollment were low quality and largest gains in reading skills for Head Start children whose prior home environment was of moderate quality.

In sum, while the literature has reached a consensus on the importance of the independent effects of the home environment and preschool education on child outcomes, less consistent are findings regarding how these sources of investment interact to shape child outcomes since evidence for boosting and compensatory relationships have both been found. In most cases, these studies are based on samples of preschool children only. Such findings have limited generalizability since they are based on small, select samples and do not consider impacts relative to children who receive parent care only. More generally, many of the studies focus on the role of cognitive stimulation in the home and overlook the effect of emotional support.

1.2 Theoretical Framework and Hypotheses

In analyses that seek to explain early child outcomes, meaningful treatment of both family and education institutions is necessary as well as consideration of how these institutions interact. Raudenbush's concept of the instructional regime provides a unified conceptual framework for explaining child outcomes by considering the full array of influences on an individual child rather than treating each source of investment in isolation (Raudenbush, 2008; Raudenbush & Eschmann, 2015). Since development and growth does not stop at the end of the school day, attempts to understand inequality in educational outcomes must take into account learning that occurs within school and non-school settings.

Theorization of processes of learning and development in early childhood requires consideration of social capital, defined as the resources that reside in relationships among people (Coleman, 1988). Social capital theory elaborates the process by which family and school institutions transmit instruction and socialization to children. In early childhood, social capital provides the means by which parents and preschool instructors transmit human capital to children, imparted primarily through nurturing, responsive interactions between parents and children. By establishing bonds that promote a sense of safety and trust, these interactions allow adults to effectively transfer knowledge to young children (Parcel & Dufur, 2001). Quality of adult-child interactions is, therefore, critical to explaining child outcomes because it conditions knowledge transfer.

There are multiple ways in which the home environment and preschool programs may interact to affect child outcomes. If a stronger home environment on kindergarten readiness is observed for children in preschool compared to their parent-care only counterparts, a synergistic relationship between the two sets of influences is at work. This relationship suggests that exposure of parents and children to the preschool program results in positive early academic outcomes. With regard to Head Start, exposure may reflect parent training programs and home visits. For both Head Start and prekindergarten programs, exposure may reflect interactions with instructors and volunteering. It may also reflect a child-centric explanation in which changes in the disposition and preferences of the child resulting from participation in the program, in turn, affect parenting approaches in the home environment. The implication here is that children from similar families who receive parent-care only are at a disadvantage because these parents are less effective in promoting stimulation and providing emotional support. Enrollment in preschool would, in this case, help parents provide more stimulating and supportive home environments.

Another possibility is stronger home environment effects observed for children in parent-care than for preschool children. This relationship would suggest that parents who keep their children at home are more effective in providing a stimulating and emotionally supportive environment. An implication of this finding may be that parents who enroll their children in preschool programs consider the preschool to be the main source of learning and development and, as a result, relinquish some responsibility to provide support and stimulation in the home. Such a relationship offer implications for greater engagement of parents by preschools.

These guiding theories reflect the complex set of influences on outcomes in early childhood. Figure 1.1 presents my conceptual model relating the joint effects of the two primary sources of investment on child outcomes. The family institution, the first influence in a child's life, has a direct effect on child outcomes varying by the quality of the home environment. The diagram conveys the moderating role of early childhood programs in which the effect of the home environment on kindergarten readiness is a function of early childhood program participation. The moderating relationship conveyed here is the key to testing whether a boosting or compensatory relationship holds. The unattached arrows in the diagram reflect the effects of unobserved characteristics that cannot be accounted for in this analysis.

Drawing from the conceptual framework specified in figure 1.1, the family institution and school institution must be considered in tandem when explaining child outcomes. I specify the following hypotheses regarding the effects of the home environment, early childhood programs, and their joint effects on kindergarten readiness.

Hypothesis 1: High quality home environments will predict higher reading and math scores at kindergarten entry controlling for the type of early childhood program attended.

For children under a specific childhood program, high quality home environments will lead to better outcomes because 1) emotional warmth and responsiveness are required for children to establish bonds and feel safe to explore and 2) the availability of stimulating activities and experiences promote cognitive and language development. As such, higher levels of emotional supportiveness and cognitive stimulation in the home will have a positive effect on learning outcomes controlling for the type of early childhood program the child attends.

Hypothesis 2: Public preschool program participation will predict higher reading and math scores at kindergarten entry than comparable children who receive parent care only, controlling for home environment.

Public preschool programs expose children from lower-income families to more structured learning environments. Children who attend Head Start or prekindergarten programs will benefit from increased instruction by trained instructors as well as holistic services provided in the case of Head Start compared to similar children who receive parent-care only within any levels of home environment. As such, attending public preschool programs will predict better reading and math scores at kindergarten entry compared to those under parental care when home environments are held constant.

Hypothesis 3: The effect of high quality home environments will be stronger for children who attend public preschool programs than comparable children who receive parent care.

This hypothesis concerns the differential effects of the home environment and early childhood programs on kindergarten readiness. Given that Head Start has several direct means of engaging parents and given the indirect means through which Head Start and prekindergarten programs may affect parents, I predict that the effect of home environment will be stronger for public preschool children than similar children who receive parent-care only.

1.3 Data and Measures

The Early Childhood Longitudinal Study-Birth Cohort (ECLS-B) study followed a nationally-representative sample of 10,688 children born in the United States in 2001. Data collections occurred at nine months, age two, preschool, and kindergarten entry. The data contain parent interviews to obtain detailed information on parent and child characteristics, measures of the home environment based on parent surveys and observations conducted by trained researchers, preschool participation, and direct child assessments.

The analytic sample consists of ~7,600 children who were observed in the kindergarten entry wave. The ECLS-B study originally intended to follow a panel of 10,688 children starting at age 9 months until kindergarten entry, but because of budgetary constraints, the sample was reduced to an 85 percent random subsample in the final wave when the sample was entering kindergarten. Children whose preschool participation could not be determined because they were not observed in the preschool wave were removed. Children who did not have kindergarten readiness outcomes because they skipped kindergarten were removed from the sample. Finally, a minimal number of observations were removed because of missing information on early childhood program in order to simplify the multiple imputation procedure described below.

Data collected during the kindergarten entry wave occurred across two consecutive school years since some sample children were not able to enter kindergarten in the first year due to age-eligibility requirements. Both sets of children are included in the sample without regard for the year they entered kindergarten, although age at kindergarten entry is included as a covariate in the modeling.

Measures used in this analysis include the outcome variables, reading and math scores at kindergarten entry, as well as three sets of predictor variables— home learning environment, early childhood program participation, and covariates. The outcome is drawn from the fall kindergarten wave only, while predictor variables were drawn from all preceding waves as necessary, taking advantage of the panel nature of the data. Unweighted summary statistics for the analytic sample are provided in appendix table A.1.

1.4.1 Outcomes

Kindergarten readiness is measured by reading and math IRT scores in the fall kindergarten wave. The reading assessment measures both language and literacy skills. The items measure letter recognition, phonological awareness, knowledge of print conventions, word matching, initial understanding, interpretation, and vocabulary. The math assessment measures skills related to number sense, counting, operations, and pattern recognition. The outcomes were standardized based on the original sample to have a mean of 0 and standard deviation of 1.

1.4.2 Home Learning Environment Measures

The home learning environment was measured based on parent-child interactions that were videotaped in the home during the data collection visit. The parent and child dyad were asked to participate in a 10 minute, semi-structured play activity called the Two Bags Task in which two bags containing a book and a game, respectively, were supplied. The Two Bags Task is an adaptation of the Three Bags Task originally developed for other large-scale child development studies (Love et al., 2002; NICHD Early Child Care Research Network, 1999). The videotaped parent-child interactions were subsequently coded by trained staff on five dimensions of parental behavior using a 7-point Likert scale ranging from very low to very high. Analyses of inter-rater reliability showed high levels of agreement.

The use of direct observations to capture parent-child interactions is subject to advantages and disadvantages (see Gardner, 2000). Observations by trained observers provide objectivity in measurement of parental behaviors in contrast to parents' self-report, which may be subject to social desirability bias or inaccuracies related to memory recall. Using scientific measures for home environment represents a crucial step forward. Moreover, parent-child interactions were observed in the home, the natural setting in which such interactions normally occur and are, therefore, may be more representative of typical interactions for that particular parent-child dyad. In addition, the interactions were systematically coded by trained coders. Concerns of direct observations may include the presence of the observer influencing the interaction as well as coder biases that would affect ratings of some groups and not others.

The scales used to rate parent-child interactions were developed for the Early Head Start Study (Fauth, Brady-Smith, and Brooks-Gunn 2003). Of the five dimensions of parent-child interactions measured, two were selected as the focus of this study— parental stimulation of cognitive development and parental emotional supportiveness. These two scales reflect positive parenting approaches, while the other three—parental intrusiveness, parental negative regard, and parental detachment— capture negative parenting approaches. Individual items for these scales were not made available in the ECLS-B restricted-use dataset.

Parental emotional supportiveness reflects the parent's emotional and physical expressions toward the child during the interaction. Emotional support provides a sense of safety from which the child can explore and play. For example, displays of enthusiasm and praise for the child's actions during the interaction would reflect higher levels of emotional support.

Parental stimulation of cognitive development reflects the parent's ability to provide developmentally-appropriate teaching, instruction, and demonstration in a way that encourages further learning and development. If the parent's efforts do not match the developmental stage of the child, they are not considered stimulating. Highly stimulating interactions may include engaging in pretend play, logical presentation of steps in an activity, and elaborating during book-reading.

Many studies have emphasized the role of cognitive stimulation in the home to estimate early childhood outcomes, which is a key aspect of the home learning environment. This study, guided by the seminal and enduring work of Bradley and Caldwell (1984) in their measurement of the home environment, incorporates measurement of the home learning environment that is multi-dimensional reflected by observers' ratings of cognitive stimulation and emotional support in the same model. This approach recognizes that not only is the availability of stimulating resources and activities is vital to child development, but that the quality of the relationship with the parent is equally important in creating conditions conducive to learning and development for young children.

1.4.3 Early Childhood Programs

Early childhood program participation is based on responses from the parent interview during the preschool wave and includes five categories: Head Start, public prekindergarten, private preschool, relative care, other non-parental care such as relative care or babysitting for example, and parental care only serving as the reference category. Children observed in multiple childcare settings were assigned to the setting where they spent a greater number of hours per week. Interpretation is focused on effects associated with Head Start and prekindergarten, but children who participated in the other programs were retained in the analytic sample in order to ensure that the analysis was not subject to sample selection.

1.4.4 Timing of Data Collection and Moderating Relationship

Timing of data collection for the preschool wave occurred from August 2005 to mid-July 2006. As such, parental reports of preschool participation reflect the child's participation from the start of the academic year. The home environment measures, however, also collected during the ECLS-B home visit reflect measurement at the time of the visit. Children and families in the study will, therefore, have had some, if varying, exposure to preschool programs at the time home visit. In the analytic sample, 72 percent of children had been in preschool for at least a month at the time of the home visit. For this reason, I consider preschool program participation to be a moderator that affects the home environment. Alternatively, given that the study is cross-sectional in nature, it is also appropriate to consider that effects of preschool programs on kindergarten may alternatively be moderated by the home environment. Both interpretations will be considered in the results.

1.4.5 Child Characteristics

Child characteristics include race/ethnicity, sex, age, low birthweight status, and childcare prior to preschool. Child's race/ethnicity consists of five categories: white, African-American, Hispanic, Asian, and other. Sex is a dichotomous variable (male=1). Age at the time of the assessment in the fall kindergarten wave is measured in months and adjusts for developmental differences. Low birthweight status is a dichotomous variable reflecting a weight of 2500g or less at the time of birth. Childcare received prior to preschool is a dichotomous measure that combines information from the 9-month wave and the 2 year old wave.

1.4.6 Family Characteristics

SES is a composite variable that combines information on household income, parental education, and parents' occupational prestige. Family structure contains three categories: two biological parents present, step-family, or single-parent family and other type. Parents' expectation of their child's educational attainment is drawn from the preschool wave. The measure ranges from 1 "less than a high school degree" to 6 "professional degree". Number of siblings is an ordinal variable measured during the preschool wave. Childcare participation prior to preschool is a dichotomous variable reflecting whether parents reported sending their child to daycare in the 9-month or two-year old wave of data collection. Region consists of four categories: northeast, Midwest, west, and south. Urbanicity is a dichotomous variable that reflects whether the child lived in an urban or rural area.

Multiple imputation was used to impute missing data as it preserves the multivariate distributions of analysis variables and reduces the risk of producing biased estimates attributed to sample selection (Schafer, 1997). Twenty multiply imputed datasets were produced using the "stratify then impute approach" in order to account for interaction terms that are required in the analysis models (Von Hippel, 2009). Mean estimates and standard errors were combined using Rubin's Rules. Survey design variables and weights adjust for nonresponse and oversampling and produce population-level summary statistics.

1.4 Methodology

Linear regression models were used to test three hypotheses regarding the effects of the home environment and early childhood programs on reading and math scores at kindergarten entry. Estimating the effect of the home environment poses considerable challenge given its non-random nature. This study used a rich set of covariates in the linear regression models help to allay concerns about the endogenous nature of the home environment. Still, it is likely that controlling for the set of covariates does not eliminate endogeneity concerns. To test hypotheses 1 and 2, I estimate the following equation:

$$Y_i = \beta_0 + \beta_1 CS_i + \beta_2 ES_i + \beta_3 EC_i + \beta_4 X_i + u_i \tag{1}$$

where Y_i is reading or math score at kindergarten entry for child i, CS_i is cognitive stimulation, ES_i is emotional support, EC_i is the early childhood program represented by four dummy variables for Head Start, prekindergarten, private preschool, and other non-parental care with parental care only as the reference category, X_i is a vector of covariates, and u_i is the error term.

Support for hypothesis 1 regarding the beneficial effects of the home environment on kindergarten readiness would be found if β_1 or β_2 is positive and statistically significant. Support for hypothesis 2 regarding the beneficial effects of public early childhood programs would be found if β_3 is positive and statistically significant.

Equation 2 is used to test hypothesis 3 regarding whether the effect of the home environment on kindergarten readiness is moderated by the effect of early childhood programs:

$$Y_i = \beta_0 + \beta_1 CS_i + \beta_2 ES_i + \beta_3 EC_i + \beta_4 X_i + \beta_5 CS_i * EC_i + \beta_6 ES_i * EC_i + u_i \qquad (2)$$

where CS_i*EC_i is the interaction of cognitive stimulation and early childhood program, and ES_i*EC_i is the interaction of emotional support and early childhood program. Support for hypothesis 3 would be found if the coefficients for the interaction terms, β_5 or β_6, are positive and statistically significant, reflecting a stronger effect of the home environment for children who attend preschool, specifically Head Start or prekindergarten, compared to children who receive parent care only.

1.5 Results

1.5.1 Descriptive Statistics

Descriptive statistics are presented by early childhood group in Table 1.1 with interpretation focused on results for children who attended the two public preschool programs—Head Start and prekindergarten—and the comparison group, parent care only children. Average reading and math scores at kindergarten entry are lowest for Head Start children, followed by parent care children, and prekindergarten children who have considerably higher reading scores. The results reflect differences in program goals and type of children served. Head Start, serving children from families in poverty, strives to meet the needs of children and families in multiple domains including health and nutrition in addition to cognitive development. Prekindergarten, serving a relatively more advantaged set of children, focuses heavily on academic preparation for kindergarten.

Summary statistics for the home environment measures by early childhood group show the same pattern. Measured on a Likert scale from 1 (very low) to 7 (very high), the mean level of cognitive stimulation is lowest for Head Start children (3.84), followed by parent care only children (3.97), and prekindergarten children (4.22). Mean levels of emotional support, generally higher than mean levels of cognitive support in the home, are lowest for Head Start children (4.07), followed by parent care only children (4.26), and prekindergarten children (4.46).

In terms of child and family characteristics, there are notable differences by early childhood group. There are higher concentrations of minority children in Head Start (65 percent) compared to just under half of parent care children (46 percent) and over one-third for prekindergarten children (37 percent). Indicators of financial wellbeing reveal a highly disadvantaged profile for Head Start children, which is expected since Head Start is a means-tested program. Head Start children have the lowest socioeconomic status (-0.61 sd) and the highest rate of poverty (49 percent). Half of Head Start children live in families with both biological parents present compared to about three-quarters of parent care children and prekindergarten children, which may help to contextualize differences in home environment measures. Having two parents in the home from birth onward likely increases the volume of parent-child interactions, offering additional opportunities for stimulation and emotional support.

Parent care children and prekindergarten are relatively more advantaged. Parent care children fare moderately better than Head Start children with a higher average SES level (-0.40 sd) and a 35 percent poverty rate. Prekindergarten children are the most financially advantaged of the three groups with a higher average SES (0.10 sd) and a 19 percent poverty rate. As for parental expectations for child's educational attainment, all parents expect their children to attain at least a bachelor's degree though average expectations for prekindergarten children are slightly higher. A greater proportion of children in Head Start and parent care children live in non-English speaking families (about one-quarter) compared to 15 percent of prekindergarten children. Finally, Head Start and prekindergarten children had higher rates of childcare prior to preschool compared to parent care children. Although childcare experiences vary considerably, it is important to include this measure as instruction and socialization attributed to childcare experiences prior to preschool may act as a confounder in the estimation of the home environment and preschool effects.

The descriptive findings provide context for understanding how early childhood groups differ in terms of composition as well as offering insights into challenges the groups face. Such differences are important to adjust for in estimating the joint effects of the home environment and preschool programs on reading scores in order to obtain a clear understanding of their effects on reading net of various family and child characteristics that affect the outcome measures.

1.5.2 Regression Results

Regression results are presented in tables 1.2 and 1.3 for models predicting reading and math scores at kindergarten entry. The tables present selected results from the full additive model (model 1) followed by results from models that include interaction terms to test whether public preschool programs moderate the effect of the home environment on kindergarten readiness (models 2 and 3). Full model results are presented in appendix tables A.2 and A.3.

Results from model 1 in table 1.2 address hypotheses 1 and 2 which specified positive effects of the home environment and public preschool programs on kindergarten readiness. The results provide partial support for these hypotheses. Cognitive stimulation is, indeed, positively associated with reading readiness (0.08 sd), though the effect of emotional support does not reach statistical significance. The effects of public preschool programs on reading vary by early childhood program. Prekindergarten program attendance is associated with increased reading scores compared to parent-care only children, while the effect of Head Start does not reach statistical significance.

Concerning the interaction effects of the home environment and preschool programs on reading at kindergarten entry, there is some evidence of a moderation relationship. Though the main effect of emotional support did not reach statistical significance, a stronger emotional support effect on reading is observed for children who attend Head Start compared to similar parent care children, net of family and child characteristics. Figure 1.2 plots the slopes for emotional support for these two groups. Additional interaction effects were tested for prekindergarten programs and home environment measures, but the effects did not reach statistical significance.

Again, this specific moderation relationship was chosen because of the timing of the data collection in which early childhood program participation reflects participation since the beginning of the academic year, while measurement of the home environment occurred over the course of the academic year, which means that most children and families in the sample had had some exposure to the preschool program at the time of the home environment data collection. It is important to consider, however, the alternative specification of the moderation relationship, i.e., the effect of preschool programs on kindergarten readiness may be moderated by the home environment. In this case, the interpretation of the results indicates that the effect of Head Start is stronger for children whose homes are more emotionally supportive. The implications for each of these sets of findings are considerably different and will be further considered in the discussion section.

Results for math readiness are presented in table 1.3. Consistent with results for the effects of the home environment on reading scores, cognitive stimulation has a positive association with math at kindergarten entry, while emotional support does not reach statistical significance. In terms of the effects of public early childhood programs, attending either prekindergarten or Head Start is not found to be statistically different from parent-care only. Results from models 2 and 3 do not provide evidence of a moderation relationship.

The results were robust to changes in the composition of the reference category. Since Head Start children face severe levels of economic disadvantage, making comparisons to an appropriate reference group is critical to accuracy of estimates. Using poverty status information, parental care children were split into higher- and lower- income groups[1] with lower-income parent care only children specified as the reference group. The results are consistent with the main results reported.

1.6 Discussion and Conclusion

A large body of literature has demonstrated the importance of the home environment and preschool programs in promoting positive child outcomes in early childhood, which is in turn critical to setting a foundation from which children can succeed in formal schooling settings and beyond. The instructional regime framework argues that in order to provide complete

[1] Multiple poverty thresholds were considered including 100 percent of the poverty line as well as 130 percent and 185 percent above. The results were consistent regardless of the poverty category used.

explanations of child outcomes, it is necessary to consider the effects of the complete set of sources of investment. The results of this study confirm the importance of including both the home environment and preschool programs in analyses of child outcomes. Some evidence was found to support the first two hypotheses regarding positive effects of the home environment and public preschool programs. Of the two dimensions of the home environment, cognitive stimulation was associated with improvements in both reading and math scores at kindergarten entry, controlling for preschool program and family and child characteristics. Of the two public early childhood programs, prekindergarten was associated with kindergarten readiness, controlling for the home environment and family and child characteristics.

In terms of findings regarding the moderating effect of preschool programs on the home environment, the results showed that the effect of emotional support on reading was stronger for Head Start children than parent-care children. This finding suggests that Head Start children benefit from more supportive parenting approaches in the home to a stronger degree than children who receive parent-care only. Such results were not observed for prekindergarten children, in contrast. It is possible that Head Start program features that focus specifically on parenting techniques promote stronger home environment effects. One implication for these findings is that Head Start parents are more effective in being emotionally supportive in comparison to parents who keep their children at home. Thus, enrollment in Head Start programs not only improves kindergarten readiness but also boosts the positive effect of emotionally supportive home environments on the same outcomes. This evidence would not be possible without modeling the preschool programs and family environment simultaneously.

The decision to specify early childhood programs as the moderator in this study was based on timing of the data collection in which home environment measures were collected after the families and children had had some exposure to the preschool program. Given that this analysis is cross-sectional in nature and that both sets of measures come from the same wave of data collection, an alternative interpretation where Head Start acts as the moderator could reasonably be considered. In this framing, the findings would reflect a stronger Head Start effect for children who come from emotionally supportive homes. Either interpretation points to a necessary consideration of the potential interaction between preschool programs and the family in early childhood research.

No differential effects were found for the other dimension of the home environment—cognitive stimulation. Given the array of stressors associated with poverty, it may be the case that low-income parents are more readily able to improve emotional support practices than levels of cognitive stimulation in the home. Purchasing books, materials, and games that promote cognitive stimulation may be more out of reach for families facing dire financial constraints.

In addition, no differential effects were found in analyses of math readiness. This finding is consistent with a stronger focus on reading in preschool classrooms whereas much less time is spent on math instruction (Early et al., 2010). It may well be the case that parents also engage children in book reading in the home compared to building numeracy and early math skills.

The findings support the value of public preschool programs and program efforts to improve parental behaviors in the home. If indeed it is the case that preschools moderate the effect of the home environment, this study reinforces findings from prior studies demonstrating Head Start's effectiveness in improving parenting skills whether through home visits, parenting skills classes, or parents' exposure to this educational setting and that similar children in parent-care only settings would benefit from Head Start enrollment. Further exploration of this question would benefit from longitudinal study.

This study has contributed to the evidence base regarding the importance of including both the home learning environment and preschool programs in tandem when estimating early childhood outcomes. Moreover, this study provides support for using an expansive scientific measurement of the home learning environment guided by child psychology and measures that are highly reliable. This study provides support for including a multi-dimensional measurement of the home environment that, in particular, goes beyond cognitive stimulation to include emotional support. By doing so, this study revealed a differential relationship between emotional support and reading outcomes for Head Start children.

The current study is subject to several limitations related to data availability and endogeneity of the home environment and preschool programs. First, measures of the home environment immediately prior to preschool are not available which is why the current study uses measures collected at one time point. It should be noted, however, that early childhood program enrollment reflects the beginning of the academic year, while the home environment data was collected over the course of the academic year according to ECLS-B data collection schedule, which means these two measures do not precisely occur at the same time, informing interpretation of the models. Estimating the effects of the home environment and preschool programs is complicated by the presence of unobserved characteristics that may confound the results. The home environment, in this case measured largely by parenting skills, may reflect unobserved characteristics that are not captured in the observed characteristics and which may affect the estimates from the models. These characteristics may include parental attitudes toward educational activities in the home, the stress level of the parents, or availability of time to provide learning supports, for example. Moreover, preschool program participation is subject to selection effects since parents enroll their children in preschool in a non-random way. Linear regression models are not suited to address such endogeneity, but the study does include a rich set of covariates that attempt to alleviate concerns related to effects of unobserved characteristics.

Future work would benefit from incorporating information on quality of the preschool program, drawing upon measures of preschool quality collected for a random subsample of children in the ECLS-B study. Including such information will further clarify conditions under which interactions between the home environment and preschool programs occur and will likely reveal stronger effects of home environments associated with high quality preschool programs. Moreover, preschool participation is treated in a nominal sense. Future analyses should consider more closely intensity and duration of participation in order to have greater leverage on understanding effects of preschool participation. Further exploration of these relationships is essential for reducing disparities in early childhood developmental experiences and resulting gaps in kindergarten readiness by family background.

Table 1.1: Weighted Descriptive Statistics for Analysis Sample

	Parent care only (n≈1,450)		Head Start (n≈1,300)		Prekindergarten (n≈1,150)		Private preschool (n≈2,900)		Other non-parental care (n≈700)	
	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error
Outcomes										
Reading at Kindergarten Entry	-0.22	(0.04)	-0.30	(0.04)	0.14	(0.05)	0.19	(0.03)	-0.28	(0.06)
Math at Kindergarten Entry	-0.19	(0.04)	-0.30	(0.04)	0.09	(0.05)	0.24	(0.03)	-0.17	(0.06)
Home Learning Environment										
Cognitive Stimulation	3.98	(0.04)	3.85	(0.04)	4.26	(0.04)	4.38	(0.03)	4.03	(0.05)
Emotional Support	4.27	(0.04)	4.09	(0.05)	4.52	(0.04)	4.61	(0.03)	4.28	(0.05)
Child Characteristics										
Race/Ethnicity										
White	0.46	(0.02)	0.28	(0.03)	0.56	(0.03)	0.68	(0.02)	0.44	(0.04)
African-American	0.12	(0.01)	0.28	(0.02)	0.17	(0.02)	0.09	(0.01)	0.17	(0.02)
Hispanic	0.34	(0.02)	0.37	(0.02)	0.20	(0.02)	0.15	(0.01)	0.32	(0.03)
Asian	0.03	(0.01)	0.01	(0.00)	0.03	(0.00)	0.03	(0.00)	0.03	(0.01)
Other	0.05	(0.01)	0.06	(0.01)	0.04	(0.01)	0.05	(0.01)	0.05	(0.01)
Male	0.49	(0.02)	0.52	(0.02)	0.51	(0.02)	0.51	(0.01)	0.55	(0.03)
Age at Kindergarten Assessment	68.47	(0.18)	67.77	(0.16)	67.15	(0.15)	68.27	(0.12)	68.63	(0.29)
SES	-0.40	(0.03)	-0.60	(0.03)	0.09	(0.03)	0.28	(0.03)	-0.28	(0.04)
Low Birthweight	0.08	(0.01)	0.09	(0.01)	0.07	(0.01)	0.07	(0.00)	0.08	(0.01)
Child Received Previous Childcare	0.14	(0.01)	0.26	(0.02)	0.37	(0.02)	0.42	(0.01)	0.32	(0.02)
Family Characteristics										
Family Structure										
Two Parent Family	0.75	(0.01)	0.50	(0.02)	0.71	(0.02)	0.78	(0.01)	0.59	(0.03)
Step-Family	0.07	(0.01)	0.08	(0.01)	0.08	(0.01)	0.04	(0.01)	0.08	(0.01)
Single-Parent Family And Other	0.18	(0.01)	0.42	(0.02)	0.22	(0.02)	0.18	(0.01)	0.33	(0.03)
Number of Siblings	1.68	(0.04)	1.42	(0.04)	1.32	(0.04)	1.26	(0.03)	1.35	(0.07)
Expectations for Child Educ Attainment	4.05	(0.06)	3.99	(0.06)	4.24	(0.05)	4.21	(0.03)	3.99	(0.05)

Non-English Speaking Household	0.27	(0.02)	0.27	(0.02)	0.15	(0.02)	0.11	(0.01)	0.20	(0.02)
Region										
Northeast	0.09	(0.01)	0.16	(0.03)	0.22	(0.02)	0.19	(0.02)	0.11	(0.02)
Midwest	0.21	(0.02)	0.17	(0.02)	0.13	(0.02)	0.25	(0.02)	0.27	(0.03)
South	0.42	(0.02)	0.41	(0.02)	0.51	(0.03)	0.31	(0.01)	0.36	(0.03)
West	0.28	(0.02)	0.26	(0.02)	0.15	(0.02)	0.24	(0.01)	0.26	(0.03)
Urban	0.80	(0.02)	0.81	(0.02)	0.84	(0.02)	0.87	(0.01)	0.82	(0.02)

Note: In accordance with NCES requirements, sample sizes have been rounded

Table 1.2: Selected Results from Linear Regression Models Predicting Reading at Kindergarten Entry

	Model 1	Model 2	Model 3
Home Learning Environment			
Cognitive Stimulation	0.077***	0.076***	0.075***
	(0.01)	(0.02)	(0.01)
Emotional Support	0.014	0.014	0.0021
	(0.01)	(0.01)	(0.02)
Early Childhood Program (Ref: Parent-Care Only)			
Head Start	0.029	0.019	-0.28*
	(0.04)	(0.12)	(0.13)
Prekindergarten	0.16***	0.16***	0.17***
	(0.04)	(0.04)	(0.04)
Private Preschool	0.083**	0.083**	0.086**
	(0.03)	(0.03)	(0.03)
Other Non-Parent Care	-0.12**	-0.12**	-0.12**
	(0.04)	(0.04)	(0.04)
Head Start * Cognitive Stimulation	--	0.0026	--
		(0.03)	
Head Start * Emotional Support	--	--	0.074*
			(0.03)

* p<0.05, ** p<0.01, *** p<0.001

Note: All models include the full set of covariates: child's race/ethnicity, gender, low birthweight indicator, previous childcare, SES, family structure, number of siblings, non-English speaking household, urbanicity, and region.

Table 1.3: Selected Results from Linear Regression Models Predicting Math at Kindergarten Entry

	Model 1	Model 2	Model 3
Home Environment			
Cognitive Stimulation	0.071***	0.074***	0.070***
	(0.01)	(0.01)	(0.01)
Emotional Support	0.012	0.012	0.0050
	(0.01)	(0.01)	(0.01)
Early Childhood Program (Ref: Parent-Care Only)			
Head Start	0.036	0.12	-0.14
	(0.04)	(0.13)	(0.13)
Private Preschool	0.057	0.056	0.059*
	(0.03)	(0.03)	(0.03)
Prekindergarten	0.070	0.069	0.071
	(0.04)	(0.04)	(0.04)
Other Non-Parental Care	-0.061	-0.061	-0.061
	(0.04)	(0.04)	(0.04)
Head Start * Cognitive Stimulation	--	-0.021	--
		(0.03)	
Head Start * Emotional Support	--	--	0.042
			(0.03)

* p<0.05, ** p<0.01, *** p<0.001

Note: All models include the full set of covariates: child's race/ethnicity, gender, low birthweight indicator, previous childcare, SES, family structure, number of siblings, non-English speaking household, urbanicity, and region

Figure 1.1: Conceptual Diagram of Family and School Institution Effects on Child Outcomes

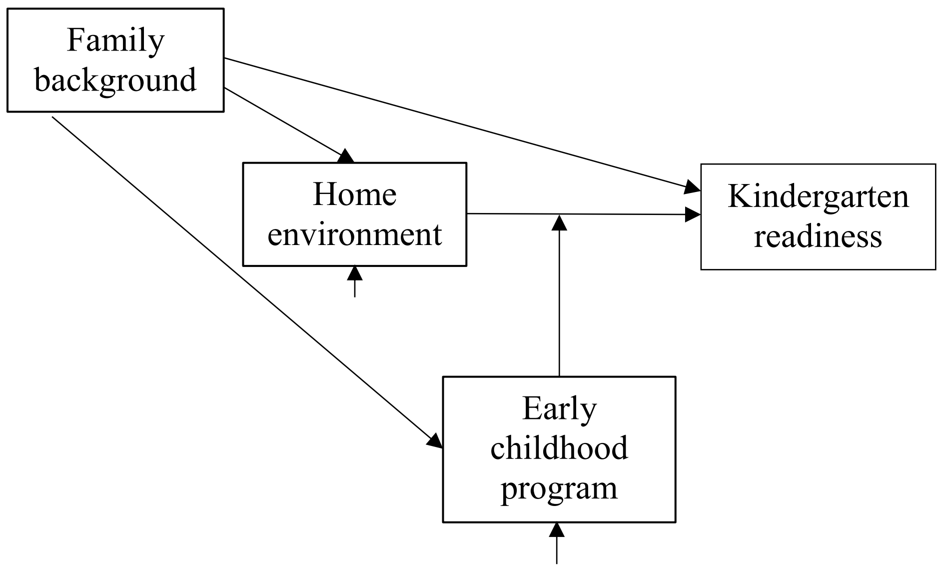
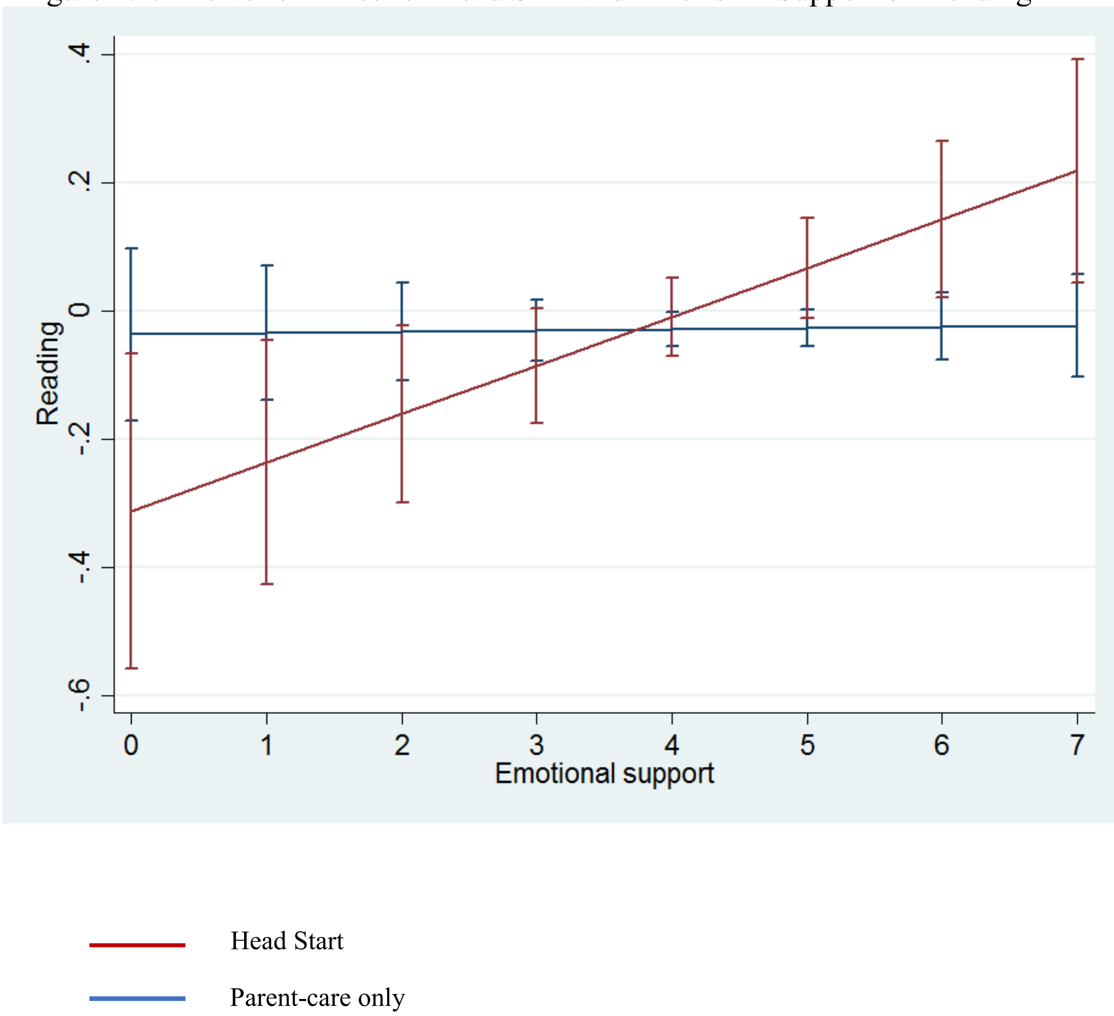


Figure 1.2: Interaction Effect of Head Start and Emotional Support on Reading



Paper 2 Inequality in Multiple Childhood Contexts: The Joint Effects of School and Family Institutions in Elementary School

2.1 Introduction

There has been long-standing debate about the relative importance of family effects and school effects on student outcomes. The Coleman Report argued that families are largely responsible for shaping inequalities in academic outcomes, downplaying the role of schools (Coleman, 1966). On the other hand, leading education scholars have argued that in the absence of schools, inequality for children from different family backgrounds would be worse (Downey, von Hippel, & Broh, 2004; Raudenbush & Eschmann, 2015). Family and school investments in children co-occur and a small but growing literature has made strides in incorporating the effects of both institutions in the same analysis in order to investigate whether and how family and school investments matter in different ways for children from different family backgrounds.

What do we risk by failing to include expansive treatments of home and school institutions in analyses of educational outcomes? Parcel and Dufur (2010), leading scholars in this area, raise several concerns. For example, leaving out measures of one institution could result in omitted variable bias, inflating the importance of variables included in the model. In addition, without both institutions included in the analysis, it is not possible to consider interactions between them. An approach that includes effects of both institutions promises to provide more complete explanations of inequality in educational outcomes, one that is especially critical in light of growing disparities in investment for children from different family backgrounds and amidst widening income-achievement gaps (Kornrich & Furstenberg, 2013; Park, Buchmann, Choi, & Merry, 2016). Considering the joint effects of both institutions on child outcomes informs these questions and suggest insights into longer-term trends in social inequality. Gaining insight into the effects of each institution as well as differential relationships are both key to understanding and ameliorating social inequality.

This study examines the joint effects of family and school institution investments in children and how they interact during the elementary school stage. Data for this study come from the Early Childhood Longitudinal Study-Kindergarten Cohort 2011, a panel dataset following children from kindergarten through fifth grade, containing rich measures of both school-level investments as well as parental investments for children in the study. I estimate growth curve models to examine the effects of school context and the home learning environment (HLE) on reading and math trajectories in elementary school when both sources of investment are included in the same model specification. In addition, the study considers a moderation relationship between the school context and the HLE. Because this study is not intended to be causal, but rather explore associations between these institutions, it is appropriate to consider each, in turn, as the moderating variable. Therefore, both interpretations are provided and future work will take further steps to clarify the nature of the moderating relationship.

Results from the study support the necessity of incorporating both family and school effects as both institutions are shown to have positive associations on growth in reading and math outcomes, net of each other's influence, suggesting that this modeling approach to estimating academic trajectories overcomes previously misspecified models without HLE. Moreover, the significant estimate for the interaction between school context and HLE suggest differential HLE effects conditional on school quality (or vice versa). Correct model specification improves confidence in the results and reassures that estimates of the school institution on academic outcomes are not subject to omitted variable bias.

2.2 Literature Review

Traditional studies of educational outcomes often focus on the role of school institutions with limited treatment of the family institution. Learning does not take place exclusively within school walls, however, and attempts to explain inequality of outcomes must take into account non-school settings, especially opportunities provided by parents in the home. While these studies include some treatment of the family, information on the HLE may be limited, especially if studies draw upon administrative datasets in which free or reduced price lunch status is the only information available on the family institution. This measure belies the extent to which parents provide supports and investments that affect educational outcomes. For example, parents with limited financial resources may structure the home environment in ways that are conducive to learning, like book reading, and setting routines, and ensuring homework has been completed. Or, parents with lots of financial resources and high educational attainment may not spend sufficient time with children, engaged in activities that promote learning.

On the other hand, studies that focus on HLE effects generally do not include sufficient characteristics of the school environment, focusing, for example, on school sector. Nevertheless, these studies provide the critical foundation from which to study the joint effects of these institutions. The literature review first provides an overview of the effects of school context on student outcomes, followed by an overview of HLE effects on student outcomes, and concludes with an overview of studies that have considered the joint effects of family and school institutions on students.

2.2.1 School Context

Educational institutions play a primary role in socialization and instruction and how well students do is linked to the organizational health and functioning of the school (Lee & Burkam, 2003). Students learn best when they feel safe and when bonds of trust are established with educators, creating a sense of attachment to the school institution and motivating them to learn (Bryk & Schneider, 2002). Successful schools have adequate funding levels that promote stable environments, have high levels of social capital between teachers and students which is requisite for students to learn, and where norms around academic excellence have been established. These dimensions of the school context have been shown to promote positive student outcomes in the school effects literature elaborated below.

The relationship between school resources and student outcomes has not reached a consensus. Recent studies have examined one dimension of school resources related to school strain and have provided consistent results (Arya Ansari & Pianta, 2018; Lowenstein et al., 2015). Schools with high levels of strain experienced problems related to teacher turnover, teacher absences, funding levels, and overcrowding. School strain undermines teacher-student relationships and the ability to foster a stable, cohesive schooling environment in which students feels attached to their schooling experience. Ansari and Pianta (2018), using ECLS-K 1998 data, demonstrated that negative school contexts, partly defined by level of school strain, does indeed undermine achievement for elementary school students.

School climate refers to the culture of the school manifested in collective norms and beliefs of school administrators and teachers as well as bonds between the various actors including personnel, teachers, and students (Bryk, A., Schneider, 2002). Stressful school contexts which are characterized by disruptive behavior and lack of discipline undermine students' ability to learn. Many studies in the education literature have focused on the link between school climate, characterized by order and safety, and positive academic outcomes and have demonstrated that it is critical to student learning (Hoy, Tarter, & Hoy, 2006; Thapa, Cohen, Guffey, & Higgins-D'Alessandro, 2013).

Finally, given that teachers play an outsized role for younger children, it is important to consider teacher-student relationships and teacher efficacy in the elementary school context. Strong bonds and high expectations demonstrate to students that they are valued and promote the perception that student learning is a goal shared by the school community. Teacher press has been defined as the alignment of school staff, teachers, and students around high academic standards. Lee and Smith (1999) found evidence of better academic outcomes in schools with higher levels of teacher press (Lee & Smith, 1999).

2.2.2 Home Learning Environment

The home learning environment is a primary setting in which children from different family backgrounds receive differential levels of investment and learning opportunities. Recent demographic studies have documented a growing disparity in levels of investment for children from different family backgrounds in recent decades (Kornrich & Furstenberg, 2013; Park et al., 2016). At the same time, income-achievement gaps have also been increasing (Park et al., 2016; Reardon, 2011).

The family's financial resources have been principally used in the literature to measure family investment, which has been shown to be strongly predictor of child wellbeing. Parents with greater financial means are able to provide more educational resources as well as enriching activities and experiences than parents with limited financial means. Moreover, higher SES parents also tend to have higher levels of educational attainment and human capital to transmit to their children. In contrast, growing up in circumstances characterized by severe material hardship is inimical to healthy child development, the effects of which are worse are most persistent when experienced during early childhood (Duncan, Kalil, & Ziol-Guest, 2013; Heckman, 2012).

Parental investments, however, may also refer to strategies used in the home environment to promote learning. There are certainly many ways, however, that parents from lower SES backgrounds support learning in the home. During the elementary school stage, this may include the availability of educational materials in the home, frequency of reading, and visits to the library. The extent to which promote reading behaviors outside of school has shown to be linked to better academic outcomes, particularly for young children (Payne, Whitehurst, Angel, & Angell, 1994). Moreover, the effects of the home literacy environment have been associated with improvements in children’s early literacy skills above and beyond the effects of family background (Aikens & Barbarin, 2008; Griffin & Morrison, 2016). In addition, high human capital levels are not necessarily sufficient to support learning and growth. Bonds between parents and children are requisite for parents to be able to create conditions conducive to learning and transmit knowledge to their children (Connell & Prinz, 2002). Particularly for lower-SES families, these strategies may offer a meaningful way to support their child’s educational achievement since they are not unreasonably burdensome on financial resources.

2.2.3 Joint Effects of the Home Learning Environment and School Institutions

Some studies have made concerted efforts to include relatively more expansive operationalizations of the HLE and school context, the results of which support the utility of such an approach. Evidence of effects, both additive and interactive, have been found, enabling a more nuanced understanding of how these institutions promote or reduce inequality. These studies have demonstrated that the relationship of family and school investments can operate in a variety of ways, for example, in a boosting or compensatory way (Parcel & Dufur, 2009, 2001). A boosting relationship refers to greater benefits observed for children in multiple advantaged contexts. Compensatory relationships refer to a relationship in which one source of investment makes up for deficiencies in the other. For example, parents may step up educational activities at home to compensate for low quality schools. Alternatively, for children who come from families with lower stores of human, social, and financial capital, schools may be disproportionately beneficial.

These studies have found evidence for both boosting and compensatory relationships depending on the types of capital. Boosting effects were found for mother's IQ score and caring teachers at school (Parcel & Dufur, 2001) as well as parent-child relationships and student-teacher relationships at school (Crosnoe & Elder, 2004). In this case, there is a synergy between family and school investments in children, such that children who have more opportunities for learning or more social capital in the home are better able to take advantage of learning opportunities at school. On the other hand, compensatory effects were found for mother's IQ score and low teacher human capital (Parcel & Dufur, 2001). In this case, parents compensated for insufficient learning opportunities in low quality schools.

These studies offer insights into ways parents may be able to structure the home learning environment to promote their child's learning beyond financial resources. However, operationalization of the home learning environment in these studies has been narrowly defined, focusing on mother's IQ score, for instance. Studies that use expansive definitions of the home learning environment and which focus on parental behaviors reflect promising steps forward. Dufur and Parcel and co-authors (2016) continue to move the literature forward by incorporating more expansive operationalizations of the home environment including and school setting. In the same spirit, the current study similarly incorporates a broader set of measures of the home learning environment centered on educational activities and practices in the home that are more mutable and that may be implemented or improved regardless of SES. The current study differs because of its focus on younger children in the elementary school stage, where parental behaviors in the home play a larger role in shaping academic outcomes than for older children.

2.3 Theoretical Framework

The instructional regime framework attempts to explain inequality in academic outcomes by considering the entire landscape of learning opportunities, which principally consists of the family and school institutions especially as it concerns young children (Raudenbush & Eschmann, 2015). This includes learning that occurs in schools, but also the sources of learning and development that occur elsewhere, principally, through the HLE. It is insufficient to consider within-school learning without considering ways in which families provide learning supports and opportunities and have varying means with which to do so.

This framework is useful for the current study as it guides expectations for the nature of the joint relationship. First, the family institution is a central source of development for children and children from different family backgrounds receive unequal levels of investment, which govern the stores of financial, human, and social capital available to children. The quality of the home context—the educational resources the child has access to, human capital of parents, the nature of the relationship between parent and child, parenting practices, rules, and routines— has great bearing on children's academic outcomes. Differences in quality of the HLE amount to a meaningful source of inequality in educational outcomes.

In addition, schools are a significant source of development and socialization for children and take on increasing importance in children's lives as they grow older. They are central settings for human and social capital investment in children. The ethos underlying the modern day public education system in the U.S. reflects a sense that every child deserves a high quality education regardless of family background. In the formal schooling stage, public schools aim to level the playing field by striving to instill a common set of higher-order thinking skills in all students. An ideal school environment—one that is cohesive, has strong leadership, an orderly environment, efficacious teachers, that meets students' needs—effectively fosters high levels of academic achievement. In reality, great variability exists in the quality and effectiveness of schools across the U.S. which contributes to inequality in child outcomes. Therefore, the quality of the schooling environment has the ability to shape child outcomes directly, independent of other influences in a child's life.

Given how important these institutions are to shaping child outcomes, how might the joint relationship of these institutions operate? Parcel, Dufur, and co-authors have provided important scholarship and theoretical frameworks for studying the joint relationship, particularly in a boosting or compensatory direction. Understanding the nature of this relationship illuminates inequality by family background. Instructional regime theory posits that a compensatory relationship is consistent with reductions in inequality since achievement gaps for children from different family backgrounds would be wider in the counterfactual world, i.e. a world in which schools did not exist. A boosting relationship refers to a scenario in which parental investments amplify the effects of strong quality schools. In this case, inequality would increase since children from wealthier backgrounds tend to have better home learning environments as well as attend better quality schools.

The present study considers whether the effects of the school context on child outcomes are moderated by the HLE. The nature of the moderating relationship determines whether social inequality is exacerbated or reduced. The processes generating the inequality under these two scenarios are dramatically different and imply different approaches by which to stem growing inequality. If a boosting relationship is observed, the process generating the inequality is driven by a complementarity between high quality schools and parental investments. In this scenario, children from higher-SES families who attend high quality schools also benefit from stronger effects of the HLE. If a compensatory relationship is found, this would suggest that schools are disproportionately beneficial for children from low-SES backgrounds, indicating that schools are performing an equalizing function and reducing inequality.

2.4 Research Questions

The following research questions guide this study:

1) What is the relationship between the home learning environment and academic achievement levels and growth when school context, student demographics, early childhood program participation, and family characteristics is controlled?
2) What is the partial association of school context with academic achievement levels and growth, when school context, student demographics, early childhood program participation, and family characteristics is controlled?
3) Does school context moderate the prediction of the home learning environment on academic achievement levels and growth, all else equal?

2.5 Hypotheses

Because child outcomes are shaped by multiple contexts, it is critical to consider both family and school institution effects. Prior studies have demonstrated these effects in isolation or while treating one institution in a meaningful way and the other in limited ways. These institutions are so critical to shaping academic outcomes that I expect they will continue to have direct impacts on academic outcomes, net of each other's influence, when they are included in the same model specification.

Hypothesis 1: The home learning environment will be positively associated with academic outcomes with school context held constant.

Hypothesis 2: School context will be positively associated with academic outcomes at any fixed level of home learning environment.

The last research question of the study refers to whether the effect of the HLE on student outcomes is moderated by school context. Previous literature has found mixed results regarding boosting and compensatory relationships. I consider the interactive effects as empirical questions. If the effect of the HLE is stronger for children in high quality schools, a boosting function relationship will be observed with an increase in inequality. If the effect of HLE is weaker in better school contexts, a compensatory relationship will be observed which will reduce inequality in outcomes for children from different family backgrounds.

2.6 Data and Measures

Data for this study come from the Early Childhood Longitudinal Study- Kindergarten Cohort 2011 (ECLS-K2011). This study tracked a cohort of 18,174 nationally-representative children from kindergarten through fifth grade beginning in fall of 2010. The study employed a complex, multi-stage sampling design. The study is a rich source of information on child experiences including the home environment, school experiences and conditions, and academic and socio-emotional learning outcomes over time. The data consist of parent interviews, child interviews, direct assessments of child academic and socio-emotional skills, teacher interviews, and school administrator interviews.

The analytic sample for the study consists of ~13,600 children who had available outcome data in spring kindergarten, spring first grade, and spring 2nd grade. The percent complete across observations was 40 percent. Missing data rates are generally higher for longitudinal data analysis due to attrition. Missing data were assumed missing at random and multiple imputation using chained-equations was used to impute 50 complete datasets. Results were combined using Rubin's Rules (Rubin, 2004).

Measures including reading and math outcomes over time and sets of explanatory variables: HLE composites, school context composites, and student and family characteristics described in further detail below. Unweighted means and standard deviations for these variables for the analytic sample are provided in appendix table B.1.

2.6.1 Outcomes

Outcomes for this study include reading and math IRT scale scores in the spring of the academic year from kindergarten through fifth grade, yielding six time points over which children are observed. IRT scale scores are suitable for longitudinal analyses and reflect growth made over time. For reference, in spring kindergarten, the reading IRT scores ranged from 33 to 134 points with an unweighted mean of 69 points and standard deviation of 15 and from 72 to 159 points in spring of fifth grade with an unweighted mean of 136 points and standard deviation of 16.

2.6.2 Home Learning Environment

Operationalization of the home learning environment was guided by a long-established literature on the home environment, principally driven by Bradley and Caldwell's measurement of the quality and quantity of support and educational activities present in the home environment (Caldwell & Bradley, 1984). The measures in this study were intentionally chosen because they reflect ways in which parents with limited financial resources may support their child's learning.

The HLE consists of three composite variables that drew from the parent-survey items across all waves of data. First, parent survey items were inconsistently measured across waves of data collection. Second, examination of the individual measures across waves showed low variation, supporting the use of time-invariant composite measures. The first HLE composite measures parent-child interactions and consists of items such as reading books with the child, telling stories, practicing reading, writing, and numbers, and visiting the library. Measures come from kindergarten, first grade, and third grade. This measure was standardized to a mean of 0 and standard deviation of 1. The reliability coefficient was 0.84. The second composite is a reflection of number of books in the home across kindergarten, first grade, and third grade with a reliability coefficient of 0.88. Because this variable contained extreme values, the variable was top-coded to 500 books. The third composite reflects learning that occurred through computer programs in the home from kindergarten, 2^{nd} grade, 3^{rd} grade, 4^{th} grade, and 5^{th} grade. This measure was standardized to a mean of 0 and standard deviation of 1 and its reliability coefficient was 0.65. See appendix table B.2 for a list of items used to construct the composites.

2.6.3 School Context

School context is measured by three time-varying composites. Two composites, school climate, and school strain, consist of items drawn from the school administrator questionnaire and the third, teacher press, consists of items drawn from the teacher survey. School climate includes items such as parental involvement, community support, and order and disruption present in the school. Reliability coefficients of the composites ranged from 0.69 to 0.73. School strain reflects teacher turnover, teacher absenteeism, student absenteeism, and funding stability. Reliability coefficients range from 0.63 to 0.73 across waves. Teacher press reflects teachers' instructional practices and attitudes towards students with reliability coefficients ranging from 0.63 to 0.83 across waves. Items include teachers' perceptions of students' family background affecting capacity to learn, the teacher ability to present things in multiple ways, and teacher satisfaction with profession, among others. See appendix table B.2 for further detail on the complete set of items used in the composite variables.

2.6.4 Covariates

The covariates in the analysis include child and family characteristics. Race/ethnicity is a four category variable with White specified as the reference category. Sex is a dichotomous variable (male=1). SES is a standardized composite variable that combines information on household income, parental education, and parents' occupational prestige. Family structure is a dichotomous variable which categorizes two-parent families vs. other types (step-family, or single-parent family and other families). Locale is a nominal variable that reflects whether the child lives in an urban area, rural area, suburban area, or town with suburban specified as the reference category. These time-invariant characteristics are drawn from the kindergarten waves. Age of the child at the time of the assessment, measured in months, is included as the time metric for the growth curve models. Early childhood program is a five category variable including parent-care only as the reference category, prekindergarten, private preschool, Head Start, and other non-parental care.

2.7 Methods

The analysis first proceeds by examining summary statistics and correlations of the analysis variables. Survey design variables and weights adjust for nonresponse and oversampling of subgroups, and produce population-level summary statistics. Growth curve models were estimated to analyze the joint effects of the HLE and school context measures on academic outcome trajectories. Panel data models are considerably more useful than cross-sectional models, enabling modeling of trajectories as opposed to outcomes measured at one time point.

The time metric for these panel data models is age measured in months with an initial starting point at spring of kindergarten entry and subsequent observations occurring every spring through fifth grade. Therefore, the panel data models include six time points and are balanced for every individual in the analytic sample using the multiply imputed data.

The model building strategy first tested for a random intercept, which indeed showed significant variability in the intercepts, followed by testing for a random slope for age in order to determine if there was variability in growth curves for individual children. The results from the likelihood ratio tests confirmed significant variability in slopes.

In order to test the first two hypotheses regarding the effects of school context and the HLE in the same model specification, the random coefficient model below with time as Level 1 and student as Level 2 was estimated. While the rate of growth is given by the linear and quadratic polynomial *Age*, I only specify differential linear age effects conditional on environment variables. This approach avoids the demand for very large data to detect the doubled number of interactive effects. With a greater amount of data, the full specification of the polynomial age differential effects could be specified. For the purposes of this study, I take an initial step in modeling linear growth rates as random, which will move forward our understanding of how linear growth rates are affected by the HLE and school context and whether they should be modeled as fixed or random. It will offer a clear foundation for future work on modeling the quadratic term as random if indeed the linear growth rate is found to be random. The model is specified as follows:

Level 1: $Y_{ti} = \pi_{0i} + \pi_{1i}Age_{ti} + \pi_{2i}Age^2{}_{ti} + \pi_{3i}SClim_{ti} + \pi_{4i}Strain_{ti} + \pi_{5i}TPress_{ti} + e_{ti}$

Level 2:

$$\begin{aligned}
\pi_{0i} &= \beta_{00} + \beta_{01}HomeP_i + \beta_{02}HomeB_i + \beta_{03}HomeC_i + \beta_{04}X_i + r_{0i} \\
\pi_{1i} &= \beta_{10} + \beta_{11}HomeP_i + \beta_{12}HomeB_i + \beta_{13}HomeC_i + r_{1i} \\
\pi_{2i} &= \beta_{20} \\
\pi_{3i} &= \beta_{30} \\
\pi_{4i} &= \beta_{40} \\
\pi_{5i} &= \beta_{50}
\end{aligned}$$

The combined model is given by:

$$\begin{aligned}
Y_{ti} = {} & \beta_{00} + \beta_{01}HomeP_i + \beta_{02}HomeB_i + \beta_{03}HomeC_i + \beta_{04}X_i + \beta_{10}Age_{ti} + \\
& \beta_{11}Age_{ti} * HomeP_i + \beta_{12}Age_{ti} * HomeB_i + \beta_{13}Age_{ti} * HomeC_i + \\
& \beta_{20}Age^2{}_{ti} + \beta_{30}SClim_{ti} + \beta_{40}Strain_{ti} + \beta_{50}TPress_{ti} + r_{0i} + r_{1i}Age_{ti} + e_{ti}
\end{aligned}$$

where Y_{it} reflects math or reading scores for child *i* at time *t*, *Age*, age in months is the time metric modeled using a quadratic specification, HLE is represented by $HomeP_i$, which refers to parent-child interactions, $HomeB_i$, which refers to number of books in the home, $HomeC_i$ which refers to computer learning in the home, X_i refers to a vector of covariates of student and family characteristics. The three time-varying school context variables are represented by $SClim_i$ which refers to school climate, $Strain_i$ which refers to school strain, and $TPress_i$ which refers to teacher press. The deviation from the intercept for child *i* is r_{0i}, and the deviation from the random slope for *Age* for child *i* is r_{1i}, and e_{ti} is the level-1 deviation of the child's outcome in time *t*. In terms of parameters, β_{00} is the average outcome in spring kindergarten and when all other covariates are set to zero, β_{10} reflects the growth rate and β_{20} reflects the acceleration in the growth rate when all other covariates in the model are set to zero.

To address hypothesis 3, the full model includes interactions between the school and home environment. The following equation specifies an interaction between $HomeB_i$, parent-child interaction, and $SClim_i$, school climate, but each combination of the home environment variables and school context variables were tested for interactions. Again, while the growth rate is modeled using a quadratic function for *Age*, only the linear term is specified as random because of the increased demands on the data to specify a differential quadratic term:

Level 1: $Y_{ti} = \pi_{0i} + \pi_{1i} Age_{ti} + \pi_{2i} Age^2{}_{ti} + \pi_{3i} SClim_{ti} + \pi_{4i} Strain_{ti} + \pi_{5i} TPress_{ti} +$
$\pi_{6i} Age_{ti} * SClim_{ti} + e_{ti}$

Level 2: $\pi_{0i} = \beta_{00} + \beta_{01} HomeP_i + \beta_{02} HomeB_i + \beta_{03} HomeC_i + \beta_{04} X_i + r_{0i}$

$$\pi_{1i} = \beta_{10} + \beta_{11} HomeP_i + \beta_{12} HomeB_i + \beta_{13} HomeC_i + r_{1i}$$

$$\pi_{2i} = \beta_{20}$$

$$\pi_{3i} = \beta_{30} + \beta_{31} HomeB_i$$

$$\pi_{4i} = \beta_{40}$$

$$\pi_{5i} = \beta_{50}$$

$$\pi_{6i} = \beta_{60} + \beta_{61} HomeB_i$$

The combined model is given by:

$$Y_{ti} = \beta_{00} + \beta_{01}HomeP_i + \beta_{02}HomeB_i + \beta_{03}HomeC_i + \beta_{04}X_i + \beta_{10}Age_{ti} +$$
$$\beta_{11}HomeP_i * Age_{ti} + \beta_{12}HomeB_i * Age_{ti} + \beta_{13}HomeC_i * Age_{ti} + \beta_{20}Age^2{}_{ti} +$$
$$\beta_{30}SClim_{ti} + \beta_{31}HomeB_i * SClim_{ti} + \beta_{40}Strain_{ti} + \beta_{50}TPress_{ti} +$$
$$\beta_{60}HomeB_i * Age_{ti} * Sclim_{ti} + r_{0i} + r_{1i}Age_{ti} + e_{ti}$$

The key parameter is β_{61}, the coefficient for the three-way interaction between age, parent-child interactions, and school climate. This interaction term tests whether the slope of growth over time and home environment is different for students in different types of schools in order to provide evidence for whether the HLE is performing a compensatory or boosting function in relation to the school context over time. Alternatively, since this study is based on observational data, the moderator relationship could also be flipped such that the school context is moderated by the HLE. Both interpretations are considered.

Estimating the effects of the HLE and school context is complicated by the presence of unobserved characteristics and selection effects. A major concern underlying analyses that attempt to estimate the effect of the family institution and school institution is the non-random nature of these institutions. Unobserved characteristics affecting the estimation of the family institution effect might include differing levels of motivation, confidence, or self-efficacy among parents. Similarly, children are not randomly enrolled in schools. Family background is highly correlated with quality of the school institution. This study, based on observational data and given the endogeneity challenges associated with the home environment and school effects, is not intended to provide causal estimates, but rather intended to provide further insights into how these complex institutions are associated. Despite these limitations, the results can, nonetheless, provide useful implications about the relationships generating inequality and implications for how to reduce such inequality.

2.8 Results

2.8.1 Descriptive Statistics

Descriptive statistics for the analytic sample by key social stratification characteristics are presented in appendix table B.3. Clear trends, consistent with documented stratification patterns, emerge. White students and Asian/Other students outperform minority students in both mean reading and math scores over time. Students in two-parent families as well as students in higher-SES families outperform their counterparts. Means of school context variables follow similar trends. School climate and school strain follow the same patterns as observed for the outcomes. With regard to teacher press, White students attend schools with higher averages, but average levels of teacher press are comparable among Asian/Other, Hispanic, and African-American students. With regard to the HLE, similar patterns are observed. One exception occurs for computer learning in the home where, notably, African-American students have the highest average level, followed by Asian/Other, White, and Hispanic students. In summary, students who enjoy more privileged positions in the social structure, as it concerns race/ethnicity, family background, and family structure, generally have access to better HLE and school environments.

Figure 2.1 presents distributions of the three school context variables and the three HLE variables. All three school context composites are standardized to a mean of zero and standard deviation of one and therefore resemble a bell curve. Parent-child interactions and computer learning are scaled to standardized normal distributions and the empirical distributions show a resemblance to a bell curve. The composite for books in the home has a mean of 78 books and a range of 0 to 500 and was mean-centered. The distribution is right-skewed, reflecting the fact that few students live in homes with hundreds of books available.

Figures 2.2 present scatterplots for the main independent variables—the HLE measures and the school context measures. As the scatterplots demonstrate, the HLE and school context measures do not have a strict one-to-one relationship, which is key for being able to estimate interactions between the home environment and school context. Some of the scatterplots show clearer relationships than others. For example, the plot for books in the home and school climate shows a positive correlation while the plot for books in the home and school strain shows a negative correlation.

Since the main research questions concern how school context and the HLE affect reading and math trajectories, figures 2.3.1 and 2.3.2 explore trends in average reading and math scores over time by school climate and books in the home. The time trend is non-linear time and is suggestive of a quadratic time trend[2]. Figure 2.3.1 presents average reading and math outcomes for school climate over time, which shows a positive relationship. Over time, the confidence interval band widens which reflects increased skill differentiation at older ages. Reading skill growth seems to occur more rapidly than math skill growth, but, for both reading and math, the rate of skill growth seems to dampen around age 8.

Figure 2.3.2 present average reading and math scores by books in the home over time. The figures show a positive relationship, indicating students in homes with more books also have higher reading and math outcomes. The confidence interval band widens over time, which again points to increased skill differentiation at older ages. The rate of growth for both read and math again seems to slow around age 8. Similar graphs are presented for the other measures of school context and HLE in appendix figure B.1.1, figure B.1.2, figure B.1.3, and figure B.1.4.

[2] Linear, quadratic, and spline time trends were tested. The model using the quadratic term was the preferred specification.

2.8.2 Model Results

Because many studies do not include expansive measures of both school and family effects in the same model specification, hypotheses 1 and 2 considers the joint effects of the HLE and school context variables with an expectation that both are positively associated with reading and math outcomes, net of the other institution. The modeling strategy began with estimation of random intercepts models, followed by a random coefficient model using maximum likelihood estimation. The likelihood ratio tests supported the use of a random coefficient model. I then proceeded to estimate my analysis models using restricted maximum likelihood estimation, including sets of explanatory variables in successive models. The variance components remained significant across model building.

Table 2.1 presents selected results from the random coefficient models which address hypotheses 1 and 2. See appendix table B.4 for complete model results. The average reading score is 68.5 points in spring of kindergarten, when the other covariates are set to reference categories and age is 8.5 years. The growth rate one year later at the end of first grade is positive. The mean math score is 48.0 points with a positive growth rate. The sign of the covariance of intercepts and slopes was different for reading and math outcomes. A negative correlation for reading indicates that children who start at a higher reading level in the spring of kindergarten improve reading skills less rapidly over time than their peers who begin at a lower level. In other words, students who started out worse appear to be catching up with students who started out with better reading skills. A positive correlation for math indicates that children who started at a lower math level in the spring of kindergarten improve math skills more rapidly over time than their peers who begin at a higher level.

The main effects of the school context variables are positively associated with outcomes. School climate has positive effects on both reading and math in spring of kindergarten, when other covariates are set to reference categories and age is 8.5 years. Schools with greater support from community and parents and more orderly environments are more conducive for student learning. Teacher press was positively associated with math, but not reading, in spring of kindergarten, setting other terms to zero. The results provide evidence for hypothesis 1 that school context is associated with better academic outcomes in spring of kindergarten, net of the HLE.

Hypothesis 2 specified positive effects of HLE on student outcomes, net of school context characteristics. The results show, indeed, HLE is positively associated with outcomes, net of school context effects. The composite for parent-child interactions is positively associated with reading outcomes, but not math, in spring of kindergarten, setting all other covariates to zero. Books in the home and computer learning are positively associated with both outcomes at the initial time point, setting all other covariates to zero. These results demonstrate the importance of the HLE to learning, suggesting students benefit from home environments where parents spend time engaged in educational activities with them as well as having access to educational resources, independent of the school environment.

2.8.3 Moderation Effects

Hypothesis 3 concerns the moderation relationship between the HLE and school context. As demonstrated above, the HLE and school context are positively associated with achievement, net of the other institution. The question of moderation goes a step further to test whether HLE effects change in better or worse school contexts, or whether the HLE effect is constant for all school context types. Table 2.2 presents selected results for the moderation analysis, demonstrating that the effect of books in the home is moderated by school climate over time. Complete model results are presented in appendix table B.5.

Before presenting the effects of the three-way interactions, I begin by presenting results for the main effects and two-way interaction effects. The main effect of school climate is negative for reading and non-significant for math in spring of kindergarten, setting other terms to zero at 8.5 years old. Over time, however, the positive two-way interaction term between school climate and age for both outcomes shows that the slope of growth is faster for students in better school environments. This result suggests that more distal school environment characteristics are more important for older children, compared to younger children, who are likely more influenced by classroom and teacher characteristics.

Turning to HLE effects, main effect results from the interactive models are consistent with the additive model results. All three HLE main effects are positively associated with the outcomes in spring of kindergarten, setting other covariates to zero at age 8.5, with the exception of the association of parent-child interactions and math. The two-way interaction for books in the home and age indicates that the slope of growth is faster for children with more learning materials in their home, setting other terms to zero and mean-centered age.

For reading outcomes, the two-way interaction between school climate and books in the home is non-significant in the spring of kindergarten, setting other terms to zero at age 8.5. The two-way interaction between school climate and books in the home is positive for math, which indicates that the association of books in the home is stronger for children who attend schools with better school climates. This reflects a complementary relationship in which children in better school environments benefit more from better HLEs at the initial time point.

Next, I present the results for the three-way interaction with age to address the question of moderation. A negative coefficient for the three-way interaction of school climate, books in the home, and age was found. For children in better schools, the growth rate is slower as books in the home increases. This suggests that the home learning environment association with learning is weaker for children in schools with better climates. This interpretation considers school climate as the moderator. However, since this analysis is not causal but, rather, associational, it is reasonable to consider an alternative interpretation in which the HLE is specified as the moderator, instead. In this case, for children with more books in the home, the growth rate is slower as the school climate effect increases. This suggests that the school climate effect on growth rates is weaker for children in homes with more educational resources.

2.9 Discussion

In order to better understand the nature of inequality of child outcomes, a focus on school influences or parental investments is insufficient. Rather, both sources of investment must be incorporated in analyses that attempt to explain academic outcomes and the processes by which gaps emerge for children from different family backgrounds. The purpose of this study was to build upon the small but growing literature that treats the HLE and school context in meaningful ways in order to explain inequality of academic outcomes in elementary school. Using rich observational data that allowed for expansive operationalizations of the HLE and school context for a nationally-representative sample of children, the results from this study are consistent with previous studies that have demonstrated the importance of the joint effects approach.

The results of this study show that when including both HLE and school context measures in the same model specification, the effects of each institution continue to be associated with academic outcomes, net of the other institution. As expected, both HLE and school context have positive impacts on elementary school outcomes. As such, the results encourage shifting the debate from the relative importance of the family vs. school institutions on child outcomes to how these co-occurring effects operate to affect inequality.

Moreover, including both institutions in the same analysis allows for additional insights regarding how they interact to affect outcomes. The results from this study have demonstrated that there is evidence to suggest a moderation relationship. This study offers a first step to better understanding the nature of the moderation relationship. Because the study is not intended to be causal, I consider alternative interpretations where school context is the moderator and HLE is considered the moderator. In the first case, children attending schools with better climates have slower growth rates as books in the home increases. This suggests that schools are performing a compensatory function since schools are more effective for children with fewer education resources in the home. In the latter case, children with more resources in the home experience slower growth rates as the effect of school climate increases. This suggests that school context matters less for children who have better educational resources in the home. The implications for inequality are different depending on which interpretation we rely. The first scenario is associated with reduced inequality because schools are serving its intended purpose of equalizing learning opportunities regardless of family background. The latter scenario is associated with increased inequality because the school institution matters less for children who come from better resourced families. These relationships are difficult to tease out because of the highly endogenous nature of the HLE and school context. Additional work is required to determine the causal direction of these relationships.

Regardless of interpretation, evidence of a differential relationship was found. This analysis provides one step in a series of analyses that could be undertaken in order to come to stronger conclusions about how family and school investments operate in tandem to affect children from different family backgrounds.

2.10 Conclusion

Given the promise of public education institutions to mitigate social inequality, it is critical to consider the ways in which public education opportunities interact with family investment to narrow skill gaps. The results will inform the social stratification trends by providing greater insight into sources of inequality across stages of childhood and how institutions responsible for child development interact to mitigate or amplify inequality. By examining these relationships for children with varying home learning environments and school contexts, I have taken an incremental step forward in understanding their associations.

In addition to finding support for the joint effects of the home learning environment and school context in the same model specification, the findings from the analysis support a moderation relationship. Estimating the effects of the family and schools is complicated by issues related to endogeneity and, as a result, strict claims about which institution should be treated as a moderator cannot be made in this study. This study, however, lays the groundwork for future work on this moderation relationship.

Limitations may include additional measurement of investments in children not captured in this study. While the measures included in this study reflect ways that lower-income families may make up for gaps in learning at school, there are other forms of investment that were not included here that may be more accessible to higher-income families, such as advanced tutoring and standardized test prep courses. These activities offer directions for future research offer a meaningful way forward to ensure that studies of joint effects capture the full range of relationships. As such, a more expansive, thorough set of measurements of the HLE may reveal additional trends about the nature of the joint effects of the HLE and school institutions.

Table 2.1: Select Estimates from Random Coefficient Models

Fixed Effect	Reading Estimate	Reading Std. Error	Math Estimate	Math Std. Error
Age	0.952***	(0.030)	-0.662***	(0.030)
Age^2	0.483***	(0.003)	0.659***	(0.003)
School Context				
School Climate	0.245***	(0.072)	0.246***	(0.071)
School Strain	-0.034	(0.064)	-0.064	(0.062)
Teacher Press	0.075	(0.066)	0.208**	(0.067)
Home Learning Environment				
Parent-Child Interactions	2.474***	(0.297)	0.825**	(0.254)
Books in Home	0.084***	(0.018)	0.137***	(0.016)
Computer Learning	0.042	(0.194)	-0.101	(0.169)
Age*Parent-Child Interactions	0.002	(0.026)	0.030	(0.024)
Age*Books in Home	0.008***	(0.001)	0.014***	(0.001)
Age*Computer Learning	-0.038*	(0.017)	-0.065***	(0.016)
Constant	68.481***	(0.401)	48.030***	(0.383)

Random Effect	Reading Variance Component	Reading 95% CI		Math Variance Component	Math 95% CI	
Constant, r_{0i}	12.019	11.819	12.223	9.351	9.161	9.545
Age, r_{1i}	0.845	0.823	0.868	0.653	0.628	0.679
Level-1 Error, e_{ti}	7.841	7.792	7.889	8.018	7.969	8.068

Note: Model includes the full set of covariates: race/ethnicity, sex, SES, family structure, locale
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.2: Select Estimates from Interactive Random Coefficient Models

Fixed Effect	Reading Estimate	Reading Std. Error	Math Estimate	Math Std. Error
Age	0.951***	(0.030)	-0.658***	(0.030)
Age^2	0.484***	(0.003)	0.659***	(0.003)
School Context				
School Climate	-0.668***	(0.152)	-0.472***	(0.127)
School Strain	-0.049	(0.064)	-0.085	(0.062)
Teacher Press	0.060	(0.067)	0.192**	(0.067)
Home Learning Environment				
Parent-Child Interactions	2.527***	(0.297)	0.862***	(0.254)
Books in Home	0.089***	(0.018)	0.133***	(0.016)
Computer Learning	0.031	(0.194)	-0.098	(0.169)
Age*Books in Home	-0.005	(0.026)	0.022	(0.024)
Age*School Climate	0.007***	(0.001)	0.014***	(0.001)
School climate*Books	-0.036*	(0.017)	-0.064***	(0.016)
Age*School Climate*Books	-0.004	(0.002)	-0.008***	(0.002)
Constant	68.466***	(0.401)	47.981***	(0.383)

Random Effect	Reading Variance Component	Reading 95% CI		Math Variance Component	Math 95% CI	
Constant, r_{0i}	12.028	11.828	12.232	9.373	9.183	9.567
Age, r_{1i}	0.842	0.820	0.864	0.646	0.621	0.672
Level-1 Error, e_{ti}	7.840	7.791	7.888	8.018	7.969	8.068

Note: Model includes the full set of covariates: race/ethnicity, sex, SES, family structure, locale
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 2.1: Distributions of Home Learning Environment Composites

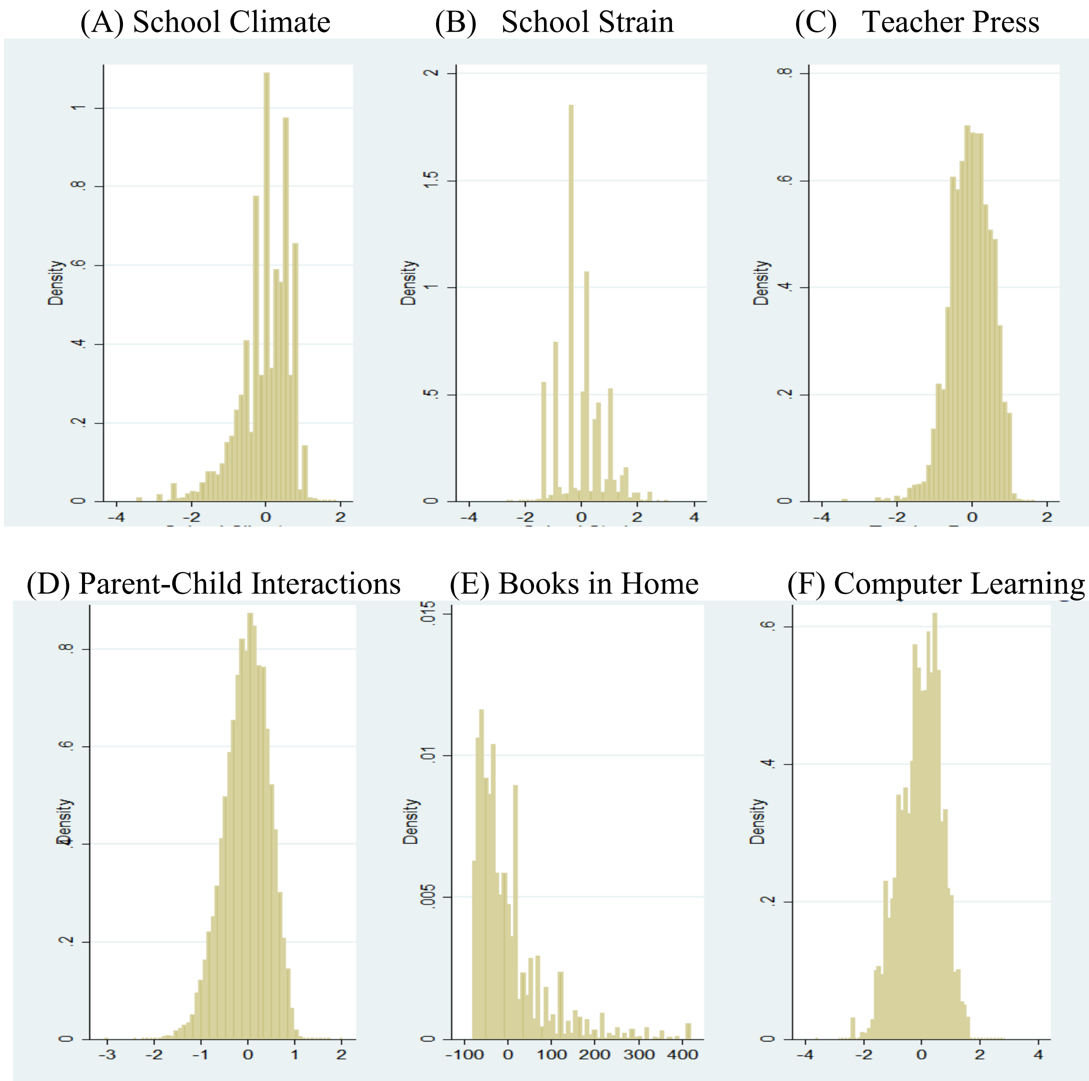


Figure 2.2: Scatterplots of Home Learning Environment Measures and School Context Measures in Kindergarten

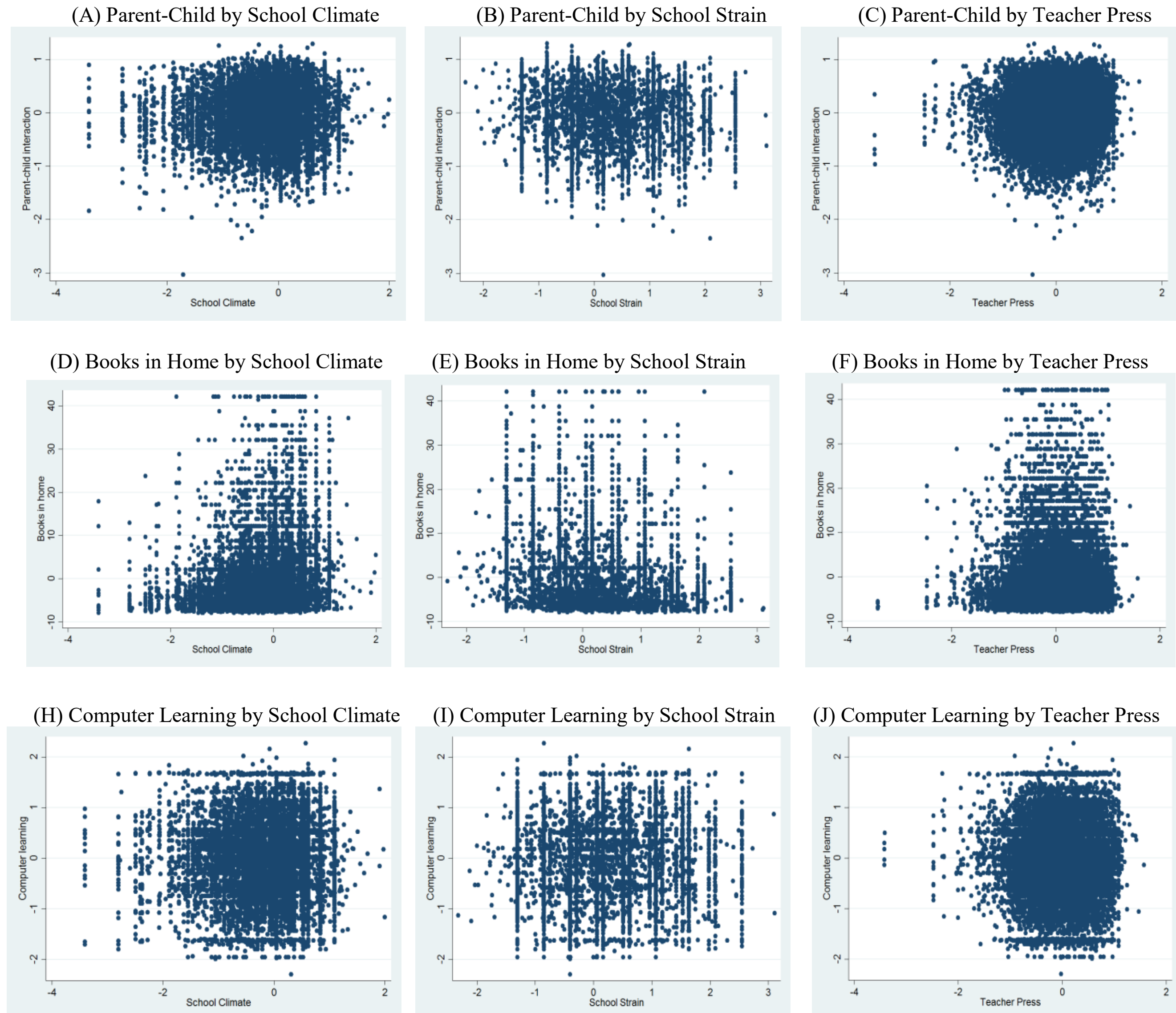


Figure 2.3.1: Mean Outcomes by School Climate over Time

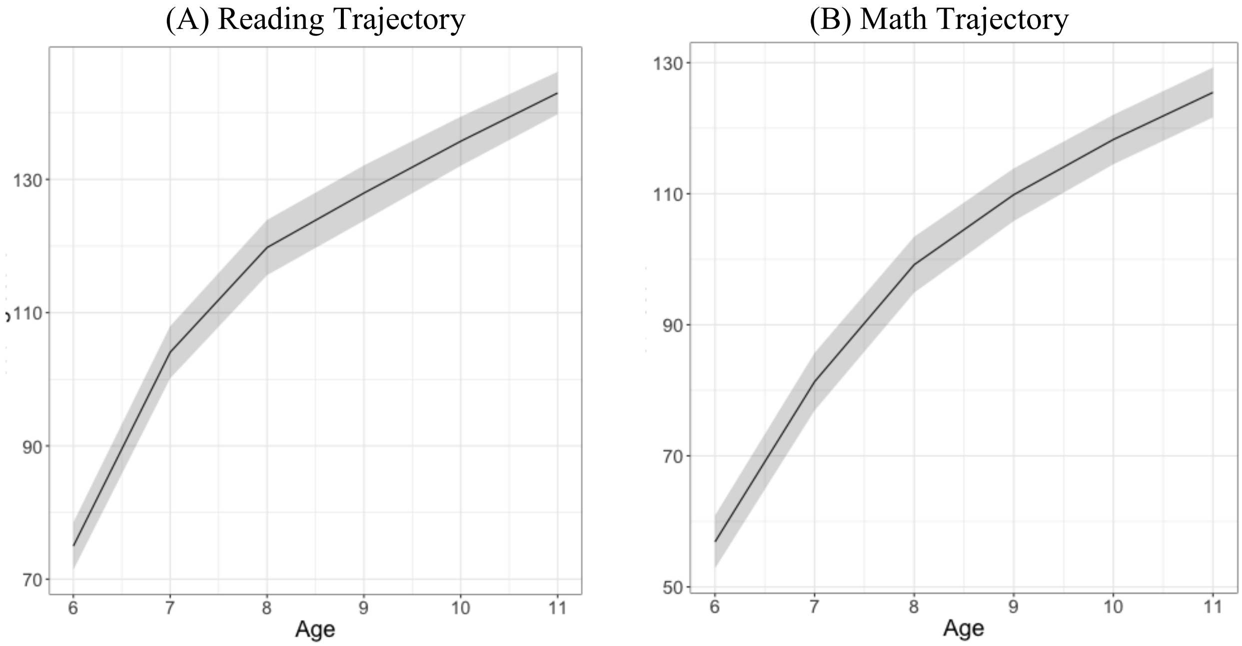
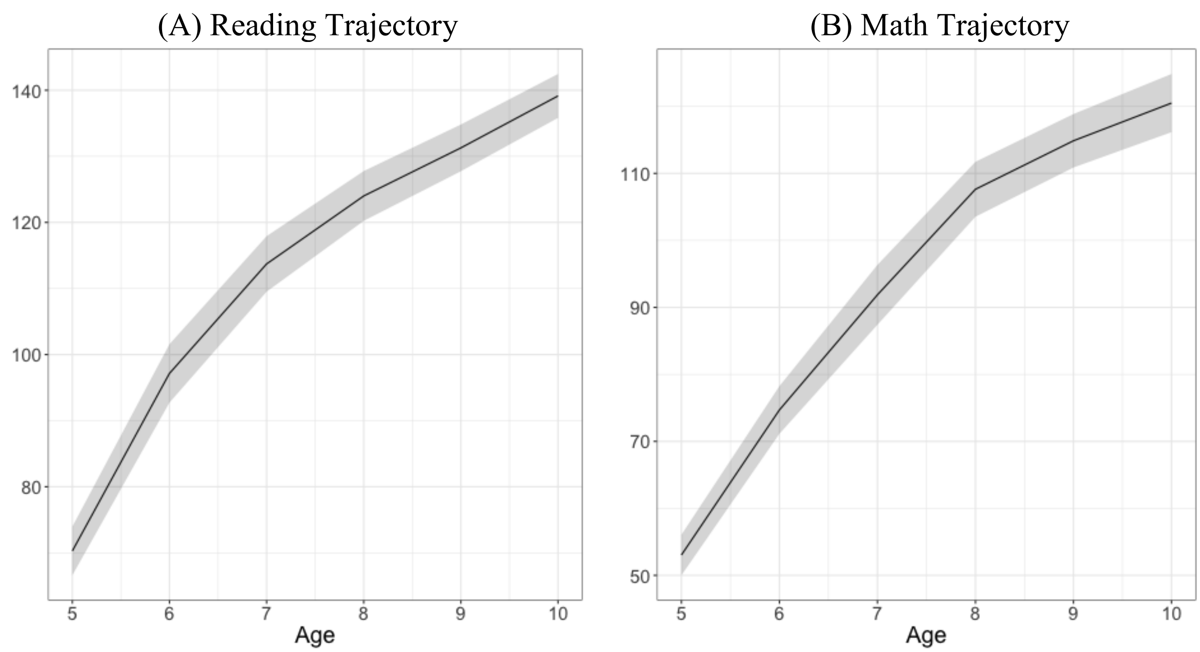


Figure 2.3.2: Mean Outcomes by Books in Home over Time



Paper 3 Early Influences on Kindergarten Readiness: An Application of Enhanced Data Integration Techniques to Address Missing-By-Design Home Environment in ECLS-K 2011

3.1 Introduction

Data combination methods have been used to combine compatible datasets from different surveys or sources in order to leverage some unique information or features to enhance the analysis. Data integration methods can enhance an analysis in a variety of ways such as correcting measurement error or combining repeated cross-sections of data to create a longitudinal dataset. Methodologically, there are different approaches to data integration relying on aggregate moments and summary statistics, for example, or merging observations at the observation level.

This study uses a novel data integration technique that applies a missing data framework to combining datasets that were not originally meant to be combined. Cross-Survey Multiple imputation (CSMI) is a departure from most of the previous data combination techniques because it uses multiple imputation principles to combine datasets at the observation level. The basic idea behind CSMI is combining a "donor" dataset that contains the full information for the analysis to impute some key missing information in a "target" dataset that contains some shared information but lacks information on key variables. In doing so, CSMI opens up a wide range of research questions that could not be previously addressed.

CSMI is subject to one important limitation that this study also addresses and attempts to rectify. While CSMI represents a step forward by developing an accessible data combination method, it relies on listwise deletion to prepare both the donor and target datasets prior to the data combination procedure, which may lead to biased results and reduced statistical power. To address this limitation, this study applies two extensions of the CSMI method that incorporate the observations that had been excluded in the original CSMI method due to listwise deletion. The two extensions of the CSMI method build upon the original method, using missing mechanism variables to predict the missing data that are re-incorporated. The Hybrid CSMI method draws upon multiply imputed data from the original approach and pairs it with chained-equation multiple imputation to impute the missing data from the observations that had been excluded by listwise deletion. The Chained-Equation Only CSMI method uses chained-equation multiple imputation to impute the missing-by-design variables as well as the additional missing data from the observations that had been excluded by listwise deletion.

The CSMI method and the two enhanced CSMI methods are applied to an empirical illustration that examines the joint effects of the home environment and preschool programs for children in the Early Childhood Longitudinal Study-Kindergarten Cohort 2011 (ECLS-K 2011). The ECLS-K 2011 dataset, the target dataset, contains all of the information for the analysis except early home environment measures, which are "borrowed" from a donor dataset— the Early Childhood Longitudinal Study-Birth Cohort which has many complementary features. This research question is one example of a variety of questions that is enabled by the use of data combination techniques.

The results show that both enhanced CSMI methods outperform the original CSMI method in terms of gains in efficiency and reduction of bias. With regard to results from the empirical illustration, differences appear in the effects of the home learning environment which were positive for the earlier ECLS-B sample, but not significant for the later cohort. Moreover, while moderation effects were found for the earlier cohort, no such moderation effects were found for the later ECLS-K sample, possibly related to effects of the Great Recession which placed strain on the effectiveness of public preschool programs. A primary contribution of this study is in the use of data combination methods, which improves the substantive model for the more recent cohort by enabling measurement of the home environment.

3.2 Data Integration Methods

3.2.1 Overview

Data combination has been used to overcome data and modeling challenges posed from using a singular dataset, or, used to enhance the analysis. Earlier approaches to data integration generally attempted to combine population-level data such as Census data and registration data with rich survey data. These methods used generalized method of moments (GMM) to constrain estimation of the analytic model based on survey data using aggregate moments from the population data (Hellerstein & Imbens, 1999; Imbens & Lancaster, 1994). A more recent set of studies adapts this method using maximum likelihood estimators (Handcock, Huovilainen, & Rendall, 2000; Rendall et al., 2008).

These studies generally place emphasis on combining sample from the same population but some studies were successful in combining samples that do not come from the same population (Handcock, Rendall, & Cheadle, 2005; Hellerstein & Imbens, 1999). One of the main advantages is that all of these studies consistently found gains in efficiencies. As more covariate information from the larger data source is included, however, these approaches become computationally intensive. In contrast, by combining datasets at the observation level, CSMI is easier to implement.

3.2.2 Within-Survey Multiple Imputation

Before elaborating the CSMI method, I first present an overview of traditional within-survey multiple imputation principles on which CSMI was developed. Multiple imputation and associated missing data theory was first developed by Rubin (1976). Multiple imputation a model-based approach to imputing missing data that creates multiple versions of the data, adding a random perturbation to each value in order to incorporate uncertainty associated with not knowing the true value. Multiple imputation has desirable properties over more rudimentary missing data handling techniques, specifically, addressing the problem of biased estimates arising from sample selection in the case of listwise deletion. It also has been shown to provide consistent estimates where other missing data handling techniques do not (van Buuren, 2012).

Traditional within-survey multiple imputation is guided by some key principles. The structure of the missing data follows one of three assumptions: 1) missing-completely-at-random (MCAR), which means that the probability of the missing data does not depend on observed or unobserved data, 2) missing-at-random assumption (MAR), which means the probability of missingness does depend on observed data but not unobserved data. Generally, the MAR assumption is considered more realistic than the MCAR assumption. 3) Finally, missing data that do not meet either of these assumption are considered missing-not-at-random (MNAR) and additional steps would be required to be able to impute data such as obtaining data that explain the missing data. If the data are assumed MAR or MCAR, the missing data mechanism is considered ignorable, a necessary assumption enabling imputation of missing data.

The appropriate multiple imputation procedure is determined by the missing data pattern. A monotone missing pattern refers to a nested missing data structure, typically observed for longitudinal datasets in which a study participant drops out of the study and never returns. Monotone multiple imputation imputes data through sequential conditional functions in which missing values for the least missing variable are imputed first. When the missing data pattern is arbitrary, chained-equation multiple imputation is used to impute data one variable at a time in an iterative way, as a function of both values from marginal distributions and successively imputed values.

After estimating the desired parameters from each imputed dataset, the estimates are combined using Rubin's Rules (Rubin, 2004). The formula to combine parameter estimates is given by,

$$\bar{Q} = \frac{1}{m}\sum_{i=1}^{m}\hat{Q}_i$$

where Q is the parameter of interest, and m is the number of multiple imputations. The total variance for the MI estimate is given by:

$$T = \frac{1}{m}\sum_{i=1}^{m}U_i + (1+\frac{1}{m})(\frac{1}{m-1}\sum_{i=1}^{m}(Q_i - \bar{Q})^2)$$

3.2.3 Cross-Survey Multiple Imputation (CSMI)

The CSMI method builds upon within-survey multiple imputation by pooling together observations from compatible samples that share some common variables and imputing information that is missing in one dataset but available in the other. There are several key principles underlying CSMI. The first principle is that the surveys are independent realizations from the same broader population. The second key principle concerns variables never jointly observed. Some previous attempts at data combination were criticized for including variables that were not jointly observed in at least one dataset. This meant the multivariate distribution of all the variables in the model was never observed, which may lead to biased estimates. To safeguard against this, CSMI requires that all variables in the substantive analytic model are available in the donor dataset so that the multivariate distribution of the analysis variables is observed.

The third principle concerns the structure of the "missing" data, or the variables not available in the target dataset that will be imputed. In traditional within-survey multiple imputation, missingness is frequently assumed MAR. Since this assumption cannot, in reality, be tested, non-response to each variable used in the analysis may, in reality, be subject to selection bias based on all other analysis variables. CSMI is, in fact, suited to meet the more stringent MCAR assumption since the information being imputed is "missing-by-design" and, therefore, non-informational for every observation.

The target dataset contains some variables common to both datasets as well as variables that are missing-by-design that were not part of data collection and, therefore, missing for all observations. The datasets are appended and values for the missing-by-design variables in the target dataset are imputed using monotone multiple imputation, drawing upon information from the multivariate joint distribution available from the donor dataset as well as information contributed from common variables in the target dataset. The multivariate distribution from the donor dataset, along with the information contributed by common variables in the target dataset, is key to imputing the missing-by-design variables in the target dataset. It is possible for correlations to vary between the two datasets but not cause notable bias or an increase in standard errors. Rendall et al. (2013) advise that sample bias resulting from the smaller of the two datasets is not problematic because the multivariate structure provided by the smaller dataset with complete information is anchored to the information provided in the larger, more representative data source.

In the literature, CSMI has been applied to a wide range of research topics (Baker, Rendall, & Weden, 2015; Brenner & Boston, 2014; Gravelle, 2018; Sintonen et al., 2016). Baker, Rendall, and Weden (2015) applied CSMI to study child obesity among immigrant children using ECLS-B and ECLS-K 1998, imputing maternal pre-pregnancy BMI for ECLS-K 1998 observations. Van Hook et al. used CSMI to combine SIPP data to estimate insurance coverage among unauthorized immigrants (Hook, Bachmeier, Coffman, & Harel, 2015). These studies demonstrate the utility of the CSMI to social science research more broadly, addressing omitted variable bias and enabling better model specification.

3.3 Enhanced CSMI Methods

A major limitation of the CSMI method results from the use of listwise deletion to prepare the donor and target datasets prior to data combination. When the donor and target datasets are appended, a monotone missing pattern results, which allows for the use of sequential multiple imputation and well-defined conditional functions. This approach is appealing since less uncertainty in the imputed values results because the missing-by-design variables are the only contributing missing data. However, listwise deletion may lead to sample selection and biased model estimates. If only a few variables are included in the model specification, sample selection from listwise deletion may not pose a large concern. When more than a few variables are included in model specification, however, as is typically the case in social science models, this concern becomes problematic. Moreover, whittling the sample may also reduce the statistical power necessary to detect relationships.

In collaborative work (Hao & D’Souza, 2019), we address the limitations of the CSMI method with two enhanced CSMI methods designed to re-incorporate observations excluded by listwise deletion: the Hybrid CSMI method and the Chained-Equation Only CSMI method. The Hybrid CSMI method is a two-step procedure which, first, draws upon the monotone multiple imputation datasets generated by the original CSMI method containing multiply imputed values for the missing-by-design variables for ECLS-K observations, and, second, combines it with the observations that were excluded by listwise deletion. In the second step of the Hybrid CSMI method, chained-equation multiple imputation is used to impute the rest of the missing values aided by missing mechanism variables. The Hybrid CSMI method combines the benefits of the monotone multiple imputation procedure with the benefits of an increased sample size.

The Chained-Equation Only CSMI method more closely resembles the standard within-survey multiple imputation procedure. This method requires appending the full donor and target analytic samples producing an arbitrary missing data pattern, which requires the use of chained-equation multiple imputation to impute the missing data, again aided by missing mechanism variables. Compared to the Hybrid CSMI method, the percentage of missing data is higher because it does not draw upon the complete datasets produced by the monotone CSMI multiple imputation procedure. This will create greater uncertainty in the estimations during the multiple imputation, which will in turn affect efficiency of estimates in the analytic models.

The three versions of CSMI are each subject to their own advantages and drawbacks as elaborated above. The original CSMI method is simplest to apply, but may produce biased estimates depending on the extent of missingness in the two samples. The enhanced CSMI methods fix the problem of missingness of the two samples but are more complicated to apply and may introduce greater variance of estimates.

3.4 Empirical Illustration

Countless research questions could be posed with the use of data combination techniques that would not be otherwise possible. The empirical example for this study examines one such research question concerning the joint effects of the home environment and early childhood programs on kindergarten readiness. Instructional regime theory (Raudenbush & Eschmann, 2015) provides a useful lens for considering the joint influence of family and school institutions on inequality in child outcomes, arguing that complete explanations require consideration of the entire landscape of learning opportunities. In early childhood, a child's instructional regime consists of the home environment and early childhood programs, and interactions between these influences may operate in boosting ways that promote inequality or in compensatory ways that reduce inequality.

For the purposes of the empirical illustration, the target dataset is ECLS-K 2011, a rich source of data on child experiences beginning in kindergarten through 5th grade. However, it lacks information on the home environment in the year preceding kindergarten, key to the research question posed. The donor dataset is ECLS-B, which contains all of the necessary information to address this research question, but follows a cohort of children that entered kindergarten over a decade ago in 2006. Examining these relationships in an up-to-date dataset such as the Early Childhood Longitudinal Study-Kindergarten Cohort 2011 (ECLS-K 2011) will yield useful insights into these relationships for contemporary children and a comparison to relationships in the earlier cohort.

The policy landscapes governing early childhood programs for ECLS-B and ECLS-K 2011 cohorts were notably different. State-funded preschool programs were experiencing a period of growth and promise in 2006. Early childhood programs expanded, received increased funding, and heightened attention toward improving quality and instituting accountability systems (Barnett, Hustedt, Hawkinson, & Kenneth B. Robin, 2006). Twenty percent of four-year olds were enrolled in prekindergarten and enrollment, part of an upward trend in recent years (Barnett et al., 2006) rising to 28% of four-year olds in 2011 (Barnett, Carolan, Fitzgerald, & Squires, 2011). The effects of the Great Recession, however, placed considerable constraints on state budgets and progress around implementation of high quality standards and programs stalled. Funding for prekindergarten decreased by $60 million nationwide and several states saw declining enrollments (Barnett et al., 2011). This strained policy context undermined the momentum of increased quality and accountability underway just five years prior.

The literature has demonstrated that prekindergarten generally provides high quality preschool programming, though the considerable amount of program heterogeneity across states in terms of enrollment, quality, and accountability, make it difficult to provide a consensus on effects. An RCT study of Tennessee’s oversubscribed prekindergarten program and found that prekindergarten students generally did not outperform control students on a range of outcomes (Lipsey, Farran, & Durkin, 2018). In contrast, an analysis of eight state prekindergarten programs using quasi-experimental methods shows that prekindergarten had positive effects on kindergarten readiness, though the magnitude of effects varied considerably across states (Barnett et al., 2018).

In addition to direct impacts on child outcomes, there is some evidence that early childhood programs may also moderate the effect of the home environment (Anders et al., 2012; Melhuish, Sylva, et al., 2008; Padilla & Ryan, 2018). Parenting quality may improve from exposure to positive child development practices provided by early childhood programs or as a result of changes in the child’s behavior and skill development from attending the program. A recent study by Padilla (2018) using ECLS-K 2011 data focuses on these relationships for immigrant children only and finds mild support for boosting effects for children of native-born Hispanic parents.

The joint relationships of public preschool programs and the home environment were likely affected by shifts in the broader national climate due to the Great Recession, whose effects were wide ranging and severe. This study is intended to provide insight into the joint effects of the home environment and public preschool programs for the ECLS-B cohort who attended preschool in 2006 *prior* to the Great Recession compared to children in the ECLS-K 2011 cohort who attended preschool in 2011, a few years *after* the Great Recession. While both datasets contain information on preschool participation in the year prior to kindergarten, only ECLS-B contains information on the home learning environment in the year prior to kindergarten. Using CSMI to impute the home learning environment for ECLS-K 2011 cohort, I will be able to make a comparison about how these relationships differed for the two cohorts.

Given prekindergarten's reputation of providing high quality preschool programming for lower-income and middle-income children, I expect that prekindergarten would continue to have positive effects even worse under conditions driven by the Great Recession. The dynamics underlying the differential relationships of public preschool programs and the home learning environment are more complicated, but are likely to differ from the earlier cohort. The compensatory nature of prekindergarten and an increasing share of disadvantaged children enrolled in prekindergarten in 2011 suggest stronger moderating effects of prekindergarten on the home environment than for similar children under parental care. However, this relationship is muddled by the effects of the Great Recession.

An important consideration to note is the imputed home learning environment values for ECLS-K 2011 observations are a function of the home learning environment variables from the

ECLS-B sample, the multivariate relationships with other variables in ECLS-B, as well as information from the common variables available in ECLS-K 2011. It is of course possible that the home learning environment in 2011 is not the same as it was in 2006. The stress placed on families as a result of the Great Recession may have undermined parenting quality in the later time period. The implication is the results for the ECLS-K 2011 provided by this analysis may present a rosier picture than reality. Ideally, I would compare the imputed home learning environment values against an external nationally-representative data source to provide a sense of how accurate the imputed values reflect reality in 2011, but no such data source exists. Nonetheless, there is some value in examining these relationships while drawing upon the older ECLS-B information, while keeping in mind the assumptions underlying the imputed home learning environment values.

3.5 Research Questions and Hypotheses

The following sets of question guide the current study. Beginning with a set of methodologically focused questions and followed by substantive questions.

3.5.1 Methodological Questions

Research Question 1: Do the enhanced CSMI methods outperform the monotone CSMI method?

Hypothesis 1: Enhanced CSMI methods should outperform the regular CSMI method in terms of efficiency gains as well as decreasing possibility of biased estimates. Enhanced CSMI methods take advantage of fuller sample sizes because they do not rely on listwise deletion in the data preparation stage. While there is tension between more missing data associated with enhanced CSMI methods which introduces additional uncertainty into the multiple imputation procedure and therefore might affect the efficiency of the estimates in the analytic models, I expect the substantial increase in sample size will outweigh the downside of the additional uncertainty.

Research Question 2: Which CSMI method performs best with regard to efficiency and bias?
Hypothesis 2: Of the enhanced CSMI methods, I expect the hybrid CSMI method to outperform the chained-equation only CSMI method. The hybrid CSMI procedure takes advantage of the monotonic multiple imputation conducted for the original CSMI procedure which has the advantage of well-defined conditional functions. It therefore has less missing data compared to the chained-equation only CSMI method.

3.5.2 Substantive Questions

Research Question 3: What are the effects of the home environment on kindergarten readiness net of early childhood programs?
Hypothesis 3: The home environment is the most important influence on child development in early childhood. Net of early childhood programs, the home environment will have positive effects on kindergarten readiness outcomes.
Research Question 4: What are the effects of prekindergarten on kindergarten readiness, net of the home environment?
Hypothesis 4: Prekindergarten has an established record of improving kindergarten readiness, in large part because of its high quality and use of standard curriculum, and in keeping with prior efforts to increase the effectiveness of the program will have a positive effect on kindergarten readiness.

Research Question 5: What are the joint effects of the home environment and prekindergarten on kindergarten readiness and how do these relationships compare to the previous ECLS cohort that experienced differing policy contexts?

Hypothesis 5: Whether the home environment is moderated by prekindergarten programs to a stronger or weaker degree in the later cohort is subject to several tensions. While prekindergarten programs in the later cohort serve a greater share of more disadvantaged children, it may be the case that a stronger moderating effect is observed. However, this effect may be undermined due to increased stressed on parents caused the Great Recession.

3.6 Data

The Donor data for this study comes from the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B) which contains all variables in the fully specified model. ECLS-B followed a cohort of 10,688 children from birth through kindergarten entry beginning in 2001. The Target data comes from Early Childhood Longitudinal Study- Kindergarten Cohort 2011 (ECLS-K 11), which is missing information on the early home environment. ECLS-K followed a cohort of 18,174 children from kindergarten entry through fifth grade beginning in 2010.

These studies were chosen for the purposes of data combination because they have a degree of complementarity. Both studies were commissioned by NCES, follow nationally-representative samples of American children. Since these studies occurred within a five year time frame of each other and therefore meet the assumption of being drawn from the same superpopulation. These studies both employed a complex, multi-stage sampling strategy intended to reflect the population of American children. They used similar survey and assessment instruments and measurement of key variables is often similar. The data used in this cross-sectional analysis intentionally capitalizes upon the point of data collection at which the studies overlap— kindergarten entry, increasing the likelihood of the compatibility of the data being combined.

3.7 Measures

Measures used in the analysis are either observed in both datasets, or only observed in the donor dataset. Variables common to both datasets in this analysis include outcome measures, early childhood program participation, all family and child characteristics, and the missing mechanism variable. The home environment variables are observed in the donor dataset, ECLS-B, and considered missing-by-design in the target dataset, the ECLS-K dataset.

Outcomes

Kindergarten readiness outcomes are available in both target and donor datasets and include reading and math IRT scores in the fall kindergarten wave. The reading assessment measures both language and literacy skills such as letter recognition, phonological awareness, knowledge of print conventions, word matching, initial understanding, interpretation, and vocabulary. The math assessment measures skills related to number sense, counting, operations, and pattern recognition.

Early Childhood Programs

Early childhood program participation is available in both target and donor datasets and is based on responses from the preschool wave parent interview the preschool wave. The variable contains five categories: Head Start, public prekindergarten, private preschool, relative care, other non-parental care such as relative care or babysitting for example, and parental care only serving as the reference category. The response rate to the parent survey in the ECLS-K fall kindergarten wave was 70 percent. Therefore, there is considerable missing data on early childhood program participation in the year prior to kindergarten.

Child Characteristics

Child characteristics include race/ethnicity, sex, age, low birthweight status. Child's race/ethnicity consists of five categories: white, African-American, Hispanic, and Asian/other. Sex is a dichotomous variable (male=1). Age at the time of the assessment in the fall kindergarten wave is measured in months. Low birthweight status is a dichotomous variable reflecting a weight of less than 2500g at the time of birth.

Family Characteristics

Family structure contains three categories: two biological or adoptive parents present, step-family, or single-parent family and other type. Poverty status is a dichotomous variable where 1 reflects poverty and 0 reflects non-poverty. Mother's education is measured in years of schooling.

Missing Mechanism

Since the enhanced CSMI methods incorporate observations with some missing data, a missing mechanism is required to predict missing data, as will be described in detail below. The missing mechanism must predict the missing data but must not be correlated with the outcome in the analysis model. The missing mechanism used in this analysis is a dummy variable for whether disruptions or interruptions occurred during child assessment. Disruptions and interruptions during the child assessment are a reflection of the conditions present during the home visit and offers explanations for poor data quality. Accordingly, it contains useful information to explain varying data quality and the absence of data.

Home Learning Environment Measures (ECLS-B only)

Home learning environment measures for the year prior to preschool are available only in the donor dataset, ECLS-B, and are missing-by-design in the target dataset, ECLS-K 2011. Two primary dimensions of the home environment are included in the analysis— parental stimulation of cognitive development and parental emotional supportiveness— obtained from observations of parent-child interactions conducted during the preschool wave. The parent and child dyad were asked to participate in a 10 minute, semi-structured activity called the Two Bags Task in which two bags containing a book and a game, respectively, were supplied. The parent-child interactions were videotaped and subsequently coded by trained staff on five dimensions of parental behavior using a 7-point Likert scale ranging from very low to very high with high levels of interrater reliability.

Parental emotional supportiveness is measured by the parent's emotional and physical expressions toward the child. Displays of enthusiasm, praise for the child's actions, and physical affection during the interaction reflect higher levels of emotional support. Parental stimulation of cognitive development reflects the parent's ability to provide developmentally-appropriate teaching, instruction, and demonstration in a way that encourages the child to progress to the next stage of development. Highly stimulating interactions may include engaging in pretend play, logical presentation of steps in an activity, and elaborating during book-reading.

3.8 CSMI Procedure

The CSMI steps are presented in brief below and then elaborated upon:

Step 1. Specify cross-survey analysis and imputation models

Step 2. Harmonize variables across datasets

Step 3. Compare multivariate distributions between donor and target datasets

Step 4. Test for survey differences using model-fit diagnostics

Step 5. Append donor and target data and impute missing-by-design variables in target dataset using monotone multiple imputation

Step 5a. For Hybrid CSMI method, draw upon monotone multiply imputed data and combine with observations that have arbitrary missing data pattern and impute missing data using chained-equation multiple imputation.

Step 5b. For Chained-Equation Only CSMI method, append donor and target data and impute missing-by-design variables as well as observations with arbitrary missing data pattern using chained-equation multiple imputation.

Step 6. Estimate analysis model using target CSMI data

Step 7. Combine estimates using Rubin's Rule

Linear regression models were used to test three hypotheses regarding the effects of the home environment and early childhood programs on reading scores at kindergarten entry. To test hypotheses 1 and 2, I estimated the following equation:

$$Y_i = \beta_0 + \beta_1 CS_i + \beta_2 ES_i + \beta_3 EC_i + \beta_4 X_i + u_i \qquad (1)$$

where Y_i is reading or math score at kindergarten entry for child i, CS_i is cognitive stimulation in the home, ES_i is emotional support in the home, EC_i is the early childhood program represented by four dummy variables for Head Start, prekindergarten, private preschool, and other non-parental care with parental care only as the reference category, X_i is a vector of covariates, and u_i is the error term.

Equation 2 was used to test hypothesis 3 regarding whether the effect of the home environment on kindergarten readiness is moderated by the effect of early childhood programs:

$$Y_i = \beta_0 + \beta_1 CS_i + \beta_2 ES_i + \beta_3 EC_i + \beta_4 X_i + \beta_5 CS_i * EC_i + \beta_6 ES_i * EC_i + u_i \qquad (2)$$

where CS_i*EC_i is the interaction of cognitive stimulation and early childhood program, and ES_i*EC_i is the interaction of emotional support and early childhood program.

The imputation models for the missing-by-design variables—cognitive stimulation and emotional support in ECLS-K 2011 are given by:

$$CS_i = \delta_0 + \delta_1 Y_i + \delta_2 EC_i + \delta_3 X_i + u_i \qquad (3)$$

$$ES_i = \delta_0 + \delta_1 Y_i + \delta_2 EC_i + \delta_3 X_i + \delta_4 CS_i + u_i \qquad (4)$$

Since listwise deletion was used to prepare the data, a nested structure of the missing pattern is created enabling the monotone multiple imputation procedure to be used.

Data preparation of the two surveys included harmonizing common variables. In addition, I compared the multivariate distributions of the two datasets by examining correlations of common variables in the two datasets and weighted univariate and bivariate distributions. This step provides an informal understanding of whether the surveys come from the same population and confirming the nature of the relationships of analysis variables.

I then tested for survey differences in a more formal way using a model-fit approach. Rendall et al. (2013) recommend comparing diagnostics of three nested models based on the pooled listwise deletion sample, given by equations 5a, 5b, and 5c:

$$Y_i = \beta_0 + \beta_1 EC_i + \beta_2 X_i + \varepsilon_i \qquad (5a)$$

$$Y_i = \beta_0 + \beta_1 EC_i + \beta_2 X_i + \beta_3 S_i + \varepsilon_i \qquad (5b)$$

$$Y_i = \beta_0 + \beta_1 EC_i + \beta_2 X_i + \beta_3 S_i + \beta_4 EC_i * S_i + \beta_5 X_i * S_i + \varepsilon_i \qquad (5c)$$

Equation 5a regresses the outcomes on the common variables— early childhood program and the covariates. Equation 5b adds a dummy term for survey membership. Equation 5c adds interaction terms between the survey membership dummy variable and the covariates, given by 5c. Model-fit is compared using AIC and BIC statistics.

Once compatibility of the surveys was demonstrated, listwise deletion was used to rid the datasets of observations with any missing values. The donor and target datasets were appended and the two home learning environment variables were then imputed sequentially using monotone multiple imputation for ECLS-K observations based on the multivariate distribution observed in the ECLS-B sample as well as using information from the common variables available in ECLS-K.

The Hybrid CSMI method combines each complete dataset produced by the monotone multiple imputation data from part 1 with the newly added observations from ECLS-K, as well as the ECLS-B observations that were discarded in part 1. In addition, the pooled dataset includes missing mechanism variables. Since the missing data pattern was no longer monotone, chained-equation multiple imputation was used to impute the rest of the missing data. The Chained-Equation Only CSMI method imputed missing-by-design data and missing data contributed from observations that would have been deleted under listwise deletion procedures using chained-equation multiple imputation on the pooled dataset aided by missing mechanism variables. Once the complete datasets were prepared using CSMI and the enhanced CSMI methods, the analysis models were estimated and the estimates were combined using Rubin's Rules.

The analysis proceeds in three parts summarized in table 3.1. Part 1 applies the original application of the CSMI method which relies on listwise deletion to create a monotonic missing pattern in the pooled dataset. The resulting sample sizes are ~5,600 observations for the ECLS-B listwise deletion sample and ~11,400 observations for the ECLS-K listwise deletion sample. Parts 2 and 3 are intended to address the problem of listwise deletion by applying the enhanced CSMI methods to increasingly larger ECLS-K samples. Part 2 applies both enhanced CSMI methods to a larger ECLS-K sample, adding children who have available outcome data, resulting in a sample size of ~16,450 observations. Part 3 applies both enhanced CSMI methods to yet a larger ECLS-K sample of ~18,150 observations— nearly the entire original sample, thereby, increasing assurance the estimates are not subject to sample selection. The drawback, however, is the additional uncertainty introduced in the multiple imputation procedure due to the addition of observations with missing outcome data. ECLS-K samples in parts 2 and 3 are compared to a larger ECLS-B sample of ~7,600 observations that was multiply imputed using regular within-sample multiple imputation.

3.9 Results

3.9.1 Results from Original CSMI Method Application

In this section, I provide a brief overview of results from the application of the original CSMI method developed by Rendall et al. (2013) before describing results from the enhanced CSMI methods in the next section. First, distributions of common variables in the donor and target samples are compared in Table 3.2. The table demonstrates very little difference in the overall means of the common variables between the two datasets. One exception occurs for the distribution of early childhood program in the 11k ECLS-K sample versus the two larger ECLS-K samples. The proportion of children in parent-care in the 11k ECLS-K sample is 0.22 compared to about 0.12 in the larger ECLS-K samples, which may reflect sample selection. Appendix tables C.1 and C.2 provide correlations of the common variables in the two samples used in part 1. Largely, the correlations across the two datasets are consistent. There were some exceptions of note. Correlations between prekindergarten and the outcome variables and between prekindergarten and Hispanic had differing signs in ECLS-B and ECLS-K samples. These correlations, less than 0.10, are weak, but the sign inconsistency is important to note.

Next, additional investigation of the comparability of the samples was conducted using a model-fit approach. Appendix table C.6 presents diagnostics from models based on the pooled listwise sample that tested for survey differences. Model 1 is the same as the analysis model but includes common variables only. Model 2 builds upon model 1 by including a dummy for survey membership. Model 3 builds upon model 2 by including interaction terms for the common variables and the survey dummy. The model fit statistics, AIC and BIC, indicate that Model 1 was least preferred. Differences in the AIC and BIC values between Models 2 and 3 were minimal. Overall, the results suggest that sample differences do not warrant an alternative approach that relies on compatible subsamples. Appendix tables C.7 and C.8 provide estimates from models predicting the missing-by-design variables after the CSMI procedure. The results show stability in the estimates from the ECLS-B models and the ECLS-K post-CSMI models. A couple of differences emerged for the emotional support results, as shown in appendix table C.8. The coefficients for two parent family, head start, and private preschool were significant in the ECLS-B sample, but did not reach significance in the ECLS-K sample.

Tables 3.3 and 3.4 present selected results from the substantive analysis models (see appendix tables C.9 and C.10 for full model results). The tables present results based on the listwise deletion 5k ECLS-B sample and the 11k CSMI ECLS-K sample. In Model 1, the additive model, the effects of public early childhood programs are consistent—the effect of prekindergarten is positive while the Head Start effect does not reach significance. With regard to the home learning environment, the effect of cognitive stimulation is positive for both samples, while emotional support is only positive and significant in the ECLS-B sample. While some evidence for moderating effects by early childhood programs was found for the ECLS-B cohort, no moderating effects were found in the ECLS-K models.

Overall, results from part 1 show consistency in the main effects of the home environment as well as the effect of prekindergarten for the two cohorts. Some differences emerged in the moderating effects of early childhood programs. Because there are potential concerns regarding sample selection, it is unclear whether these results would be observed when the full samples are used. The enhanced CSMI methods used in parts 2 and 3 provide the means by which to adjudicate whether sample selection has affected the results.

3.9.2 Results from Enhanced CSMI Methods Application

In this section, I present results using the enhanced CSMI methods for two increasingly larger ECLS-K 2011 samples. By comparing the results from the two enhanced CSMI approaches for the two larger samples, it will provide insights on how these methods perform compared to the original CSMI method and how trade-offs regarding sample size and prevalence of missing data affect gains in efficiency and bias of estimates.

Weighted univariate statistics are presented in Table 3.2 for the 7k ECLS-B sample, 16k ECLS-K sample, and 18k ECLS-K sample. There is a considerable amount of consistency in the distributions. The main difference, mentioned in part 1, is that the distribution of early childhood programs is different in the ECLS-B sample versus the larger ECLS-K samples, suggesting that parent-care declined in the newer ECLS-K samples and that more children entered preschool programs, particularly Head Start. Supplemental tables C.3, C.4, and C.5 present correlations of common variables in the full ECLS-B sample and the two larger ECLS-K samples. Overall, the correlations are consistent across samples.

Appendix tables C.11 and C.12 compare results from models that predict the missing-by-design variables after CSMI to models based on ECLS-B data, providing an initial assessment of the quality of the CSMI data. First, the standard errors from the CSMI estimates are consistently larger than standard errors from the models based on ECLS-B data, reflecting the random perturbations applied to imputed values. Generally, there is consistency in the coefficients across models. However, many of the estimates from the 18k Hybrid ECLS-K sample do not reach significance.

3.9.2.1 Do Enhanced CSMI Methods Outperform the Original CSMI Method?

The criteria by which to determine whether the enhanced CSMI methods outperform the original CSMI method include consistency of estimates and gains in efficiency. Tables 3.5 and 3.6 present selected results for the full models estimating math and reading scores based on the 16k and 18k ECLS-K samples using the two enhanced CSMI methods. See appendix tables C.13 and C.14 for complete model results. With regard to math results presented in table 3.5, the effect of Head Start on math is positive and significant for the two larger ECLS-K samples, but does not reach significance in the 11k ECLS-K sample. It is possible the reduced sample size prevents the effect of Head Start from being detected. The prekindergarten effect is positive and significant across all samples and methods. With regard to the missing-by-design home environment variables, the results are mostly uniform. The effect of cognitive stimulation on both reading and math is positive and significant for all samples and models with the exception of the 18k Hybrid ECLS-K sample. The effect of emotional support is not significant in any of the models.

Next, I consider efficiency of estimates across CSMI methods. Figures 3.1 and 3.2 present ratios of select standard errors for the enhanced CSMI methods compared to the original CSMI method in order to determine whether expected gains in efficiency are realized by using increased sample sizes. Ratios (standard errors from enhanced CSMI models / standard errors from original CSMI models) under 1 indicate that the standard errors of the enhanced CSMI methods are more efficient. See appendix tables C.15 and C.16 for exact ratios. The results indeed demonstrate that all enhanced CSMI methods outperform the monotone CSMI method in terms of efficiency. These gains in efficiency occur in spite of the uncertainty introduced by the additional missing data present in the observations that were previously excluded using the listwise deletion approach.

3.9.2.2 Which Enhanced CSMI Method Performs Best?

Having demonstrated that enhanced CSMI methods outperform the original CSMI method, I now consider which enhanced CSMI method performs best. As shown in Tables 3.5 and 3.6, there is considerably consistency in the sign and significance of the estimates with the exception of cognitive stimulation on reading in the 18k Hybrid ECLS-K sample. In terms of gains of efficiencies, shown in Figures 3.1 and 3.2, while all enhanced CSMI methods outperform the original CSMI method, there is no clear distinction between the enhanced CSMI methods and samples. The 18k Hybrid CSMI method produced slightly larger gains in efficiencies. Since the efficiency gains are generally similar and most of the estimates are stable, the 18k hybrid method, in theory, must be preferred. By virtue of including all children in the sample, this sample precludes sample selection bias. Interpretations of the final substantive analyses in the following section are based on the 18k Hybrid ECLS-K sample.

3.9.2.3 Substantive Results Using the 18k Hybrid ECLS-K Method

Returning to the empirical illustration of this study, I now consider the joint effects of the home environment and early childhood programs in an up-to-date nationally, representative sample. Tables 3.7 and 3.8 present math and reading results from models based on the 7k ECLS-B sample and the 18k Hybrid ECLS-K sample. Complete model results are presented in Supplemental Tables C.17 and C.18. Several differences emerged in the results for the two cohorts of children. Model 1, the additive model, shows the effect of Head Start on math readiness is non-significant in the 7k ECLS-B models whereas its effect is positive in the 18k Hybrid ECLS-K models. The effect of prekindergarten is positive for both samples. With regard to reading results, the effects of the public early childhood programs are consistent across samples.

With regard to the home learning environment, while the effect of cognitive stimulation is positive on math and reading for the 7k ECLS-B sample, it is not significant in the 18k Hybrid ECLS-K sample. The effect of emotional support does not reach statistical significance.

There were differences in the two samples with regard to the moderation analyses. Moderation effects were observed for the ECLS-B sample but not for the 18k Hybrid ECLS-K sample. The effect of emotional support on kindergarten readiness for children attending Head Start was stronger than for children in parent-care only. The effect of cognitive stimulation on kindergarten readiness was stronger for prekindergarten children than for children in parent-care only. No such moderation effects were found in the ECLS-K CSMI models.

3.10 Discussion

The purpose of this study was to apply a data combination technique, CSMI, in order to be able to address research questions that would not be otherwise answerable. CSMI combines data at the observation level, borrowing information from a complete dataset in order to impute missing-by-design information in a second dataset. CSMI was applied to an empirical illustration that examined the joint effects of the home environment and early childhood programs on kindergarten readiness.

One drawback of the CSMI method is that it relies on listwise deletion to create a monotonic missing pattern necessary for the multiple imputation procedure, raising concerns about sample selection and biased estimates. I use two additional enhanced CSMI methods to address these concerns— a Hybrid CSMI method and a Chained-Equation Only CSMI method which combines elements of the monotonic and chained-equation multiple imputation methods in order to be able to preserve fuller sample sizes. These enhanced CSMI methods were also applied to two increasingly larger ECLS-K samples in order to address sample selection concerns from the original CSMI method.

The results of the analysis revealed that the enhanced CSMI methods outperformed the original CSMI method. The Hybrid CSMI method using nearly all children in the ECLS-K sample was chosen for the substantive analysis because it demonstrated slightly better efficiency gains and ensures the least biased estimates. The substantive analysis was intended to provide an informal comparison of these relationships for two cohorts of children who attended early childhood programs in the context of considerably different policy landscapes and economic conditions due to the Great Recession. While children in the ECLS-B study attended preschool in an era of growth and improvement in publicly-funded early childhood programs, children in the ECLS-K 2011 study attended preschool during a strained policy climate, leading to stagnation in standards and accountability.

The substantive results indicated that the joint effects, indeed, varied for the two cohorts with worse implications for the later cohort. First, the effect of prekindergarten was consistently positive for both cohorts. The effect of Head Start on math was positive in the ECLS-K sample but was not significant for the earlier cohort. The home learning environment effects also showed some differences across the two cohorts. While cognitive stimulation was positively associated with kindergarten readiness for the earlier cohort, it was not significant for the later cohort. Moreover, while the effect of the home learning environment was found to be moderated by public early childhood programs for the earlier cohort, there were no such moderation effects for the later cohort. That no moderation effects were observed for the later cohort suggests that the effectiveness of public early childhood programs may have been constrained by the Great Recession.

3.11 Conclusion

Data combination techniques are useful in allowing researchers to conduct analyses that would otherwise not be possible. In the case of this study, cross-survey multiple imputation was used to combine data from two nationally-representative studies of early childhood, which allowed for an analysis of the joint effects of the home environment and early childhood programs on kindergarten readiness for two cohorts of children attending kindergarten under very different national climates and policy conditions. Substantively, while the study shows that the home environment and early childhood programs promote kindergarten readiness under both sets of conditions, early childhood programs were more effective moderators of the home environment in the earlier time period, characterized by an era of growth and improvement in early childhood programs.

Moreover, this study offers evidence that, while the original application of CSMI is a useful approach for addressing omitted variable bias and imputing information for missing-by-design variables, its main shortcoming is its reliance on listwise deletion in the data preparation stage which raises concerns about sample selection. Enhanced versions of CSMI methods work toward addressing this shortcoming, but further development of these methods must be undertaken. Future directions include using simulations to determine the threshold at which listwise deletion becomes problematic in order to determine when enhanced CSMI methods may be necessary.

Table 3.1: Summary of Analysis Plan

	ECLS-B Sample (Donor)	ECLS-K Sample (Target)
Part 1 Analysis		
Original CSMI Method	~5,600	~11,400
Part 2 Analysis (Enhanced CSMI)		
Hybrid CSMI Method	~7,600	~16,450
Chained-Equation Only CSMI Method	~7,600	~16,450
Part 3 Analysis (Enhanced CSMI)		
Hybrid CSMI Method	~7,600	~18,150
Chained-Equation Only CSMI Method	~7,600	~18,150

Note: Sample sizes rounded to nearest 50 in accordance with NCES regulations.

Table 3.2. Weighted Descriptive Statistics of Common Variables in Donor and Target Datasets

	Part 1 Analysis				Parts 2 and 3 Analysis					
	5k ECLS-B (Donor)		11k ECLS-K (Target)		7k ECLS-B (Donor)		16k ECLS-K (Target)		18k ECLS-K (Target)	
	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error
Reading	43.84	0.27	53.29	0.13	43.73	0.25	52.42	0.09	52.24	0.09
Math	44.10	0.19	35.27	0.12	43.86	0.18	34.39	0.09	34.08	0.10
Race/Ethnicity										
White	0.55	0.01	0.56	0.00	0.52	0.01	0.48	0.00	0.47	0.00
Hispanic	0.25	0.01	0.22	0.00	0.25	0.01	0.25	0.00	0.25	0.00
African-American	0.14	0.01	0.12	0.00	0.15	0.01	0.13	0.00	0.13	0.00
Other Race	0.06	0.00	0.09	0.00	0.07	0.00	0.14	0.00	0.15	0.00
Male	0.51	0.01	0.51	0.01	0.51	0.01	0.51	0.00	0.51	0.00
Low Birthweight	0.07	0.00	0.09	0.00	0.08	0.00	0.09	0.00	0.09	0.00
Age at Kindergarten	6.81	0.01	6.75	0.00	6.81	0.01	6.75	0.00	6.74	0.00
Poverty	0.23	0.01	0.24	0.01	0.25	0.01	0.28	0.00	0.28	0.00
Mother's Education	13.67	0.05	14.09	0.03	13.49	0.05	13.83	0.02	13.82	0.03
Two Parent Family	0.69	0.01	0.72	0.00	0.70	0.01	0.70	0.00	0.72	0.00
Early Childhood Program										
Parent-Care	0.21	0.01	0.22	0.00	0.22	0.01	0.12	0.00	0.11	0.00
Head Start	0.16	0.01	0.13	0.00	0.16	0.01	0.20	0.00	0.20	0.00
Prekindergarten	0.15	0.01	0.16	0.00	0.15	0.01	0.17	0.00	0.16	0.00
Private Preschool	0.38	0.01	0.37	0.00	0.37	0.01	0.39	0.00	0.38	0.00
Other Non-Parental Care	0.10	0.01	0.12	0.00	0.11	0.01	0.13	0.00	0.15	0.00

Table 3.3: Selected Estimates of Math Scores from 5k ECLS-B Models and 11k ECLS-K CSMI Models

	5k ECLS-B						11k ECLS-K					
	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error
Early Childhood Program (Ref: Parent-care Only)												
Head Start	-0.07	0.402	1.136	1.385	-0.082	0.402	0.226	0.343	1.875	1.577	0.228	0.344
Prekindergarten	1.822***	0.412	1.807***	0.412	-1.251	1.571	0.898***	0.323	0.886***	0.324	0.944	1.801
Private Preschool	1.552***	0.345	1.533***	0.345	1.584***	0.345	3.614***	0.276	3.590***	0.278	3.613***	0.277
Other Care	-0.878*	0.467	-0.884*	0.467	-0.874*	0.467	0.058	0.354	0.052	0.355	0.058	0.354
Home Learning Environment												
Cog Stimulation	0.835***	0.155	0.884***	0.164	0.722***	0.164	0.890***	0.249	0.944***	0.263	0.893***	0.251
Emo Support	0.278*	0.161	0.282*	0.161	0.295*	0.161	0.265	0.255	0.269	0.254	0.265	0.256
Head Start * Emo	--	--	-0.307	0.338	--	--	--	--	-0.415	0.388	--	--
Prek * Cog	--	--	--	--	0.721**	0.356	--	--	--	--	-0.011	0.419

* p<0.05, ** p<0.01, *** p<0.001

Note: The full set of covariates are included in the models: race/ethnicity, sex, poverty, mother's education, family structure, low birthweight and age at assessment. Estimates are not presented here. See appendix table C.9 for full model results.

Table 3.4: Selected Estimates of Reading Scores from 5k ECLS-B 5k Models and 11k ECLS-K CSMI Models

	5k ECLS-B						11k ECLS-K					
	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error
Early Childhood Program (Ref: Parent-Care Only)												
Head Start	-0.058	0.581	-4.158*	2.186	-0.072	0.581	-0.012	0.358	0.808	1.693	-0.012	0.359
Prekindergarten	4.394***	0.595	4.427***	0.595	0.715	2.273	0.777**	0.335	0.773**	0.335	0.552	1.588
Private Preschool	2.675***	0.498	2.723***	0.499	2.713***	0.499	2.993***	0.288	2.985***	0.288	2.997***	0.29
Other Care	-1.668**	0.676	-1.662**	0.676	-1.664**	0.676	-0.635*	0.369	-0.636*	0.369	-0.635*	0.369
Home Learning Environment												
Cog Stimulation	1.138***	0.223	1.128***	0.223	1.002***	0.238	0.811***	0.212	0.813***	0.212	0.803***	0.22
Emo Support	0.395*	0.233	0.237	0.246	0.414*	0.233	0.289	0.205	0.315	0.207	0.289	0.205
Head Start * Emo	--	--	0.990*	0.509	--	--	--	--	-0.196	0.392	--	--
Prek * Cog	--	--	--	--	0.863*	0.514	--	--	--	--	0.053	0.37

* p<0.05, ** p<0.01, *** p<0.001

Note: The full set of covariates are included in the models: race/ethnicity, sex, poverty, mother’s education, family structure, low birthweight and age at assessment. Estimates are not presented here. See appendix table C.10 for full model results

Table 3.5: Estimates of Math Scores by CSMI Method and ECLS-K Samples

	11k Monotone ECLS-K		16k Hybrid ECLS-K		16k Chained-Equation ECLS-K		18k Hybrid ECLS-K		18k Chained-Equation ELCS-K	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Early Childhood Program (Ref: Parent-Care Only)										
Head Start	0.226	0.343	0.556*	0.286	0.492*	0.291	0.603**	0.277	0.666**	0.295
Prekindergarten	0.898***	0.323	1.390***	0.301	1.360***	0.292	1.421***	0.289	1.377***	0.304
Private Preschool	3.614***	0.276	3.919***	0.245	3.837***	0.241	3.950***	0.25	3.858***	0.246
Other Non-Parental Care	0.058	0.354	0.04	0.299	0.07	0.314	-0.035	0.312	0.041	0.304
Home Learning Environment										
Cognitive Stimulation	0.890***	0.249	0.850***	0.22	1.032***	0.234	0.186	0.157	1.014***	0.217
Emotional Support	0.265	0.255	0.266	0.225	0.178	0.226	0.048	0.158	0.149	0.228

* p<0.05, ** p<0.01, *** p<0.001

Note: The full set of covariates are included in the models: race/ethnicity, sex, poverty, mother's education, family structure, low birthweight and age at assessment. Estimates are not presented here. See appendix table C.11 for full model results.

Table 3.6. Estimates of Reading Scores by CSMI Method and ECLS-K Samples

	11k Monotone ECLS-K		16k Hybrid ECLS-K		16k Chained-Equation ECLS-K		18k Hybrid ECLS-K		18k Chained-Equation ELCS-K	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Early Childhood Program (Ref: Parent-Care Only)										
Head Start	-0.012	0.358	0.239	0.292	0.175	0.294	-0.113	0.284	-0.043	0.294
Prekindergarten	0.777**	0.335	1.317***	0.298	1.278***	0.292	1.533***	0.298	1.424***	0.327
Private Preschool	2.993***	0.288	3.369***	0.248	3.316***	0.247	3.511***	0.255	3.398***	0.26
Other Care	-0.635*	0.369	-0.356	0.303	-0.391	0.31	-0.469	0.33	-0.395	0.3
Home Learning Environment										
Cognitive Stimulation	0.811***	0.212	0.747***	0.184	0.829***	0.19	0.186	0.146	0.798***	0.194
Emotional Support	0.289	0.205	0.289	0.181	0.156	0.197	0.085	0.15	0.169	0.209

* p<0.05, ** p<0.01, *** p<0.001

Note: The full set of covariates are included in the models: race/ethnicity, sex, poverty, mother's education, family structure, low birthweight and age at assessment. Estimates are not presented here. See appendix table C.12 for full model results.

Table 3.7: Selected Estimates of Math Scores from 7k ECLS-B Models and Preferred 18k ECLS-K Hybrid CSMI Models

	7k ECLS-B						18k Hybrid ECLS-K					
	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error
Early Childhood Program (Ref: Parent-care Only)												
Head Start	0.088	0.364	-1.573	1.439	0.075	0.364	0.603**	0.277	1.098	1.359	0.602**	0.277
Prekindergarten	1.963***	0.389	1.977***	0.389	-1.648	1.57	1.421***	0.289	1.421***	0.289	1.29	1.269
Private Preschool	1.865***	0.323	1.885***	0.322	1.904***	0.323	3.950***	0.25	3.949***	0.25	3.951***	0.25
Other Care	-0.714*	0.428	-0.713*	0.428	-0.708*	0.428	-0.035	0.312	-0.035	0.312	-0.035	0.312
Home Learning Environment												
Cog Stimulation	0.849***	0.151	0.844***	0.151	0.721***	0.161	0.186	0.157	0.187	0.157	0.182	0.158
Emo Support	0.231	0.172	0.165	0.174	0.247	0.171	0.048	0.158	0.066	0.163	0.048	0.158
Head Start * Emo	--	--	0.401	0.333	--	--	--	--	-0.113	0.301	--	--
Prek * Cog	--	--	--	--	0.853**	0.36	--	--	--	--	0.031	0.293

* p<0.05, ** p<0.01, *** p<0.001

Note: The full set of covariates are included in the models: race/ethnicity, sex, poverty, mother's education, family structure, low birthweight and age at assessment. Estimates are not presented here. See appendix table C.13 for full model results.

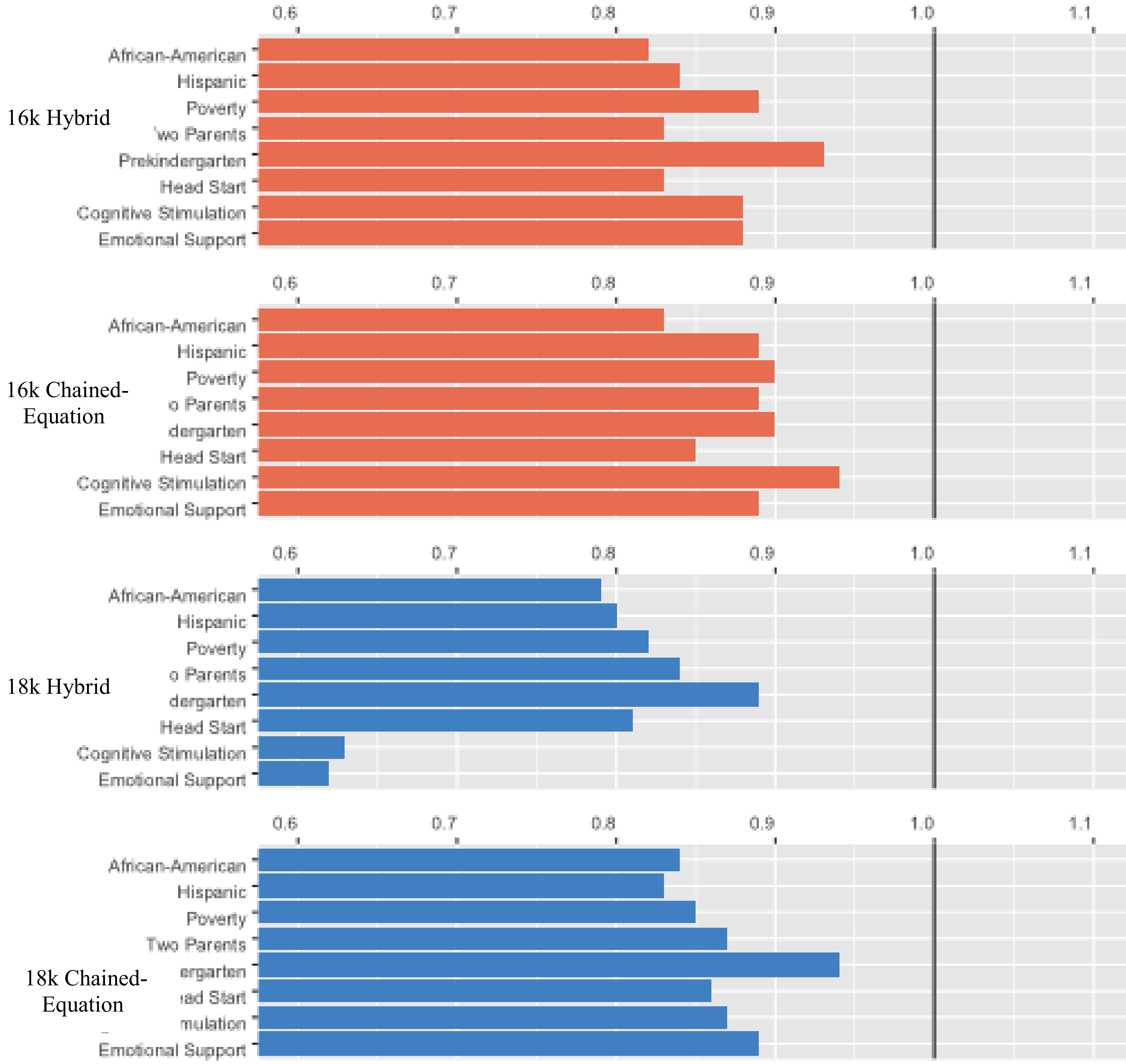
Table 3.8: Selected Estimates of Reading Scores from ECLS-B 7k Models and Preferred 18k ECLS-K Hybrid CSMI Models

	7k ECLS-B						18k Hybrid ECLS-K					
	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error
Early Childhood Program (Ref: Parent-Care Only)												
Head Start	-0.044	0.526	-3.944**	2.011	-0.059	0.526	-0.113	0.284	-0.018	1.44	-0.113	0.284
Prekindergarten	4.387***	0.555	4.418***	0.555	0.179	2.196	1.533***	0.298	1.533***	0.298	1.414	1.388
Private	2.815***	0.464	2.863***	0.464	2.861***	0.464	3.511***	0.255	3.511***	0.255	3.512***	0.255
Other care	-1.790***	0.612	-1.788***	0.611	-1.783***	0.611	-0.469	0.33	-0.468	0.33	-0.468	0.33
Home Learning Environment												
Cog Stimulation	1.152***	0.218	1.140***	0.218	1.002***	0.227	0.186	0.146	0.186	0.146	0.182	0.15
Emo Support	0.322	0.241	0.166	0.25	0.34	0.24	0.085	0.15	0.089	0.165	0.085	0.15
Head Start * Emo	--	--	0.941**	0.471	--	--	--	--	-0.021	0.321	--	--
Prek * Cog	--	--	--	--	0.993**	0.502	--	--	--	--	0.028	0.32

* p<0.05, ** p<0.01, *** p<0.001

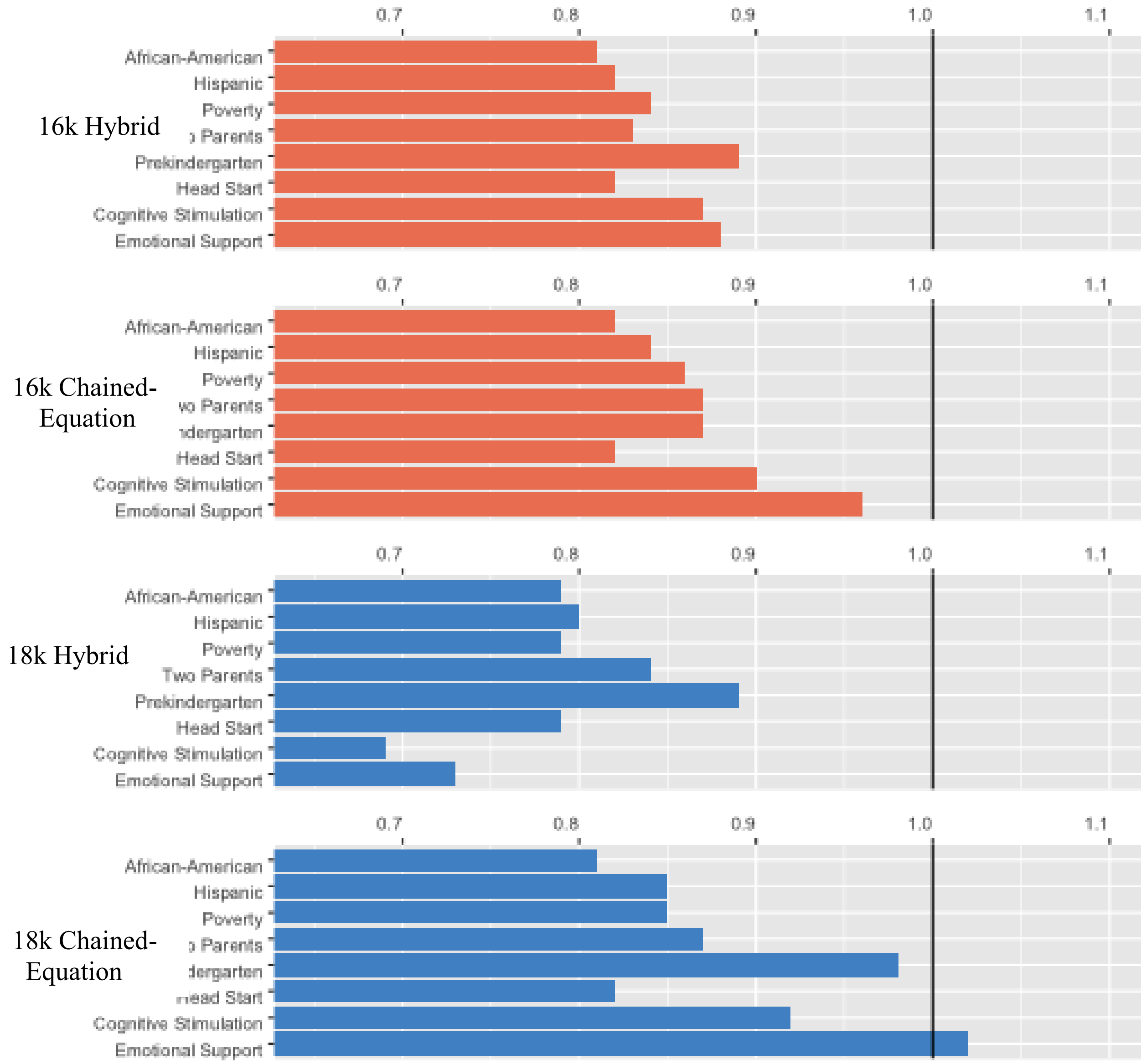
Note: The full set of covariates are included in the models: race/ethnicity, sex, poverty, mother's education, family structure, low birthweight and age at assessment. Estimates are not presented here. See appendix table C.14 for full model results.

Figure 3.1: Selected Ratios of Standard Errors of Coefficients from ECLS-K Original CSMI Models and Enhanced CSMI Models for Math (Ratios = Enhanced Std Errors / Original Std Errors)



Note: See appendix table C.15 for full set of results.

Figure 3.2: Selected Ratios of Standard Errors of Coefficients from ECLS-K Original CSMI Models and Enhanced CSMI Models for Reading (Ratios = Enhanced Std Errors / Original Std Errors)



Note: See appendix table C.16 for full set of results

Appendix A

Appendix For Paper 1

Table A.1: Unweighted Summary Statistics

	Mean	Std. Error
Reading Outcome at Kindergarten Entry	-0.023	(0.01)
Home Environment		
Cognitive Stimulation	4.141	(0.01)
Emotional Support	4.370	(0.01)
Early Childhood Group		
Parent Care Only	0.194	(0.00)
Head Start	0.171	(0.00)
Private	0.384	(0.01)
Prekindergarten	0.153	(0.00)
Other Non-Parental Care	0.098	(0.00)
Race/Ethnicity		
White	0.406	(0.01)
African-American	0.158	(0.00)
Hispanic	0.202	(0.00)
Asian	0.111	(0.00)
Other	0.123	(0.00)
Male	0.508	(0.01)
Age At Kindergarten Entry (Months)	68.13	(0.05)
SES	-0.013	(0.01)
Poverty Status	0.247	(0.00)
Low Birthweight	0.250	(0.00)
Family Structure		
Two Parent Family	0.708	(0.01)
Step-Family	0.059	(0.00)
Single-Parent Family And Other	0.233	(0.00)
Number Of Siblings	0.678	(0.01)
Expectations For Child Educational Attainment	4.189	(0.01)
Prior Childcare	0.308	(0.01)
Non-English Speaking Household	1.469	(0.00)
Region		
Northeast	0.147	(0.00)
Midwest	0.228	(0.00)
South	0.359	(0.01)
West	0.266	(0.01)
Urban	0.832	(0.00)

Table A.2: Full Model Results from Linear Regression Models Predicting Reading at Kindergarten Entry

	Model 1	Model 2	Model 3	Model 4	Model 5
Home Learning Environment					
Cognitive Stimulation	0.077***	0.076***	0.071***	0.075***	0.077***
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Emotional Support	0.014	0.014	0.015	0.0021	0.014
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Early Childhood Program (Ref: Parent-Care Only)					
Head Start	0.029	0.019	0.029	-0.28*	0.029
	(0.04)	(0.12)	(0.04)	(0.13)	(0.04)
Prekindergarten	0.16***	0.16***	0.015	0.17***	0.16
	(0.04)	(0.04)	(0.16)	(0.04)	(0.16)
Private Preschool	0.083**	0.083**	0.085**	0.086**	0.083**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Other Non-Parental Care	-0.12**	-0.12**	-0.12**	-0.12**	-0.12**
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Race/Ethnicity (Ref: White)					
African-American	-0.068	-0.068	-0.068	-0.066	-0.068
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Hispanic	-0.16***	-0.16***	-0.16***	-0.16***	-0.16***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Asian	0.33***	0.33***	0.33***	0.33***	0.33***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Other Race	-0.067*	-0.067*	-0.067*	-0.068*	-0.067*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Male	-0.16***	-0.16***	-0.16***	-0.16***	-0.16***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Age at Kindergarten Entry	0.065***	0.065***	0.065***	0.065***	0.065***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Low Birthweight	-0.21***	-0.21***	-0.21***	-0.21***	-0.21***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Education Expectations	0.085***	0.085***	0.085***	0.084***	0.085***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Prior Childcare	-0.0032	-0.0033	-0.0030	-0.0040	-0.0032
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
SES	0.33***	0.33***	0.33***	0.33***	0.33***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Family Structure (Ref: Two-Parent Family)					
Step-Family	-0.13**	-0.13**	-0.13**	-0.13**	-0.13**
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)

Single-Parent Family And Other	-0.091**	-0.091**	-0.091**	-0.093***	-0.091**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Number of Siblings	-0.088***	-0.088***	-0.088***	-0.088***	-0.088***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Non-English Speaking Household	-0.069	-0.069	-0.069	-0.070	-0.069
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Region (Ref: Northeast)					
Midwest	0.0068	0.0068	0.0070	0.0067	0.0069
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
South	0.12***	0.12***	0.12***	0.12***	0.12***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
West	0.027	0.027	0.027	0.028	0.027
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Urbancity	0.025	0.026	0.026	0.026	0.026
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Head Start * Cognitive stimulation		0.0026			
		(0.03)			
Prekindergarten * Cognitive stimulation			0.035		
			(0.04)		
Head Start * Emotional support				0.074*	
				(0.03)	
Prekindergarten * Emotional support					0.0012
					(0.04)
Constant	-5.00***	-5.00***	-4.98***	-4.94***	-5.00***
	(0.19)	(0.19)	(0.19)	(0.19)	(0.19)

* p<0.05, ** p<0.01, *** p<0.001
Note: Standard errors given in parentheses.

Table A.3: Full Model Results from Linear Regression Models Predicting Math at Kindergarten Entry

	Model 1	Model 2	Model 3	Model 4	Model 5
Home Learning Environment					
Cognitive Stimulation	0.071***	0.074***	0.063***	0.070***	0.071***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Emotional Support	0.012	0.012	0.013	0.0050	0.0094
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Early Childhood Program (Ref: Parent-Care Only)					
Head Start	0.036	0.12	0.035	-0.14	0.036
	(0.04)	(0.13)	(0.04)	(0.13)	(0.04)
Prekindergarten	0.057	0.056	0.060*	0.059*	0.058
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Private Preschool	0.070	0.069	-0.16	0.071	-0.0075
	(0.04)	(0.04)	(0.16)	(0.04)	(0.17)
Other Non-Parental Care	-0.061	-0.061	-0.061	-0.061	-0.061
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Race/Ethnicity (Ref: White)					
African-American	-0.24***	-0.24***	-0.24***	-0.24***	-0.24***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Hispanic	-0.25***	-0.25***	-0.25***	-0.25***	-0.25***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Asian	0.25***	0.25***	0.25***	0.25***	0.25***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Other Race	-0.14***	-0.14***	-0.14***	-0.14***	-0.14***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Male	-0.070***	-0.070***	-0.070***	-0.070***	-0.070***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Age at Kindergarten Entry	0.067***	0.067***	0.067***	0.067***	0.067***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Low Birthweight	-0.33***	-0.33***	-0.33***	-0.33***	-0.33***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Education Expectations	0.081***	0.081***	0.080***	0.080***	0.081***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Prior Childcare	0.017	0.017	0.017	0.017	0.017
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
SES	0.35***	0.35***	0.35***	0.35***	0.35***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Family Structure (Ref: Two-Parent Family)					
Step-Family	-0.14**	-0.14**	-0.14**	-0.14**	-0.14**
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)

Single-Parent Family And Other	-0.081**	-0.081**	-0.082**	-0.082**	-0.081**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Number of Siblings	-0.060***	-0.060***	-0.060***	-0.060***	-0.060***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Non-English Speaking Household	-0.0019	-0.0026	-0.0017	-0.0023	-0.0019
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Region (Ref: Northeast)					
Midwest	0.025	0.025	0.025	0.025	0.025
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
South	0.0069	0.0068	0.0067	0.0074	0.0069
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
West	0.0044	0.0042	0.0034	0.0050	0.0041
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Urbancity	0.0083	0.0083	0.0092	0.0087	0.0084
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Head Start * Cognitive stimulation		-0.021			
		(0.03)			
Prekindergarten * Cognitive stimulation			0.055		
			(0.04)		
Head Start * Emotional support				0.042	
				(0.03)	
Prekindergarten * Emotional support					0.017
					(0.04)
Constant	-5.01***	-5.03***	-4.98***	-4.98***	-5.00***
	(0.18)	(0.19)	(0.18)	(0.18)	(0.18)

* p<0.05, ** p<0.01, *** p<0.001
Note: Standard errors given in parentheses.

Appendix B

Appendix For Paper 2

Table B.1: Unweighted Univariate Summary Statistics

	Mean	Std. Dev	Min	Max
Reading IRT Score				
Spring Kindergarten	69.18	14.78	32.58	133.54
Spring 1st Grade	94.52	17.93	32.74	139.21
Spring 2nd Grade	111.95	17.15	45.91	146.42
Spring 3rd Grade	120.48	15.66	65.54	156.47
Spring 4th Grade	128.91	15.03	73.40	155.49
Spring 5th Grade	135.90	15.77	72.35	159.01
Math IRT Score				
Spring Kindergarten	50.23	13.44	11.75	112.54
Spring 1st Grade	72.39	15.80	12.27	138.92
Spring 2nd Grade	89.92	18.23	18.24	139.10
Spring 3rd Grade	103.64	18.11	43.41	147.89
Spring 4th Grade	112.18	18.02	25.73	147.90
Spring 5th Grade	119.38	17.98	26.76	148.04
Age In Months (Spring Kindergarten)	73.47	4.43	52.21	99.45
Teacher Press				
Spring Kindergarten	0.00	0.54	-3.42	2.59
Spring 1st Grade	0.00	0.54	-2.21	2.40
Spring 2nd Grade	-0.01	0.55	-3.31	2.21
Spring 3rd Grade	0.00	0.59	-2.66	2.44
Spring 4th Grade	0.00	0.55	-2.56	2.38
Spring 5th Grade	-0.01	0.53	-2.68	2.34
School Climate				
Spring Kindergarten	0.01	0.66	-3.73	2.60
Spring 1st Grade	-0.01	0.63	-3.69	2.56
Spring 2nd Grade	-0.02	0.66	-4.08	2.51
Spring 3rd Grade	-0.01	0.68	-4.14	2.59
Spring 4th Grade	-0.02	0.78	-4.87	3.18
Spring 5th Grade	-0.03	0.82	-4.63	3.32
School Strain				
Spring Kindergarten	0.01	0.80	-3.34	3.59
Spring 1st Grade	0.01	0.81	-2.89	4.12
Spring 2nd Grade	0.00	0.52	-2.11	2.52
Spring 3rd Grade	0.01	0.80	-3.24	4.14
Spring 4th Grade	0.02	0.81	-3.25	3.86

Spring 5th Grade	0.02	0.77	-2.78	3.88
Home Learning Environment				
Parent-Child Interaction	-0.01	0.47	-3.03	2.08
Books In Home	0.45	7.89	-7.90	42.10
Computer Learning	0.00	0.70	-3.06	2.93
Race/Ethnicity				
White	0.48	0.50		
Hispanic	0.27	0.44		
African-American	0.11	0.32		
Other	0.14	0.35		
Male	0.51	0.50		
SES	-0.07	0.82	-3.09	2.60
Two-Parent Family	0.70	0.46		
Locale				
City	0.33	0.47		
Suburban	0.36	0.48		
Town	0.08	0.27		
Rural	0.23	0.42		

Table B.2: Items Used in Composite Variables

Composite	Item
Home Learning Environment	
Parent-child interactions (alpha=0.84)	Read books to child
	How often tell stories
	How often reads picture books
	How often practice read write numbers
	Number of days family eats dinner together
	Visited a bookstore/library
	Frequently read outside of school
	Frequently looks picture books outside school
Books in Home (alpha=0.88)	How many books child has
Computer Learning (alpha=0.76)	Use computer to read stories
	Use computer for drawing/art
	Use computer to learn skills
	How often child plays computer program that teaches math or reading skills
School Context	
School Climate (alphas~0.69 to 0.73)	Parents active in programs
	Community support
	How often classroom disorder problem
	How often physical conflict prob
	Problem with aggressive/disruptive behavior
School Strain (alphas~0.63 to 0.73)	Problem with student absenteeism
	Problem with teachers absenteeism
	Problem with student tardiness
	Reduction in teaching staff
	Funding levels decreased
	Problem with teacher turnover
	Salaries decreased
	Problem with overcrowding
Teacher Press (alphas~0.63 to 0.83)	Get through to students
	Change approach
	Try different method
	Little to do for high achievement
	Work to make lessons
	Waste of time to do best
	Factors beyond control
	Learning related to family
	Habits reduce chance of success
	Increase retention
	Know techniques to avoid disruptions
	Teacher enjoys present teaching job
	Teacher makes difference in children lives

Teacher would choose teaching again
Consensus on expectations
Child misbehavior interfere with teaching

Table B.3: Weighted Means or Proportions by Key Social Stratification Characteristics

	African-American	Hispanic	Asian/Other	White	Other Family	Two-Parent	Low	Non-Low
Reading								
Spring Kindergarten	67.32	65.82	72.69	71.96	66.32	71.25	61.94	70.79
Spring 1st Grade	91.65	89.82	98.11	98.90	90.85	97.41	83.76	96.93
Spring 2nd Grade	107.97	107.41	115.20	116.18	107.98	114.65	101.09	114.13
Spring 3rd Grade	115.48	116.45	124.04	124.86	116.88	123.18	111.06	122.61
Spring 4th Grade	123.66	125.36	132.41	132.82	125.13	131.45	119.15	130.90
Spring 5th Grade	130.40	132.27	139.79	139.99	132.15	138.54	126.50	137.94
Math								
Spring Kindergarten	44.92	46.50	53.14	54.11	46.79	52.47	43.08	51.79
Spring 1st Grade	65.38	67.67	75.74	77.45	68.67	74.99	63.58	74.33
Spring 2nd Grade	79.99	84.60	95.52	95.41	84.63	93.00	79.39	91.96
Spring 3rd Grade	93.42	99.35	108.78	109.23	98.39	107.01	93.30	105.91
Spring 4th Grade	100.98	107.98	117.47	117.67	106.68	115.44	102.00	114.27
Spring 5th Grade	108.28	114.97	124.55	124.62	113.65	122.48	108.42	121.36
School Climate								
Spring Kindergarten	-0.17	-0.13	0.10	0.13	-0.14	0.08	-0.29	0.06
Spring 1st Grade	-0.16	-0.15	0.05	0.13	-0.13	0.07	-0.30	0.05
Spring 2nd Grade	-0.28	-0.15	0.01	0.11	-0.17	0.04	-0.36	0.02
Spring 3rd Grade	-0.38	-0.11	0.06	0.10	-0.19	0.05	-0.29	0.01
Spring 4th Grade	-0.35	-0.09	0.01	0.07	-0.19	0.03	-0.31	0.00
Spring 5th Grade	-0.44	-0.07	0.05	0.05	-0.23	0.03	-0.31	-0.01
School Strain								
Spring Kindergarten	0.17	0.27	-0.11	-0.12	0.17	-0.05	0.37	-0.03
Spring 1st Grade	0.08	0.13	-0.03	-0.08	0.10	-0.04	0.23	-0.03
Spring 2nd Grade	0.10	0.08	-0.04	-0.05	0.08	-0.03	0.18	-0.02
Spring 3rd Grade	0.19	0.11	-0.05	-0.07	0.12	-0.03	0.20	-0.01
Spring 4th Grade	0.35	0.21	-0.09	-0.14	0.19	-0.05	0.41	-0.03
Spring 5th Grade	0.41	0.21	-0.05	-0.10	0.25	-0.02	0.39	0.01
Teacher Press								
Spring Kindergarten	-0.03	0.02	-0.01	0.01	-0.04	0.02	-0.04	0.01
Spring 1st Grade	0.00	-0.04	-0.02	0.07	-0.03	0.05	-0.07	0.03
Spring 2nd Grade	-0.06	0.00	-0.02	0.02	-0.05	0.02	-0.09	0.01
Spring 3rd Grade	-0.17	-0.02	-0.07	0.06	-0.11	0.04	-0.11	0.01

Spring 4th Grade	-0.12	-0.07	0.01	0.04	-0.09	0.02	-0.15	0.00
Spring 5th Grade	-0.15	-0.05	0.03	0.06	-0.07	0.03	-0.14	0.02
Home Learning Environment								
Parent-Child Interaction	-0.01	-0.11	0.05	0.03	-16.43	12.40	-44.26	9.93
Books In The Home	-30.69	-30.19	-0.01	30.93	-0.07	0.04	-0.33	0.04
Computer Learning	0.16	-0.15	0.08	0.03	-0.06	0.01	-0.15	0.01

Table B.4: Complete Estimates from Random Coefficient Models

	Reading		Math	
Fixed Effect	Estimate	Std. Error	Estimate	Std. Error
Age	0.952***	(0.030)	-0.662***	(0.030)
Age^2	0.483***	(0.003)	0.659***	(0.003)
School Context				
School Climate	0.245***	(0.072)	0.246***	(0.071)
School Strain	-0.034	(0.064)	-0.064	(0.062)
Teacher Press	0.075	(0.066)	0.208**	(0.067)
Home Learning Environment				
Parent-Child Interactions	2.474***	(0.297)	0.825**	(0.254)
Books In Home	0.084***	(0.018)	0.137***	(0.016)
Computer Learning	0.042	(0.194)	-0.101	(0.169)
Age*Parent-Child Interactions	0.002	(0.026)	0.030	(0.024)
Age*Books In Home	0.008***	(0.001)	0.014***	(0.001)
Age*Computer Learning	-0.038*	(0.017)	-0.065***	(0.016)
Baseline Age	0.297***	(0.024)	0.606***	(0.024)
Race/Ethnicity (Ref: White)				
Hispanic	-1.492***	(0.303)	-2.327***	(0.295)
African-American	-1.826***	(0.394)	-4.055***	(0.383)
Asian/Other	2.574***	(0.337)	1.988***	(0.327)
Male	-2.658***	(0.215)	0.567**	(0.208)
SES	5.566***	(0.169)	4.816***	(0.165)
Two-Parent Family	1.767***	(0.263)	1.651***	(0.256)
Early Childhood Program (Ref: Parent)				
Head Start	0.092	(0.384)	0.621	(0.367)
Prekindergarten	1.016**	(0.394)	0.801*	(0.377)
Private Preschool	2.518***	(0.329)	2.413***	(0.329)
Other Non-Parental Care	0.701	(0.405)	0.757	(0.389)
Locale (Ref: Suburban)				
City	-0.060	(0.263)	0.844***	(0.255)
Town	-0.111	(0.426)	0.168	(0.413)
Rural	0.168	(0.294)	0.104	(0.286)
Constant	68.481***	(0.401)	48.030***	(0.383)

Random Effect	Variance Component	95% CI		Variance Component	95% CI	
Constant, r_{0i}	12.019	11.819	12.223	9.351	9.161	9.545
Age, r_{1i}	0.845	0.823	0.868	0.653	0.628	0.679
Level-1 Error, e_{ti}	7.841	7.792	7.889	8.018	7.969	8.068

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

Table B.5: Complete Estimates from Interactive Random Coefficient Models

	Reading		Math	
Fixed Effect	Estimate	Std. Error	Estimate	Std. Error
Age	0.951***	(0.030)	-0.658***	(0.030)
Age^2	0.484***	(0.003)	0.659***	(0.003)
School Context				
School Climate	-0.668***	(0.152)	-0.472***	(0.127)
School Strain	-0.049	(0.064)	-0.085	(0.062)
Teacher Press	0.060	(0.067)	0.192**	(0.067)
Home Learning Environment				
Parent-Child Interactions	2.527***	(0.297)	0.862***	(0.254)
Books In Home	0.089***	(0.018)	0.133***	(0.016)
Computer Learning	0.031	(0.194)	-0.098	(0.169)
Baseline Age	0.297***	(0.024)	0.605***	(0.024)
Race/Ethnicity (Ref: White)				
Hispanic	-1.536***	(0.304)	-2.420***	(0.295)
African-American	-1.897***	(0.395)	-4.258***	(0.384)
Asian/Other	2.565***	(0.337)	1.972***	(0.327)
Male	-2.660***	(0.215)	0.568**	(0.208)
SES	5.606***	(0.170)	4.903***	(0.165)
Two-Parent Family	1.794***	(0.263)	1.708***	(0.256)
Early Childhood Program (Ref: Parent)				
Head Start	0.083	(0.385)	0.607	(0.368)
Prekindergarten	1.023**	(0.394)	0.826*	(0.378)
Private Preschool	2.557***	(0.329)	2.485***	(0.330)
Other Non-Parental Care	0.695	(0.405)	0.752	(0.390)
Locale (Ref: Suburban)				
City	-0.082	(0.263)	0.813**	(0.256)
Town	-0.109	(0.427)	0.199	(0.413)
Rural	0.174	(0.294)	0.145	(0.286)
Age*Parent-Child Interactions	-0.005	(0.026)	0.022	(0.024)
Age*Computer Learning	-0.036*	(0.017)	-0.064***	(0.016)
Age*Books In Home	0.007***	(0.001)	0.014***	(0.001)
Age*School Climate	0.108***	(0.016)	0.091***	(0.014)
School Climate*Books	0.029	(0.020)	0.054***	(0.016)
Age*School Climate*Books	-0.004	(0.002)	-0.008***	(0.002)
Constant	68.466***	(0.401)	47.981***	(0.383)

Random Effect	Variance Component	95% CI		Variance Component	95% CI	
Constant, r_{0i}	12.028	11.828	12.232	9.373	9.183	9.567
Age, r_{1i}	0.842	0.820	0.864	0.646	0.621	0.672
Level-1 Error, e_{ti}	7.840	7.791	7.888	8.018	7.969	8.068

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

Figure B.1.1: Mean Outcomes by School Strain over Time

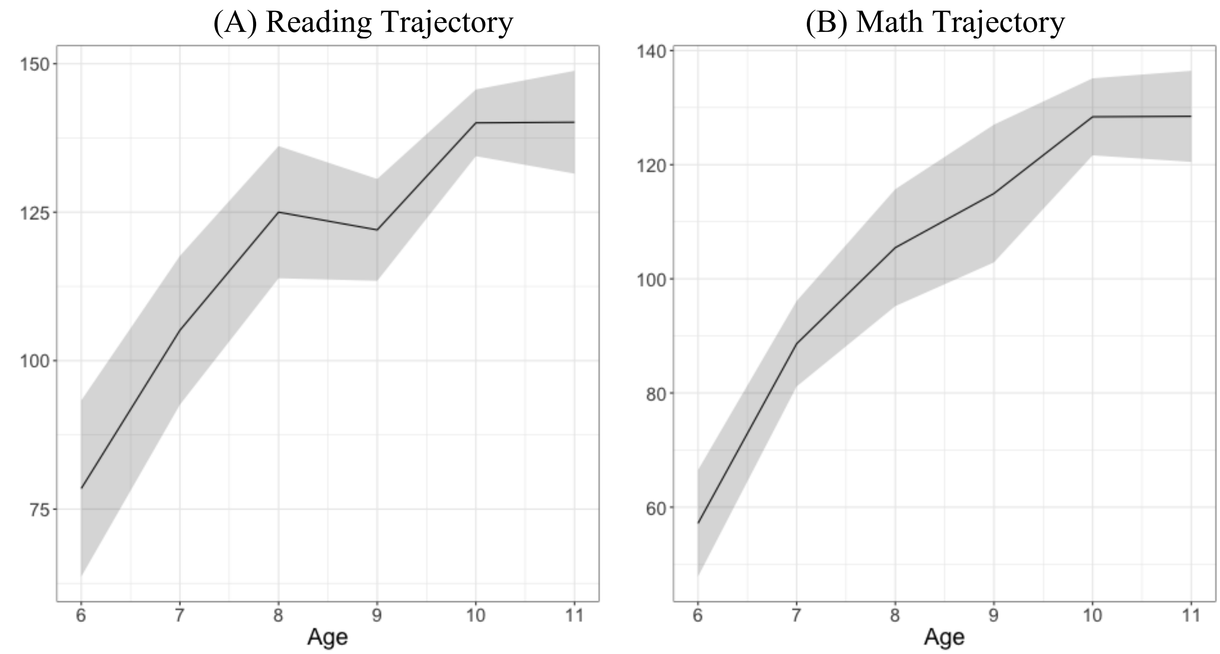


Figure B.1.2: Mean Outcomes by Teacher Press over Time

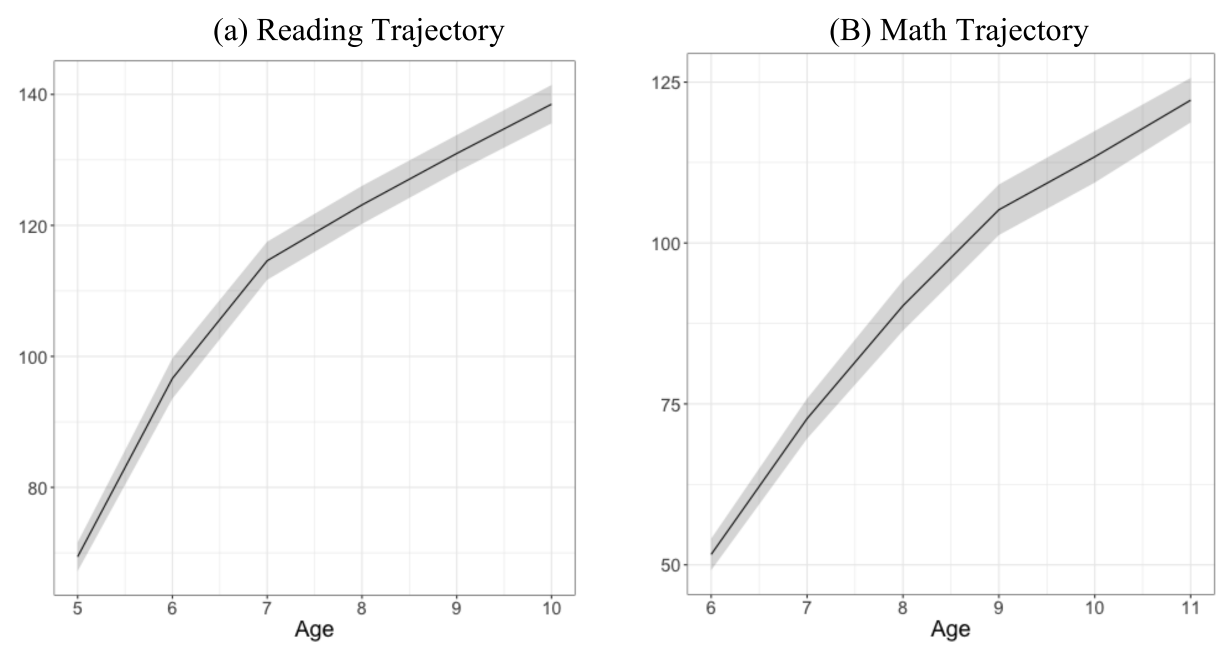


Figure B.1.3: Mean Outcomes by Parent-Child Interactions over Time

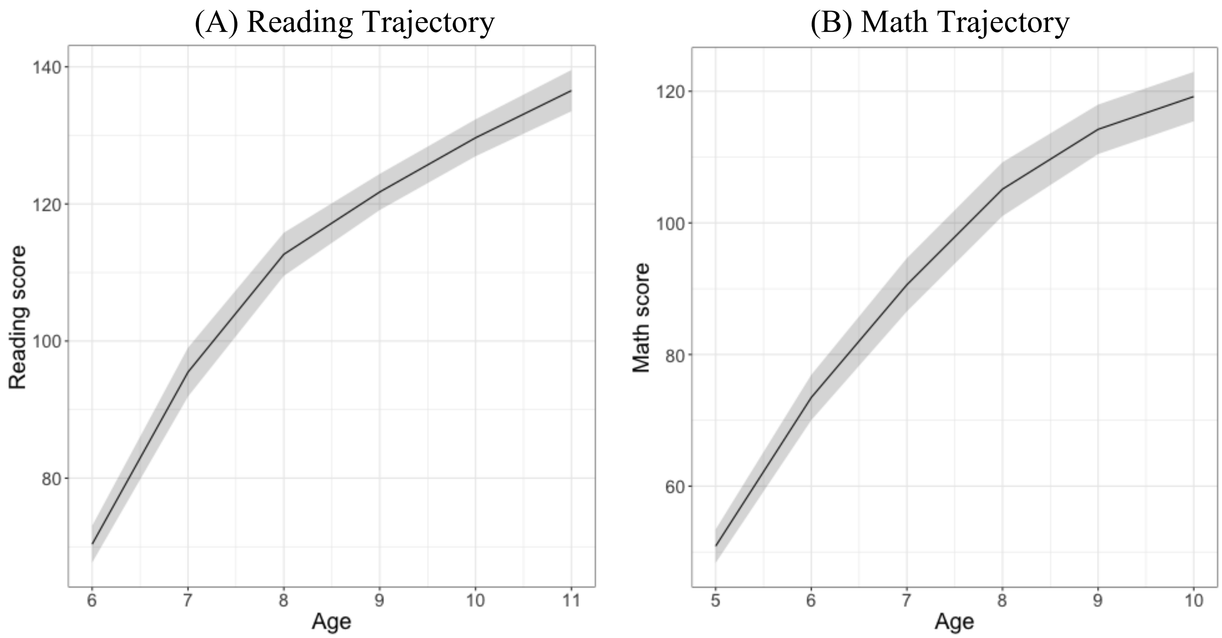
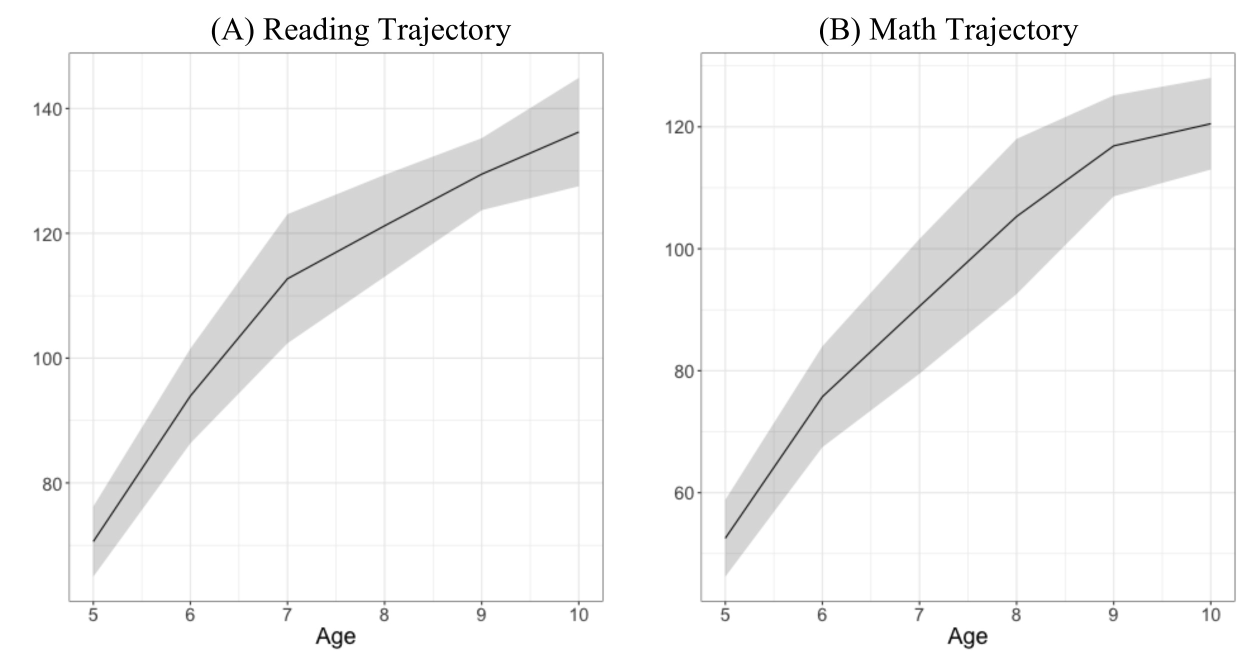


Figure B.1.4: Mean Outcomes by Computer Learning over Time



Appendix C

Appendix For Paper 3

Table C.1: Correlations of Common Variables in 5k ECLS-B Sample

	Reading	Math	White	Hispanic	African-American	Other Race	Male	Low Birthweight	Age	Poverty
Reading	1									
Math	0.8034*	1								
White	0.1166*	0.1836*	1							
Hispanic	-0.1718*	-0.1815*	-0.4484*	1						
African-American	-0.1127*	-0.1847*	-0.3772*	-0.2244*	1					
Other Race	0.1336*	0.1264*	-0.4352*	-0.2589*	-0.2179*	1				
Male	-0.0627*	-0.0221	-0.0092	0.0114	-0.0184	0.0165	1			
Low Birthweight	-0.0931*	-0.1515*	0.0856*	-0.0292*	0.0871*	-0.1555*	-0.0495*	1		
Age	0.2629*	0.2781*	0.0798*	-0.0387*	-0.0166	-0.0441*	0.0201	0.0622*	1	
Poverty	-0.2821*	-0.3125*	-0.2388*	0.1255*	0.2422*	-0.0536*	0.0026	0.0270*	-0.0128	1
Mother Ed	0.3145*	0.3209*	0.2097*	-0.2468*	-0.1044*	0.0873*	0.0099	-0.0212	0.0296*	-0.3512*
Two Parent	0.2030*	0.2282*	0.1774*	0.004	-0.3137*	0.0629*	0.0184	-0.0635*	-0.0176	-0.3535*
Parent Care	-0.0976*	-0.0836*	-0.0168	0.0894*	-0.0381*	-0.0354*	-0.0262*	-0.0032	0.0398*	0.1337*
Head Start	-0.1608*	-0.1708*	-0.2017*	0.0942*	0.1853*	-0.0157	-0.0137	-0.0002	-0.0378*	0.2591*
Prek	0.0851*	0.0434*	0.0044	-0.0417*	0.0273*	0.0121	-0.0014	0.0064	-0.0844*	-0.0674*
Private Preschool	0.1869*	0.2034*	0.1951*	-0.1452*	-0.1425*	0.0366*	0.0157	-0.0009	0.0287*	-0.2493*
Other Care	-0.0818*	-0.0640*	-0.0521*	0.0544*	0.0196	-0.0088	0.0277*	-0.0019	0.0506*	-0.0083

* p<0.05, ** p<0.01, *** p<0.001

	Mother Ed	Two Parent	Parent Care	Head Start	Prek	Private Preschool	Other Care
Mother’s Ed	1						
Two Parent	0.1882*	1					
Parent Care	-0.1769*	0.0126	1				
Head Start	-0.2044*	-0.1979*	-0.2149*	1			
Prek	0.0571*	0.0196	-0.2052*	-0.1902*	1		
Private Preschool	0.2944*	0.1613*	-0.3898*	-0.3613*	-0.3450*	1	
Other Care	-0.0656*	-0.0585*	-0.1570*	-0.1455*	-0.1389*	-0.2639*	1

* p<0.05, ** p<0.01, *** p<0.001

Table C.2: Correlations of Common Variables in 11k ECLS-K Sample

	Reading	Math	White	Hispanic	African-American	Other Race	Male	Low Birthweight	Age	Poverty
Reading	1									
Math	0.7589*	1								
White	0.1122*	0.2035*	1							
Hispanic	-0.1903*	-0.2310*	-0.5696*	1						
African-American	-0.0617*	-0.1357*	-0.3909*	-0.1913*	1					
Other Race	0.1269*	0.1119*	-0.4135*	-0.2024*	-0.1389*	1				
Male	-0.0476*	0.0222*	0.0109	0.006	-0.0046	-0.0193*	1			
Low Birthweight	-0.0391*	-0.0696*	-0.0519*	-0.0063	0.0774*	0.011	-0.0268*	1		
Age	0.1793*	0.2484*	0.1088*	-0.0713*	-0.0138	-0.0608*	0.0768*	0.0250*	1	
Poverty	-0.2494*	-0.2930*	-0.2914*	0.2522*	0.1728*	-0.0425*	-0.008	0.0390*	0.006	1
Mother Ed	0.3070*	0.3313*	0.2432*	-0.3010*	-0.0526*	0.0595*	0.0014	-0.0228*	-0.0073	-0.3786*
Two Parent	0.1732*	0.2005*	0.1594*	-0.0226*	-0.2788*	0.0572*	-0.002	-0.0268*	-0.0311*	-0.3067*
Parent Care	-0.1027*	-0.1239*	-0.0818*	0.1076*	-0.0071	-0.0041	0.0001	0.0142	0.0279*	0.1448*
Head Start	-0.1102*	-0.1333*	-0.1573*	0.0927*	0.1364*	-0.0106	0.0139	0.0311*	-0.0234*	0.1975*
Prek	-0.0220*	-0.0326*	-0.0275*	0.0272*	0.0297*	-0.0211*	0.0112	-0.0067	0.0128	0.0049
Private Preschool	0.2401*	0.2798*	0.2356*	-0.2140*	-0.1301*	0.0375*	-0.0009	-0.0306*	-0.0168	-0.2921*
Other Care	-0.0916*	-0.0885*	-0.0567*	0.0583*	0.0304*	-0.0166	-0.0257*	0.0033	-0.0004	0.0451*

* p<0.05, ** p<0.01, *** p<0.001

	Mother Ed	Two Parent	Parent Care	Head Start	Prek	Private Preschool	Other Care
Mother's Ed	1						
Two Parent	0.1813*	1					
Parent Care	-0.1784*	0.0227*	1				
Head Start	-0.1505*	-0.1286*	-0.2012*	1			
Prek	-0.0126	-0.0540*	-0.2256*	-0.1645*	1		
Private Preschool	0.3266*	0.1825*	-0.4128*	-0.3010*	-0.3376*	1	
Other Care	-0.0941*	-0.1097*	-0.1924*	-0.1402*	-0.1573*	-0.2878*	1

* p<0.05, ** p<0.01, *** p<0.001

Table C.3: Correlations of Common Variables in 7k ECLS-B Sample

	Reading	Math	White	Hispanic	African-American	Other Race	Male	Low Birthweight	Age	Poverty
Reading	1									
Math	0.8021*	1								
White	0.1023*	0.1531*	1							
Hispanic	-0.1718*	-0.1810*	-0.4162*	1						
African-American	-0.1115*	-0.1750*	-0.3581*	-0.2175*	1					
Other Race	0.1400*	0.1445*	-0.4575*	-0.2779*	-0.2391*	1				
Male	-0.0653*	-0.0216	-0.0125	0.0103	-0.0121	0.0151	1			
Low Birthweight	-0.0926*	-0.1532*	0.1194*	-0.0316*	0.0862*	-0.1828*	-0.0429*	1		
Age	0.2503*	0.2614*	0.0843*	-0.0351*	-0.0101	-0.0558*	0.0254*	0.0523*	1	
Poverty	-0.2824*	-0.3075*	-0.2207*	0.1277*	0.2250*	-0.0587*	0.0045	0.0197	-0.0067	1
Mother Ed	0.3153*	0.3188*	0.1894*	-0.2350*	-0.0932*	0.0832*	-0.001	-0.0134	0.0206	-0.3514*
Two Parent	0.2064*	0.2283*	0.1442*	-0.0046	-0.3043*	0.0991*	0.0081	-0.0690*	-0.0330*	-0.3278*
Parent Care	-0.0904*	-0.0862*	-0.0226*	0.0916*	-0.0350*	-0.0306*	-0.0302*	-0.0034	0.0405*	0.1292*
Head Start	-0.1597*	-0.1657*	-0.1803*	0.0833*	0.1717*	-0.0175	-0.0084	0.0026	-0.0324*	0.2451*
Prek	0.0838*	0.0380*	0.0109	-0.0436*	0.0356*	-0.002	0.0022	0.0004	-0.0801*	-0.0575*
Private Preschool	0.1878*	0.2067*	0.1715*	-0.1355*	-0.1402*	0.0501*	0.0146	0.0011	0.0230*	-0.2530*
Other Care	-0.0866*	-0.0599*	-0.0356*	0.0473*	0.0155	-0.0168	0.0241*	-0.001	0.0465*	0.0014

* p<0.05, ** p<0.01, *** p<0.001

	Mother Ed	Two Parent	Parent Care	Head Start	Prek	Private Preschool	Other Care
Mother's Ed	1						
Two Parent	0.1811*	1					
Parent Care	-0.1698*	0.0196	1				
Head Start	-0.1985*	-0.1936*	-0.2226*	1			
Prek	0.0530*	0.018	-0.2084*	-0.1930*	1		
Private Preschool	0.2946*	0.1545*	-0.3875*	-0.3588*	-0.3360*	1	
Other Care	-0.0692*	-0.0556*	-0.1614*	-0.1495*	-0.1399*	-0.2602*	1

* p<0.05, ** p<0.01, *** p<0.001

Table C.4: Correlations of Common Variables in 16k ECLS-K Sample

	Reading	Math	White	Hispanic	African-American	Other Race	Male	Low Birthweight	Age	Poverty
Reading	1									
Math	0.7675*	1								
White	0.1323*	0.2137*	1							
Hispanic	-0.1986*	-0.2410*	-0.5513*	1						
African-American	-0.0640*	-0.1272*	-0.3773*	-0.2257*	1					
Other Race	0.1200*	0.1177*	-0.3851*	-0.2303*	-0.1576*	1				
Male	-0.0543*	0.0071	0.011	0.0021	0.0001	-0.0187*	1			
Low Birthweight	-0.0408*	-0.0664*	-0.0501*	-0.0097	0.0714*	0.0142	-0.0323*	1		
Age	0.1793*	0.2406*	0.1015*	-0.0712*	0.0038	-0.0614*	0.0653*	0.0213*	1	
Poverty	-0.2554*	-0.2919*	-0.2987*	0.2423*	0.1828*	-0.0510*	0.0016	0.0442*	0.0106	1
Mother Ed	0.3142*	0.3398*	0.2547*	-0.3024*	-0.0548*	0.0637*	-0.0009	-0.0301*	-0.0109	-0.3827*
Two Parent	0.1729*	0.1960*	0.1524*	-0.0076	-0.2906*	0.0755*	-0.0019	-0.0320*	-0.0410*	-0.3035*
Parent Care	-0.1172*	-0.1414*	-0.0774*	0.1059*	-0.0105	-0.0101	0.0019	0.0209*	0.0221*	0.1496*
Head Start	-0.1071*	-0.1290*	-0.1581*	0.0866*	0.1390*	-0.0165*	0.0250*	0.0257*	-0.0121	0.1978*
Prek	-0.0062	-0.0126	-0.0135	0.0091	0.0304*	-0.0218*	0.0123	-0.0044	0.0023	-0.0143
Private Preschool	0.2550*	0.2975*	0.2408*	-0.2093*	-0.1328*	0.0443*	-0.0166*	-0.0351*	-0.0146	-0.2993*
Other Care	-0.0916*	-0.0910*	-0.0602*	0.0602*	0.0212*	-0.0091	-0.0184*	0.0005	0.003	0.0390*

* p<0.05, ** p<0.01, *** p<0.001

	Mother Ed	Two Parent	Parent Care	Head Start	Prek	Private Preschool	Other Care
Mother's Ed	1						
Two Parent	0.1776*	1					
Parent Care	-0.1874*	0.0291*	1				
Head Start	-0.1430*	-0.1267*	-0.2311*	1			
Prek	0.0041	-0.0439*	-0.2315*	-0.1739*	1		
Private Preschool	0.3343*	0.1811*	-0.3972*	-0.2984*	-0.2988*	1	
Other Care	-0.0877*	-0.1116*	-0.2140*	-0.1607*	-0.1610*	-0.2762*	1

* p<0.05, ** p<0.01, *** p<0.001

Table C.5: Correlations of Common Variables in 18k ECLS-K Sample

	Reading	Math	White	Hispanic	African-American	Other Race	Male	Low Birthweight	Age	Poverty
Reading	1									
Math	0.7643*	1								
White	0.1277*	0.2116*	1							
Hispanic	-0.1967*	-0.2401*	-0.5462*	1						
African-American	-0.0657*	-0.1299*	-0.3663*	-0.2274*	1					
Other Race	0.1247*	0.1213*	-0.3885*	-0.2412*	-0.1617*	1				
Male	-0.0547*	0.0119	0.0146*	-0.0021	0.0007	-0.0187*	1			
Low Birthweight	-0.0423*	-0.0721*	-0.0509*	0.0018	0.0566*	0.0153*	-0.0322*	1		
Age	0.1691*	0.2314*	0.0977*	-0.0678*	0.0035	-0.0579*	0.0625*	0.0166*	1	
Poverty	-0.2581*	-0.2966*	-0.2912*	0.2457*	0.1676*	-0.0519*	0.0006	0.0309*	0.0056	1
Mother Ed	0.3123*	0.3414*	0.2523*	-0.3014*	-0.0546*	0.0669*	0.0051	-0.0245*	-0.0142	-0.3746*
Two Parent	0.1669*	0.1970*	0.1384*	0.0033	-0.2907*	0.0792*	0.0006	-0.0255*	-0.0454*	-0.2866*
Parent Care	-0.1159*	-0.1320*	-0.0774*	0.1080*	-0.0085	-0.0155*	0.0085	0.0223*	0.0163*	0.1334*
Head Start	-0.1123*	-0.1342*	-0.1742*	0.0969*	0.1252*	0.0067	0.014	0.0274*	-0.0097	0.2093*
Prek	-0.0011	-0.0116	-0.0002	0.0039	0.0294*	-0.0327*	0.0076	-0.0113	0.0074	-0.0221*
Private Preschool	0.2548*	0.2933*	0.2469*	-0.2142*	-0.1241*	0.0340*	-0.0112	-0.0338*	-0.0127	-0.2937*
Other Care	-0.0853*	-0.0839*	-0.0568*	0.0524*	0.0175*	-0.0011	-0.0181*	0.001	0	0.0362*

* p<0.05, ** p<0.01, *** p<0.001

	Mother Ed	Two Parent	Parent Care	Head Start	Prek	Private Preschool	Other Care
Mother's Ed	1						
Two Parent	0.1700*	1					
Parent Care	-0.1779*	0.0378*	1				
Head Start	-0.1592*	-0.1127*	-0.2436*	1			
Prek	0.0068	-0.0521*	-0.2284*	-0.1770*	1		
Private Preschool	0.3350*	0.1679*	-0.3919*	-0.3038*	-0.2848*	1	
Other Care	-0.0758*	-0.1053*	-0.2181*	-0.1690*	-0.1585*	-0.2719*	1

* p<0.05, ** p<0.01, *** p<0.001

Table C.6: AIC and BIC Comparisons for Models Testing Survey Differences

	Model 1	Model 2	Model 3
Math			
AIC	128,215	125,875	125,812
BIC	128,324	125,991	126,021
Reading			
AIC	133,571	131,254	131,068
BIC	133,679	131,370	131,278

Table C.7: Estimates of Cognitive Stimulation from 5k ECLS-B Model and 11k ECLS-K CSMI Model

	5k ECLS-B		11k ECLS-K		
	Estimate	Std. Error	Estimate	Std. Error	Std. Error Percent Change
Race/Ethnicity (Ref: White)					
Hispanic	-0.649***	0.07	-0.717***	0.133	0.90
African-American	-0.497***	0.08	-0.483***	0.122	0.53
Asian/Other	-0.427***	0.068	-0.401***	0.118	0.74
Male	0.041	0.05	0.041	0.056	0.12
Low Birthweight	0.05	0.058	0.074	0.091	0.57
Age at Kindergarten	-0.069	0.061	-0.067	0.101	0.66
Poverty	-0.454***	0.07	-0.549***	0.094	0.34
Mother's Education	0.123***	0.01	0.114***	0.015	0.50
Two Parent Family	0.086	0.06	0.07	0.097	0.62
Early Childhood Program (Ref: Parent Care)					
Head Start	-0.009	0.085	0.032	0.157	0.85
Prekindergarten	0.180**	0.087	0.206*	0.108	0.24
Private Preschool	0.197***	0.073	0.201*	0.107	0.47
Other Care	0.069	0.098	0.117	0.158	0.61
Reading	0.014***	0.002	0.014***	0.003	0.50

* p<0.05, ** p<0.01, *** p<0.001

Table C.8: Estimates of Emotional Support from 5k ECLS-B Model and 11k ECLS-K CSMI Model

	5k ECLS-B		11k ECLS-K		
	Estimate	Std. Error	Estimate	Std. Error	Std. Error Percent Change
Race/Ethnicity (Ref: White)					
Hispanic	-0.537***	0.071	-0.600***	0.104	0.46
African-American	-0.716***	0.081	-0.661***	0.118	0.46
Asian/Other	-0.460***	0.069	-0.377***	0.101	0.46
Male	-0.052	0.05	-0.028	0.058	0.16
Low Birthweight	0.001	0.059	0.031	0.093	0.58
Age at Kindergarten	-0.084	0.062	-0.069	0.079	0.27
Poverty	-0.392***	0.071	-0.511***	0.092	0.30
Mother's Education	0.099***	0.01	0.088***	0.015	0.50
Two Parent Family	0.118*	0.061	0.091	0.082	0.34
Early Childhood Program (Ref: Parent Care)					
Head Start	-0.174**	0.087	-0.177	0.122	0.40
Prekindergarten	0.135	0.088	0.024	0.147	0.67
Private Preschool	0.177**	0.074	0.069	0.085	0.15
Other Care	0.007	0.101	-0.025	0.16	0.58
Reading	0.010***	0.002	0.008***	0.003	0.50

* p<0.05, ** p<0.01, *** p<0.001

Table C.9: Estimates of Math Scores from 5k ECLS-B Models and 11k ECLS-K Monotone CSMI Models

	5k ECLS-B						11k ECLS-K						M1 Std. Error Ratio (ECLS-K / ECLS-B)
	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error	
Race/Ethnicity (Ref: White)													
Hispanic	-2.447***	0.333	-2.442***	0.333	-2.458***	0.333	-2.227***	0.274	-2.229***	0.274	-2.227***	0.274	0.82
African-American	-2.306***	0.378	-2.302***	0.378	-2.304***	0.377	-2.019***	0.338	-2.020***	0.337	-2.019***	0.338	0.89
Asian/Other	1.131***	0.323	1.145***	0.323	1.129***	0.323	2.823***	0.302	2.827***	0.303	2.822***	0.302	0.93
Male	-0.824***	0.234	-0.825***	0.234	-0.818***	0.234	0.035	0.192	0.036	0.192	0.034	0.192	0.82
Birthweight	-3.530***	0.273	-3.527***	0.273	-3.540***	0.273	-2.329***	0.33	-2.328***	0.33	-2.331***	0.33	1.21
Age at Kindergarten	6.827***	0.274	6.827***	0.274	6.829***	0.274	6.796***	0.222	6.796***	0.222	6.797***	0.222	0.81
Poverty	-3.465***	0.327	-3.476***	0.328	-3.467***	0.327	-2.812***	0.268	-2.821***	0.268	-2.811***	0.268	0.82
Mother's Education	0.565***	0.048	0.565***	0.048	0.567***	0.048	0.716***	0.04	0.715***	0.04	0.716***	0.04	0.83
Two Parent Family	2.033***	0.283	2.017***	0.284	2.023***	0.283	2.295***	0.238	2.288***	0.238	2.296***	0.238	0.84
Early Childhood Program (Ref: Parent-Care Only)													
Head Start	-0.07	0.402	1.136	1.385	-0.082	0.402	0.226	0.343	1.875	1.577	0.228	0.344	0.85
Prekindergarten	1.822***	0.412	1.807***	0.412	-1.251	1.571	0.898***	0.323	0.886***	0.324	0.944	1.801	0.78
Private Preschool	1.552***	0.345	1.533***	0.345	1.584***	0.345	3.614***	0.276	3.590***	0.278	3.613***	0.277	0.80
Other Care	-0.878*	0.467	-0.884*	0.467	-0.874*	0.467	0.058	0.354	0.052	0.355	0.058	0.354	0.76
Cognitive Stimulation	0.835***	0.155	0.884***	0.164	0.722***	0.164	0.890***	0.249	0.944***	0.263	0.893***	0.251	1.61
Emotional Support	0.278*	0.161	0.282*	0.161	0.295*	0.161	0.265	0.255	0.269	0.254	0.265	0.256	1.58
Head Start * Emotional	--	--	-0.307	0.338	--	--	--	--	-0.415	0.388	--	--	1.15
Prek * Cognitive	--	--	--	--	0.721**	0.356	--	--	--	--	-0.011	0.419	1.18

* p<0.05, ** p<0.01, *** p<0.001

Table C.10: Estimates of Reading Scores from 5k ECLS-B Models and 11k ECLS-K Monotone CSMI Models

	5k ECLS-B						11k ECLS-K						M1 Std. Error Ratio (ECLS-K / ECLS-B)
	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error	
Race/Ethnicity (Ref: White)													
Hispanic	-1.742***	0.481	-1.773***	0.481	-1.756***	0.481	-0.699**	0.285	-0.699**	0.285	-0.700**	0.285	0.59
African-American	0.199	0.546	0.22	0.546	0.202	0.546	0.983***	0.352	0.980***	0.352	0.983***	0.352	0.64
Asian/Other	3.286***	0.467	3.253***	0.467	3.283***	0.467	3.973***	0.314	3.975***	0.315	3.973***	0.314	0.67
Male	-2.262***	0.339	-2.272***	0.339	-2.255***	0.339	-1.476***	0.2	-1.475***	0.2	-1.476***	0.2	0.59
Birthweight	-3.025***	0.395	-3.029***	0.395	-3.037***	0.395	-1.391***	0.345	-1.390***	0.345	-1.392***	0.345	0.87
Age at K	9.326***	0.396	9.327***	0.396	9.329***	0.396	5.223***	0.231	5.222***	0.231	5.223***	0.231	0.58
Poverty	-4.354***	0.473	-4.342***	0.473	-4.357***	0.473	-2.485***	0.282	-2.488***	0.282	-2.484***	0.282	0.60
Mother's Educ	0.850***	0.069	0.850***	0.069	0.852***	0.069	0.740***	0.041	0.739***	0.041	0.740***	0.041	0.59
Two Parent Family	2.961***	0.41	2.990***	0.41	2.950***	0.41	2.235***	0.248	2.232***	0.247	2.235***	0.248	0.60
Early Childhood Program (Ref: Parent-Care Only)													
Head Start	-0.058	0.581	-4.158*	2.186	-0.072	0.581	-0.012	0.358	0.808	1.693	-0.012	0.359	0.62
Prekindergarten	4.394***	0.595	4.427***	0.595	0.715	2.273	0.777**	0.335	0.773**	0.335	0.552	1.588	0.56
Private Preschool	2.675***	0.498	2.723***	0.499	2.713***	0.499	2.993***	0.288	2.985***	0.288	2.997***	0.29	0.58
Other Care	-1.668**	0.676	-1.662**	0.676	-1.664**	0.676	-0.635*	0.369	-0.636*	0.369	-0.635*	0.369	0.55
Cog Stimulation	1.138***	0.223	1.128***	0.223	1.002***	0.238	0.811***	0.212	0.813***	0.212	0.803***	0.22	0.95
Emotional Support	0.395*	0.233	0.237	0.246	0.414*	0.233	0.289	0.205	0.315	0.207	0.289	0.205	0.88
Head Start * Emo	--	--	0.990*	0.509	--	--	--	--	-0.196	0.392	--	--	0.77
Prek * Cognitive	--	--	--	--	0.863*	0.514	--	--	--	--	0.053	0.37	0.72

* p<0.05, ** p<0.01, *** p<0.001

Table C.11: Estimates of Cognitive Stimulation from ECLS-B 7k Model and ECLS-K CSMI Models

	7k ECLS-B		11k ECLS-K			16k Hybrid ECLS-K			16k Chained-Equation ECLS-K			18k Hybrid ECLS-K			18k Chained-Equation ECLS-K		
	Estimate	Std. Err.	Estimate	Std. Err.	Std. Error % Δ	Estimate	Std. Err.	Std. Error % Δ	Estimate	Std. Err.	Std. Error % Δ	Estimate	Std. Err.	Std. Error % Δ	Estimate	Std. Err.	Std. Err. % Δ
Race/Ethnicity (Ref: White)																	
Hispanic	-0.618***	0.07	-0.717***	0.13	1.02	-0.723***	0.12	0.79	-0.647***	0.09	0.41	-0.073	0.05	-0.20	-0.658***	0.07	0.11
Af-Amer	-0.500***	0.07	-0.483***	0.12	0.65	-0.479***	0.11	0.51	-0.433***	0.11	0.47	-0.087	0.07	-0.07	-0.454***	0.10	0.27
Asian/Oth	-0.412***	0.07	-0.401***	0.12	0.82	-0.393***	0.01	0.51	-0.342***	0.09	0.37	-0.062	0.06	-0.03	-0.356***	0.10	0.20
Male	0.023	0.04	0.041	0.06	0.22	0.039	0.05	0.09	0.029	0.05	0.13	0.021	0.04	-0.13	0.024	0.06	0.33
Low Birth	0.073	0.05	0.074	0.09	0.69	0.071	0.07	0.35	0.063	0.09	0.63	0.008	0.06	0.17	0.054	0.08	0.43
Age at K	-0.069	0.06	-0.067	0.10	0.68	-0.07	0.09	0.48	-0.055	0.08	0.43	-0.015	0.05	-0.22	-0.052	0.08	0.33
Poverty	-0.442***	0.06	-0.549***	0.09	0.49	-0.550***	0.08	0.30	-0.506***	0.08	0.32	-0.057	0.05	-0.14	-0.480***	0.07	0.16
Mother Ed	0.122***	0.01	0.114***	0.02	0.50	0.116***	0.01	0.30	0.121***	0.01	0.20	0.015*	0.01	-0.10	0.125***	0.01	0.10
Two Parent	0.075	0.06	0.07	0.10	0.70	0.08	0.09	0.51	0.085	0.07	0.28	0.036	0.05	-0.09	0.09	0.08	0.40
Early Childhood Program (Ref: Parent-Care Only)																	
Head Start	-0.012	0.08	0.032	0.16	1.07	0.048	0.13	0.76	0.019	0.10	0.42	-0.007	0.07	-0.04	0.021	0.11	0.43
Prek	0.186**	0.08	0.206*	0.11	0.37	0.220**	0.10	0.24	0.204**	0.10	0.32	0.066	0.08	-0.03	0.176	0.12	0.47
Private	0.216***	0.07	0.201*	0.11	0.55	0.206**	0.10	0.36	0.230***	0.08	0.22	0.035	0.06	-0.14	0.210**	0.10	0.39
Other Care	0.088	0.09	0.117	0.12	0.78	0.123	0.14	0.53	0.097	0.11	0.31	0.045	0.07	-0.17	0.074	0.12	0.38
Reading	0.013***	0.00	0.014***	0.00	0.50	0.014***	0.00	0.50	0.014***	0.00	0.50	0.004*	0.00	0.00	0.013***	0.00	0.50

* p<0.05, ** p<0.01, *** p<0.001

Table C.12: Estimates of Emotional Support from ECLS-B 7k Model and ECLS-K CSMI Models

	7k ECLS-B		11k ECLS-K			16k Hybrid ECLS-K			16k Chained-Equation ECLS-K			18k Hybrid ECLS-K			18k Chained-Equation ECLS-K		
	Estimate	Std. Error	Estimate	Std. Error	Std. Error % Δ	Estimate	Std. Error	Std. Error % Δ	Estimate	Std. Error	Std. Error % Δ	Estim ate	Std. Error	Std. Error % Δ	Estimate	Std. Error	Std. Error % Δ
Race/Ethnicity (Ref: White)																	
Hispanic	-0.525***	0.066	-0.600***	0.104	0.58	-0.611***	0.092	0.39	-0.602***	0.092	0.39	-0.053	0.06	-0.09	-0.572***	0.072	0.09
Afr-Am	-0.658***	0.075	-0.661***	0.118	0.57	-0.635***	0.103	0.37	-0.579***	0.09	0.20	-0.097	0.067	-0.11	-0.591***	0.096	0.28
Asian/Other	-0.417***	0.062	-0.377***	0.101	0.63	-0.380***	0.085	0.37	-0.361***	0.094	0.52	-0.052	0.062	0.00	-0.371***	0.086	0.39
Male	-0.041	0.045	-0.028	0.058	0.29	-0.028	0.052	0.16	-0.019	0.056	0.24	0.008	0.044	-0.02	-0.041	0.057	0.27
Low Birth	0.028	0.056	0.031	0.093	0.66	0.039	0.073	0.30	0.012	0.074	0.32	-0.003	0.069	0.23	0.01	0.085	0.52
Age at K	-0.074	0.059	-0.069	0.079	0.34	-0.074	0.071	0.20	-0.054	0.08	0.36	-0.018	0.048	-0.19	-0.055	0.084	0.42
Poverty	-0.428***	0.064	-0.511***	0.092	0.44	-0.517***	0.082	0.28	-0.516***	0.081	0.27	-0.059	0.054	-0.16	-0.509***	0.074	0.16
Mother's Ed	0.097***	0.01	0.088***	0.015	0.50	0.090***	0.014	0.40	0.094***	0.012	0.20	0.011	0.01	0.00	0.095***	0.011	0.10
Two Parent	0.107*	0.06	0.091	0.082	0.37	0.086	0.075	0.25	0.099	0.062	0.03	0.049	0.059	-0.02	0.086	0.077	0.28
Early Childhood Program (Ref: Parent-Care Only)																	
Head Start	-0.165**	0.08	-0.161	0.134	0.68	-0.147	0.117	0.46	-0.123	0.11	0.38	-0.051	0.079	-0.01	-0.125	0.106	0.33
Prek	0.132	0.084	0.136	0.125	0.49	0.149	0.11	0.31	0.154	0.096	0.14	0.044	0.068	-0.19	0.139	0.099	0.18
Private	0.170**	0.073	0.118	0.083	0.14	0.123	0.077	0.05	0.136*	0.072	-0.01	0.023	0.058	-0.21	0.131	0.082	0.12
Oth Care	-0.04	0.097	0.019	0.147	0.52	0.011	0.126	0.30	-0.043	0.108	0.11	0.02	0.077	-0.21	-0.05	0.106	0.09
Reading	0.010***	0.002	0.010***	0.003	0.50	0.010***	0.002	0.00	0.009***	0.003	0.50	0.003	0.002	0.00	0.009***	0.003	0.50

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

Table C.13: Estimates of Math Scores by CSMI Methods and ECLS-K Sample

	11k Monotone CSMI		16k Hybrid CSMI		16k Chained-Equation CSMI		18k Hybrid CSMI		18k Chained-Equation CSMI	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Race/Ethnicity (Ref: White)										
Hispanic	-2.227***	0.274	-2.260***	0.229	-2.252***	0.244	-2.570***	0.22	-2.184***	0.228
African-American	-2.019***	0.338	-1.770***	0.277	-1.793***	0.28	-1.950***	0.268	-1.726***	0.283
Asian/Other	2.823***	0.302	2.716***	0.255	2.716***	0.259	2.830***	0.258	3.024***	0.262
Male	0.035	0.192	-0.185	0.164	-0.169	0.162	-0.188	0.165	-0.18	0.164
Low Birthweight	-2.329***	0.33	-2.105***	0.29	-2.026***	0.301	-1.870***	0.284	-1.883***	0.287
Age at Kindergarten	6.796***	0.222	6.573***	0.187	6.550***	0.191	6.561***	0.184	6.563***	0.186
Poverty	-2.812***	0.268	-2.582***	0.238	-2.576***	0.241	-2.898***	0.221	-2.620***	0.227
Mother's Education	0.716***	0.04	0.693***	0.036	0.687***	0.036	0.754***	0.034	0.682***	0.037
Two Parent Family	2.295***	0.238	2.259***	0.198	2.256***	0.211	2.461***	0.2	2.383***	0.207
Early Childhood Program (Ref: Parent-Care Only)										
Head Start	0.226	0.343	0.556*	0.286	0.492*	0.291	0.603**	0.277	0.666**	0.295
Prekindergarten	0.898***	0.323	1.390***	0.301	1.360***	0.292	1.421***	0.289	1.377***	0.304
Private Preschool	3.614***	0.276	3.919***	0.245	3.837***	0.241	3.950***	0.25	3.858***	0.246
Other Care	0.058	0.354	0.04	0.299	0.07	0.314	-0.035	0.312	0.041	0.304
Home Learning Environment										
Cognitive Stimulation	0.890***	0.249	0.850***	0.22	1.032***	0.234	0.186	0.157	1.014***	0.217
Emotional Support	0.265	0.255	0.266	0.225	0.178	0.226	0.048	0.158	0.149	0.228

* p<0.05, ** p<0.01, *** p<0.001

Table C.14: Estimates of Reading Scores by CSMI Methods and ECLS-K Sample

	11k Monotone CSMI		16k Hybrid CSMI		16k Chained-Equation CSMI		18k Hybrid CSMI		18k Chained-Equation CSMI	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Race/ethnicity (ref: White)										
Hispanic	-0.699**	0.285	-0.895***	0.235	-0.925***	0.24	-1.331***	0.229	-0.997***	0.243
African-American	0.983***	0.352	0.746***	0.284	0.686**	0.288	0.369	0.277	0.547*	0.286
Asian/Other	3.973***	0.314	3.032***	0.262	2.997***	0.267	2.531***	0.266	2.718***	0.257
Male	-1.476***	0.200	-1.548***	0.166	-1.529***	0.166	-1.532***	0.172	-1.543***	0.173
Low Birthweight	-1.391***	0.345	-1.671***	0.297	-1.643***	0.302	-1.800***	0.298	-1.864***	0.306
Age at Kindergarten	5.223***	0.231	5.009***	0.193	4.975***	0.192	4.975***	0.191	4.984***	0.196
Poverty	-2.485***	0.282	-2.348***	0.236	-2.342***	0.243	-2.519***	0.223	-2.307***	0.24
Mother's Education	0.740***	0.041	0.724***	0.036	0.729***	0.036	0.794***	0.035	0.741***	0.036
Two Parent Family	2.235***	0.248	2.097***	0.206	2.077***	0.215	1.946***	0.209	1.902***	0.216
Early Childhood Program (Ref: Parent-Care Only)										
Head Start	-0.012	0.358	0.239	0.292	0.175	0.294	-0.113	0.284	-0.043	0.294
Prekindergarten	0.777**	0.335	1.317***	0.298	1.278***	0.292	1.533***	0.298	1.424***	0.327
Private Preschool	2.993***	0.288	3.369***	0.248	3.316***	0.247	3.511***	0.255	3.398***	0.26
Other Care	-0.635*	0.369	-0.356	0.303	-0.391	0.31	-0.469	0.33	-0.395	0.3
Home Learning Environment										
Cognitive Stimulation	0.811***	0.212	0.747***	0.184	0.829***	0.19	0.186	0.146	0.798***	0.194
Emotional Support	0.289	0.205	0.289	0.181	0.156	0.197	0.085	0.15	0.169	0.209

* p<0.05, ** p<0.01, *** p<0.001

Table C.15: Ratio of Standard Errors for Math Results by CSMI Method and ECLS-K Sample

	Std. Error Ratio 16k Hybrid / Monotone	Std. Error Ratio 16k Chained / Monotone	Std. Error Ratio 18k Hybrid / Monotone	Std. Error Ratio 18k Chained / Monotone
Race/Ethnicity (Ref: White)				
Hispanic	0.84	0.89	0.80	0.83
African-American	0.82	0.83	0.79	0.84
Asian/Other	0.84	0.86	0.85	0.87
Male	0.85	0.84	0.86	0.85
Low Birthweight	0.88	0.91	0.86	0.87
Age at Kindergarten	0.84	0.86	0.83	0.84
Poverty	0.89	0.90	0.82	0.85
Mother's Education	0.90	0.90	0.85	0.93
Two Parent Family	0.83	0.89	0.84	0.87
Early Childhood Program (Ref: Parent-Care Only)				
Head Start	0.83	0.85	0.81	0.86
Prekindergarten	0.93	0.90	0.89	0.94
Private Preschool	0.89	0.87	0.91	0.89
Other Care	0.84	0.89	0.88	0.86
Home Learning Environment				
Cognitive Stimulation	0.88	0.94	0.63	0.87
Emotional Support	0.88	0.89	0.62	0.89

Table C.16: Ratio of Standard Errors for Reading Results by CSMI Method and ECLS-K Sample

	Std. Error Ratio 16k Hybrid / Monotone	Std. Error Ratio 16k Chained / Monotone	Std. Error Ratio 18k Hybrid / Monotone	Std. Error Ratio 18k Chained / Monotone
Race/Ethnicity (Ref: White)				
Hispanic	0.82	0.84	0.80	0.85
African-American	0.81	0.82	0.79	0.81
Asian/Other	0.83	0.85	0.85	0.82
Male	0.83	0.83	0.86	0.87
Low Birthweight	0.86	0.88	0.86	0.89
Age at Kindergarten	0.84	0.83	0.83	0.85
Poverty	0.84	0.86	0.79	0.85
Mother's Education	0.88	0.88	0.85	0.88
Two Parent Family	0.83	0.87	0.84	0.87
Early Childhood Program (Ref: Parent-Care Only)				
Head Start	0.82	0.82	0.79	0.82
Prekindergarten	0.89	0.87	0.89	0.98
Private Preschool	0.86	0.86	0.89	0.90
Other Care	0.82	0.84	0.89	0.81
Home Learning Environment				
Cognitive Stimulation	0.87	0.90	0.69	0.92
Emotional Support	0.88	0.96	0.73	1.02

Table C.17: Estimates of Math Scores from 7k ECLS-B Models and Preferred ECLS-K 18k Hybrid CSMI Models

	7k ECLS-B						18k Hybrid ECLS-K					
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Race/Ethnicity (Ref: White)												
Hispanic	-2.373***	0.328	-2.383***	0.328	-2.387***	0.328	-2.570***	0.22	-2.569***	0.22	-2.569***	0.22
African-American	-2.028***	0.361	-2.023***	0.361	-2.030***	0.362	-1.950***	0.268	-1.950***	0.268	-1.950***	0.268
Asian/Other	1.458***	0.292	1.447***	0.292	1.451***	0.292	2.830***	0.258	2.831***	0.258	2.831***	0.258
Male	-0.825***	0.22	-0.826***	0.22	-0.816***	0.22	-0.188	0.165	-0.188	0.165	-0.188	0.165
Low Birthweight	-3.657***	0.257	-3.657***	0.257	-3.666***	0.257	-1.870***	0.284	-1.870***	0.284	-1.870***	0.284
Age at Kindergarten	6.976***	0.253	6.979***	0.253	6.977***	0.253	6.561***	0.184	6.562***	0.184	6.562***	0.184
Poverty	-3.322***	0.292	-3.316***	0.292	-3.321***	0.292	-2.898***	0.221	-2.899***	0.221	-2.898***	0.221
Mother's Education	0.591***	0.044	0.591***	0.044	0.591***	0.044	0.754***	0.034	0.754***	0.034	0.754***	0.034
Two Parent Family	2.288***	0.253	2.298***	0.253	2.277***	0.253	2.461***	0.2	2.461***	0.2	2.461***	0.200
Early Childhood Program (Ref: Parent-Care Only)												
Head Start	0.088	0.364	-1.573	1.439	0.075	0.364	0.603**	0.277	1.098	1.359	0.602**	0.277
Prekindergarten	1.963***	0.389	1.977***	0.389	-1.648	1.57	1.421***	0.289	1.421***	0.289	1.29	1.269
Private Preschool	1.865***	0.323	1.885***	0.322	1.904***	0.323	3.950***	0.25	3.949***	0.25	3.951***	0.250
Other Care	-0.714*	0.428	-0.713*	0.428	-0.708*	0.428	-0.035	0.312	-0.035	0.312	-0.035	0.312
Home Learning Environment												
Cognitive Stimulation	0.849***	0.151	0.844***	0.151	0.721***	0.161	0.186	0.157	0.187	0.157	0.182	0.158
Emotional Support	0.231	0.172	0.165	0.174	0.247	0.171	0.048	0.158	0.066	0.163	0.048	0.158
Head Start*Emotional	--	--	0.401	0.333	--	--	--	--	-0.113	0.301	--	--
Prek*Cognitive	--	--	--	--	0.853**	0.36	--	--	--	--	0.031	0.293

* p<0.05, ** p<0.01, *** p<0.001

Table C.18: Estimates of Reading Scores from 7k ECLS-B Models and Preferred 18k ECLS-K Hybrid CSMI Models

	7k ECLS-B						18k Hybrid CSMI ECLS-K					
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Race/Ethnicity (Ref: White)												
Hispanic	-1.658***	0.471	-1.683***	0.471	-1.676***	0.472	-1.331***	0.229	-1.331***	0.228	-1.331***	0.229
African-American	0.241	0.508	0.254	0.509	0.239	0.509	0.369	0.277	0.37	0.277	0.369	0.277
Asian/Other	3.265***	0.417	3.239***	0.417	3.257***	0.417	2.531***	0.266	2.531***	0.266	2.531***	0.266
Male	-2.346***	0.311	-2.350***	0.311	-2.336***	0.31	-1.532***	0.172	-1.532***	0.172	-1.532***	0.172
Low Birthweight	-3.172***	0.371	-3.173***	0.371	-3.182***	0.371	-1.800***	0.298	-1.800***	0.298	-1.801***	0.298
Age at Kindergarten	9.511***	0.363	9.517***	0.363	9.512***	0.363	4.975***	0.191	4.975***	0.192	4.975***	0.192
Poverty	-4.366***	0.422	-4.351***	0.422	-4.365***	0.421	-2.519***	0.223	-2.519***	0.223	-2.519***	0.223
Mother's Education	0.898***	0.065	0.899***	0.065	0.899***	0.065	0.794***	0.035	0.794***	0.035	0.794***	0.035
Two Parent Family	3.120***	0.368	3.145***	0.369	3.108***	0.368	1.946***	0.209	1.947***	0.209	1.946***	0.209
Early Childhood Program (Ref: Parent-Care Only)												
Head Start	-0.044	0.526	-3.944**	2.011	-0.059	0.526	-0.113	0.284	-0.018	1.44	-0.113	0.284
Prekindergarten	4.387***	0.555	4.418***	0.555	0.179	2.196	1.533***	0.298	1.533***	0.298	1.414	1.388
Private Preschool	2.815***	0.464	2.863***	0.464	2.861***	0.464	3.511***	0.255	3.511***	0.255	3.512***	0.255
Other Care	-1.790***	0.612	-1.788***	0.611	-1.783***	0.611	-0.469	0.33	-0.468	0.33	-0.468	0.33
Home Learning Environment												
Cognitive Stimulation	1.152***	0.218	1.140***	0.218	1.002***	0.227	0.186	0.146	0.186	0.146	0.182	0.15
Emotional Support	0.322	0.241	0.166	0.25	0.34	0.24	0.085	0.15	0.089	0.165	0.085	0.15
Head Start * Emotional	--	--	0.941**	0.471	--	--	--	--	-0.021	0.321	--	--
Prek * Cognitive	--	--	--	--	0.993**	0.502	--	--	--	--	0.028	0.32

* p<0.05, ** p<0.01, *** p<0.001

BIBLIOGRAPHY

Aikens, N. L., & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology*, *100*(2), 235–251. https://doi.org/10.1037/0022-0663.100.2.235

Anders, Y., Rossbach, H.-G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S., … von Maurice, J. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly*, *27*(2), 231–244. https://doi.org/10.1016/j.ecresq.2011.08.003

Ansari, A., & Crosnoe, R. (2015). Immigration and the interplay of parenting, preschool enrollment, and young children's academic skills. *Journal of Family Psychology*, *29*(3), 382–393. https://doi.org/10.1037/fam0000087.Immigration

Ansari, Arya, & Pianta, R. C. (2018). The role of elementary school quality in the persistence of preschool effects. *Children and Youth Services Review*, *86*, 120–127. https://doi.org/10.1016/j.childyouth.2018.01.025

Baker, E. H., Rendall, M. S., & Weden, M. M. (2015). Epidemiological Paradox or Immigrant Vulnerability ? Obesity Among Young Children of Immigrants, 1295–1320. https://doi.org/10.1007/s13524-015-0404-3

Barnett, Jung, K., Friedman-Krauss, A., Frede, E. C., Nores, M., Hustedt, J. T., … Daniel-Echols, M. (2018). State Prekindergarten Effects on Early Learning at Kindergarten Entry : An Analysis of Eight State Programs. *AERA Open*, *4*(2). https://doi.org/10.1177/2332858418766291

Barnett, W. S., Carolan, M. E., Fitzgerald, J., & Squires, J. H. (2011). *The state of preschool 2011*.

Barnett, W. S., Hustedt, J. T., Hawkinson, L. E., & Kenneth B. Robin. (2006). *The State of Preschool 2006*.

Bianchi, S. M., Cohen, P. N., Raley, S., & Nomaguchi, K. (2004). Inequality in Parental Investment in Child-Rearing: Expenditures, Time, and Health. In *Social inequality* (pp. 189–219).

Bornstein, M. H., & Tamis-lemonda, C. S. (1989). Maternal responsiveness and cognitive development in children. *New Directions for Child and Adolescent Development*, (43), 49–61.

Bradley, R. H., Caldwell, B. M., Rock, S. L., Ramey, C. T., Barnard, K. E., Gray, C., … Johnson, D. L. (1989). Home environment and cognitive development in the first 3 years of life: A collaborative study involving six sites and three ethnic groups in North America. *Developmental Psychology*, *25*(2), 217–235. https://doi.org/10.1037/0012-1649.25.2.217

Bradley, R. H., Caldwell, B. M., Rock, S. L., Ramey, C. T., Barnard, K. E., Gray, C., … Ohnson, D. L. (1989). Home environment and cognitive development in the first 3 years of life: A collaborative study involving six sites and three ethnic groups in North America. *Developmental Psychology*, *25*(2), 217–235. https://doi.org/10.1037/0012-1649.25.2.217

Bradley, R. H., Mckelvey, L. M., & Whiteside-Mansell, L. (2011). Does the Quality of Stimulation and Support in the Home Environment Moderate the Effect of Early

Education Programs? *Child Development*, *82*(6), 2110–2122. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed10&NEWS=N&AN=22004351

Brenner, P. S., & Boston, M. (2014). Testing the Veracity of Self-Reported Religious Practice in the Muslim World, *92*(November 2013), 1009–1037. https://doi.org/10.1093/sf/sot120

Bryk, A., Schneider, B. (2002). *Trust in Schools: A Core Resource for Improvement: A Core Resource for Improvement*. Russell Sage Foundation.

Caldwell, B., & Bradley, R. (1984). Home Observation for Measurement of the Environment (HOME) - Revised Edition. University of Arkansas, Little Rock.

Castro, D. C., Bryant, D. M., Peisner-feinberg, E. S., & Skinner, M. L. (2004). Parent involvement in Head Start programs: the role of parent, teacher and classroom characteristics. *Early Childhood Research Quarterly*, *19*(3), 413–430. https://doi.org/10.1016/j.ecresq.2004.07.005

Coleman, J. S. (1966). *Equality of Educational Opportunity (COLEMAN) Study (EEOS)*. Ann Arbor, MI.

Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, *94*, S95–S120. https://doi.org/10.2307/2780243

Connell, C. M., & Prinz, R. J. (2002). The Impact of Childcare and Parent–Child Interactions on School Readiness and Social Skills Development for Low-Income African American Children. *Journal of School Psychology*, *40*(2), 177.

Crane, J. (1996). Effects of home environment, SES, and maternal test scores on mathematics achievement. *Journal of Educational Research*, *89*(5), 305–314.

Cristofaro, T. N. 1., & Tamis-LeMonda, C. S. (2012). Mother-child conversations at 36 months and at pre-kindergarten: Relations to children's school readiness. *Journal of Early Childhood Literacy*, *12*(1), 68–97. https://doi.org/10.1177/1468798411416879

Crosnoe, R., & Elder, G. H. J. (2004). Family Dynamics, Supportive Relationships, and Educational Resilience During Adolescence. *Journal of Family Issues*, *25*(5), 571–602. https://doi.org/10.1177/0192513X03258307

Currie, J. (2001). Early Childhood Education Programns. *Journal of Economic Perspectives*, *15*(2), 213–238.

Downey, D. B., von Hippel, P. T., & Broh, B. A. (2004). Are Schools the Great Equalizer? Cognitive Inequality during the Summer Months and the School Year. *American Sociological Review*, *69*(5), 613–635. https://doi.org/10.1177/000312240406900501

Duncan, G. J., Kalil, A., & Ziol-Guest, K. (2013). Early childhood poverty and adult achievement, employment and health. *Family Matters*, (93), 27–35. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=ofm&AN=93275208&site=ehost-live&scope=site

Duncan, G. J., & Magnuson, K. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps. *Future of Children*, *15*(1), 35–54.

Duncan, G. J., & Murnane, R. J. (2016). Rising Inequality in Family Incomes and Children's Educational Outcomes. *RSF*.

Early, D. M., Iruka, I. U., Ritchie, S., Barbarin, O. A., Winn, D.-M. C., Crawford, G. M.,

… Pianta, R. 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. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ874840&site=ehost-live&scope=site

Feller, A., Grindal, T., Miratrix, L., & Page, L. C. (2016). Compared to what? Variation in the impacts of early childhood education by alternative care type. *The Annals of Applied Statistics*, *10*(3), 1245–1285. https://doi.org/10.1214/16-AOAS910

Gardner, F. (2000). Methodological Issues in the Direct Observation of Parent – Child Interaction : Do Observational Findings Reflect the Natural Behavior of Participants ? *Clinical Child and Family Psychology Review*, *3*(3).

Gauthier, A. H., Smeedeng, T. M., & Furstenberg, F. F. (2004). Are Parents Investing Less Time in Children? Trends in Selected Industrialized Countries. *Population and Development Review*, *30*(4), 647–671.

Gormley, W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology*, *41*(6), 872–884. https://doi.org/10.1037/0012-1649.41.6.872

Gormley, W. T., Phillips, D., & Anderson, S. (2018). The Effects of Tulsa's Pre-K Program on Middle School Student Performance. *Journal of Policy Analysis & Management*, *37*(1), 63–87. https://doi.org/10.1002/pam.22023

Gravelle, T. B. (2018). Politics, time, space, and attitudes toward US–Mexico border security. *Political Geography*, *65*, 107–116.

https://doi.org/10.1016/j.polgeo.2018.05.012

Griffin, E. A., & Morrison, F. J. (2016). The Unique Contribution of Home Literacy Environment to Differences in Early Literacy Skills, *4430*(November). https://doi.org/10.1080/0300443971270119

Handcock, M. S., Huovilainen, S. M., & Rendall, M. S. (2000). Combining registration-system and survey data to estimate birth probabilities. *Demography*, *37*(2), 187–192.

Handcock, M. S., Rendall, M. S., & Cheadle, J. E. (2005). Improved Regression Estimation of a Multivariate Relationship with Population Data on the Bivariate Relationship. *Sociological Methodology*, *35*(1), 291–334.

Hao, L., & Yeung, W. J. (2015). Parental Spending on School-Age Children: Structural Stratification and Parental Expectation. *Demography*, 835–860. https://doi.org/10.1007/s13524-015-0386-1

Heckman, J. J. (2012). The case for investing in young children. *Defending Early Childhood*, 235–242.

Hellerstein, J. K., & Imbens, G. W. (1999). Imposing moment restrictions from auxiliary data by weighting. *Review of Economics and Statistics*, *81*(1), 1–14.

Hook, J. Van, Bachmeier, J. D., Coffman, D. L., & Harel, O. (2015). Can We Spin Straw Into Gold ? An Evaluation of Immigrant Legal Status Imputation Approaches. *Demography*, *52*, 329–354. https://doi.org/10.1007/s13524-014-0358-x

Hoy, W. K., Tarter, C. J., & Hoy, A. W. (2006). Academic optimism of schools: A force for student achievement. *American Educational Research Journal*, *43*(3), 425–446.

Imbens, G. W., & Lancaster, T. (1994). Combining Micro and Macro Data in

Microeconometric Models. *Review of Economic Studies*, *61*(4), 655–680.

Kena, G., Hussar, W., McFarland, J., de Brey, C., Musu-Gillette, L., Wang, X., … Barmer, A. (2016). *The Condition of Education 2016. NCES 2016-144*.

Kornrich, S., & Furstenberg, F. (2013). Investing in Children: Changes in Parental Spending on Children, 1972-2007. *Demography*, *50*, 1–23. https://doi.org/10.1007/sl3524-012-0146-4

Lareau, A. (2000). *Home Advantage: Social class and parental intervention in elementary education*. Rowman & Littlefield Publishers.

Lee, R., Zhai, F., Brooks-Gunn, J., Han, W.-J., & Waldfogel, J. (2014). Head start participation and school readiness: evidence from the Early Childhood Longitudinal Study-Birth Cohort. *Developmental Psychology*, *50*(1), 202–215. https://doi.org/10.1037/a0032280

Lee, V. E., & Burkam, D. T. (2003). Dropping Out of High School: The Role of School Organization and Structure. *American Educational Research Journal*, *40*(2), 353–393.

Lee, V. E., & Smith, J. B. (1999). Social support and achievement for young adolescents in Chicago: The role of school academic press. *American Educational Research Journal*, *36*(4), 907–945.

Lipsey, M. W., Farran, D. C., & Durkin, K. (2018). Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly*, *45*, 155–176. https://doi.org/10.1016/j.ecresq.2018.03.005

Love, J. M., Kisker, E. E., Ross, C. M., Schochet, P. Z., Brooks-Gunn, J., Paulsell, D., … Brady-Smith, C. (2002). *Making a Difference in the Lives of Infants and Toddlers and their Families: The Impacts of Early Head Start. Executive Summary.* Washington, DC US.

Love, John M., Kisker, E. E., Ross, C., Raikes, H., Constantine, J., Boller, K., … Vogel, C. (2005). The Effectiveness of Early Head Start for 3-Year-Old Children and Their Parents: Lessons for Policy and Programs. *Developmental Psychology*, *41*(6), 885–901. https://doi.org/10.1037/0012-1649.41.6.885

Lowenstein, A. E., Wolf, S., Gershoff, E. T., Sexton, H. R., Raver, C. C., & Aber, J. L. (2015). The stability of elementary school contexts from kindergarten to third grade ☆. *Journal of School Psychology*, *53*, 323–335. https://doi.org/10.1016/j.jsp.2015.05.002

Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). Does Prekindergarten Improve School Preparation and Performance? *Economics of Education Review*, *26*(1), 33–51. https://doi.org/http://www.elsevier.com/wps/find/journaldescription.cws_home/743/description#description

Magnuson, K. A., & Waldfogel, J. (2005). Preschool Child Care and Parents’ Use of Physical Discipline. *Infant and Child Development*, *14*, 177–198.

McLaughlin, A. E., Campbell, F. A., Pungello, E. P., & Skinner, M. (2007). Depressive symptoms in young adults: the influences of the early home environment and early educational child care. *Child Development*, *78*(3), 746–756.

https://doi.org/10.1111/j.1467-8624.2007.01030.x

Melhuish, E. C., Phan, M. B., Sammons, P., Siraj-blatchford, I., & Taggart, B. (2008). Effects of the Home Learning Environment and Preschool Center Experience upon Literacy and Numeracy Development in Early Primary School, *64*(1), 95–115.

Melhuish, E. C., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., Phan, M. B., & Malin, A. (2008). Preschool influences on mathematics achievement. *Science*, *321*(5893), 1161–1162. https://doi.org/10.1126/science.1158808

Miller, E., Farkas, G., Vandell, D. L., & Duncan, G. J. (2014). Do the Effects of Head Start Vary by Parental Preacademic Stimulation? *Developmental Psychology*, *85*(4), 1385–1400. https://doi.org/10.1111/cdev.12233

Network, N. E. C. C. R. (2003). Does Quality of Child Care Affect Child Outcomes at Age 4 1/2? *Developmental Psychology*, *39*(3), 451–469. https://doi.org/10.1037/0012-1649.39.3.451

NICHD Early Child Care Research Network. (1999). Child Care and Mother-Child Interaction in the First 3 Years of Life. *Developmental Psychology*, *35*(6), 1399–1413.

Nievar, M. A., Van Egeren, L. A., & Pollard, S. (2010). A meta-analysis of home visiting programs: Moderators of improvements In maternal behavior. *Infant Mental Health Journal*, *31*(5), 499–520.

Padilla, C. M., & Ryan, R. M. (2018). School readiness among children of Hispanic immigrants and their peers: The role of parental cognitive stimulation and early care and education. *Early Childhood Research Quarterly*.

https://doi.org/10.1016/j.ecresq.2018.04.008

Parcel, T. L., & Dufur, M. (2009). Family and school capital explaining regional variation in math and reading achievement. *Research in Social Stratification and Mobility*, *27*(3), 157–176. https://doi.org/10.1016/j.rssm.2009.04.003

Parcel, T. L., & Dufur, M. J. (2001). Capital at Home and at School : Effects on Student Achievement. *Social Forces*, *79*(3), 881–911.

Parcel, T. L., Dufur, M. J., & Zito, R. C. (2010). Capital at Home and at School : A Review and Synthesis. *Journal of Marriage and Family*, *72*(4), 828–846. https://doi.org/10.1111/j.l741-3737.2010.00733.x

Park, H., Buchmann, C., Choi, J., & Merry, J. J. (2016). Learning Beyond the School Walls : Trends and Implications, (May), 1–22. https://doi.org/10.1146/annurev-soc-081715-074341

Payne, A. C., Whitehurst, G. J., Angel, A. L., & Angell, A. L. (1994). The role of home literacy environment in the development of language ability in preschool children from low-income families. *Early Childhood Research Quarterly*, *9*(3–4), 427–440. https://doi.org/10.1016/0885-2006(94)90018-3

Philips, D. A., & Dodge, K. (2017). The Current State of Scientific Knowledge on Pre-Kindergarten Effect, 106. Retrieved from https://www.brookings.edu/wp-content/uploads/2017/04/duke_prekstudy_final_4-4-17_hires.pdf

Pinto, A. I., Pessanha, M., & Aguiar, C. (2013). Effects of home environment and center-based child care quality on children’s language, communication, and literacy outcomes. *Early Childhood Research Quarterly*, *28*(1), 94–101.

https://doi.org/10.1016/j.ecresq.2012.07.001

Raudenbush, S. W. (2008). Advancing Educational Policy by Advancing Research on Instruction. *American Educational Research Journal*, *45*(1), 206–230.

Raudenbush, S. W., & Eschmann, R. D. (2015). Does Schooling Increase or Reduce Social Inequality? *Annual Review of Sociology*, *41*, 443–470. https://doi.org/10.1146/annurev-soc-071913-043406

Reardon, S. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In R. M. Murnane & G. Duncan (Eds.), *Whither Opportunity* (pp. 91–116). New York, NY: Russell Sage Foundation Press.

Rendall, M. S., Admiraal, R., Derose, A., Digiulio, P., Handcock, M. S., & Racioppi, F. (2008). Population constraints on pooled surveys in demographic hazard modeling. *Statistical Methods and Applications*, *17*(4), 519–539. https://doi.org/10.1007/s10260-008-0106-8

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581–592.

Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons.

Sammons, P., Anders, Y., Sylva, K., & Melhuish, E. (2009). Children's Cognitive Attainment and Progress in English Primary Schools During Key Stage 2: Investigating the potential continuing influences of preschool education. *Frühpädagogische Förderung in Institutionen*, (March), 179–198. https://doi.org/10.1007/978-3-531-91452-7

Sayer, L. C., Bianchi, S. M., & Robinson, J. P. (2004). Are Parents Investing Less in Children? Trends in Mothers ' and Fathers ' Time with Children. *American Journal of Sociology*, *110*(1), 1–43.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*.

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

Sintonen, S., Tarkiainen, A., Cadogan, J. W., Kuivalainen, O., Lee, N., & Sundqvist, S. (2016). Cross-country cross-survey design in international marketing research: The role of input data in multiple imputation. *International Marketing Review*, *33*(3), 454–482.

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A Review of School Climate Research. *Review of Educational Research*, *83*(3), 357–385. https://doi.org/10.3102/0034654313483907

van Buuren, S. (2012). *Flexible Imputation of Missing Data*. CRC Press.

Von Hippel, P. T. (2009). How to Impute Interactions, Squares, and other Transformed Variables. *Sociological Methodology*, *1*, 265–291.

Webster-Stratton, C., & Hammond, M. (1998). Conduct Problems and Level of Social Competence in Head Start Children: Prevalence, Pervasiveness, and Associated Risk Factors. *Clinical Child and Family Psychology Review*, *1*(2), 101–124.

Curriculum Vitae

Date of birth: August 8, 1983
Place of birth: Toronto, Canada

EDUCATION

2013- Ph.D. Candidate (Expected 2019), Sociology, Johns Hopkins University
Dissertation: Effects of Family and School Institutions on Child Outcomes from Early Childhood through Elementary School
Committee: Lingxin Hao (chair), Julia Burdick-Will, Andrew Cherlin

2013 M.A., Sociology, University of California Los Angeles

2005 B.A., Sociology with Honors, University of Maryland College Park

FELLOWSHIPS, GRANTS, AND HONORS

2017 Martin Levin Methodology Fellowship (JHU)

2017 Peabody Scholar in Sociology of Education (JHU)

2013 NSF Graduate Research Fellowship Program Honorable Mention

2012 Graduate Summer Research Mentorship Grant (UCLA)

2011-13 Eugene Cota-Robles Fellowship (UCLA)

PUBLICATIONS AND REPORTS

Grigg, J., Connolly, F., D'Souza, S., and Cronister, C. (2016). "Early Education Data Collaborative: Kindergarten Readiness for Baltimore's High School Class of 2027." BERC, Baltimore, MD.

Connolly, F., Grigg, J., Cronister, C., and D'Souza, S. (2015). "Born in Baltimore." Baltimore Education Research Consortium, Baltimore, MD

Estacion, A., Cotner, B.A., D'Souza, S., Smith, C.A.S., & Borman, K.M. (2011). "Who enrolls in dual enrollment and other acceleration programs in Florida high schools?" (Issues & Answers Report, REL 2012–No. 119). Washington, DC: U.S. Department of Education, IES, National Center for Education Evaluation and Regional Assistance, REL Southeast.

Estacion, A., D'Souza, S., Bozick, R. (2011). "Characteristics of career academies in 12 Florida school districts." (Issues & Answers Report, REL 2011–No.106). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast.

Hartman, J., Wilkins, C., Gregory, L., Feagans-Gould, L., & D'Souza, S. (2011). "Adaptation of the Consortium on Chicago School Research on-track for high school graduation indicator for Texas." (Issues & Answers Report, REL 2010–No. 100). Washington, DC: U.S. Dept of Education, IES, National Center for Education Evaluation and Regional Assistance, REL

Southwest.

Xu, Z., Hannaway, J., & D'Souza, S. (2009). Student transience in North Carolina: The effect of mobility on student outcomes using longitudinal data. CALDER Working Paper #22

MANUSCRIPTS IN PROGRESS

D'Souza, Stephanie. "The Joint Effects of the Home Environment and Early Childhood Programs on Kindergarten Readiness: Evidence from ECLS-B." Submitted for review to *Sociology of Education*

Hao, Lingxin and Stephanie D'Souza. "Confronting Data Challenges in Research on Early Conditions and Later Outcome Trajectories: Methods of Combining Samples of Panel Data." Will be submitted for review to *Journal of Educational Effectiveness*.

Hao, Lingxin, and Stephanie D'Souza. "The Initialization and Persistence of STEM Majoring among First-Generation Students." Pathways through College Research Network.

Hao, Lingxin and Stephanie D'Souza. "Labor Market Activities and Wealth of Immigrant and Native-born Americans."

Parker, Carolyn, Jeffrey Grigg, Stephanie D'Souza, Charlie Mitchell, Erica Smith, and Yolanda Abel. "Enhanced Interest and Awareness in Science and Engineering with Students Placed at Risk After One Year of a STEM Intensive Partnership." Under review, *Journal of Education for Students Placed at Risk*

Grigg, Jeffrey, Faith Connolly, Stephanie D'Souza, and Charlie Mitchell. "How Early Childhood Programs Interact with Early Life Circumstances: Evidence from Baltimore."

RESEARCH EXPERIENCE

2017-19 Research Assistantship, Professor Lingxin Hao, JHU

2013-17 STEM Achievement in Baltimore Schools, Research Assistant, JHU

2013-15 Early Education Data Collaborative, Research Assistant, BERC

2011 Early Identification of At-risk Students, Research Analyst, Impaq International

2009-11 High School Reform Studies, Research Officer I, Academy for Educational Development

2007-09 Student Transience CALDER study, Research Assistant, The Urban Institute

2005 Kids Count Databook, Research Intern, Population Reference Bureau