THE ROLE OF DOSAGE AND QUALITY OF HEAD START EXPERIENCES IN THE DEVELOPMENT OF SELF-REGULATION

A Dissertation

Presented to

The Faculty of the Curry School of education

University of Virginia

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

---

by

Carolina Melo Hurtado, M.A.

B.A. Pontificia Universidad Católica de Chile, 2003

M.A. Ed. Pscyh. University of Arizona, 2005

May 1st, 2017

Department of Curriculum and Instruction

Curry School of Education

University of Virginia

APPROVAL OF THE DISSERTATION

This dissertation, The Role of Dosage and Quality of Head Start Experiences in the Development of Self-Regulation, has been approved by the Graduate Faculty of the Curry School of Education in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

____
Robert C. Pianta, Ph.D., Chair

____
Peter Youngs Ph.D. Co-Chair

____
Jennifer LoCasale-Crouch

____
Natalia Palacios

Date______________________________________

Abstract

The present study examined how dosage and quality of preschool experiences in Head Start (HS) are associated with cognitive and behavioral self-regulation skills in kindergarten. The study included 2,383 children from the Head Start Family and Children Experiences Survey (2009). Using multiple regression (OLS), with multiple imputation methods to address missing data, and propensity score matching to address selection bias, this study examined how the number of hours a week in HS, absenteeism, and number of years (starting at 3-years of age versus at 4-years of age) was related to self-regulation. The study also examined how the quality of classroom experiences, conceptualized as domain-general and domain-specific aspects of teacher-child interactions, was related to self-regulation; and how the quality of teacher-child interactions moderated the relation between dosage and self-regulation. There we two main findings. First, an additional year in HS was the only form of dosage that was significantly associated to self-regulation in kindergarten. Children that attended one more year of HS scored 0.30 points higher on the cognitive self-regulation measure and were scored 0.32 standard deviations higher in behavioral self-regulation, as reported their teachers in kindergarten. Second, quality of domain-general of teacher-child interactions (Responsive Teaching) moderated the relation between hours a week in HS and cognitive self-regulation. In other words, children who participated in classrooms with higher-quality of teacher-child interactions benefited the most from more hours a week in HS. Findings from this study contribute to the growing body of evidence about how dosage and quality of early childhood education experiences relate to child development. Results support the importance of the investment in early childhood education amount and quality for the development of self-regulation.

Dedication

I dedicate my dissertation to my husband Eduardo Katz and my children, Matteo (12), Pedro (9), Emma (7), and Tomás (4). Thanks for inspiring me, giving me the strength for each day of this journey, for providing me with my daily dose of oxitocyn, and the perspective necessary to do this work. Without your love and support I could have not done it.

Acknowledgements

This dissertation would have not bee possible without the support of so many around me.

First, I would like to thank my mentor Bob Pianta for his guidance these years, for his support and encouragement, for helping me find the confidence and determination needed to do this work, for always holding me to high expectations, and providing me with amazingly helpful and timely feedback.

I also would like to thank my whole committee, for their support and feedback in this process. Peter, thank you for helping me navigate the intricate world of program requirements, for being such a great example of a teacher, and for being always available to provide a word of support and encouragement. Natalia, thank you for joining the committee and providing great perspectives and feedback. Jennifer, thank you for being such an amazing informal mentor, and always reminding me to "take a step back" and see the bigger picture.

I would like to thank Arya Ansari, Jorge Miranda, Justin Doromal and Shannon Reiley for being a great resource and support for all my statistical questions, and for being always willing to have a chat with me about the exciting world of statistics. I could have never finished this dissertation is such a timely manner without you.

Thanks, Bridget Hamre, Jennifer LoCasale-Crouch, and Jessica Vick.Whittaker for welcoming me into your lab. I have enjoyed being part of such a great group of scholars, and I appreciate having that space to develop my thinking and learn from all of you.

Special thanks to Marcia Invernizzi, an incredible scholar, mentor, and human being. Thank you for supporting me in a difficult process of doctoral discerning during my first year. Thank you for encouraging me to pursue my interests and goals in such a generous and supportive manner.

Thanks to Sonia Cabell for leading the writing group, which I valued all these years immensely, and helped me improved my writing beyond what I would have imagined possible. Thanks for your constant support encouragement and your always-positive attitude.

I also want to give special thanks to my two Chilean friends Pilar Alamos and Francisca Romo. My experience in this journey was so much more enjoyable because you were here. Thanks for being such great friends and support group. I'm not sure I could have finished this doctorate without you! Also, thanks to Cathy (our honorary Chilean!), Christina, Ann, Chelsea, Helyn, Jeff, Justin, Shannon. Thank you all for adopting me in your cohorts!

I would like to thank special people who in different ways inspired me to pursue the doctorate.

Thank you, Pelusa Orellana, for being such a great colleague, mentor, and inspiration. Without your example and encouragement, I might have never embarked in this crazy idea of getting a doctorate while raising four children!

Thank you, Jota Hurtado, for encouraging me to go for more so many years ago, when I was an undergraduate. Thanks for your all your love and support.

Thanks to my parents for always believing in me. For always having high expectations and lots of love.

Finally, none of this could have happened without the love and support of my husband and children. Thanks to my husband, Eduardo Katz, for his encouragement and support all these years, for picking me up when I thought I could not do it (which was countless times), for taking the load when things got crazy, and for being such a great dad to our kids in this intense time of our life. Thanks to my kids, who have grown among papers and gotten used to falling asleep to the sound of the keyboard. For being such great helpers, and giving me so much perspective and inspiration.

This dissertation and my experience these years have been possible thanks to the support of Universidad de Los Andes, and "Becas Chile para doctorado en el extranjero" from Conycit, Chile.

Table of Contents

List of Figures

List of Tables

Introduction

Self-regulation is an important skill (Weissberg, Durlak, Domitrovich & Gullota, 2015), which is related to success in school, health and general well-being (Becker, Miao, Duncan, & McClelland, 2014; Moffitt et al., 2011; Morrison, Cameron Ponitz, & McClelland, 2010; Raver et al., 2011; Ursache, Blair, & Raver, 2012). The ability to monitor and regulate behavior, emotion, and cognition to accomplish one's goal, and to adapt to the cognitive and social demands of specific situations develops considerably during the preschool years. Thus, it is well established that the quality of preschool experiences during these years has a major role in its development (Morris et al., 2012; Raver, 2014; Raver et al., 2011; Rimm-Kaufman, Curby, Grimm, Nathanson, & Brock, 2009).

Less is known, however, about the effects of the amount of preschool experiences, also called *dosage*, on the development of self-regulation. Recent studies have found mixed evidence regarding the associations between dosage of preschool and socioemotional skills. A set of studies, for example, have found that more hours a week in preschool can have adverse effects in a range of socioemotional outcomes that are related to self-regulation (Huston, Bobbitt, & Bentley, 2015; Loeb, Bridges, Bassok, Fuller, & Rumberger, 2007; McCartney et al., 2010; Skibbe, Connor, Morrison, & Jewkes, 2011). While research examining the number of years that children attend pre-school, has found mostly positive relations to child academic and socioemotional outcomes (Moore et al., 2015; Wen, Leow, Hahs-Vaughn, Korfmacher, & Marcus, 2012; Youn, 2016a). These studies nonetheless have not examined the relation between number of years and/or number of hours as week in pre-school, and self-regulatory skills directly.

A recent report by the Nobel laureate James Heckman (2016) shows evidence of the benefits of high-quality early childhood programs that start from birth to age five for disadvantaged children. The authors emphasize the role of preschool in promoting the development of socioemotional skills, which in turn result in better long-term life outcomes, with investment returns of as much as 13% per dollar invested. Even though the data from this study comes from small-scale programs in the 1970s, it shows an important proof of concept: high-quality intense dosage ECE programs can have a significant impact the development of socioemotional skills and long-term outcomes.

Given the mixed evidence regarding the effects of dosage of preschool experiences on the development of socioemotional skills, and the consistent evidence about the importance of quality experiences in preschool for the development of self-regulation, the next step is to examine under which conditions more years and more hours in preschool can be beneficial for the development of socioemotional skills. Evidence about the importance of quality of early childhood experience would suggest that more years and more hours of ECE, should be beneficial only if the experiences are of an adequate level of quality (McCartney et al., 2010). However, this specific question has yet to be examined.

The present study will examine the relation between dosage in Head Start (HS), classroom quality, and self-regulation. To do this, I will first examine the associations between dosage, conceptualized as number of hours a week, and number of years in HS, and children's self-regulation. Then I will discuss absenteeism as a moderator of dosage, and classroom quality as a moderator of the relations between dosage and self-regulation. I hypothesize that more hours a week and more years in HS will have a positive and stronger association to self-regulation in kindergarten only when classroom have certain minimum quality features. Furthermore, it is possible that classrooms that are characterized by low-quality experiences will have a negative association with children's self-regulation.

Literature Review

1. Self-Regulation

1.a Conceptualizing Self-Regulation

Self-regulation is the ability to monitor and regulate behavior, emotion, and cognition to accomplish one's goal, and to adapt to the cognitive and social demands of specific situations (Berger, 2011; Blair, Berry, & Friedman, 2012). Self-regulatory skills develop across lifespan but have a substantial growth during the preschool and early schooling years (Kochanska, Coy, & Murray, 2001). During these years, children move from a model of co-regulation, where the child's regulation of behavior is largely dependent on the significant adults in their life, to a model of self-regulation, in which the child starts to try out his skills in their daily interactions. Daily preschool and school activities require putting self-regulation skills into practice, like taking turns in play and conversations, focusing and shifting attention, sitting quietly in circle time, following directions with multiple steps, solving simple problems. These new demands from the context rely on set of self-regulatory skills, like being able to inhibit automatic responses or reactions in favor of less natural while more adaptive responses, being able to keep pieces of information in short-term memory, focusing attention on relevant stimuli and shift attention when needed. As such, self-regulation has been identified as one of the foundational processes implicated in children's readiness for school (Becker, Miao, Duncan, & McClelland, 2014; Morrison, Cameron Ponitz, & McClelland, 2010; Raver et al., 2011; Ursache, Blair, & Raver, 2012), and has been associated with long-term academic and behavioral outcomes (Denham et al., 2013).

Even though there is some disagreement in the fields of psychology and education about the conceptualization of self-regulation (see Jones, Zaslow, Darling-Churchill, & Halle, 2016), there is consensus in that it is a multidimensional construct, which is conceptually useful to describe as having three overlapping and interrelated domains: Cognitive, Emotional and Behavioral (Berger, Kofman, Livneh, & Henik, 2007; Calkins & Williford, 2009; Murray, Rosanbalm, Christopoulos, & Hamoudi, 2015)

Cognitive self-regulation includes the abilities to focus and redirect attention, to have cognitive flexibility, and inhibit impulses, skills that often fall under the domain of Executive Functions (Halle et al., 2016). Emotional self-regulation refers to the ability to manage and modulate strong or unpleasant feelings, which partly relies on cognitive regulatory processes such as inhibitory control (Calkins, 1997). Both cognitive and emotional self-regulation are important and necessary skills for exerting behavioral regulation (Blair & Dennis, 2015), which is the ability to organize and monitor behavior, including compliance to adult demands and directives, delaying gratification, inhibiting impulsive responses (Kuczynski & Kochanska, 1995). Skills such as persistence, and organizing cognitive skills to solve problems and direct behavior towards a goal are also part of behavioral self-regulation (Berger, 2011; 2015; Murray et al., 2015; Smith-Donald, Raver, Hayes, & Richardson, 2007).

The present study focuses particularly on cognitive and behavioral domains of self-regulation, as defined above. Consistently with recent recommendations about the measurement of self-regulation (Campbell, 2016), this study includes more than one type of measure of self-regulation, including a direct measure, and teacher-reported measures. More specifically, the study includes a direct measure of cognitive self-regulation, which is part of the established Preschool Self-Regulation Assessment (PSRA; Smith-Donald, Raver, Hayes, & Richardson, 2007), and a teacher-reported measure of children behavioral problems and learning behavior in the classroom. Cognitive self-regulation skills in preschool have been found to predict a range of academic outcomes (Blair et al., 2015; Cameron Ponitz, McClelland, Matthews, & Morrison, 2009) with a particularly strong association to mathematic performance (Espy, 2004). Behavioral self-regulation, on the other hand, predicts a range of academic (Cameron Ponits et al., 2009; Morrison et al., 2010), social and behavioral outcomes (Blair & Raver, 2015; Lonigan et al., 2017; Rimm-Kaufman, et al., 2009).

Additionally, consistently with the theoretical model from the Collaborative for Academic, Social, and Emotional Learning (CASEL), this study acknowledges self-regulation as a core competency involved in the development of socio-emotional skills. Children need to manage and integrate their cognition, affect, and behavior to deal effectively with daily tasks and challenges, both at school and in life (Weissberg, Durlak, Domitrovich & Gullota, 2015).

1. b. Self-regulation and its Causal Relation to Child Outcomes.

Self-regulatory skills have been found to be strongly related to many aspects of life, from school readiness and school success, to mental health, public safety and job success (Best, Miller, & Naglieri, 2011; Brock, Rimm-Kaufman, Nathanson, & Grimm, 2009; Espy et al., 2004; McClelland, Acock, Piccinin, Rhea, & Stallings, 2012; Moffitt et al., 2011). As discussed above, self-regulation has been associated with a wide variety of outcomes in school settings, including better social competencies, less behavioral problems, engagement, and math and reading achievement (Blair & Razza, 2007; Raver et al., 2012).

However, most of the studies examining the relation between self-regulation and other outcomes use correlation models, with only a few studies examining causal relations (Blair, McKinnon, & Investigators, 2016; Willoughby, Kupersmidt, & Voegler-Lee, 2012). Because it is unfeasible to randomize self-regulation skills in children, the closest mechanism to examine actual near-causal relations between self-regulation and child outcomes are studies that randomize classrooms to interventions that promote self-regulation. By randomizing classrooms to interventions for improving self-regulation, when the interventions are effective in boosting children's skills in these domains, different methodological and statistical mechanisms can be utilized to examine causality i.e., fixed effects (Willoughby et al., 2012), instrumental variables (Gennetian, Magnuson, & Morris, 2008), or mediation models (Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008; see Raver et al., 2011).

Several randomized control trials (RCT) have examined the effects of improved self-regulation on academic and behavioral outcomes. For example, the Chicago School Readiness Project (CSRP) used self-regulation as an Instrumental Variable (IV) to test the hypothesized mechanism of self-regulation mediating gains in academic outcomes (Raver, 2012). Researchers randomly assigned nine preschool sites to a comprehensive intervention service aimed at supporting children's self-regulation, and nine sites as a control group. The study found that children in the intervention group displayed significant increases in executive functions, such as inhibitory control, and attention, and that these mediated gains in vocabulary, letter-naming and early math skills (Raver et al., 2011). They also found that children had less behavioral problems, and reduced sad and withdrawn behavior (Raver et al., 2012; 2008).

Another RCT that supports the relation between self-regulation and school readiness is REDI (or Research-Based, Developmentally Informed). REDI focused on the regulation of behavior and awareness of emotions through language and literacy experiences for preschoolers and emphasized reading and conversation between teachers and children. The results of this study support the mediating role of self-regulation, in particular of executive function skills in children's vocabulary, phonological and print awareness (Bierman et al., 2008).

Similarly, results from a RCT that examined the impact of Tools of the Mind, a curriculum designed to enhance children's EF within an instructional context (Barnett et al., 2008), showed that children who received the curriculum improved significantly more, compared to their peers in the control group, in executive functions, reasoning abilities, attention, reading, vocabulary and mathematics by the end of kindergarten, and continued to increase in first grade (Blair & Raver, 2014). Interestingly, some of the effects found in this study (for inhibitory control, vocabulary and reasoning abilities) were greater for children in higher-poverty schools. Additionally, the study found evidence or reduced cortisol levels for children in high-risk schools that were in the treatment group, indicating a significant impact on stress response in comparison with those children in the control group. These findings highlight the importance of self-regulatory skills as a way of closing the achievement gap for this group of children (Blair & Raver, 2015). Similar evidence has been shown in 4Rs (Reading, Writing, Respect, and Resolution), an intervention with school-age children that targeted children's effective modulation of emotion and behavior in peer interactions (Jones, Brown, & Lawrence Aber, 2011). Children in the intervention group showed improved socio-emotional abilities and academic achievement, with a differential positive benefit for those students at initial higher risk.

In sum, results from these randomized controlled studies support the claim that better self-regulation skills lead to children's gains in socio-emotional and academic outcomes. Additionally, higher levels of self-regulation seem to have clear differential benefits for children in high poverty families and schools (Blair & Raver, 2014; Jones et al., 2011). Last but not least, these studies also present evidence of the malleability of self-regulatory skills, and support the notion that high-quality ECE environments can have an effect on the development of self-regulation.

2. Dosage: Amount of Early Childhood Education experiences

In spite of some recent criticism of the methods and cautions regarding generalization of the studies examining the effects of preschool on long-term outcomes (Farran & Lipsey, 2016), there seems to be consensus among the research community regarding the benefits of early childhood education for school readiness and other long life outcomes (Campbell et al., 2014; Heckman & Raut, 2016; Schweinhart, 1993; Yoshikawa, Weiland, & Brooks-Gunn, 2013). Researchers also agree in that these benefits seem particularly salient and strong for children from low-income families.

Nonetheless, evidence is less consistent regarding the optimal amount of preschool. The question of how much preschool is the "right" amount of preschool to produce the most optimal cost-benefit relation is a somewhat less studied one, with significant policy implications. Some studies have explored this question of amount or "dosage" of preschool, concluding that greater dosage, in general, has positive benefits for academic outcomes by kindergarten (Xue, Burchinal, Auger, & Tien, 2016), but shows mixed associations to socioemotional outcomes (Huston, Bobbitt, & Bentley, 2015). What is the optimal amount of preschool for promoting the development of self-regulation, in particular, and under which conditions is still unclear.

Predominant conceptualizations of dosage of ECE in the study of the quantity or amount of preschool and child outcomes include: attendance/absenteeism, the number of hours a week, and attending one versus two years. There is evidence supporting the notions that, less absence and a greater number of years (two versus one) in ECE are related to better academic and behavioral outcomes (Arbour et al., 2016; Xue et al., 2016), while the number of hours a week is mostly related to externalizing behavior (Huston et al., 2015).

For the purpose of this study I include all three conceptualizations in the analyses. More specifically, I will examine the differences between starting Head Start at two versus at three-years of age on self-regulation skills, the relation between number of hours a week and self-regulation skills, and the role of absenteeism as a moderator of these relations.

The particular nature of the Family and Children Experiences Survey (FACES) 2009 data, which included two cohorts of children that were attending Head Start for the first time, presents an excellent opportunity to focus on the question of dosage as the number of years in HS that is otherwise difficult to do. All children were assessed on their first year of attendance to HS and in four time points, while some children attended two years, others only one year. Additionally, the question of the optimal number of years of Head Start (HS) for the most cost-benefit investments, in particular is one of important policy implications. Research on this topic can contribute to the knowledge base by providing evidence that supports or not the funding for an additional year of HS.

In the next sections I summarize the evidence about the different forms of preschool dosage and its associations to child outcomes. I then discuss some methodological consideration in the study of dosage as one versus two years of ECE, and the evidence from studies examining the question of one versus two years, specifically for Head Start (HS).

2.a. Dosage of Early Childhood Education Experiences and Child Outcomes

Less absence in preschool has been found to predict better academic and behavioral outcomes, often moderating the “effects” of preschool on child outcomes (Arbour et al., 2016). For example, Ansari & Purtell (forthcoming) examined the implications 3- and 4-year-old’s absences from Head Start using nationally representative data from the FACES 2009 cohort, for their early academic learning, and found that children who attended more days of school showed significantly greater gains in math and literacy during the preschool years than those who missed more days. Furthermore, this study found that excessive absenteeism detracted from the benefits of high-quality preschool on child learning. Xue and colleagues (2016) showed similar findings in a meta-analysis examining absenteeism, among other conceptualizations of dosage, in eight different data sets of preschool children and found that fewer absences were associated with stronger gains in literacy and mathematics skills.

More time in preschool contexts, conceptualized as more hours a week, has shown mixed associations to child outcomes, with children that spend more hours in center care settings sometimes more academically ready in kindergarten (Skibbe, Connor, Morrison, & Jewkes, 2011), while displaying more behavioral problems than their peers who have attended ECE for less hours a day (Huston et al., 2015; McCartney et al., 2010; NICHD Early Child Care Research Network, 2004). For instance, Loeb et al., (2007) using nationally representative data from Early Childhood Longitudinal Study Kindergarten Cohort (ECLS-K) found that children spending more than 30 hours a week in center care setting had lower self-control and higher externalizing behavior. However, Magnuson et al., (2007), in a study with the same data set made the nuance that the negative associations did not hold for children in Head Start centers. Similarly, Skibbe et al., (2011) found that neither the first nor the second year of preschool predicted self-regulation skills, while more years in ECE was associated to better decoding and letter knowledge gains. Nonetheless, the study used a small sample (N=76) and included only 10% of HS children. Taken together, these findings suggest that negative effects of more hours a week at center care are less clear for disadvantaged children. One possible interpretation of these results considers that disadvantaged children are likely to live in a chaotic family environment, which in turn may explain that they benefit from participation in care settings that present a more predictable environment. Additionally, McCarteny et al., (2010) found that more center care hours had a stronger negative association to socioemotional outcomes in low-quality settings. This finding suggests that quality of classroom experiences may be a key moderator of the effects of dosage on children's development of self-regulation.

Lastly, evidence systematically supports the benefits of attending preschool for two years, in contrast to one, with results consistently showing better academic outcomes in kindergarten for those children who start preschool a year earlier (at three years of age), and few mixed results for socioemotional outcomes. Research has shown that students who attend two years of preschool have significantly stronger literacy and numeracy skills (Domitrovich et al., 2013; Reynolds, 1995; Wen, Leow, Hahs-Vaughn, Korfmacher, & Marcus, 2012; Xue et al., 2016; Youn, 2016a), and socioemotional outcomes (Moore et al., 2015; Wen et al., 2012; Youn, 2016b) that those who attend only one year. Whereas, some studies have found negative effects of starting preschool earlier on socioemotional outcomes (Loeb, Bridges, Bassok, Fuller, & Rumberger, 2007; McCartney et al., 2010), and effects tend to fade out in time (Reynolds, 1995).

Reynolds, A. J. (1995), for example, conducted one of the first studies that examined the topic of dosage as one versus two years in early childhood education, including a sample of 757 low-income black children from the federally funded Child Parent Center preschool program. His findings indicated that there were significant advantages in academic skills for those children that attended two years of preschool versus one, at the beginning of kindergarten. However, these benefits for disappeared into elementary grades, with no significant differences between grades 1 and 6th.

In contrast, a newer study found evidence supporting the benefits of an additional year of preschool on life-long outcomes. Arteaga and colleagues (2014), presented evidence from the first study examining longer-term effects of preschool dosage on development, which supported that higher dosage of ECE related positively to adult outcomes. The authors, who conducted this study with data from The Chicago Longitudinal Study, found that children who attended two years of preschool, instead of one, were significantly less likely to have received special education, to be abused or neglected or to commit crimes by age 26.

I will further discuss the relation between dosage, conceptualized as one versus two years, and child outcomes in the context of Head Start, given that a majority of the studies that have examined this question have used data from this program. However, I would first like to review some methodological considerations in the study of dosage as number of years.

2.b. Methodological Considerations in the Study of Dosage and Child outcomes

One of the challenges in studying effects of dosage, as number of years on child outcomes, is how to draw causal inferences about the relation between dosage and child outcomes in the absence of randomization. Very few studies are able to randomize children into one versus two years of pre-school, therefore the majority of studies are correlational. The problem with the correlational design is selection bias. In a randomized design the probability of a child receiving the treatment (in this case attending two year of preschool) is by definition 50%, and there is expected equivalence in terms of re-treatment characteristics (Austin, 2011), while in a non-randomized design children can vary in their probability to attend two years of preschool and in their re-treatment characteristics. The samples of children that attend two years may be very different to the sample of children that attend one year, in ways that may also be related to the outcome of study (Zanutto, Lu, & Hornik, 2005).

In the literature examining ECE dosage and child outcomes there are two methods that researchers mainly use to deal with this problem: *Regression Discontinuity* (RD) and *Propensity Score Matching* (PSM). Regression Discontinuity uses the age cut-off strategy to study the differences in an additional year that occur naturally because of some children not having the strict cut-age to move to kindergarten (see Jenkins, Farkas, Duncan, Burchinal, & Vandell, 2016). This method examines the differences in the outcomes for these groups that are close to the cut-off age, and who are statistically non-different, but who fall into the treatment (o non-treatment) category by being in either side of the cutoff age. This method however, is mostly used to examine the effects of preschool programs (schooling effects) and not as much in the research about dosage. Even though RD has strong internal validity (Cook, Shadish, & Wong, 2008), it has reduced external validity, because the sample of analysis is reduced to those individuals who are only a few months away from the age cut-off, generalizations to the population are limited.

Another, more commonly used, statistical approach to minimize selection bias and improve the ability to make causal inferences from non-randomized studies of ECE dosage is *Propensity Score Matching* (PSM) (Rosenbaum & Rubin, 1984). PSM is a method that can greatly reduce selection bias, by matching participants with comparable demographic and other pre-treatment characteristics on their probability to participate in the treatment, independently of their actual participation in treatment. PMS presents an advantage over just controlling for preexisting differences using covariates in the model, by significantly reducing selection bias and providing more robust estimates(Shadish, Cook, & Campbell, 2002). Additionally, this method allows for the use of a large number of covariates, which in a mathematical estimation produce one single score, this is an advantage in comparison to using multiple covariates in a model, which sometimes can result in a failure to converge (Wen et al., 2012).

Propensity score matching has become one of the preferred methods of researchers studying dosage of ECE and child outcomes in non-randomized samples (see Domitrovich et al., 2013; Wen et al., 2012; Xue et al., 2016; Youn, 2016a). The propensity scores are typically estimated using logistic regression, with a dichotomous variable on the left indicating the treatment status (in this case attending two years of ECE) computed as a function of a set of covariates. After the scores are estimated there are different ways to match the samples, however one of the prefer methods in the dosage literature has been nearest-neighbor-matching (Shadish et al., 2002), with replacement.

Nearest neighbor matching is a method were each subject in the treatment condition is matched with the individual in the control group whose score is closest to him (Shadish et al., 2002). Matching with replacement means that subjects in the control group can be matched to more than one subject in the intervention group, as opposed to only one subject, when is done without replacement. Matching with replacement reduces bias and improves the quality of the matching, by ensuring that the matches are the closest in terms of their score, representing a better counterfactual. However matching with replacement can also reduce variance, in cases when the propensity score distributions are too different between the control and intervention group, by reducing the actual sample used to build the counterfactual group (A Smith & E Todd, 2005). Therefore, the choice of using matching with or without replacement is a tradeoff that needs to be considered.

For example, Youn (2016) used propensity score methods with nearest neighbor matching and with replacement, to show the impact of an additional preschool year in academic outcomes. The author used 38 variables in the construction of the propensity score and found that using PSM significantly reduced bias. Similarly, Xue et al., (2016) in a meta-analysis using several preschool data sets, used PSM including a large set of covariates, used NN matching with replacement and showed that the children attending to school for an additional year presented significant higher academic scores in kindergarten. Following these researchers approach, and given that I'm interested in producing the less biased estimate of dosage effects –which will allow for a closer to causal inference- I will use NN matching with replacement. A full description of the model and further justification can be found in the Analytic Plan section under Methods.

2.c. Two versus One year of Head Start and Child Outcomes

Because most of the studies comparing one versus two years of attendance to preschool have used Head Start data, in this section, I examine the literature regarding the associations between attending one or two years and child outcomes, including self-regulation, for children attending these type of programs.

HS is a program from the United States Department of health and Human Services in existence since 1965, which provides early childhood services to promote school readiness for low-income children birth to age-five. HS provides funds and regulations for programs that are run locally and depend on the respective school district Office of Head Start (2012). Historically, evidence regarding its effectiveness has been mixed, with several studies showing positive significant benefits for children who participated in a HS program by kindergarten (Aikens, Klein, Tarullo, & West, 2013; Bloom & Weiland, 2015), and non significant longitudinal effects (Puma et al., 2012). The Head Start Impact Study, which examined longitudinal effects for children attending Head Start, between 2002–2003 or 2003–2004 school years, found no significant effects by third grade between children randomly assigned to Head Start and those who were in the control group (Puma et al., 2012).

Studies examining the difference between attending two versus one year of HS have also shown mixed results. The Impact Study (Puma et al., 2012), for example also found that those who attended for two years (beginning at 3-years of age) showed almost no significant differences in the cohort who started HS at 4-years of age, except for higher scores in parent-reported measures of socioemotional skills and positive approaches to learning by 3rd grade.

More recent studies, however, have found evidence of positive associations between an additional year of HS and child academic (Domitrovich et al., 2013; Xue et al., 2016; Youn, 2016a) and socioemotional outcomes by kindergarten (J. E. Moore et al., 2015; Wen et al., 2012; Youn, 2016b). Youn, M., (2016), for instance, examined the impact of Head Start duration on children's language and mathematics skills, using the Family and Children Experiences Survey (FACES, 2009). As mentioned above, the researcher used propensity score matching in order to provide a more rigorous approach to controlling for selection by accounting for pre-existing differences, and found that children who attended two years of Head Start had significantly higher performance in math and language outcomes by the end preschool, with a slight reduction of the effects sizes by the end of kindergarten. Similarly, Xue et al., (2016) used the same Head Start data and reported that children who were exposed to two years of Head Start had significantly greater language and literacy skills, upon exiting preschool and at the end of Kindergarten, than those who attended for only one year.

Regarding socioemotional outcomes, three recent studies have found favorable associations of attending an additional year of HS and socioemotional and behavioral outcomes. None of these studies, however, have examined specifically the association to self-regulation. Wen and colleagues (2012) for example, using data from Family and Child Experience Survey (FACES 2003), found significant favorable differences for children that attended two years of Head Start in contrast to those who attended only one year, in all teacher-reported measures of social outcomes by Kindergarten, including the Preschool Learning Behavior Scale (PLBS; McDermott, Green, Francis, & Stott, 2000) and items from Personal Maturity Scale (Alexander & Entwisle, 1988) and the Social Skills Rating System (Elliot & Gresham, 1987). In a similar vein, Moore and colleagues (2015) found that children who attended two years to a Head Start program, which used a curriculum that targeted socioemotional development (PATHS; Kusché & Greenberg, 1994), were rated by teachers as having higher levels of emotional knowledge in kindergarten and having made more gains in social competence over the preschool years. Likewise, Youn (2016) using nationally representative data from FACES 2009, showed that children who attended two years of Head Start had significantly greater levels of classroom cooperative behaviors, better approaches to learning and less behavioral problems than those who attended one year, as reported by their teacher in kindergarten.

Positive associations between more time on Head Start and socioemotional outcomes are consistent with prior evidence that has indicated that negative associations between amount of ECE and socioemotional outcomes are non-significant for low-income students, and have stronger negative associations for white non-Hispanic children (Huston et al., 2015). However, it is still unclear why there are some negative associations between amount of ECE and socioemotional outcomes, and the associations between dosage and self-regulation more specifically using direct measures.

In sum, even though there is ample evidence supporting the importance of preschool, and there seems to be a general agreement regarding the benefit of an additional year of HS, there is some mixed evidence regarding the effects of preschool over time and particularly about dosage and socioemotional outcomes. One hypothesis that could explain this inconsistency is that preschool might also need to be of a certain minimum quality during the two years. As evidenced by McCartney and colleagues (2010), quality of preschool moderated the adverse effects of dosage on socioemotional outcomes. Similarly, Currie and Thomas (2000) examined fade-out effects and concluded that the benefits disappear for black children because most of the Head Start black children attend low-quality public schools. But after controlling for school quality, they find significantly positive effects of the Head Start Preschool Program.

The present study aims to examine the role of classroom quality as a moderator of the relation between ECE dosage and the development of self-regulation. To do this, it is important first to discuss how we conceptualize classroom quality and to examine the literature regarding the relation between classroom quality and the development of self-regulation in the context of ECE.

3. Classroom Quality

Quality in ECE is generally studied from three approaches: structural quality, general features of the classroom environment, and teacher-child interactions (Pianta, Downer, & Hamre, 2016). Structural quality includes elements such as, student-teacher ratios, teacher's training, and duration of the day; these are easily measured and regulated by policy. General features of classroom environment are most widely measured in ECE using the Early Childhood Environmental Rating Scale –Revised Edition (ECRS-R), which assesses aspects such a hygiene, space and furnishing, program structure, and interactions in the classroom. Recent research examining these approaches to quality has found that improvements to these general classroom features in ECERS-R do not predict child outcomes after controlling for background characteristics (Sabol & Pianta, 2014). Similarly, Burchinal and colleagues (2016) in a meta-analysis found that this general measure of classroom quality did not predict significant gains in child outcomes when more specific measures of interactions were included. Other structural factors such as teacher-child ratio and teacher training have not shown strong associations to child development. Besides, most preschool programs already adhere to the recommended teacher-child ratio of 1:20 children, and other basic structural features, which are the recommended by the evidence (Pianta et al., 2016).

On the other hand, research that examines quality with a focus on teacher-child interactions in the classroom has shown that teacher-child interactions predict a range of academic and socioemotional child outcomes (Burchinal et al., 2008; Curby, Brock, & Hamre, 2013; Hatfield, Hestenes, Kintner-Duffy, & O'Brien, 2012; Mashburn et al., 2008; Morris et al., 2012; Rimm-Kaufman, Curby, Grimm, Nathanson, & Brock, 2009). This approach to quality is based on the Teach Through Interactions Framework (TTI) (Hamre, Pianta, Mashburn, & Downer, 2007), which describes three domains of interactions that have strong theoretical and empirical support. In this framework, the quality of the interactions between teacher and children is the primary mechanism thorough which classroom experiences promote children development and learning (Downer, Sabol, & Hamre, 2010).

A widely used tool for measuring teacher-child interactions is the Classroom Assessment Scoring System (CLASS: Pianta, LaParo, & Hamre, 2008). Experimentally controlled studies have show that students in classrooms that are characterized by high-quality teacher-child interactions, as measured by the CLASS, learn more than their peers in classrooms characterized by low-quality teacher-child interactions (Araujo, Carneiro, Cruz-Aguayo, & Schady, 2014; Kane, McCaffrey, Miller, & Staiger, 2013). Moreover, since 2011, CLASS is used as a measure of quality to assess Head Start programs around the country (Office of Head Start, 2012).

There is growing consensus in the field that quality of teacher-child interactions is a central feature of classroom quality, which predicts a variety of child outcomes and particularly important for children from low-income backgrounds (Hamre & Pianta, 2001; 2005; Yoshikawa et al., 2013). Evidence from a recent meta-analysis that examined features of quality in Early Childhood Education found that teacher-child interactions measured by the CLASS were a stronger predictor of child outcomes than other measures of general quality features for low-income contexts (Burchinal et al., 2016).

In this study, I use teacher-child interactions quality as a variable to examine how the quality of the classroom experiences relate to self-regulation in preschool and determine if it moderates the relation between dosage and self-regulation.

3.a. Role of Teacher-Child Interactions in Promoting and Inhibiting the Development of Self-Regulatory Skills

Once children start attending preschool and kindergarten there are new adults in their lives that take on an important role in their development. It is well established that teacher-child relationships and interactions are important for child's development and learning (Baker, Grant, & Morlock, 2008; Blair et al., 2016; Hatfield et al., 2012; Johnson, Seidenfeld, Izard, & Kobak, 2012; Pianta, Hamre, & Stuhlman, 2003). Furthermore, there is evidence that high-quality preschool and kindergarten experiences can counteract some of the negative effects of harsh parenting and living in low-income households (Hamre & Pianta, 2005; Raver et al., 2008). For instance, Hamre and Pianta (2005) conducted a longitudinal study with children identified as at-risk by the kindergarten teachers. The authors found that children in classrooms with high-quality emotional and instructional supports, performed as well as their low-risk peers by the end of first grade, on a series of academic and behavioral outcomes. As Obradovic (2016) describes, supportive educational contexts, like those in the classroom with supportive teacher-child interactions, may be especially helpful for children that are physiologically reactive. These environments may support children to regulate their behavior through modulating their arousal and providing them an opportunity to level up in their developmental trajectory.

There are different theoretical frameworks that provide an approach to better understand the relevance of teacher-child interactions and relationships in the development of self-regulation, as well as other developmental outcomes. On one hand, the research in this field draws from attachment theory (Ainsworth, Blehar, Waters, &

Wall, 1978), which provides a framework to understand the importance of early relationships between children and their caregivers, and the different types of experiences that lead to different attachment models (Sabol & Pianta, 2012). Consistent with this theory, having a close and positive relationship with an adult outside the home may be particularly helpful for children who live in higher risk contexts. These positive relationships present an opportunity for the child to reorganize attachment models and thus promote healthy development for this group of children (Downer, Sabol, & Hamre, 2010). Experts have argued that the quality of teacher-child relationships appears to be a protective factor for children in high-risks contexts, if not the single most important protective factor (Burchinal, Peisner-Feinberg, & Pianta, 2002; Hamre & Pianta, 2001; Howes et al., 2008). For instance, Burchinal, et al., (2002) explored the statistical interactions between parental attitudes and teacher-child relationships and found that, for children who had more authoritarian parents, close relationship between teacher and child, as reported by the teacher, lead to significantly more reading gains compared to children who did not have a close relationship with their teacher.

The development systems theory (DST) (Bronfenbrenner & Morris, 1998), also know as ecologically oriented systems theory, also sheds light in the study of classroom experiences. This theory provides a framework to organize the complexity of teacher-child interactions and its study. DST posits that children are embedded in multilevel systems, and these systems reciprocally influence one another at the individual, family, classroom, and community level (Sabol & Pianta, 2012). Teacher-child interactions, for example, are the product of child and teacher level characteristic, which interact bidirectionally (Pianta et al., 2003). To illustrate, Portilla et al., (2014), found empirical evidence that children's low self-regulation at kindergarten predicted greater levels of conflict with teachers at the end of the year and into first grade. While most research has focused on the effect of teacher quality on child's outcomes, DST helps us situate the findings in a larger framework that emphasizes the bi-directionality of relations between systems.

Consistently, research highlights the importance of quality of teacher-child interactions in the classroom for the development of children's self-regulation (Hamre & Pianta, 2005; Hamre, Hatfield, Pianta, & Jamil, 2014; Johnson et al., 2012; Morris et al., 2012; Raver et al., 2011; Rimm-Kaufman et al., 2009). Evidence shows that classrooms characterized by responsive teachers, and warmth interactions that provide clear behavioral expectations, proactive behavior management, predictive classrooms routines, in which children spend most of their time engaged in meaningful activities, and have teachers that promote children engagement, result in better development of self-regulatory skills (Morris et al., 2012; Raver et al., 2011; Raver, Blair, Garrett-Peters, Family Life Project Key Investigators, 2014; Rimm-Kaufman et al., 2009). For instance, a study with 341 preschool children found that the more positively engaged with teachers the children were, the more gains they showed in EF skills (Williford, Vick Whittaker, Vitiello, & Downer, 2013). Similarly, another study found that classroom quality in preschool, measured by teacher-child interactions, predicted adaptive behavior in kindergarten, including self-regulatory skills (Rimm-Kaufman et al., 2009). In other words, children who were enrolled in classrooms, in which teacher-child interactions supported classroom organization, displayed significantly higher levels of cognitive and behavioral self-regulation.

International evidence also supports the critical role of teacher-child interactions for child's development of self-regulatory skills, indicating that these are principles that go beyond cultures and matter for child development in different contexts (LoCasale-Crouch et al., 2016). In Ecuador, for example, a study that randomized 15,000 children to teachers within schools to measure the teacher effect, found that a one standard deviation increase in quality of teacher-child interactions in kindergarten, resulted in as much as 0.6 of a standard deviation on average in a range of academic outcomes and inhibitory control as measured at the end of the year grade (Araujo et al., 2014). Similarly, a study in Chile with 1,868 children from public preschools, found that better quality of teacher-child interactions predicted gains in early literacy skills and inhibitory control (Leyva et al., 2015). Other studies supporting the importance of teacher-child interactions have been done in Finland (Pakarinen et al., 2010), Germany (Suchodoletz, Fäsche, Gunzenhauser, & Hamre, 2014) and China (Hu, Dieker, Yang, & Yang, 2016).

Nonetheless, some studies have found only modest linear relations between quality and child's self-regulation and behavioral outcomes (Hatfield et al., 2016), and in some cases none at all (Burchinal, et al., 2016). For example, a meta-analysis examining the relation between quality of teacher-child interaction and a range of child outcomes found no linear association between quality of teacher-child interactions and children's behavioral and social outcomes in Head Start. This study, however, did not examine cognitive self-regulation as an outcome.

In sum, there is a large body of evidence suggesting that quality of teacher-child interactions is related to children's self-regulatory skills. And as shown by studies that have used an experimental design, higher quality of teacher-child interactions result in better self-regulatory skills in children (Araujo et al., 2014). However, some studies have not found strong associations between teacher-child interactions and self-regulation. And studies examining the quality of teacher-child interactions in Head Start have not found an association between teacher-child interaction quality and behavioral and social outcomes, and only found a significant modest association between teacher-child interaction and language outcomes (Burchinal et al., 2016). Therefore it is important to examine the relation between teacher-child interactions and self-regulation in the context of sefl-regulation.

3.b. Classroom Quality as a moderator of dosage effects for gains in Self-Regulation.

As previously discussed there are some inconsistent results regarding the associations between an additional year of preschool and child outcomes. It is still unclear whether under which conditions an additional year of ECE can have positive effects on the development of socioemotional outcomes, and more specifically for self-regulation. Consistently with the literature, mixed results of the relation of dosage to socioemotional outcomes could be explained by the lack of high-quality teacher-child interactions during these years. According to the evidence only 27% of a child's time in ECE is spent in interactions with an adult on a typical day, and the quality of teacher-child interactions in ECE tends to be very low in the area of instructional support and only moderately positive in terms of emotional support and classroom organization (Pianta et al., 2016).

Recent evidence from a study that examined the relation between dosage of Head Start and child outcomes, however, indicated that quality of classroom experiences did not moderate the “effects” of dosage on child academic outcomes. Namely, children benefited from an additional year of preschool irrespective of quality classroom experiences (Xue et al., 2016). However, this same study found no associations between one additional year of preschool and teacher-reported social outcomes for children attending Head Start. It is not clear if the results would be the same for self-regulation outcomes, including direct measures of inhibitory control. The most novel contribution of the present study is the examination of quality of teacher-child interactions as a moderator between dosage and self-regulation.

The Present Study

The main purpose of the present study is to examine the relation between dosage and quality of preschool experiences and self-regulation, with a special focus on the role of quality of classroom experiences (teacher-child interactions) as a moderator between dosage and child outcomes. I first examined the associations between dosages, conceptualized as one versus two years of attendance to Head Start, and number of hours a week and child’s self-regulation, including absence as covariate in the first model, and later as a moderator of dosage. Then I examined the associations between quality of classroom experiences and children’s self-regulation, and finally, I examined classroom quality as a moderator of the relations between dosage and self-regulation.

Research questions

1.a. Do children that attend Head Start for two years (starting at three years of age) have significant better self-regulation in kindergarten than those who attended one year (starting at four years of age)? Does attendance moderate the relation between number of years in HS and self-regulation?

1.b. Does the number of hours-a-week in Head Start significantly predict self-regulation in kindergarten? Does attendance moderate de relation between number of hours a week in HS and self-regulation?

2. Does the quality of teacher-child interaction in Head Start predict self-regulation skills in Kindergarten? Do specific domains of Teacher-Child Interactions differentially predict gains in self-regulation?

3. Does quality of teacher-child interaction moderate dosage relation to self-regulation?

I hypothesize that: 1. Children that attend two years of Head Start will have significantly better self-regulatory skills by spring of kindergarten, more hours a week is less clear, based on prior evidence about non-academic outcomes. Furthermore, absences should moderate the relation between amount of preschool experiences and self-regulation. Children who have more absences could have fewer benefits from more years in HS. 2. Quality of teacher-child interactions will predict self-regulatory skills in kindergarten, in particular the factor of Positive Management and Routines is expected to be significantly associated with self-regulation in kindergarten, as it has been evidenced in Hamre et al., (2014). 3. It is possible that more years in Head Start and/or more hours-a-week, will have a positive and stronger association to self-regulation in kindergarten in classrooms with higher quality of teacher-child interactions.

Methods

Participants

The present study uses secondary data from the Family and Children Experiences Survey (FACES, 2009), which followed children attending Head Start since 3 years of age and 4 years of age to kindergarten. This sample represents the population of children who entered Head Start at the fall of 2009, excluding children who were already in their second year of Head Start (which is estimated to be approximately 30% of the Head Start population). It includes 60 programs, with two centers per program, and up to three classrooms per center, 486 classrooms, with a total of about 2,383 children across all programs in fall 2009. Faces did not follow children who left the program before Kindergarten. The data includes four waves of data collection—fall and spring of children's first Head Start year, spring of the second Head Start year for children who were 3 years old as of the local school district's kindergarten cut-off date for 2009, and spring of the children's kindergarten year.

Measures

To measure self-regulation used one direct measure together with a set of teacher-reported measures of self-regulatory skills.

Direct Measure of Self-Regulation.

Pencil Tap (Diamond & Taylor, 1996; Smith-Donald et al., 2007) is an adapted version of the peg tap (Diamond & Taylor, 1996). In this test the child is required to do the opposite of what the assessor does, to tap with the pencil one time when the assessor taps twice and viceversa. Pencil Tapping provides an objective assessment of children's cognitive self-regulation, particularly inhibitory control, which has been shown to relate to young children's development in mathematics, vocabulary, and literacy (Blair and Razza 2007; Espy et al. 2004; McClelland et al. 2007) The task also assesses working memory and attention (United States Department of Health and Human Services. Administration for Children and Families. Office of Planning, Research and Evaluation, n.d.) Pencil tap task was not performed with 3years old. There is data available from pencil tap for Fall 2009, Spring 2010, Spring 2011, Spring 2012. The peg-tapping task has an internal reliability (alpha) of 0.82 in preschool and 0.75 in kindergarten (Blair & Razza, 2007).

This variable was transformed into a categorical variable to deal with the negatively skweeness of its distribution in this sample, which was slightly above the limit of tolerance by Curran, West & Finch (1996) (Skeewness= $|2.24|$). The transformed variables included three categories: a low-range (values under 60% of correct responses), a mid-range (values between 60% and 89% of correct responses) and a high-range (values of 90%, or more, correct responses). The transformations resulted in a skeewness value of $|0.31|$, which is well below the acceptable level

Teacher-Reported Measures Self-Regulation.

I used a composite variable of teacher reported measures of child's behavior related to behavioral self-regulation. The composite included Child's approaches to learning scale, and the Problem behavior scale inversed. Items in both scales were inter-correlated at an alpha level of .90. The composite was created by standardizing the scales scores, and averaging across scales.

Child's Approaches to Learning Scale (U.S. Department of Education 2002). The scale assesses child's motivation, attention, organization, persistence, and independence in learning. The scale has established reliability (alpha = 0.89), and has demonstrated relations with academic achievement in elementary school (G. J. Duncan et al., 2007). I will use all the items except the ones pertaining to motivation, a construct, which is less related directly to self-regulation.

Problem Behaviors Scale. This scale comes from an abbreviated adaptation of the Personal Maturity Scale (Entwisle, Alexander, Pallas, & Cadigan, 1987) and from the Behavior Problems Index (BPI, Peterson & Zill, 1986). It measures negative child behaviors associated with learning problems and later grade retention. The Personal Maturity Scale has an internal reliability ranging from 0.74 to 0.85. The internal consistency of the BPI total score ranged from 0.88 to 0.89 in the National Health Interview Survey and the National Longitudinal Study of Youth (Berry, Bridges, & Zaslow, 2004). This scale was inversed prior to being standardized.

Classroom Quality.

The Classroom Assessment Scoring System (CLASS; Pianta et al., 2008) Is a validated classroom observation tool that assesses teacher-child interactions across 10 distinct dimensions (Mashburn et al., 2008). Previous research demonstrates that these dimensions are organized into three broad domains (Hamre et al., 2014): Emotional Support, Classroom and Instructional Support. Each domain includes dimensions, which are scored on a 7-point scale, with 1-2 representing low scores, 3-5 representing moderate scores, and 6-7 representing high scores.

For this study will examine a bifactor approach to interpret the CLASS, which presents a revised conceptualization of the CLASS scores that has been previously validated and shown to better fit the data than the 3 domain factor solution (Hamre et al., 2014). Bifactor scores include three uncorrelated factors that allow for a better interpretation of results, given that there are no issues of multicollinearity in the prediction models (Reise, Moore, & Haviland, 2010). This model includes a general factor of teacher-child interactions (Responsive Teaching) and two domain-specific factors, one related to positive management and routines (PMR), and one related to the cognitive domain (CF). These factors have been associated with growth in multiple developmental domains.

Covariates.

I have classified existing covariates in three categories: parent/family context, child-level, program/teacher level. I examined all existing variables in each of these categories and selected a number of variables that, from a theoretical perspective, are the most relevant. The variables selected relate either to the outcome, or to the dependent variables, either from a theoretical and/or empirical perspective (see table 1). More details about covariate selection are specified below in the analytic plan section about Propensity Score Matching.

Analytic plan

I first describe the procedures for creating the latent factor scores of CLASS using the bifactor approach and examining its fit, in order to use these factors in subsequent analyses. Next I discuss the predictive models estimated using Ordinary Least Square (OLS) and Propensity Score Matching (PSM) techniques to answer each of the research questions.

Creating CLASS orthogonal factors.

Given that the CLASS domains are strongly correlated it is important to examine an alternative analytic approach that allows examining differential aspects of interactions and child development. Using a bifactor approach can allow us to continue to expand our understanding of teacher-child interactions and their specific relation to child outcomes. In the present study I used Mplus 7.4 to fit a bifactor model of the CLASS as specified in Hamre's, et al., (2014) paper (see figure 1.), including none general domain (Responsive Teaching) and two domain-specific factors (Positive management and routines and Cognitive Facilitation). The fit statistics (RMSEA, SRMR, and CFI) were then compared with the three-domain model and a one-domain model, which were estimated using Confirmatory Factor Analysis. The orthogonal factor scores created in the bifactor model were then imported into Stata 14. SE, and examined concurrently in the predictive models for each outcome.

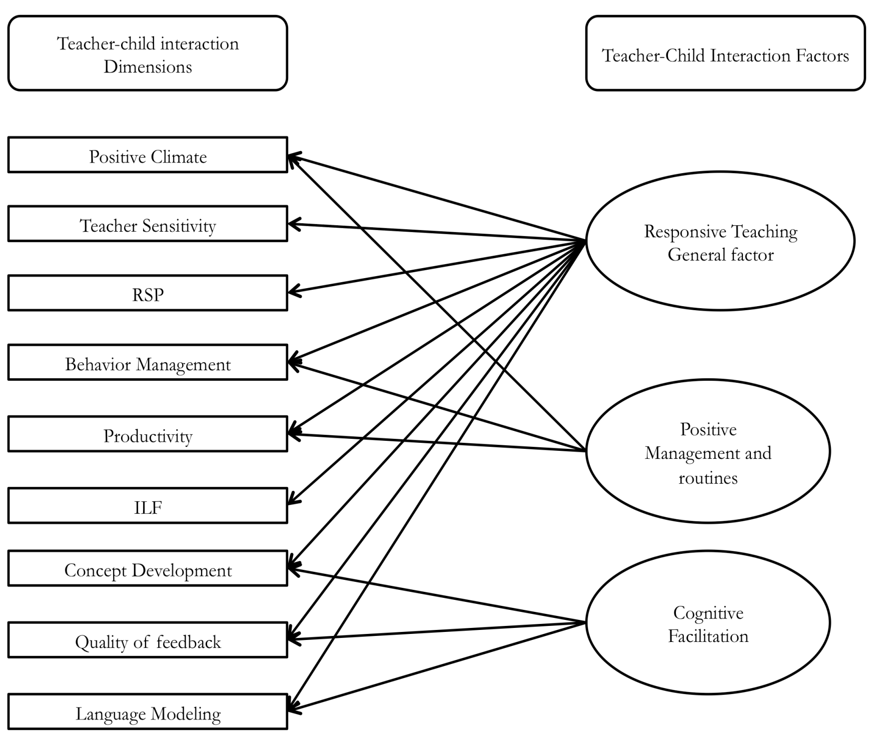
Missingness and nesting.

To address issues of missing data (20% on average) multiple imputation with chained equations was used (White, Royston, & Wood, 2011). I imputed a total of 20 data sets, following the recommendations of White et al., (2011), who argues that the number of imputations should be at least equivalent to the percentage of missing data. I included all the variables of the model in the imputation processes except number of years in Head Start. This variable was not imputed means that only the children who had information about whether they attended one or two year were considered in the analyses.

I accounted for nesting of children in classrooms by using robust standard errors clustered at the classroom level, similarly to other studies using this sample and data set have used (see Ansari, Purtell, & Gershoff, 2016). Like multilevel modeling, clustered robust standard errors correct for the shared variance of observations due to children being in the same classrooms.

For answering all the questions I run two sets of analyses: Ordinary-least-squares (OLS), and OLS with Propensity score matching to reduce selection bias in the comparison between number of years in HS.

Figure 1. Bifactor Model of the CLASS: General and Domain-specific Factors of Teacher-Child Interactions



Ordinary-least-squares (OLS).

Ordinary-least-squares (OLS) regression analyses were conducted, using Stata/SE 14.2, to examine the relation between three forms of dosage (number of years, number of hours a week, and absenteeism), quality of HS experiences and children's self-regulation in kindergarten. Three models were examined. The first model included the main effects of dosage and quality variables and the covariates on self-regulation. Each outcome variable was examined on a separate model, that is: teacher-reported self-regulation, and the direct measure. The main-effects model was specified as shown below:

Self-Regulation= β Constant + β Covariates + β CLASS factors + β N of years in HS + β N of hours a week + β Absenteeism+ Ɛ

In a second model, to examine how absenteeism moderated the relation between the main dosage variables and self-regulation, I added interaction terms between number of years and absenteeism and between numbers of hours a week and absenteeism. The model was estimated as shown:

Self-Regulation= β Constant + β Covariates + β CLASS factors + β N of years in HS + β N of hours a week + β Absenteeism + β N of years in HS*Absenteeism + β N of hours a week*Absenteeism+ Ɛ

A third model included main-effects and interaction terms between the two main dosage variables (number of years, number of hours a week) and the CLASS factors, to examine how quality of classroom experiences moderated the relation between dosage and self-regulation. The model estimated is shown below:

Self-Regulation= β Constant + β Covariates + β CLASS factors + β N of years in HS + β N of hours a week + β Absenteeism + β N of years in HS*CLASS factors + β N of hours a week* CLASS factors + Ɛ.

Propensity Score Matching (PSM).

The sample of children was not randomized to attend one or two years of preschool, and thus there might be important family characteristics that may be driving this decision, which leads to possibly important selection bias while comparing the group of children who attended one year versus the group of children that attended two years. To reduce selection bias I used *Propensity Score Matching* (PSM) (Rosenbaum & Rubin, 1984).

PSM reduces selection bias by adjusting for variables that relate to children and contextual characteristics that could be correlated to the decision of some families to enroll their children in preschool at 3 versus 4 years old. This quasi-experimental technique intend to imitate random assignment by matching cases (in the one condition) with possible counterfactual (in the other condition) as defined by their propensity to select into treatment, in this case attending 2 years of HS. This probability is defined by a set of covariates that are used in the estimation of the *propensity score* (*p)*. The propensity score in this case indicates a child's propensity to attend two years of HS as a function of the included covariates.

A benefit of propensity score matching according to Rosenbaum & Rubin (1984) is that matching on a propensity score is equivalent to matching on the full vector of covariates, which simplifies the process of matching by reducing its dimensionality. Also, PSM is preferable to controlling for covariates in an OLS because it clarifies the region of common support including only those cases that have a valid counterfactual according to the propensity scores, and leaving out from the effect estimation those cases that do not have a match. Also it does not make assumptions about the functional forms of the regression lines, and it does not rely on extrapolating for estimating effects. PSM tries to imitate random assignment, by ignorably assigning subjects to treatment conditions given their value on the propensity score, providing an unbiased estimate of the treatment and control effects at *p*.

One of the limitations of using propensity score matching is that the population of causal inference is limited to the region of common support. However research has shown that with in this region estimates are reliable and comparable to other quasi-experimental methods (Cook et al., 2008), and when covariates are carefully chosen it can greatly reduce biased in comparison to controlling for covariates in a simple OLS.

There are four steps in utilizing PSM: estimating the propensity score, matching individuals, checking for balance, estimating the treatment effect (Lanza, Moore, & Butera, 2013). Even though attending one or two years of HS is not exactly a treatment that includes an intervention, it will be defined as such for the purpose of the PSM analyses.

For the first step I used logistic regression for predicting the propensity of selecting into treatment (T_i) as a function of the chosen set of covariates (*X*) and generating a *p* score for each individual (p_i), where $p_i = Pr(T_i = 1|\mathbf{X})$. I used a set of child family-level covariates, school-level covariates, as well as child school readiness measures at the beginning of Head Start, selected through the mechanisms described below.

When utilizing PSM strategies one of the most important decision researchers must make is the selection of covariates, even more important than the matching strategy (Cook, Steiner, & Pohl, 2009; Steiner, Cook, Shadish, & Clark, 2010). This is because the propensity scores are created based on the covariates selected. If the propensity score matching procedure is to eliminate as much selection bias as possible, it needs to include all relevant preexisting sources of variation between groups, which is done by including covariates that relate to the outcome and the treatment condition, and presumably to unobservable characteristic of individuals.

The literature describes four broad approaches to covariate selection (Steiner, 2012), which are sometime used in combinations: 1) Causal Structural Model, 2) Covariates that work in general, 3) Empirical Tests, 4) Kitchen Sink. I will use a mix of strategy number 1 and 2.

The Causal Structural Model (Pearl, 2009) is a selection method that recommends knowing for every covariate if it is (1) correlated with treatment, (2) correlated with outcome, (3) correlated with neither; and knowing the strength and direction of those correlations. Some of the advantages is that it offers clear guidance as to what to include, and pushes the researchers to think carefully through what threatens bias and the ability to eliminate it.

The second approach is to include covariates that work in general (Steiner, Cook, Shadish, & Clark, 2010). This selection method is based on available theories that in general suggest that certain type of covariates can threaten selection bias, by relating to the outcome or condition. This approach helped me to determine some additional constructs that may not be empirically linked to the outcome or condition but that are generally important in any educational study and that can fall under three general domains: family-level, child-level, and program-level.

For each covariate I have examined how they correlated to the outcome (self-regulation) and the condition (dosage) (see table 1). All covariates were measured at baseline (Fall 2009).

For matching the sample of children in each condition (those who attended one versus those who attended two years of Head Start), I used the nearest-neighbor matching technique, with replacement, using a caliper of 0.1, which allows enough overlap between the two groups, reducing this way the selection bias in over 98% (Stuart, 2010). 1:1 Nearest neighbor matching selects, for each individual i in the treated group, the individual in the control group with the smallest distance from individual i . This technique has been described as one of the most simple and common methods for matching and when using a small caliper optimally reducing selection bias. Furthermore using nearest-neighbor matching with replacement can often reduce bias because individuals in the control group that look similar to those in the treatment group can be used multiple times (Stuart, 2010).

The next step in PSM is checking balance of the matched samples. I did this by using two of the methods recommended by Harder (2010), checking mean differences for all the covariates across each group, and comparing the ratio of the variances of the propensity score in the treated and control groups

Once the samples were matched and adequate balance was achieved, then I added the outcomes variables back to the data set and estimated the effects of two versus one year of HS. I run the same models as in the OLS regressions using the matched samples, that is: one model including the main effects of dosage and quality on self-regulation, a second model adding interactions between dosage and absence, to examine how absenteeism moderated the relation between dosage and self-regulation. A third model examined quality of teacher-child interactions as moderator between dosage and self-regulation.

Results

Bifactor Model of the CLASS

The first step in this study was to examine the extent to which the bifactor model of teacher-child interactions fitted this sample. The model included one general factor (Responsive Teaching) and two domain-specific factors (positive management and Routines, and Cognitive Facilitation), as shown in figure 1. Following the procedures in Hamre at al., (2014) and consistently with Reise's (2010) recommendations, all the factors were constrained to be uncorrelated, and one of the loading in each factor was set to 1, and the error terms were constrained to be uncorrelated. The model showed adequate fit based on examined fit statistics, comparative fit index (CFI= .94), root mean square error of approximation (RMSEA= .13), and standardized root mean square residual (SMSEA= .05).

To compare the bifactor model with the traditional three-domain solution and a one-domain solution, first a three-factor confirmatory analyses were estimated (CFA), then and a one-factor CFA. Results from these models show worst fit than the bifactor solution (see table 2). The orthogonal estimated factor scores of the CLASS were then incorporated in the predictive models, without the problem of collinearity among the factors.

Ordinary-least-Square (OLS) Results

The main purpose of this study was to examine how the amount and quality of preschool experiences in Head Start relate to self-regulation in kindergarten. Three predictive multiple regression models were estimated for each outcome separately, while controlling for a set of family, child and school characteristics, and the age at kindergarten testing. Descriptive statistics of all the variables in the models, including outcomes, can be found in table 3. The table shows the means for each variable in the sample.

The first model examined self-regulation as a function of the amount and quality of HS experiences , while controlling for the set of selected covariates. The model predicted 14% of the variance in the outcome. Results from multiple linear regression indicated that a larger number of years in HS was positively associated with cognitive self-regulation in kindergarten, as measured by pencil tap (R^2= .14, F(38, 350.4)= 2.42 , *p*=.001). Children's pencil tap score increased 0.20 points for the additional year in HS, after controlling for age at the moment of the test and baseline characteristics. Table 6 shows the results for all the models.

In the model examining teacher-reported behavioral self-regulation as a function of amount and quality of HS experiences, multiple linear regression results also indicated that number of years in HS was the only variable significantly associated with self-regulation in kindergarten (R^2= 22, F(40, 405.3)= 9.94, *p*<.001). Children's self-regulation as reported by their teacher increased 0.3 standard deviations for an additional year in HS, after controlling for age in kindergarten an all baseline characteristics.

None of the estimated factor scores of average quality of teacher-child interactions in HS predicted self-regulation in kindergarten. I went back to check the quality of teacher-child interaction in the proximal year of Kindergarten, to rule out the possibility that using an average of quality between two years for one of the cohorts of children, might be obscuring significant results. Results examining the quality of teacher-child interactions in the proximal year showed no significant relation to cognitive self-regulation (β=.047, p=.31), or to behavioral self-regulation in kindergarten (β=-.021, p=64). Using proximal scores was not different from using an average measure of HS teacher-child interaction quality in the prediction of self-regulation in kindergarten.

The second model examined absenteeism as a moderator between dosage variables and self-regulation. The model included all the independent variables, the interaction terms for absenteeism and all the selected covariates, and examined the outcomes separately. Multiple linear regression results indicated only a marginally significant interaction between number of hours a week and number of days absent on Pencil Tap (R^2= .14, F(40, 355.3)= 2.35, p<.10) (see figure 2). No significant interactions were identified.

The last model examined quality of teacher-child interactions in HS as a moderator between amount of HS experiences and self-regulation in kindergarten. The model included the independent variables, the covariates, and the interaction terms for the CLASS factors. Outcomes were examined separately as in previous models. Multiple linear regression results showed a significant interaction between number of hours a week and responsive teaching (R^2= .14, F(44, 364.7)= 2.30, p= .05), indicating that the relation between hours a week and the direct measure of self-regulation is stronger in classrooms with higher quality of responsive teaching (see figure 3). Additionally a significant interaction was found for the teacher-reported measure of behavioral self-regulation between number of years in HS and Cognitive Facilitation ($\beta = .31$, $p < .05$), indicating that children in classrooms with lower quality of Cognitive Facilitation benefited more from two years of HS, than children who attended one year of HS (see figure 4). No other significant interactions were identified.

Propensity Score Matching

I used PSM to address potential selection bias as a function of pre-existing differences between children that attended one year of HS and those who attended two years of HS. The samples of children were matched using nearest neighbor matching and only those who had a comparable match were included in the analysis, this is referred to as the region of common support (Harder, 2010). The region of common support included 1,653 children, excluding 730 children from the whole sample who had no match.

Mean comparison between children that were not in the region of common support and those who were, revealed that children who did not have a match were significantly younger in kindergarten, all of the, attended two years of HS, scored significantly lower in HS Math and Literacy baseline test, and had mothers with significantly lower educational level (see table 4).

To assess the balance of the matched samples, means differences between the samples were examined. As can be seen in table 5, all the covariates used in the PSM were balanced across samples, with no covariate significantly predicting condition (one versus two years of HS). Table 7, in the appendix, shows the variable means differences in the groups prior to matching.

The PSM models replicated the OLS models specifications, examining outcomes separately, and including the full set of covariates in order to adjust for any potential remaining bias from previous measure characteristics.

The Results of the PSM models can be found in table 6. Results were consistent with what was previously found in the OLS analyses regarding the significant relation between number of years in HS and self-regulation in kindergarten for the Pencil Tap measure (R^2= .17, $F(38, 355.3)$= 1.70, p<.01) and teacher-reported measure of self-regulation (R^2= .26, $F(38, 385.4)$= 5,07, p<.01). With an additional year in HS resulting in 0.30 more points in pencil tap, and 0.32 additional standard deviations in the teacher reported measure of self-regulation.

Number of days absent did not significantly moderate number of year in HS or hours a week for none of the two outcomes.

In the models examining classroom quality as a moderator of the relation between dosage and self-regulation, the Cognitive Facilitation factor of teacher-child interactions only marginally moderated the relation between number of years in HS and the teacher-reported measure of self-regulation (p= .06) (see figure 5.). No other significant interactions were found.

It is worth noting that all the models estimated using PSM explained a larger percentage of the variance than the OLS estimates did.

Table 1. Correlations Among Covariates and Dependent Variables

Variable	Responsive teaching	Positive management and	Cognitive facilitation	Number of yea in HS
Family Demographic Information				
Income/poverty ratio	0.01	-0.03	0.01	0.02
Household size	0.05*	0.03	-0.02	-0.01
Parents born in the U.S.?	0.10***	0.01	-0.06**	0.01
Home language/ non-English	0.08***	0.03	-0.10***	-0.05*
Mother's age	0.06**	0.03	-0.02	0.01
Mother's education	0.04	0.04*	0.01	0.06**
Mother's employment status	0.05*	0.06**	-0.01	0.03
Father's employment status	0.07**	0.02	-0.04*	0.00
Parent depression symptoms	0.00	0.01	0.06**	0.01
Family Activities				
Parents read to the child 3 times a week or more	0.06**	-0.01	0.04	-0.03
Number of parent-child activities in the past week	-0.02	-0.01	0.03	0.03
Time child spent watching TV	-0.05*	0.02	0.06**	0.00
Hours of sleep at night	0.07**	0.01	0.01	0.00
Child care before or after HS	0.01	0.02	0.06**	0.00
Child Characteristics				
Child's gender/Male	-0.03	0.00	0.01	0.03
Child's race	0.03	-0.01	-0.10***	-0.02
Child's age in month at Kindergarten assessment	-0.01	0.00	0.03	0.17***
ECLS–B Mathematics T-score at baseline	0	-0.01	0.11***	-0.43***
ECLS–B Letter-Sound Knowledge T-score at baseline	-0.01	-0.02	0.12**	-0.15***
HS Program/Teacher Characteristics				
Type of curriculum	-0.1	0.10***	0.04	-0.02
Teacher-child ratio	-0.21***	0.00	0.03	-0.07**
Number of children in classroom	-0.01***	-0.01	0.08***	-0.17***
Teacher depressive symptoms	0.06**	0.05*	-0.04	0.06**

*<.05 ** <.01 ***<.001

Table 2. Comparison between Bifactor Model of the CLASS, Three-Domain and One-Domain CFA

	Bifactor model			CFA- Three factors			CFA - One factor
Domains	General domain Responsive Teaching β (B)	Positive Management and Routines β (B)	Cognitive Facilitation β (B)	Emotional support β (B)	Classroom organization β (B)	Instructional support β (B)	β (B)
CLASS dimensions							
Positive climate	0.70 (1.00)	0.35 (1.00)		0.74 (1.00)			0.76 (1.00)
Teacher sensitivity	0.92 (1.34)			0.92 (1.27)			0.85 (1.13)
Regard for student perspectives	0.82 (1.19)			0.81 (1.27)			0.76 (1.02)
Behavior management	0.60 (0.92)	0.52 (1.62)			0.72 (1.00)		0.67 (0.96)
Productivity	0.62 (1.05)	0.43 (1.48)			0.74 (1.12)		0.69 (1.09)
Instructional learning formats	0.70 (1.23)				0.74 (1.16)		0.70 (0.85)
Concept development	0.60 (0.75)		0.60 (1.00)			0.83 (1.00)	0.64 (0.87)
Quality of feedback	0.87 (0.77)		0.87 (1.50)			0.93 (1.17)	0.63 (0.89)
Language modeling	0.51 (0.97)		0.51 (1.07)			0.76 (1.10)	0.63 (1.02)
Variances	0.18	0.04	0.14	0.20	0.22	0.27	0.21
Model Fit							
CFI	0.94			0.91			0.73
RMSEA	0.13			0.15			0.24
SRMR	0.05			0.05			0.09

Table 3. Means and Standard Deviations of the Covariates and Variables for the whole sample

Variable	Means and Standard Deviations (n=2,383)
Family-level Covariates	
Family Demographic Information	
Income/poverty ratio	2.8 (1.79)
Household size	4.6 (1.62)
Parents born in the U.S.?	
Both parents born in US	0.62
One parent born outside US	0.09
Both parents born outside US	0.29
Home language/ non-English	0.29
Mother's age	28.7 (5.92)
Mother's education	
Less than high school diploma	0.36
High school diploma	0.34
Some vocational/tech-associate degree	0.24
Bachelor degree or higher	0.06
Mother's employment status	
Working full-time	0.26
Working part-time	0.21
Looking for work	0.22
Not in labor force	0.31
Father's employment status	
Working full-time	0.53
Working part-time	0.18
Looking for work	0.18
Not in labor force	0.11
Parent depression symptoms	1.60 (0.93)
Family Activities	
Parents read to the child 3 times a week or more	0.75
Number of parent-child activities in the past week	11.28 (2.07)
Time child spent watching TV	2.77 (0.86)
Hours of sleep at night	10.43 (0.92)
Child care before or after HS	0.38 (0.49)
Child characteristics	
Child's gender/Male	0.50
Child's race	
White	0.20
Black	0.32
latino	0.40
American Indian or other	0.08
*Child's age in month at Kindergarten assessment	71.2 (3.85)
ECLS–B Mathematics T-score at baseline	49.45 (9.81)
ECLS–B Letter-Sound Knowledge T-score at baseline	44.31 (10.8)
HS Program/Teacher Characteristics	
Type of curriculum	
Creative curriculum	0.56
High scope	0.17
Other (montessori, high reach or Scholastic)	0.12
Locally created	0.15
Teacher-child ratio	8.44 (2.23)
Number of children in classroom	17.12 (2.22)
Teacher depressive symptoms	1.49 (0.76)
Predictor Variables	
Quality of teacher-child interactions in HS	
Responsive teaching	0.0 (0.40)
Positive management and routines	0.00 (0.15)
Cognitive facilitation	-0.00 (0.37)
Amount of Head Start	
Number of hours a week	25.17 (10.14)
Absenteeism	6.19 (5.93)
Child Outcomes	
Pencil Tap	2.27 (0.76)
Teacher-reported Self-regulation	-0.02 (0.95)

Note: Means for categorical variables are proportions. Standard deviations as shown in parenth

Table 4. Means and Standard Deviations for Samples in the Region of Common Support (ROCS) and outside the ROCS.

Variable	Means and Standard Deviations	Region of Common Support	Not in the Region of Common Support	*p*
Family-level Covariates				
Family Demographic Information				
Income/poverty ratio		2.76 (1.77)	2.91 (1.87)	>.250
Household size		4.61 (1.61)	4.52 (1.55)	>.250
Parents born in the U.S.?				
Both parents born in US		0.59	0.66	
One parent born outside US		0.08	0.10	>.250
Both parents born outside US		0.33	0.24	0.02
Home language/ non-English		0.34	0.25	0.02
Mother's age		29.1 (5.94)	28.47 (5.94)	0.14
Mother's education				
Less than high school diploma		0.39	0.28	0.00
High school diploma		0.34	0.39	0.01
Some vocational/tech-associate degree		0.22	0.26	0.02
Bachelor degree or higher		0.05	0.07	0.11
Mother's employment status				
Working full-time		0.27	0.26	>.250
Working part-time		0.21	0.23	>.250
Looking for work		0.21	0.22	>.250
Not in labor force		0.31	0.29	>.250
Father's employment status				
Working full-time		0.54	0.50	>.250
Working part-time		0.16	0.23	0.17
Looking for work		0.19	0.16	>.250
Not in labor force		0.11	0.11	>.250
Parent depression symptoms		1.6 (0.92)	1.62 (0.96)	>.250
Family Activities				
Parents read to the child 3 times a week or more		0.74	0.73	>.250
Number of parent-child activities in the past week		11.21 (2.13)	11.47 (2.03)	0.12
Time child spent watching TV		2.78 (0.84)	2.76 (0.88)	>.250
Hours of sleep at night		10.42 (0.92)	10.46 (0.97)	>.250
Child care before or after HS		0.39	0.39	>.250
Child characteristics				
Child's gender/Male		0.5	0.53	>.250
Child's race				
White		0.19	0.21	>.250
Black		0.3	0.36	>.250
Latino		0.44	0.33	0.13
American Indian or other		0.07	0.10	>.250

*Child's age in month at Kindergarten assessment	72.03 (3.86)	71.32 (3.56)	0.01
ECLS–B Mathematics T-score at baseline	51.71 (9.36)	40.23 (6.99)	0.00
ECLS–B Letter-Sound Knowledge T-score at baseline	45.27 (10.75)	40.26 (10.29)	0.00
HS Program/Teacher Characteristics			
Type of curriculum			
Creative curriculum	0.55	0.58	
High scope	0.15	0.15	>.250
Other (Montessori, high reach or Scholastic)	0.12	0.11	>.250
Locally created	0.18	0.16	>.250
Teacher-child ratio	8.63 (2.30)	8.03 (2.19)	>.250
Number of children in classroom	17.55 (2.0)	15.99 (2.24)	0.00
Teacher depressive symptoms	1.47 (0.74)	1.66 (0.87)	0.00
Predictor Variables			
Quality of teacher-child interactions in HS			
Responsive teaching	-0.00 (0.40)	0.02 (0.42)	>.250
Positive management and routines	0.00 (0.15)	0.00 (0.17)	>.250
Cognitive facilitation	0.00 (0.38)	-0.03 (0.31)	0.18
Amount of Head Start			
Number of hours a week	25.53 (10.07)	24.19 (10.28)	0.07
Absenteeism	6.02 (5.78)	5.80 (5.59)	0.62
Child Outcomes			
Pencil Tap	2.29 (0.75)	2.16 (0.75)	0.03
Teacher-reported Self-regulation	0.02 (0.94)	-0.15 (0.95)	0.016

Note: Means for categorical variables are proportions. Standard deviations as shown in parentheses.

Table 5. Mean Differences for Matched Samples

Variable	One year in HS (n=772)	Two years in HS (n=799)	Differences between cohorts (*p*)
Family-level Covariates			
Family Demographic Information			
Income/poverty ratio	2.8 (1.85)	2.78 (1.75)	>.250
Household size	4.59 (1.60)	4.61 (1.60)	>.250
Parents born in the U.S.?			
Both parents born in US	0.24	0.21	>.250
One parent born outside US	0.09	0.09	>.250
Both parents born outside US	0.33	0.32	>.250
Home language/ non-English	0.34	0.32	>.250
Mother's age	29.21 (6.10)	29.06 (5.93)	>.250
Mother's education			
Less than high school diploma	0.37	0.36	>.250
High school diploma	0.35	0.36	>.250
Some vocational/tech-associate degree	0.22	0.23	>.250
Bachelor degree or higher	0.06	0.06	>.250
Mother's employment status			
Working full-time	0.27	0.26	>.250
Working part-time	0.21	0.21	>.250
Looking for work	0.21	0.21	>.250
Not in labor force	0.32	0.31	>.250
Father's employment status			
Working full-time	0.53	0.53	>.250
Working part-time	0.18	0.18	>.250
Looking for work	0.18	0.18	>.250
Not in labor force	0.11	0.11	>.250
Parent depression symptoms	1.63 (0.93)	1.61 (0.93)	>.250
Family Activities			
Parents read to the child 3 times a week or more	0.73	0.73	>.250
Number of parent-child activities in the past week	11.20 (2.16)	11.26 (2.17)	>.250
Time child spent watching TV	2.77 (0.84)	2.77 (0.84)	>.250
Hours of sleep at night	10.43 (0.89)	10.42 (0.94)	>.250
Child care before or after HS	0.39	0.40	>.250
Child Characteristics			
Child's gender/Male	0.51	0.51	>.250
Child's race			
White	0.19	0.19	>.250
Black	0.31	0.32	>.250
latino	0.43	0.42	>.250
American Indian or other	0.07	0.07	>.250
*Child's age in month at Kindergarten assessment	70.76 (4.03)	72.67 (3.64)	<.001
ECLS–B Mathematics T-score at baseline	47.80 (8.17)	47.91 (7.73)	>.250
ECLS–B Letter-Sound Knowledge T-score at baseline	43.50 (10.29)	43.52 (10.38)	>.250

HS Program/Teacher Characteristics			
Type of curriculum			
Creative curriculum	0.56	0.54	>.250
High scope	0.16	0.15	>.250
Other (Montessori, High Reach or Scholastic)	0.12	0.12	>.250
Locally created	0.16	0.17	>.250
Teacher-child ratio	8.44 (2.28)	8.43 (2.09)	>.250
Number of children in classroom	17.15 (2.24)	17.15 (2.08)	>.250
Teacher depressive symptoms	1.50 (0.74)	1.49 (0.74)	>.250
*Predictor variables:			
Quality of teacher-child interactions in HS			
Responsive teaching	-0.01 (0.41)	0.03 (0.38)	0.104
Positive management and routines	-0.01 (0.16)	-0.00 (0.15)	>.250
Cognitive facilitation	-0.02 (0.36)	-0.02 (0.36)	>.250
Amount of Head Start			
Number of hours a week	23.74 (10.39)	25.87 (9.81)	<.001
Absenteeism	6.06 (5.14)	6.08 (6.56)	>.250
*Child Outcomes			
Pencil Tap	2.08 (0.77)	2.34 (0.74)	<.001
Teacher-reported Self-regulation	-0.26 (0.99)	0.08 (0.92)	<.001

Note: Means for categorical variables are proportions. Standard deviations are shown in parentheses.
*Not included in the matching

Table 6. Results from OLS and PSM Predictive Models

	Pencil Tap						Teac	
	OLS model			PSM model			OLS mod	
Variables	*B*	*SE B*	*p*	*B*	*SE B*	*p*	*B*	*SE B*
Dosage								
Number of year in HS	0.20 **	0.06	0.001	0.23 *	0.06	0.000	0.30 **	0.05
Number of hours a week	0.00	0.00	0.258	0.00	0.00	0.681	0.00	0.00
Days Absent	0.00	0.00	0.837	0.00	0.01	0.659	0.00	0.00
Quality								
Responsive teaching	0.09	0.07	0.165	0.06	0.09	0.496	-0.02	0.05
Positive management and rountines	0.05	0.17	0.753	0.02	0.18	0.922	-0.08	0.16
Cognitive facilitation	0.01	0.07	0.885	0.01	0.09	0.874	-0.07	0.06
Dosage interactions								
Number of year in HS x Days absent	0.01	0.01	0.295	0.01	0.01	0.263	0.01	0.01
Number of hours a week x Days absent	0.00	0.00	0.079	0.00	0.00	0.217	0.00	0.00
Quality Interactions								
N. of years in HS x Responsive teaching	-0.18	0.11	0.094	-0.12	0.15	0.439	0.02	0.12
N. of years in HS x Pos. Mng. and routines	0.08	0.36	0.833	0.07	0.46	0.873	0.15	0.30
N. of years in HS x Cognitive facilitation	-0.02	0.12	0.847	0.00	0.17	0.978	0.31 **	0.11
N. of Hrs. a week x Responsive teaching	0.01 *	0.01	0.042	0.01	0.01	0.154	0.00	0.01
N. of Hrs. A week x Pos. Mng.and routines	-0.02	0.02	0.154	-0.01	0.02	0.554	0.01	0.02
N. of Hrs. A week x Cognitive facilitation	0.00	0.01	0.944	0.00	0.01	0.746	0.00	0.01

**p* < .05. ***p*< .01

Figure 2. Two-way Interaction Between Number of Hours a Week and Absences: Results from OLS Estimations with Pencil Tap as Outcome

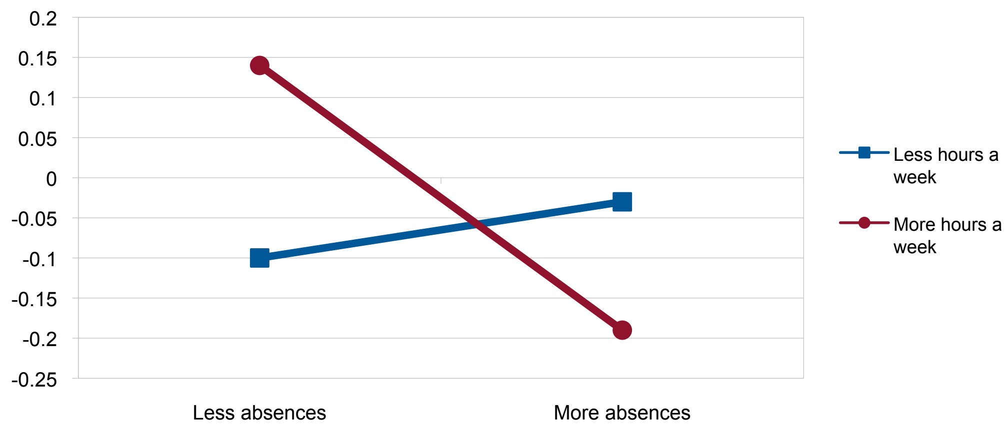


Figure 3. Two-way Interaction Between Number of Hours a week and Responsive Teaching: Results from OLS Estimations with Pencil Tap as Outcome

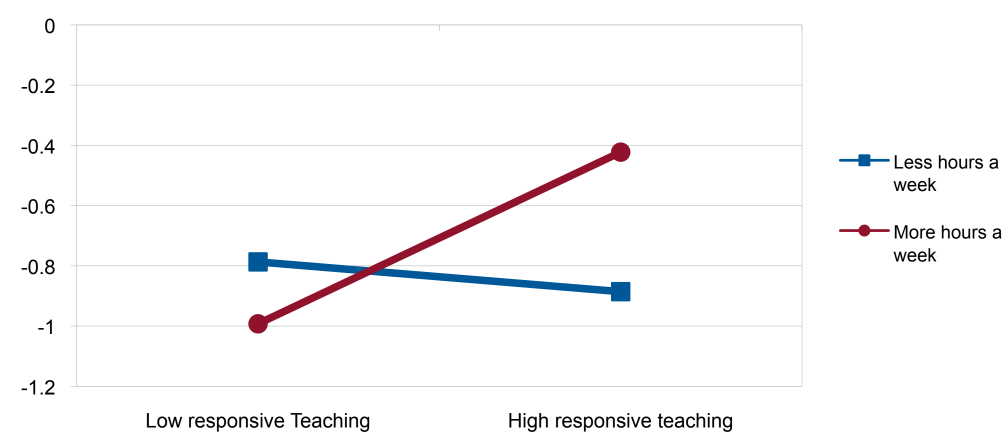


Figure 4. Two-way Interaction Between Number of Years in HS and Cognitive Facilitation: Results from OLS Estimations with Teacher-reported Measure of Self-regulation

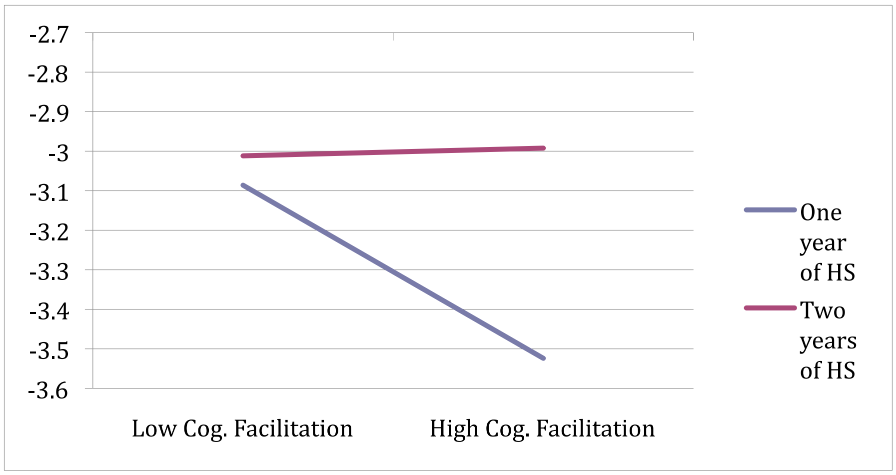
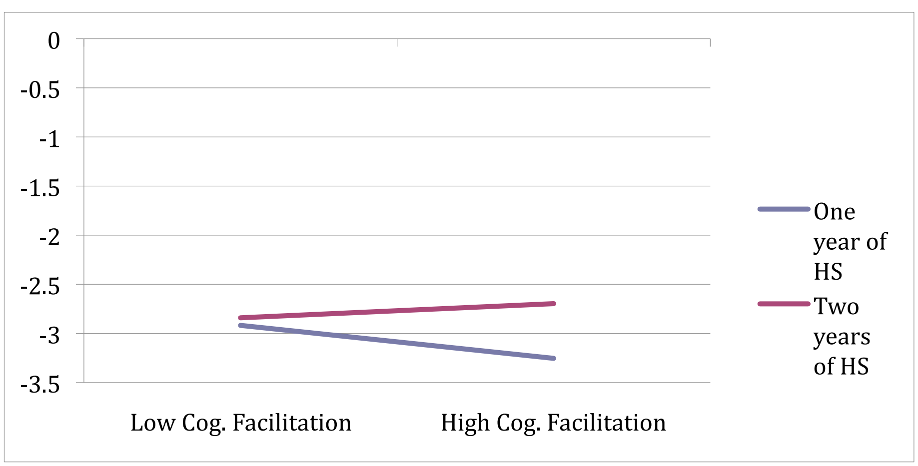


Figure 5. Marginally significant Two-way Interaction Between Number of Years in HS and Cognitive Facilitation: results from PSM Estimations with Teacher-reported Self-regulation as the Outcome



Discussion

The current study examined the associations between amount and quality of Head Start preschool experiences and children's cognitive and behavioral self-regulation in Kindergarten. I first examined the relation between dosage conceptualized as number of years and hours a week and children's self-regulation in kindergarten, and the role of absenteeism as a moderator of dosage. I then examined the relation between the quality of classroom experiences, conceptualized as quality of teacher-child interactions, and its role as a moderator between dosage and children's self-regulation. Because research that examines dosage of ECE experiences, and quality as a moderator, on larger scale, using rigorous methods that reduce selection bias is relatively new (Zaslow, Anderson, & Redd, 2016), and because the evidence regarding its relation to self-regulation is scant, this study can provide a novel contribution to the field.

Results indicate that overall, children benefit from attending one more year of Head Start. Children, who attended two years of HS instead of one, had consistently higher cognitive and behavioral self-regulation in kindergarten, even after controlling for a large set of covariates and their age in kindergarten. Results were consistent across direct and teacher-reported measures of self-regulation, and across OLS and PSM estimations, with PSM estimation resulting in larger effects.

Absenteeism did not significantly moderate the relation between number of years in HS and self-regulation. Marginally significant results only appeared in the context of OLS estimations for absenteeism as moderator between number of hours a week and cognitive self-regulation, but disappeared after the samples were matched using PSM. In other words, the relation between number of years in HS and number of hours in HS did not significantly change as a function of children's parent-reported absences.

Classroom quality in HS did not significantly predict self-regulation in Kindergarten. However, one factor of classroom quality significantly moderated the relation between number of hours a week in HS and cognitive self-regulation. The Responsive Teaching domain-specific factor of teacher-child interactions significantly moderated the relation between number of hours-a-week in HS and self-regulation in Kindergarten. Indicating that children in classrooms that had higher quality of Responsive Teaching benefited significantly more from more hours a week in HS, than children in classrooms characterized by lower quality of Responsive Teaching.

A brief review of the study's main hypotheses and results in relation to the existing literature is presented below. Following this section, limitations will be discussed, and implications for future research addressed.

Hypothesis 1.

It was proposed that children that attend two years of Head Start would have significantly better self-regulatory skills by spring of kindergarten. While more hours a week presented a less clear benefit. Results from this study support the notion of a positive relation between greater number of years in HS and self-regulation, including cognitive and behavioral self-regulatory skills, through direct and teacher reported measures. Furthermore, results were consistently significant across methods of estimation (OLS and PSM), with larger effect sizes in the context of Propensity Score matching.

However, results did not support a significant relation between number of hours a week, and/or absences and self-regulation, independently of the analytical approach (OLS or PSM). This lack of significant results for other forms of dosage could be because these relied in parent-reports. Given that responses rely on the ability of parents to remember number of days absent during a whole semester and accurately report the number of hours that the child attends each week, it is possible that measures are less reliable. For example, parents might report children attending six hours a week, however, parents might be late most mornings or might only bring the child for half a day.

Findings from this study are consistent with prior research that has found somewhat mixed evidence with regard to how the amount of preschool experiences relate to outcomes (see Huston et al., 2015). Number of years in preschool has been previously found to relate to positive academic (Domitrovich et al., 2013; Xue et al., 2016; Youn, 2016a) and behavioral outcomes (J. E. Moore et al., 2015; Wen et al., 2012; Youn, 2016b). However, more hours a week in pre-school have been previously found to relate to classroom behavior problems (Huston et al., 2015; McCartney et al., 2010; NICHD Early Child Care Research Network, 2004). In the present study, nonetheless, there was no evidence of a negative relation between more hours a week and children self-regulation in kindergarten. This finding stresses the importance of continuing to examine the contexts in which more hours a week can be beneficial for child development of self-regulation. I discuss classroom quality as a moderator of the relation between number of hours a week and children's self-regulation in hypothesis 3.

A second part of this hypothesis proposed that absences should moderate the relation between amount of HS preschool experiences and self-regulation. More specifically, it was hypothesized that children who have more absences would have fewer benefits from more years in HS.

Findings from the OLS estimations indicated that more hours a week in HS was marginally significantly more beneficial to children's cognitive self-regulation when students had fewer absences. Children, who attended for more hours a week and had the most absences, seem to benefit the least, while children who attended for larger amount of hours a week and had fewer absences benefited the most. The direction of these findings is consistent with prior evidence that has supported the notion that there are greater benefits from preschool experiences when there is less absenteeism (Arbour et al., 2016; Xue et al., 2016). However, findings were only marginally significant and did not extend to other form of dosage, namely number of years, nor to the outcome of behavioral self-regulation.

In short, results were consistent with previous literature that provides evidence of the benefit of an additional year of pre-school for academic and behavioral outcomes. The present findings contribute to the knowledge base by including cognitive and behavioral self-regulation as an outcome that is positively related to an additional year of preschool. Results were stronger in the context of PSM, which is a more precise method of estimation, given that it eliminates selection bias (Harder et al., 2010). No other forms of dosage were significantly related to self-regulation in this sample, and absenteeism did not significantly moderate the relation between number of years and self-regulation, however it did moderate the relation between number of hours a week in HS and cognitive self-regulation.

Hypothesis 2.

The second main hypothesis proposed was that quality of teacher-child interactions would predict self-regulatory skills in kindergarten. In particular, it was expected that the factor of Positive Management and Routines would be significantly associated with self-regulation in kindergarten, as it has been evidenced in Hamre et.al., (2014). Hamre and colleagues validated a bifactor approach to the CLASS (Pianta, la Paro, & Hamre, 2008), and examined domain-general and domain-specific relations between CLASS orthogonal factors of teacher-child interactions and child outcomes. Their study found that the domain-specific factor related to Positive Management and Routines was significantly associated to executive functions, a set of skills related to self-regulation of behavior and cognition.

The present study was able to replicate the bifactor structure of the CLASS, with a general factor of responsive teaching, and two domain-specific factors of teacher child interactions. The model presented an adequate fit, based on fit statistics, and a better fit than the traditional three-domain structure. This bifactor structure allowed including the domain-general and domain-specific estimated factor scores of teacher-child interactions in a single model to examine possible differential relations between factors and children's self-regulation. This study failed to replicate Hamre's (2014) findings in regard to the relation between CLASS factors and self-regulation. None of the estimated factor scores was significantly related to children's self-regulation. This lack of association between teacher-child interaction quality and children's self-regulation could be due to overall low-quality of teacher-child interactions in the context of HS. These findings are consistent with findings from the meta-analysis by Burchinal and colleagues (2016) including data from FACES 2009, which concluded that there was no association between quality of teacher-child interaction domains and children's behavioral and social outcomes for HS. However, it is also possible that there may be a non-linear relation between quality of teacher-child interactions and self-regulation.

A large body of prior research has supported the relation between teacher-child interactions and children's self-regulation (Hamre et al., 2014; Hamre & Pianta, 2005; Johnson et al., 2012; Morris et al., 2012; Raver et al., 2011; Rimm-Kaufman et al., 2009). Teacher-child interactions that are characterized by responsive teachers, positive and clear behavioral expectations, proactive behavior management and predictive classroom routines, are consistently associated with better child self-regulation skills. Nonetheless, several studies have found that relations between teacher-child interactions and child outcomes are not always linear. In fact, a number of studies, examining teacher-child interactions in relation to children's self-regulation, have found that these have stronger associations at higher levels of quality (Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Hatfield, Burchinal, Pianta, & Sideris, 2016; Weiland, Ulvestad, Sachs, & Yoshikawa, 2012). For example, Weiland et al. (2012) found no significant linear associations between quality of teacher-child interactions and inhibitory control –a key skill in self-regulation-, but significant associations in quadratic associations, with stronger associations at higher levels of quality. Similarly, Hatfield at and colleagues (2016) found evidence of quality teacher-child interactions in the domains of emotional support and classroom organization relating more strongly to self-regulation skills at higher ranges of quality.

In the light of the evidence discussed above, it is possible that the lack of significant findings in the present study is due to the low quality of teacher-child interactions in this sample, and/or a possible non-linear relation between teacher-child interactions and child's self-regulation. Consistently with the evidence, it is reasonable to expect that significant relations between teacher-child interactions and self-regulation be only found at higher levels of teacher-child interaction quality.

Hypothesis 3.

The third hypothesis proposed that quality of teacher-child interactions would moderate the relation between dosage and children's self-regulation. With more years in Head Start and/or more hours-a-week, having a positive and stronger association to cognitive and behavioral self-regulation in kindergarten, in classrooms with higher quality of teacher-child interactions.

Taken together, research which has found stronger association between teacher-child interactions and self-regulation at higher levels of quality (Burchinal et al., 2010; Hatfield et al., 2016; Weiland et al., 2012), and the mixed evidence about the relation between dosage of preschool experiences and child outcomes (see Huston et al., 2015), research suggests that the extent of the benefits of more amount of preschool may be influenced by the quality of the experiences during that time.

Results from this study supported this hypothesis for one form of dosage: number of hours a week. During the first set of analyses using OLS, results revealed a significant interaction between more hours a week and Responsive Teaching factor of classroom quality for cognitive self-regulation. Children in classrooms with higher quality of

Responsive Teaching benefited more from attending more hours a week than those in low-quality classrooms.

It is important to note that because propensity score matching was used to match children on number of years in HS, and it is not suitable to be used with continuous variables, the results from OLS estimations are the pertinent ones to examine for the questions regarding this form of dosage. Furthermore, the sample from PSM as detailed in the results section has significant differences with the children that were not included in the analyses. Children that were outside of the region of common support had mothers with significantly lower educational backgrounds and were significantly younger in Kindergarten. Maternal educational level has been systematically associated to child's health and development (Schultz, 2002), further more, evidence suggests that maternal education can be used as a proxy for socioeconomic level (Bradley & Corwyn, 2002). Therefore it is reasonable to assume that children who were outside of the region of common support had lower socioeconomic backgrounds.

It is possible that the quality of teacher-child interactions is even more important for the group of children with no matches, who came from more disadvantaged contexts. Previous studies have emphasized the importance of quality of preschool experiences for more disadvantaged children, with this group of children benefiting the most of high-quality teacher-child interactions (Hamre & Pianta, 2005). Because the relation between quantity and quality of preschool experiences is largely dependent on the counterfactual experience for a particular child, high-quality interactions in disadvantaged settings are possibly even more important than in less disadvantaged contexts. For example, if a child spend fewer hours a week at a preschool, but every day when he goes back home there is a parent with whom he spends quality time, it presents a different counterfactual experience than attending fewer hours but staying at an aftercare or going back home to a chaotic or stressful family environment.

These findings are consistent with previous research examining the relation between number of hours a week in preschool and child outcomes, which has pointed at the role of quality of experiences as a moderator for children living in poverty. McCartney and colleagues (2010), for example, found that the quality of classroom experiences moderated the adverse effects of more hours of preschool on socioemotional outcomes. In other words, more dosage in classrooms with higher quality did not relate negatively to socioemotional outcomes, while more hours in low-quality classrooms resulted in negative child outcomes.

Children living in poverty face a particular set of challenges that can harm the development of self-regulation. In particular, children's development of self-regulation is highly sensitive to toxic stress exposure, which is characteristic of poverty–living conditions due to economic hardship, food insecurity, unpredictability and other daily hassles (Blair, 2010; Hamoudi, Murray, Sorensen, & Fontaine, 2015). Research has shown that the relation between stress and attention takes the shape of an inverted U, where some middle levels of stress result in enhanced self- regulation and functioning, while very low or very high levels of stress result in poorer self-regulation (Hamoudi et al., 2015). Chronic or prolonged stimulation of stress results in concentrations of stress hormones in the brain, which inhibit the functioning of higher order skills.

There is evidence, however, indicating that high-quality preschool and experiences can counteract some of the negative effects, described above, which result from living in chaotic and stressful environments associated households (Hamre & Pianta, 2005; Raver et al., 2008). For instance, Hamre and Pianta (2005) conducted a longitudinal study with children identified as at-risk by the kindergarten teachers. The authors found that children in classrooms with a high quality of emotional and instructional supports, performed as well as their low-risk peers by the end of first grade, on a series of academic and behavioral outcomes. As Obradovic (2016) also describes, supportive educational contexts, like those in the classroom with supportive teacher-child interactions, may be especially helpful for children that are physiologically reactive.

Findings from the present study, highlight the relevance of high-quality experiences in preschool for children from disadvantaged backgrounds, supporting the benefits of more hours a week in preschools with high-quality of Responsive Teaching. Furthermore, these findings may have important policy implications, suggesting the importance of increasing number of hours jointly with improving the quality of experiences to support the development of self-regulation.

The next question examined classroom quality as a moderator between number of years in HS and self-regulation. Results from OLS estimations showed a significant interaction between number of years in HS and the Cognitive Facilitation factor of teacher-child interactions. Interestingly, this interaction showed a relation in the inverse direction to the hypothesized, with children in classroom with higher quality of Cognitive Facilitation benefiting less from an additional year of HS than children in lower-quality classrooms. However, these results did not remain to be significant at a $p<.05$ level after matching the samples using PSM. The lack of significant results in the matched samples when using PSM, suggests that it is likely that the previously significant interactions in the OLS context were due to pre-existing differences between the children or mere chance. As it can be seen in table 3, the group who attended two years had significantly lower quality of Cognitive Facilitation during preschool than those children who attended one year. When the estimations were computed using the matched sample from PSM, the significant differences in quality of teacher-child interaction between children, who attended one year, versus those who attended two years, were no longer present (see table 4). Even though the children were not matched on any of the predictor variables, quality of teacher-child interactions was balanced across groups as a result of using PSM to match samples on pre-existing family and school characteristics.

The lack of a significant moderation of classroom quality, in the relation between number of years in HS and self-regulation, is consistent with a previous study that examined the relation between dosage and quality for academic outcomes using data from FACES 2009 (Xue et al., 2016). Xue and colleagues (2016) found that classroom quality did not significantly moderate the relation between number of years in HS and academic outcomes. It is possible that there is an accumulative effect of quality that is not being captured when using an average of the years, nor when using the proximal year. Future studies should consider examining the relation between cumulative high-quality teacher-child interactions in preschool experiences and child outcomes.

In sum, finding from this study suggest that children's development of self-regulation benefits from attending an additional year of HS, and that children only benefit from attending more hours a week when in classroom with higher quality of Responsive Teaching. These findings are important because they provide evidence of how amount and quality of preschool experiences relate to self-regulation for children living in disadvantaged contexts. As Blair and Raver (2015) have argued, self-regulation may be a primary mechanism through which poverty affects school success, thus, focusing in ways to promote self-regulation in preschool is of great importance.

Limitations and Directions for Future Research

Limitations of this study are reviewed with regard to each of the hypotheses tested. Regarding the first hypothesis of determining the relation between forms of dosage of HS to children's self-regulation, it is important to note there was no school-reported information about children's attendance and absences. Therefore, the study relied on parent-reported measures of absenteeism and amount of hours a week, which as discussed above may be less reliable. Future research should aim to include school reported measures of preschool attendance and absenteeism.

Another limitation in the study of dosage is the relative scarcity of information about counterfactual experiences. For example, it is difficult to know what the children who started at 3-years of age would have had experienced in terms of child-care instead of HS during that year. Similarly, the FACES 2009 data does not provide information about the type of child-care that the children had the previous years of life. Future studies examining preschool experiences should ideally include information about possible counterfactuals to preschool, or at least information about child-care in the previous years of life.

Regarding the second hypothesis about the relation between teacher-child interaction CLASS orthogonal factors and child's self-regulation. It is important to note that the analyses were performed using estimated factor scores, which do not necessarily behave like true factor scores(see Skrondal & Laake, 2001). Estimated factor scores are the factor scores that Mplus generates when modeling latent variables, and it is a common practice to use factor scores in subsequent analyses. However, analyses using true factor scores in the context of Mplus could possibly yield different results. Future analyses using bifactor CLASS scores should replicate the analyses using true factor scores and compare results to those with estimated factor scores, to rule out possible not significant results due to the nature of the factors.

Future analyses should also examine non-linear relations between CLASS factors and children's self-regulation. As mentioned in the discussion section, several studies have found that relation between teacher-child interaction and children's self-regulation is stronger at higher quality of interactions (Burchinal et al., 2010; Hatfield et al., 2016; Weiland et al., 2012). It is possible that significant associations between teacher-child interactions and children's self-regulation may be present in higher ranges of quality for this sample. Additionally, future studies should examine the cumulative effects of quality of preschool experiences, by examining how do children who are exposed to two years of high-quality preschool experiences, differ from those who had a mixed quality of experiences or two years of low quality.

Overall, the information gathered in this study is not intended to generalize to all the National Head Start population. Even though the study used a large sample of HS children from the FACES 2009 survey, the use of propensity score matching reduced this sample to include in the analyses only children that had a match in the other condition. That is, only children that attended either one or two years of HS and had a match in the other group. Even though the propensity score matching method greatly reduces selection bias, resulting in increased internal validity, it does reduce external validity (Lanza et al., 2013). Furthermore, the study did not include the use of sample weights to recreate a nationally representative sample in the OLS estimations.

Additionally, it is important to keep in mind that the design of the present study is of correlational nature; consequently, it does not warrant causality in the relation between predictor and outcome variables. Because PSM only accounts for observed covariates, factors that may affect the assignment to each condition and the outcome but may not be observed cannot be accounted included in the matching process. Therefore, even though PSM greatly reduces selection bias, it is not equivalent to randomized designs, where unobservable are accounted for in the procedure of randomization. Furthermore, because children were matched on their probability to attend one versus two years of HS, results from PSM are only relevant in the comparison between the group of children who attended one versus the group of children who attended two years, and does not reduce selection bias between children who attended more or less hours a week.

Future studies should continue to examine other measures of self-regulation, including all of its domains and different forms of assessments, such as direct measures, teacher/parent –reported and classroom observations. As Williford and Vick Whittaker suggest (in Cambell et al., 2016), measures of self-regulation within the classroom context can provide researchers with a more accurate representation of child’s self-regulatory skills under more familiar circumstances.

Conclusions

The purpose of the present study was to examine the relation between amount and quality of Head Start experiences and children’s self-regulation in kindergarten. Results supported the benefits of an additional year in Head Start for the development of cognitive and behavioral self-regulation. This study contributes to the growing body of evidence that has already found benefits of an additional year of preschool for academic and other socioemotional outcomes. Furthermore, findings indicate that children benefit from more hours a week in Head Start only when in classrooms with high-quality Responsive teaching, highlighting the importance of investing in quality in tandem with coverage.

Results from this study have important policy implications, by providing rigorous evidence that supports the investment in an additional year of HS, and the importance of investing in improvements of teacher quality for the development of children's self-regulation.

References

A Smith, J., & E Todd, P. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics*, *125*(1–2), 305–353.

Aikens, N., Klein, A. K., Tarullo, L., & West, J. (2013). *Getting Ready for Kindergarten: Children's Progress During Head Start, FACES 2009 Report*. Mathematica Policy Research.

Ainsworth, M., Blehar, M. C., Waters, E., & Wall, S. (1978). Patterns of attachment. Hills-dale. *NJ Eribaum*.

Ansari, A., Purtell, K., & Gershoff, E. (2016). Classroom Age Composition and the School Readiness of 3- and 4-Year-Olds in the Head Start Program. *Psychological Science*, *27*(1), 53–63. http://doi.org/10.1177/0956797615610882

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2014). *A Helping Hand? Teacher Quality and Learning Outcomes in Kindergarten.* Wasington, DC: Teacher Quality and Learning ….

Arbour, M., Yoshikawa, H., Willett, J., Weiland, C., Snow, C., Mendive, S., et al. (2016). Experimental Impacts of a Preschool Intervention in Chile on Children's Language Outcomes: Moderation by Student Absenteeism. *Journal of Research on Educational Effectiveness*, *9*(sup1), 117–149. http://doi.org/10.1080/19345747.2015.1109013

Arteaga, I., Humpage, S., Reynolds, A. J., & Temple, J. A. (2014). One year of preschool or two: Is it important for adult outcomes? *Economics of Education Review*, *40*, 221–

237. http://doi.org/10.1016/j.econedurev.2013.07.009

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, *46*(3), 399–424. http://doi.org/10.1080/00273171.2011.568786

Baker, J. A., Grant, S., & Morlock, L. (2008). The teacher-student relationship as a developmental context for children with internalizing or externalizing behavior problems. *School Psychology Quarterly*, *23*(1), 3–15. http://doi.org/10.1037/1045-3830.23.1.3

Barnett, W. S., Jung, K., Yarosz, D. J., Thomas, J., Hornbeck, A., Stechuk, R., & Burns, S. (2008). Educational effects of the Tools of the Mind curriculum: A randomized trial. *Early Childhood Research Quarterly*, *23*(3), 299–313. http://doi.org/10.1016/j.ecresq.2008.03.001

Becker, D. R., Miao, A., Duncan, R., & McClelland, M. M. (2014). Behavioral self-regulation and executive function both predict visuomotor skills and early academic achievement. *Early Childhood Research Quarterly*, *29*(4), 411–424. http://doi.org/10.1016/j.ecresq.2014.04.014

Berger, A. (2011). Self-regulation: Brain, cognition, and development. American Psychological Association.

Berger, A. (2015). Introduction. In *Self-regulation: Brain, cognition, and development.* (pp. 3–17). Washington: American Psychological Association. http://doi.org/10.1037/12327-001

Berger, A., Kofman, O., Livneh, U., & Henik, A. (2007). Multidisciplinary perspectives on attention and the development of self-regulation. *Progress in Neurobiology*, *82*(5), 256–286. http://doi.org/10.1016/j.pneurobio.2007.06.004

Berry, D. J., Bridges, L. J., & Zaslow, M. J. (2004). Early childhood measures profiles. *Washington, DC: Child Trends*.

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*(4), 327–336. http://doi.org/10.1016/j.lindif.2011.01.007

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. E. (2008). Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI program. *Development and Psychopathology*, *20*(03), 1–23. http://doi.org/10.1017/S0954579408000394

Blair, C., & Dennis, T. (2015). An optimal balance: The integration of emotion and cognition in context. In *Child Development at the Intersection of Emotion and Cognition* (pp. 17–35). Washington: American Psychological Association. http://doi.org/10.1037/12059-002

Blair, C., & Raver, C. (2014). Closing the Achievement Gap through Modification of Neurocognitive and Neuroendocrine Function: Results from a Cluster Randomized Controlled Trial of an Innovative Approach to the Education of Children in Kindergarten. *PloS One*, *9*(11), e112393. http://doi.org/10.1371/journal.pone.0112393

Blair, C., & Raver, C. (2015). School Readiness and Self-Regulation: A Developmental Psychobiological Approach. *Annual Review of Psychology*, *66*(1), 711–731. http://doi.org/10.1146/annurev-psych-010814-015221

Blair, C., & Razza, R. P. (2007). Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten. *Child Development*, *78*(2), 647–663. http://doi.org/10.2307/4139250?ref=search-gateway:12d6d27230f26eb1aa002634d78c0b46

Blair, C., Berry, D. J., & Friedman, A. H. (2012). The development of self-regulation in infancy and early childhood. Guilford Press.

Blair, C., McKinnon, R. D., & Investigatorsbcd, T. F. L. P. (2016). Moderating effects of executive functions and the teacher-child relationship on the development of mathematics ability in kindergarten. *Learning and Instruction*, *41*(c), 85–93. http://doi.org/10.1016/j.learninstruc.2015.10.001

Bloom, H. S., & Weiland, C. (2015). Quantifying variation in Head Start effects on young children's cognitive and socio-emotional skills using data from the National Head Start Impact Study. *Available at SSRN 2594430*.

Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., & Grimm, K. J. (2009). The contributions of 'hot' and 'cool' executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Childhood Research Quarterly*, *24*(3), 337–349. http://doi.org/10.1016/j.ecresq.2009.06.001

Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes.

Burchinal, M. R., Peisner-Feinberg, E., & Pianta, R. (2002). Development of academic skills from preschool through second grade: Family and classroom predictors of developmental trajectories. *Journal of School Psychology*, *40*(5), 415–436. http://doi.org/10.1016/S0022-4405(02)00107-3

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting Child Outcomes at the End of Kindergarten from the Quality of Pre-Kindergarten Teacher–Child Interactions and Instruction. *Applied Developmental Science*, *12*(3), 140–153. http://doi.org/10.1080/10888690802199418

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

Burchinal, M., Xue, Y., Auger, A., & Tien, H. C. (2016). II. QUALITY THRESHOLDS, FEATURES, AND DOSAGE IN EARLY CARE AND EDUCATION: METHODS. *Monographs of the …*. http://doi.org/10.1111/mono.v81.2/issuetoc

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., & Pan, Y. (2014). Early childhood investments substantially boost adult health. *Science*, *343*(6178), 1478–1485. http://doi.org/10.1126/science.1248429

Campbell, S. B., Denham, S. A., Howarth, G. Z., Jones, S. M., Whittaker, J. V., Williford, A. P., et al. (2016). Commentary on the review of measures of early childhood social and emotional development: Conceptualization, critique, and recommendations, 1–23. http://doi.org/10.1016/j.appdev.2016.01.008

Cook, T. D., Shadish, W. R., & Wong, V. C. (2008). Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within‐study comparisons. *Journal of Policy Analysis and Management*, *27*(4), 724–750.

Curby, T. W., Brock, L. L., & Hamre, B. K. (2013). Teachers“ Emotional Support

Consistency Predicts Children”s Achievement Gains and Social Skills. *Early Education & Development*, *24*(3), 292–309. http://doi.org/10.1080/10409289.2012.665760

Currie, J. (2000). School Quality and the Longer-Term Effect of Head Start. *Journal of Human Resources*, *35*(4), 755. http://doi.org/10.2307/146372

Denham, S. A., Kalb, S., Way, E., Warren-Khot, H., Rhoades, B. L., & Bassett, H. H. (2013). Social and emotional information processing in preschoolers: Indicator of early school success? *Early Child Development and Care*, *183*(5), 667–688. http://doi.org/10.1080/03004430.2012.682728

Diamond, A., & Taylor, C. (1996). Development of an Aspect of executive Control: Development of the Abilities to remember What I Said and to “Do as I Say Not as I Do.” *Developmental Psychobiology*, (29), 315–334.

Domitrovich, C. E., Morgan, N. R., Moore, J. E., Cooper, B. R., Shah, H. K., Jacobson, L., & Greenberg, M. T. (2013). One versus two years: Does length of exposure to an enhanced preschool program impact the academic functioning of disadvantaged children in kindergarten? *Early Childhood Research Quarterly*, *28*(4), 704–713. http://doi.org/10.1016/j.ecresq.2013.04.004

Downer, J., Sabol, T. J., & Hamre, B. (2010). Teacher–Child Interactions in the Classroom: Toward a Theory of Within- and Cross-Domain Links to Children's Developmental Outcomes. *Early Education & Development*, *21*(5), 699–723. http://doi.org/10.1080/10409289.2010.497453

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428.

Entwisle, D. R., Alexander, K. L., Pallas, A. M., & Cadigan, D. (1987). The emergent academic self-image of first graders: Its response to social structure. *Child Development*, 1190–1206.

Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, *26*(1), 465–486. http://doi.org/10.1207/s15326942dn2601_6

Farran, D. C., & Lipsey, M. W. (2016). Evidence for the benefits of state prekindergarten programs: Myth & misrepresentation. *Behavioral Science & Policy*, *2*(1), 9–18. http://doi.org/10.1353/bsp.2016.0002

Gennetian, L. A., Magnuson, K., & Morris, P. A. (2008). From statistical associations to causation: What developmentalists can learn from instrumental variables techniques coupled with experimental data. *Developmental Psychology*, *44*(2), 381–394. http://doi.org/10.1037/0012-1649.44.2.381

Hamre, B. K., & Pianta, R. C. (2001). Early Teacher-Child Relationships and the Trajectory of Children's School Outcomes through Eighth Grade. *Child Development*, *72*(2), 625–638. http://doi.org/10.1111/1467-8624.00301

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development*. http://doi.org/10.1111/j.1467-8624.2005.00889.x

Hamre, B. K., Pianta, R. C., Mashburn, A. J., & Downer, J. T. (2007). Building a science of classrooms: Application of the CLASS framework in over 4,000 US early childhood and elementary classrooms. *Downloaded on March*.

Hamre, B., Hatfield, B., Pianta, R., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: associations with preschool children's development. *Child Development*, *85*(3), 1257–1274. http://doi.org/10.1111/cdev.12184

Harder, V. S., Stuart, E. A., & Anthony, J. C. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychological Methods*, *15*(3), 234.

Hatfield, B. E., Burchinal, M. R., Pianta, R. C., & Sideris, J. (2016). Thresholds in the association between quality of teacher–child interactions and preschool children's school readiness skills. *Early Childhood Research Quarterly*, *36*, 561–571. http://doi.org/10.1016/j.ecresq.2015.09.005

Hatfield, B. E., Hestenes, L. L., Kintner-Duffy, V. L., & O'Brien, M. (2012). Classroom Emotional Support predicts differences in preschool children's cortisol and alpha-amylase levels. *Early Childhood Research Quarterly*, *28*(2), 347–356. http://doi.org/10.1016/j.ecresq.2012.08.001

Heckman, J. J., & Raut, L. K. (2016). Intergenerational long-term effects of preschool-structural estimates from a discrete dynamic programming model. *Journal of Econometrics*, *191*(1), 164–175. http://doi.org/10.1016/j.jeconom.2015.10.001

Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-Kindergarten programs. *Early Childhood Research Quarterly*, *23*(1), 27–50. http://doi.org/10.1016/j.ecresq.2007.05.002

Hu, B. Y., Dieker, L., Yang, Y., & Yang, N. (2016). The quality of classroom experiences in Chinese kindergarten classrooms across settings and learning activities: Implications for teacher preparation. *Teaching and Teacher Education*, *57*(C), 39–50. http://doi.org/10.1016/j.tate.2016.03.001

Huston, A. C., Bobbitt, K. C., & Bentley, A. (2015). Time spent in child care: How and why does it affect social development? *Developmental Psychology*, *51*(5), 621–634. http://doi.org/10.1037/a0038951

Jenkins, J. M., Farkas, G., Duncan, G. J., Burchinal, M., & Vandell, D. L. (2016). Head Start at Ages 3 and 4 Versus Head Start Followed by State Pre-K: Which Is More Effective? *Educational Evaluation and Policy Analysis*, *38*(1), 88–112. http://doi.org/10.3102/0162373715587965

Johnson, S. R., Seidenfeld, A. M., Izard, C. E., & Kobak, R. (2012). Can classroom emotional support enhance prosocial development among children with depressed caregivers? *Early Childhood Research Quarterly*, *28*(2), 282–290. http://doi.org/10.1016/j.ecresq.2012.07.003

Jones, S. M., Brown, J. L., & Lawrence Aber, J. (2011). Two-Year Impacts of a Universal School-Based Social-Emotional and Literacy Intervention: An Experiment in Translational Developmental Research. *Child Development*, *82*(2), 533–554. http://doi.org/10.1111/j.1467-8624.2010.01560.x

Jones, S. M., Zaslow, M., Darling-Churchill, K. E., & Halle, T. G. (2016). Assessing early childhood social and emotional development: Key conceptual and measurement issues, 1–7. http://doi.org/10.1016/j.appdev.2016.02.008

Kane, T. J., McCaffrey, D. F., Miller, T., & Staiger, D. O. (2013). Have We Identified

Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment. Research Paper. MET Project. *Bill & Melinda Gates Foundation*.

Kochanska, G., Coy, K. C., & Murray, K. T. (2001). The Development of Self–Regulation in the First Four Years of Life. *Child Development*, *72*(4), 1091–1111. http://doi.org/10.1111/1467-8624.00336

Lanza, S. T., Moore, J. E., & Butera, N. M. (2013). Drawing Causal Inferences Using Propensity Scores: A Practical Guide for Community Psychologists. *American Journal of Community Psychology*, *52*(3-4), 380–392. http://doi.org/10.1007/s10464-013-9604-4

Leyva, D., Weiland, C., Barata, M., Yoshikawa, H., Snow, C., Treviño, E., & Rolla, A. (2015). Teacher-Child Interactions in Chile and Their Associations With Prekindergarten Outcomes. *Child Development*, *86*(3), 781–799. http://doi.org/10.1111/cdev.12342

LoCasale-Crouch, J., Vitiello, G., Hasbrouck, S., Aguayo, Y. C., Schodt, S. C., Hamre, B., et al. (2016). Cómo medir lo que importa en las aulas de primera infancia: un enfoque sobre las interacciones educadora-niño Measuring What Matters in Early Childhood …. *Pensamiento Educativo. Revista De Investigación Educacional Latinoamericana*. http://doi.org/10.7764/PEL.53.1.2016.6

Loeb, S., Bridges, M., Bassok, D., Fuller, B., & Rumberger, R. W. (2007). How much is too much? The influence of preschool centers on children's social and cognitive development. *Economics of Education Review*, *26*(1), 52–66. http://doi.org/10.1016/j.econedurev.2005.11.005

Lonigan, C.J., Spiegel, J.A., Goodrich, J.M. et al. *Journal of Abnormal Child Psychology*

(2017). doi:10.1007/s10802-016-0260-7

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

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., et al. (2008). Measures of Classroom Quality in Prekindergarten and Children's Development of Academic, Language, and Social Skills. *Child Development*, *79*(3), 732–749. http://doi.org/10.1111/j.1467-8624.2008.01154.x

McCartney, K., Burchinal, M., Clarke-Stewart, A., Bub, K. L., Owen, M. T., Belsky, J., NICHD Early Child Care Research Network. (2010). Testing a series of causal propositions relating time in child care to children's externalizing behavior. *Developmental Psychology*, *46*(1), 1–17. http://doi.org/10.1037/a0017886

McClelland, M. M., Acock, A. C., Piccinin, A., Rhea, S. A., & Stallings, M. C. (2012). Relations between preschool attention span-persistence and age 25 educational outcomes. *Early Childhood Research Quarterly*, *28*(2), 314–324. http://doi.org/10.1016/j.ecresq.2012.07.008

Moffitt, T. E., Arseneault, L., Belsky, D., Hancox, R. J., Harrington, H., Houts, R., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety, *108*(7), 1–6. http://doi.org/10.1073/pnas.1010076108/-/DCSupplemental/sapp.pdf

Moore, J. E., Cooper, B. R., Domitrovich, C. E., Morgan, N. R., Cleveland, M. J., Shah, H., et al. (2015). The effects of exposure to an enhanced preschool program on the social-emotional functioning of at-risk children. *Early Childhood Research Quarterly*, *32*, 127–138. http://doi.org/10.1016/j.ecresq.2015.03.004

Morris, P., Lloyd, C. M., Millenky, M., Leacock, N., Raver, C., & Bangser, M. (2012). Using Classroom Management to Improve Preschoolers' Social and Emotional Skills: Final Impact and Implementation Findings from the Foundations of Learning Demonstration in Newark and Chicago. *Mdrc*.

Morrison, F. J., Cameron Ponitz, C., & McClelland, M. M. (2010). Self-regulation and academic achievement in the transition to school. (pp. 203–224). Washington: American Psychological Association. http://doi.org/10.1037/12059-011

Murray, D. W., Rosanbalm, K. D., Christopoulos, C., & Hamoudi, A. (2015). *Self-regulation and toxic stress: Foundations for understanding self-regulation from an applied developmental perspective* (No. OPRE Report 2015-21). Washington D.C.

NICHD Early Child Care Research Network. (2004). Type of child care and children's development at 54 months. *Early Childhood Research Quarterly*, *19*(2), 203–230. http://doi.org/10.1016/j.ecresq.2004.04.002

Obradović, J. (2016). Physiological Responsivity and Executive Functioning: Implications for Adaptation and Resilience in Early Childhood. *Child Development Perspectives*, *10*(1), 65–70. http://doi.org/10.1111/cdep.12164

Pakarinen, E., Kiuru, N., Lerkkanen, M.-K., Poikkeus, A.-M., Siekkinen, M., & Nurmi, J.-E. (2010). Classroom organization and teacher stress predict learning motivation in kindergarten children. *European Journal of Psychology of Education*, *25*(3), 281–300. http://doi.org/10.1007/s10212-010-0025-6

Peterson, J. L., & Zill, N. (1986). Marital disruption, parent-child relationships, and behavior problems in children. *Journal of Marriage and the Family*, 295–307.

Pianta, R. C., Hamre, B., & Stuhlman, M. (2003). Relationships between teachers and children. *Handbook of Psychology*. http://doi.org/10.1002/0471264385.wei0710

Pianta, R. C., la Paro, K. M., & Hamre, B. K. (2008). Classroom Assessment Scoring System (CLASS). Brookes Pub.

Pianta, R., Downer, J., & Hamre, B. (2016). Quality in Early Education Classrooms: Definitions, Gaps, and Systems. *The Future of Children*, *26*(2), 119–138.

Portilla, X. A., Ballard, P. J., Adler, N. E., Boyce, W. T., & Obradović, J. (2014). An Integrative View of School Functioning: Transactions Between Self-Regulation, School Engagement, and Teacher-Child Relationship Quality. *Child Development*, n/a–n/a. http://doi.org/10.1111/cdev.12259

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., et al. (2012). Third Grade Follow-Up to the Head Start Impact Study: Final Report. OPRE Report 2012-45. *Administration for Children & Families*.

Raver, C. (2012). Low-income children's self-regulation in the classroom: scientific inquiry for social change. *The American Psychologist*, *67*(8), 681–689. http://doi.org/10.1037/a0030085

Raver, C., Blair, C., Garrett-Peters, P., Family Life Project Key Investigators. (2014). Poverty, household chaos, and interparental aggression predict children's ability to recognize and modulate negative emotions. *Development and Psychopathology*, *27*(03), 695–708. http://doi.org/10.1017/S0954579414000935

Raver, C., Carter, J. S., McCoy, D. C., Roy, A., Ursache, A., & Friedman, A. (2012). Testing models of children's self-regulation within educational contexts: implications for measurement. *Advances in Child Development and Behavior*, *42*(4), 245–270. http://doi.org/10.1111/j.1467-9507.2011.00608.x

Raver, C., Jones, S. M., Li Grining, C., Zhai, F., Bub, K., & Pressler, E. (2011). CSRP's Impact on Low-Income Preschoolers' Preacademic Skills: Self-Regulation as a Mediating Mechanism. *Child Development*, *82*(1), 362–378. http://doi.org/10.1111/j.1467-8624.2010.01561.x

Raver, C., Jones, S. M., Li-Grining, C. P., Metzger, M., Champion, K. M., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly*, *23*(1), 10–26. http://doi.org/10.1016/j.ecresq.2007.09.001

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor Models and Rotations: Exploring the Extent to Which Multidimensional Data Yield Univocal Scale Scores. *Journal of Personality Assessment*, *92*(6), 544–559. http://doi.org/10.1080/00223891.2010.496477

Reynolds, A. J. (1995). One year of preschool intervention or two: Does it matter? *Early Childhood Research Quarterly*, *10*(1), 1–31. http://doi.org/10.1016/0885-2006(95)90024-1

Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., & Brock, L. L. (2009). The contribution of children's self-regulation and classroom quality to children's adaptive behaviors in the kindergarten classroom. *Developmental Psychology*, *45*(4), 958–972. http://doi.org/10.1037/a0015861

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association*, *79*(387), 516–524. http://doi.org/10.1080/01621459.1984.10478078

Sabol, T. J., & Pianta, R. C. (2012). Recent trends in research on teacher–child relationships. *Attachment & Human Development*, *14*(3), 213–231. http://doi.org/10.1080/14616734.2012.672262

Sabol, T. J., & Pianta, R. C. (2014). Do Standard Measures of Preschool Quality Used in Statewide Policy Predict School Readiness? *Education Finance and Policy*, *9*(2), 116–164. http://doi.org/10.1162/EDFP_a_00127

Schweinhart, L. J. (1993). Significant Benefits: The High/Scope Perry Preschool Study through Age 27. Monographs of the High/Scope Educational Research Foundation, No. Ten. High/Scope Educational Research Foundation, 600 North River Street, Ypsilanti, MI 48198-2898 ($25).

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston, MA, US: Houghton, Mifflin and Company.

Skibbe, L. E., Connor, C. M., Morrison, F. J., & Jewkes, A. M. (2011). Schooling effects on preschoolers' self-regulation, early literacy, and language growth. *Early Childhood Research Quarterly*, *26*(1), 42–49. http://doi.org/10.1016/j.ecresq.2010.05.001

Skrondal, A., & Laake, P. (2001). Regression among factor scores. *Psychometrika*, *66*(4), 563–575. http://doi.org/10.1007/BF02296196

Smith-Donald, R., Raver, C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research. *Early Childhood Research Quarterly*, *22*(2), 173–187. http://doi.org/10.1016/j.ecresq.2007.01.002

Stuart, E. A. (2010). Matching Methods for Causal Inference: A Review and a Look

Forward. *Statistical Science*, *25*(1), 1–21. http://doi.org/10.1214/09-STS313

Suchodoletz, von, A., Fäsche, A., Gunzenhauser, C., & Hamre, B. K. (2014). A typical morning in preschool: Observations of teacher–child interactions in German preschools. *Early Childhood Research Quarterly*, *29*(4), 509–519. http://doi.org/10.1016/j.ecresq.2014.05.010

United States Department of Health and Human Services. Administration for Children and Families. Office of Planning, Research and Evaluation. (n.d.). Head Start Family and Child Experiences Survey (FACES): 2009 Cohort [United States]. http://doi.org/10.3886/ICPSR34558.v1

Ursache, A., Blair, C., & Raver, C. (2012). The Promotion of Self-Regulation as a Means of Enhancing School Readiness and Early Achievement in Children at Risk for School Failure. *Child Development Perspectives*, *6*(2), 122–128. http://doi.org/10.1111/j.1750-8606.2011.00209.x

Weiland, C., Ulvestad, K., Sachs, J., & Yoshikawa, H. (2012). Associations between classroom quality and children's vocabulary and executive function skills in an urban public prekindergarten program. *Early Childhood Research Quarterly*, *28*(2), 199–209. http://doi.org/10.1016/j.ecresq.2012.12.002

Wen, X., Leow, C., Hahs-Vaughn, D. L., Korfmacher, J., & Marcus, S. M. (2012). Are two years better than one year? A propensity score analysis of the impact of Head Start program duration on children's school performance in kindergarten. *Early Childhood Research Quarterly*, *27*(4), 684–694. http://doi.org/10.1016/j.ecresq.2011.07.006

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in Medicine*, *30*(4), 377–399.

Williford, A. P., Vick Whittaker, J. E., Vitiello, V. E., & Downer, J. T. (2013). Children's Engagement within the Preschool Classroom and Their Development of Self-Regulation. *Early Education & Development*, *24*(2), 162–187. http://doi.org/10.1080/10409289.2011.628270

Willoughby, M. T., Kupersmidt, J. B., & Voegler-Lee, M. E. (2012). Is preschool executive function causally related to academic achievement? *Child Neuropsychology*, *18*(1), 79–91. http://doi.org/10.1080/09297049.2011.578572

Xue, Y., Burchinal, M., Auger, A., & Tien, H. C. (2016). IV. TESTING FOR DOSAGE-OUTCOME ASSOCIATIONS IN EARLY CARE AND EDUCATION. *Monographs of the …*. http://doi.org/10.1111/mono.v81.2/issuetoc

Yoshikawa, H., Weiland, C., & Brooks-Gunn, J. (2013). Investing in our future: The evidence base on preschool education. *Ann Arbor*.

Youn, M. (2016a). One Year or Two? The impact of Head Start enrollment duration on academic achievement. *KEDI Journal of Educational Policy*.

Youn, M. (2016b). THE EFFECTS OF HEAD START DURATION ON THE BEHAVIORAL COMPETENCE OF SOCIALLY DISADVANTAGED CHILDREN. *Journal of Community Psychology*, *44*(8), 980–996. http://doi.org/10.1002/jcop.21822

Zanutto, E., Lu, B., & Hornik, R. (2005). Using propensity score subclassification for multiple treatment doses to evaluate a national antidrug media campaign. *Journal of Educational and Behavioral Statistics*, *30*(1), 59–73.

Zaslow, M., Anderson, R., & Redd, Z. (2016). I. QUALITY THRESHOLDS,

FEATURES, AND DOSAGE IN EARLY CARE AND EDUCATION: INTRODUCTION AND LITERATURE REVIEW. *Monographs of the ….* http://doi.org/10.1111/mono.v81.2/issuetoc

Appendix A

Table 7. Mean Differences Between Groups Before Propensity score Matching

Variable	One year in HS (n=1,197)	Two years in HS (n=1,148)	Differences between cohorts (*p*)
Family-level Covariates			
Family Demographic Information			
Income/poverty ratio	2.7 (1.78)	2.81 (1.79)	>.250
Household size	4.6 (1.62)	4.59 (1.61)	>.250
Parents born in the U.S.?			
Both parents born in US	0.58	0.31	<.001
One parent born outside US	0.08	0.09	>.250
Both parents born outside US	0.34	0.3	0.043
Home language/ non-English	0.34	0.3	0.044
Mother's age	29.1 (6.00)	28.9 (5.93)	>.250
Mother's education			
Less than high school diploma	0.41	0.33	<.001
High school diploma	0.33	0.37	0.004
Some vocational/tech-associate degree	0.21	0.24	0.007
Bachelor degree or higher	0.05	0.06	0.054
Mother's employment status			
Working full-time	0.27	0.25	>.250
Working part-time	0.21	0.22	>.250
Looking for work	0.21	0.22	>.250
Not in labor force	0.31	0.31	>.250
Father's employment status			
Working full-time	0.54	0.52	>.250
Working part-time	0.15	0.19	0.098
Looking for work	0.20	0.18	>.250
Not in labor force	0.11	0.11	>.250
Parent depression symptoms	1.60 (0.91)	1.61 (0.93)	>.250
Family Activities			
Parents read to the child 3 times a week or more	0.75	0.73	0.207
Number of parent-child activities in the past week	11.18 (2.11)	11.36 (2.13)	0.168
Time child spent watching TV	2.78 (0.83)	2.77 (0.85)	>.250
Hours of sleep at night	10.41 (0.90)	10.43 (0.95)	>.250

Child care before or after HS	0.39 (0.49)	0.39 (0.49)	>.250
Child characteristics			
Child's gender/Male	0.48	0.52	0.201
Child's race			
White	0.19	0.19	>.250
Black	0.29	0.33	>.250
latino	0.46	0.40	0.174
American Indian or other	0.06	0.08	0.184
*Child's age in month at Kindergarten assessment	71.57 (3.95)	72.31 (3.66)	<.001
ECLS–B Mathematics T-score at baseline	54.38 (9.50)	45.86 (8.27)	<.001
ECLS–B Letter-Sound Knowledge T-score at baseline	46.57 (10.84)	42.68 (10.46)	<.001
HS Program/Teacher Characteristics			
Type of curriculum			
Creative curriculum	0.54	0.56	0.139
High scope	0.15	0.15	>.250
Other (montessori, high reach or Scholastic)	0.12	0.12	>.250
Locally created	0.19	0.17	0.139
Teacher-child ratio	8.78 (2.43)	8.32 (2.12)	<.001
Number of children in classroom	17.83 (1.89)	18.84 (2.18)	<.001
Teacher depressive symptoms	1.45 (0.71)	1.54 (0.80)	0.006
Predictor Variables			
Quality of teacher-child interactions in HS			
Responsive teaching	-0.02 (0.42)	0.03 (0.39)	0.003
Positive management and routines	0.00 (0.16)	0.00 (0.16)	>.250
Cognitive facilitation	0.03 (0.41)	-0.03 (0.32)	<.001
Amount of Head Start			
Number of hours a week	25.29 (10.24)	25.41 (9.97)	>.250
Absenteeism	5.98 (5.16)	6.00 (6.32)	>.250
Child Outcomes			
Pencil Tap	2.26(0.76)	2.29 (0.74)	>.250
Teacher-reported Self-regulation	-0.03 (0.96)	0.02 (0.93)	>.250

Note: Means for categorical variables are proportions. Standard deviations as shown in parentheses.

Memo 1: Outcome variables

Part1: Descriptives

From the original 3,365 sample in Fall 2009, there are a total of 2,383 children who have outcome data in K. From this sample there are 1,220 children who attended 1 year of HS and 1,163 children who attended 2 years.

The outcome measures include:

Penciltapcat: Direct Measure. Categorical variable representing ranges of percentage of time child correctly taps. The original continuous variable was skewed slightly above the recommended limit of tolerance by Curran, West & Finch (1996) (Skeewness= ❘2.24❘).

Therefore I transformed it into a categorical variable with three levels.

The variable includes values for each child at their kindergarten year

Values 1(low)= under 60%, 2(mid)=60%-89%, 3(high)=90%-100%

Reliability .85

Missing 466/2,383

PencilTapcat	Freq.	Percent	Cum.
low	208	18.46	18.46
mid	488	43.30	61.76
high	431	38.24	100.00
Total	1,127	100.00	

The transformations resulted in a skeewness value of $|0.31|$, which is well below the acceptable level.

BProbK: Teacher reported behavior problems—total score (Westat)

Includes values for each child at their Kindergarten year

Values 0-30

Reliability .86

Missing 659/2,383

AppchlK_nm: Teacher reported approaches to learning, excluding motivation item

Includes values for each child at their kindergarten year

Values 1-4

Reliability .93

Missing .: 660/2,383

Inter-item reliability between the scales (Alpha):

Average inter-item covariance: .6298049

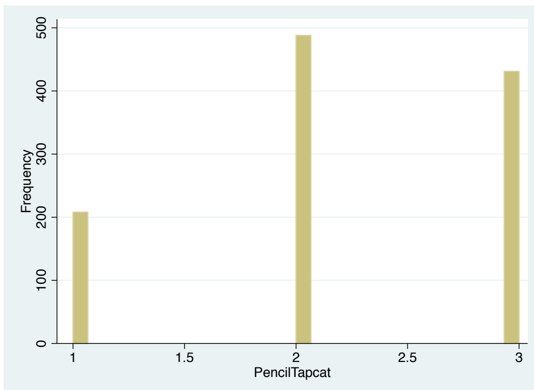
Number of items in the scale: 19

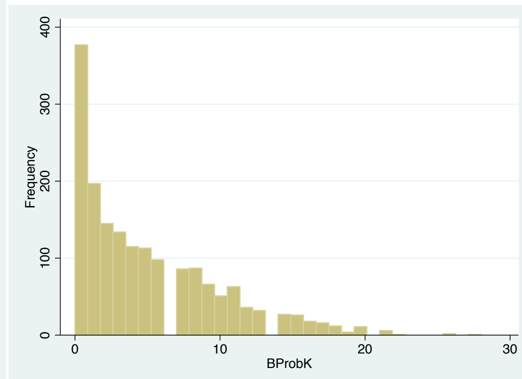
Scale reliability coefficient: 0.8958

Table 1. Summary statistics for continuous outcome variables

Variable	Obs	Mean	Std. Dev.	Min	Max
BProbK	1,724	4.936485	4.934344	0	28
AppchlK_nm	1,723	2.852815	.8032577	.2	4

Distributions





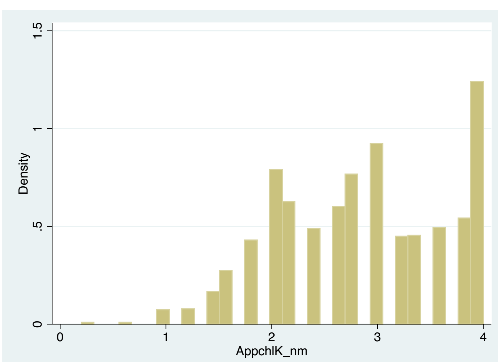


Table 2. Pair-wise correlations among the outcome variables

	Pencil~t	BProbK	Appchl~m
PencilTapcat	1.0000		
BProbK	-0.2329 0.0000	1.0000	
AppchlK_nm	0.2688 0.0000	-0.7319 0.0000	1.0000

Table 3. Outcome variables N and mean by number of years in HS.

Summary statistics: N, mean

by categories of: numyearsinHS

numyearsinHS	Pencil~t	BProbK	Appchl~m	AttenK	ImpConK
1	590 2.172881	832 5.012115	831 2.821661	987 24.13475	987 19.23708
2	537 2.225326	892 4.865942	892 2.881839	934 23.00857	934 18.7666
Total	1127 2.19787	1724 4.936485	1723 2.852815	1921 23.58719	1921 19.00833

Part 2:

Challenges and How I will Address Them

1. **Multiple measures of similar constructs**: Having so many outcomes (5 in this case) presents a challenge for the analyses because it reduces power and it is redundant for the interpretation of results. Additionally, having multiple outcomes increases the probability of Type 1 error.

   This challenge can be solved through data reduction. There are conceptual overlaps and high significant correlations between the two teacher-reported measures r= -.73, p < .01. , and between the two assessor-reported measures r=.89, p < .01.. Therefore, it seems important to reduce the outcomes. One way to reduce is to use a composite score for teacher reported and one for assessor reported. This will result in 3 outcomes rather than 5. The teacher reported and assessor reported measures would provide information about children's behavioral regulation, while the direct measure would provide information about children's cognitive self-regulation skills. The assessor reported measures are more strongly correlated to the direct measure than with the teacher reported, therefore I think they should not be analyzed as part of the same composite.

   Furthermore, because the direct measure provides information about a distinct construct I think it is important to keep the direct measure separate as an outcome. Also, the mean differences for the two cohorts in this particular measure seem to go in the opposite direction than it does in the other measures.

Creating the composite variables:

For creating the composite variable for teacher-reported measures it is important to first reverse code the behavioral problems scale. Then results from both scales need to bee standardized given that they are in different scales (1 to 30 versus 1 to 4), and later averaged.

The same procedure of standardizing variables needs to be done for the assessor reported measures before averaging given that measures are in different scales.

2. **Missing data:**

There is on average about 20% of the data missing for each variable in the target sample. Therefore it will be important to use some method to deal with missingness. A good alternative is to use Multiple Imputation with Chained Equations (MICE). This technique is as effective a Maximum Likelihood estimation (FIML) however it has fewer restrictions regarding the structure of the models. The process of multiple imputation involves the creation of multiple data sets were the data is filled for the missing cases according to observed values for the given individuals and the relations to observed data of other participants. The analyses of multiply imputed data take into account the uncertainty in the imputations and yield accurate standard errors (White, Royston, & Wood, 2011).

Conclusion

The existing 3 outcome variables should be reduced to 2 by using one direct measure of cognitive self-regulation, and a composite measure of teacher reported behavioral self-regulation. This will reduce type 1 error at the same time it will increase the power to find real significant results.

Missing data will be address by using MICE in the analyses.

References

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in Medicine*, *30*(4), 377–399.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*(1), 16-29.

Memo 2: Dosage variables

Variables

numyearsinHS: This a categorical variable that reflects the number of years that the child was in Head Start, for the sample that has some kindergarten data available.

Values: 1= one year in HS, 2= two years in HS.

Table1. Descriptives for Number of years in Head Start

numyearsinHS	Freq.	Percent	Cum.
1	1,220	51.20	51.20
2	1,163	48.80	100.00
Total	2,383	100.00	

HalfFulDay: Child's program type. Parent reported. Categorical variable.

Values: 1: Part-time HS, 2: Full-time HS

Missing: 321/2,107

Table 2. Descriptives for Program Type

HalfFullDay	Freq.	Percent	Cum.
1	1,221	42.59	42.59
2	1,646	57.41	100.00
Total	2,867	100.00	

Table 3. Number of children who attended one and two years of HS by type of program

numyearsin HS	HalfFullDay 1	2	Total
1	452	603	1,055
2	417	590	1,007
Total	869	1,193	2,062

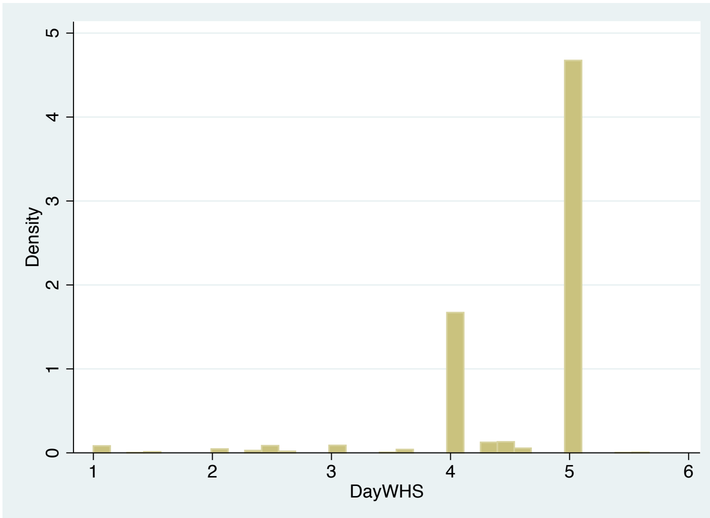
DayWHS: Number of day a week child attended HS as reported by parents.

This measure is an average of the reported number of days child attended HS in fall 2009 and spring 2010 for those children who attended 1 year, and an average across fall 2009, spring 2010 and spring 2011 for children who attended 2 years of HS. Values could range from 1 to 7

Missing= 59/ 2,324

Table 4. Average number of days a week children attended HS

Variable	Obs	Mean	Std. Dev.	Min	Max
DayWHS	2,324	4.587349	.7276891	1	5.666667



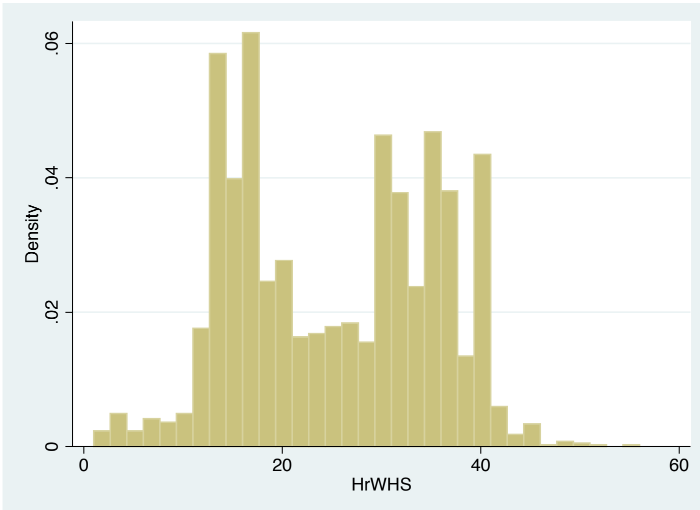
HrWHS: Number of hours a week child attended HS as reported by parents.

This is an average of the number of hours that the child attended HS. For children who attended one year it is an average between parent reported number of hours in fall 2009 and spring 2010. For children who attended 2 years of HS it is an average of number of hours in fall 2009, spring 2010, and spring 2011.

Missing: 65/ 2,318

Table 5. Average number of hours a week children attended HS

Variable	Obs	Mean	Std. Dev.	Min	Max
HrWHS	2,318	25.3301	10.10014	1	56



P2C05: Number of days absent as reported by parents in spring 2010.

Values ranged from 0 to 90 days.

Missing: 357/ 2,026

Table 6. Number of days absent in HS as reported by parents

Variable	Obs	Mean	Std. Dev.	Min	Max
P2C05	2,552	6.175157	5.90101	0	90

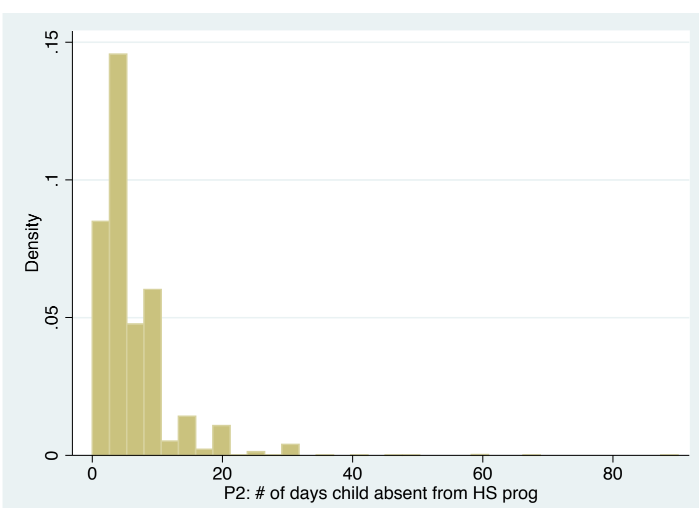


Table 7. Pair-wise correlations between number of years, days a week, hours a week, and days absent.

	numyea~S	HalfFu~y	DayWHS	HrWHS	P2C05
numyearsinHS	1.0000				
HalfFullDay	0.0145	1.0000			
	0.5102				
DayWHS	-0.0080	0.4434	1.0000		
	0.7003	0.0000			
HrWHS	0.0057	0.7183	0.6362	1.0000	
	0.7844	0.0000	0.0000		
P2C05	-0.0032	-0.0472	0.1088	0.0255	1.0000
	0.8843	0.0249	0.0000	0.2520	

Table 8. Number of days a week and number of hours a week per program type.

HalfFullDay	DayWHS	HrWHS
1	4.36613	17.60107
2	4.875803	31.75014
Total	4.661009	25.78433

Table 9. Number of hours a week as predicted by number of days a week and Program type (Hal/Full day)

HrWHS	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
numyearsinHS	.2652598	.2726589	0.97	0.331	-.269503	.8000226
HalfFullDay	10.99981	.3107751	35.39	0.000	10.39029	11.60933
DayWHS	6.221524	.2740115	22.71	0.000	5.684108	6.758939
P2C05	-.0130058	.023654	-0.55	0.583	-.0593981	.0333866
_cons	-20.96282	1.201864	-17.44	0.000	-23.32002	-18.60562

Part 2:

Challenges and How I will Address Them

1. Multiple measures of similar constructs

There are five variables in the data set that could be conceptualized as dosage. However, the variables of number of hours a week and number of days a week and type of program (Half/full day) are strongly correlated (see table 7). Furthermore number of days a week and type of program predict number of hours a week, as shown in the regression (see table 9), which is also evidence of a significant correlation between these variables. Considering the above, I propose choosing one of the three variables as a measure of "amount of preschool during the school year". From a theoretical perspective the number of hours a week is the one that makes most sense, given that is the construct most commonly used in the type of research that explores dosage together with number of years and absences.

The absence variable was not significantly correlated with any of the other dosage variables, with the exception of days a week, however this correlation was very small. Therefore absence should be treated separately.

2. How to conceptualize dosage

During my defense the committee expressed concern about the conceptualization of dosage and encouraged me to expand the construct to include other measures. After examining the available variables related to dosage I see a couple of possible routes:

a. Use number of years as main measure of dosage and examine how hours a week and absence moderate the relation to self-regulation.
b. Examine the three variables (number of years, hours a week, and absences) as predictors of self-regulation. Together they would constitute "dosage" however given the different nature of each construct I would examine them as independent predictors in the same model. I would use them in the same model to account for any shared variance.
c. Examine number of years and hours a week as two independent predictors of self-regulation in the same model, but use absences as a moderator between each of the above variables and the outcome. This is my preferred option.

I prefer the option of examining dosage conceptualized as both, number of years and number of hours a week, and then examine how absences moderate the relation between these variables and the outcome. From a theoretical perspective this approach is aligned with the existing literature about dosage in preschool, which has traditionally examined number of years and number of hours a week as separate constructs (see Xue, Burchinal, Auger, & Tien, 2016).

Regarding absences, some studies examine how they relate to outcomes and the effects of chronic absence on child development; additionally, absence is commonly used as a moderator between the treatment and child outcomes (see Arbour et al., 2016). In the present study I'm interested in examining the relation between the amount of Head Start and self-regulation, expanding my previous conceptualization of dosage to include the number of year the child attended HS, and the number of hours a week as a measure of "amount of preschool within a year". Furthermore, absences could change the relation between these dosage variables and self-regulation, a child could have attended two years and for a certain amount of hours a week, but if he was absent for more than 18 days a year, which is considered chronic absenteeism (Chen & Rice, 2016), prior evidence suggest that child development may be affected. Therefore examining absenteeism as a moderator would be important.

Conclusion

After examining the available variables for the dosage construct, I have decided that I will expand the conceptualization of dosage from one to two years, to include number of hours a week, as a measure of amount of HS experiences within a year. Given that this variable is highly and significantly correlated with number of days a week and type of program (full/half day) I will not include these other variables. Additionally, and based on prior evidence about absenteeism, I will use absenteeism as a moderator of the relation between dosage and self-regulation.

References

Arbour, M., Yoshikawa, H., Willett, J., Weiland, C., Snow, C., Mendive, S., et al. (2016). Experimental Impacts of a Preschool Intervention in Chile on Children's

Language Outcomes: Moderation by Student Absenteeism. *Journal of Research on Educational Effectiveness*, *9*(sup1), 117–149. http://doi.org/10.1080/19345747.2015.1109013

Chen, P., & Rice, C. (2016). Showing up Matters: Newark Chronic Absenteeism in the Early Years. *Advocates for Children of New Jersey*.

Xue, Y., Burchinal, M., Auger, A., & Tien, H. C. (2016). IV. TESTING FOR DOSAGE-OUTCOME ASSOCIATIONS IN EARLY CARE AND EDUCATION. *Monographs of the …*. http://doi.org/10.1111/mono.v81.2/issuetoc

Memo 3: Quality variables

Variables

Classroom Assessment Scoring System (CLASS; Pianta et al., 2008) Assesses teacher-child interactions across 10 distinct dimensions: Emotional Support (CLASS_ES), Classroom Organization (CLASS_CO) and Instructional Support (CLASS_IS). Each domain includes dimensions, which are scored on a 7-point scale, with 1-2 representing low scores, 3-5 representing moderate scores, and 6-7 representing high scores.

Domain-level.

This set of variables was computed by averaging at the domain level variables for each child. For the cohort who attended 2 years of HS the average includes Spring 2010 and Spring 2011, for the cohort who attended one year of HS the variable reflects CLASS scores from Spring 2010.

Missing: All the variables are missing observations in the same proportion 166/2,217

CLASS_ES This variable reflects the average Emotional Support score in during HS for each child.

CLASS_CO This variable reflects the average Classroom organization score in during HS for each child.

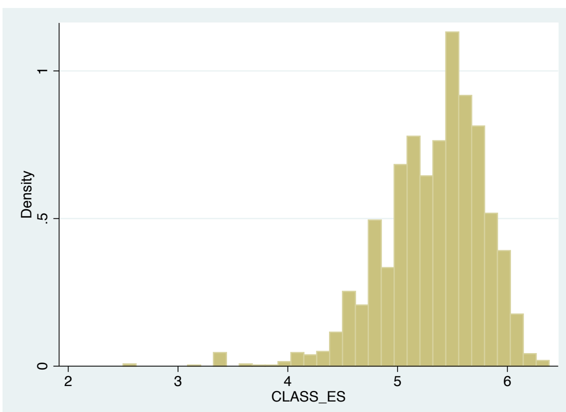
CLASS_IS This variable reflects the average Instructional Support score in during HS for each child.

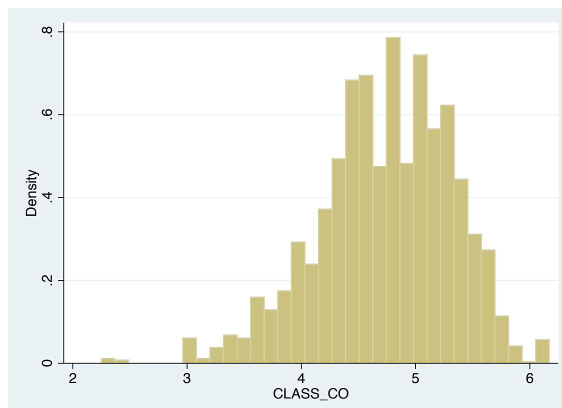
ClassPK is an overall HS CLASS score for each child.

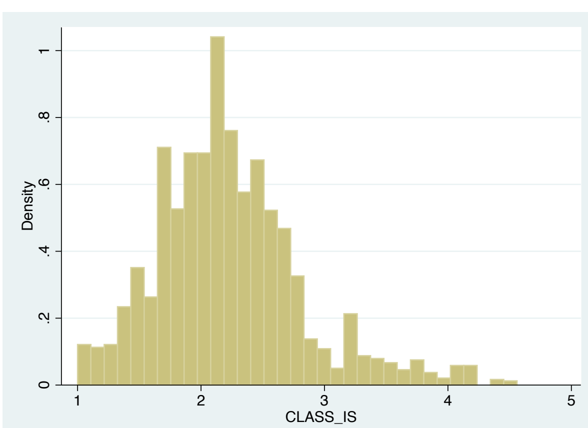
Table 1. Mean, Standard Deviation and Range of CLASS Domain scores in HS.

Variable	Obs	Mean	Std. Dev.	Min	Max
CLASS_ES	2,217	5.325965	.4595117	2.5	6.38
CLASS_CO	2,217	4.737582	.5853358	2.25	6.17
CLASS_IS	2,217	2.225683	.598107	1	4.56
CLASSPK	2,217	4.09641	.4673354	2.053333	5.49

Distributions:







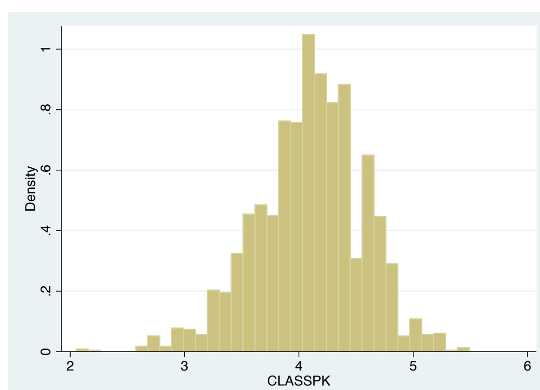


Table 2. Piece-wise correlations between CLASS Domains and overall score.

	CLASS_ES	CLASS_CO	CLASS_IS	CLASSPK
CLASS_ES	1.0000			
CLASS_CO	0.7238	1.0000		
	0.0000			
CLASS_IS	0.5556	0.5132	1.0000	
	0.0000	0.0000		
CLASSPK	0.8670	0.8737	0.8230	1.0000
	0.0000	0.0000	0.0000	

Dimension-Level.

This set of variables was computed in the same way as the domain-level by averaging at the dimension-level each variable for each child. For the cohort who attended 2 years of HS the average includes Spring 2010 and Spring 2011, for the cohort who attended one year of HS the variable reflects CLASS scores from Spring 2010.

Missing: All the variables are missing observations in the same proportion 166/2,217

Variables are: Positive Climate (CLASS_PC), Negative Climate (CLASS_NC), Teacher Sensitivity (CLASS_TS), Regard for Student Perspectives (CLASS_SP), Behavior Management (CLASS_BM), Productivity (CLASS_PR), Instructional Learning Formats (CLASS_ILF), Concept Development (CLASS_CD), Quality of Feedback (CLASS_QF), Language Modeling (CLASS_LM).

Table 3. Mean, Standard Deviation and Range of CLASS Dimension scores in HS.

Variable	Obs	Mean	Std. Dev.	Min	Max
CLASS_PC	2,217	5.32631	.6018093	2.33	7
CLASS_NC	2,217	1.229707	.4287268	1	5.67
CLASS_TS	2,217	4.683978	.6112241	2.67	6.33
CLASS_SP	2,217	4.520171	.6129796	2	6.25
CLASS_BM	2,217	5.115724	.6540138	2.5	7
CLASS_PR	2,217	4.980232	.719717	2	7
CLASS_ILF	2,217	4.116511	.7374791	1.75	6
CLASS_CD	2,217	2.054926	.6216781	1	4.33
CLASS_QF	2,217	2.21463	.6485799	1	5
CLASS_LM	2,217	2.406737	.7438151	1	5

Table 4. Piece-wise Correlation of CLASS Dimensions, Domains, and Overall Score.

	CLASS_PC	CLASS_NC	CLASS_TS	CLASS_SP	CLASS_BM	CLASS_PR	CLASS~LF

CLASS_PC	1.0000						
CLASS_NC	-0.4097	1.0000					
	0.0000						
CLASS_TS	0.6611	-0.4002	1.0000				
	0.0000	0.0000					
CLASS_SP	0.5376	-0.3771	0.7752	1.0000			
	0.0000	0.0000	0.0000				
CLASS_BM	0.5974	-0.3500	0.5469	0.4319	1.0000		
	0.0000	0.0000	0.0000	0.0000			
CLASS_PR	0.5826	-0.3491	0.5307	0.5273	0.5908	1.0000	
	0.0000	0.0000	0.0000	0.0000	0.0000		
CLASS_ILF	0.4626	-0.1367	0.6413	0.5512	0.5037	0.5191	1.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
CLASS_CD	0.4945	-0.1577	0.4319	0.3937	0.3834	0.4392	0.4105
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CLASS_QF	0.5097	-0.2536	0.4522	0.3453	0.3721	0.3539	0.3419
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CLASS_LM	0.3707	-0.3323	0.4992	0.4522	0.3309	0.3709	0.4191
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CLASS_ES	0.8219	-0.6262	0.9007	0.8555	0.6026	0.6242	0.5804
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

CLASS_CO	0.6554	-0.3307	0.6903	0.6084	0.8264	0.8480	0.8204
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CLASS_IS	0.5092	-0.2843	0.5203	0.4490	0.4047	0.4339	0.4398
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CLASSPK	0.7603	-0.4646	0.8053	0.7259	0.7151	0.7437	0.7204
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

	CLASS_CD	CLASS_QF	CLASS_LM	CLASS_ES	CLASS_CO	CLASS_IS	CLASSPK
CLASS_CD	1.0000						
CLASS_QF	0.7762	1.0000					
	0.0000						
CLASS_LM	0.5849	0.7164	1.0000				
	0.0000	0.0000					
CLASS_ES	0.4731	0.4913	0.5159	1.0000			
	0.0000	0.0000	0.0000				
CLASS_CO	0.4950	0.4273	0.4513	0.7238	1.0000		
	0.0000	0.0000	0.0000	0.0000			
CLASS_IS	0.8694	0.9274	0.8765	0.5556	0.5132	1.0000	
	0.0000	0.0000	0.0000	0.0000	0.0000		
CLASSPK	0.7326	0.7351	0.7314	0.8670	0.8737	0.8230	1.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	

Part 2:

Challenges and How I will Address Them

Domain-specific Relations to Outcomes

One of my interests in this study is to examine the domain-specific relations between teacher-child interactions and child's self-regulation. Prior research has shown evidence that certain aspects of these interactions tend to be more closely related to self-regulation outcomes. More specifically, evidence supports that classrooms characterized by high Emotional Supports and Classroom Organization, results in better development of self-regulatory skills (Morris et al., 2012; Raver, Blair, Garrett-Peters, Family Life Project Key Investigators, 2014; Raver et al., 2011; Rimm-Kaufman, Curby, Grimm, Nathanson, & Brock, 2009). Nonetheless, it is difficult from a methodological point of view to attribute the improvements in child development to specific domains of teacher-child interactions with certainty, given that the domains are strongly correlated (as can be seen in tables 2 and 4).

The problem of using correlated factors in a model to examine differential relations between the domains and child outcomes is twofold: on one hand when predictors are highly correlated estimations are less likely to provide significant results, on the other hand when estimation do yield significant results multicollinearity due to the interrelatedness of the domains makes it difficult to interpret the results. One solution is to examine each domain in a separate model, however by doing this is not possible to rule out that other domains of teach-child interactions are also responsible for the changes in the outcome. Alternatively, one could use one overall score (CLASSPK), by averaging across domains, however this approach would not allow to examining domain-specific relations between teacher-child interactions and self-regulation.

A possible solution is to use a bifactor approach (Reise, Moore, & Haviland, 2010). This analytic approach uses dimension-level scores to generate orthogonal latent factors, one general factor, and a number of domain-specific factors that allow for a better interpretation of results, given that there are no issues of multicollinearity in the prediction models. Hamre et al., (2014) have validated this approach to the CLASS including one general factor, and two domain-specific factors.

For the present study I propose the examination of the bifactor analytic approach. If this model fits the data then I would proceed to include all latent factors in the estimation model for each outcome.

If the data does not fit the bifactor structure, then I will use the three traditional domains and an overall score. I would use the three domains in the same estimation model, and the overall score in a separate model. I will do this to contrast the coefficients and its significance for the three domains and the coefficient for the overall domain as predictors.

Conclusion

Given that the CLASS domains are strongly correlated it is important to examine an alternative analytic approach that allow examining differential aspects of interactions and child development. Using a bifactor approach would allow us to continue to expand our understanding of teacher-child interactions and their specific relation to child outcomes. In the absence of this bifactor structure in the data, the more traditional three-domain approach will be examined.

References

Hamre, B., Hatfield, B., Pianta, R., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: associations with preschool children's development. *Child Development*, *85*(3), 1257–1274. http://doi.org/10.1111/cdev.12184

Morris, P., Lloyd, C. M., Millenky, M., Leacock, N., Raver, C., & Bangser, M. (2012). Using Classroom Management to Improve Preschoolers' Social and Emotional Skills: Final Impact and Implementation Findings from the Foundations of Learning Demonstration in Newark and Chicago. *Mdrc*.

Raver, C., Blair, C., Garrett-Peters, P., Family Life Project Key Investigators. (2014). Poverty, household chaos, and interparental aggression predict children's ability to recognize and modulate negative emotions. *Development and Psychopathology*, *27*(03), 695–708. http://doi.org/10.1017/S0954579414000935

Raver, C., Jones, S. M., Li Grining, C., Zhai, F., Bub, K., & Pressler, E. (2011). CSRP’s Impact on Low-Income Preschoolers’ Preacademic Skills: Self-Regulation as a Mediating Mechanism. *Child Development*, *82*(1), 362–378. http://doi.org/10.1111/j.1467-8624.2010.01561.x

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor Models and Rotations: Exploring the Extent to Which Multidimensional Data Yield Univocal Scale Scores. *Journal of Personality Assessment*, *92*(6), 544–559. http://doi.org/10.1080/00223891.2010.496477

Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., & Brock, L. L.

(2009). The contribution of children's self-regulation and classroom quality to children's adaptive behaviors in the kindergarten classroom. *Developmental Psychology*, *45*(4), 958–972. http://doi.org/10.1037/a0015861

Memo 4: Covariates

I have classified existing covariates in three categories: parent/family context, child-level, program/center level. I examined all existing variables in each of these categories and selected a number of variables that from a theoretical perspective are the most relevant. The variables selected relate either to the outcome, or to the treatment condition, either from a theoretical or empirical perspective.

When utilizing PSM strategies one of the most important decision researchers must make is the selection of covariates, even more important than the matching strategy (Cook, Steiner, & Pohl, 2009; Steiner, Cook, Shadish, & Clark, 2010). This is because the propensity scores are created based on the covariates selected. If the propensity score matching procedure is to eliminate as much selection bias as possible, it needs to include all relevant preexisting sources of variation between groups, which is done by including covariates that relate to the outcome and the treatment condition, and presumably to unobservable characteristic of individuals.

The literature describe four broad approaches to covariate selection (Steiner, 2012), which are sometime used in combinations: 1) Causal Structural Model, 2) Covariates that work in general, 3) Empirical Tests, 4) Kitchen Sink. I will use a mix of strategy number 1 and 2.

The Causal Structural Model (Pearl, 2009) is a selection method that recommends knowing for every covariate if it is (1) correlated with treatment, (2) correlated with outcome, (3) correlated with neither; and knowing the strength and direction of those correlations. Some of the advantages is that it offers clear guidance as to what to include, and pushes the researchers to think carefully through what threatens bias and the ability to eliminate it.

The second approach is to include covariates that work in general (Steiner, Cook, Shadish, & Clark, 2010). This selection method is based on available theories that in general suggest that certain type of covariates can threaten selection bias, by relating to the outcome or condition. This approach helped me to determine some additional constructs that may not be empirically linked to the outcome or condition but that are generally important in any educational study and that can fall under three general domains: family-level, child-level, and program-level.

For each covariate I have examined how they correlate to the outcome (self-regulation) and the condition (dosage). Below is a list of the selected variables in each of the three categories and their descriptive information. All covariates described below were measured at baseline (Fall 2009).

1. Family-Level covariates

1.a. Demographic information.

- Poverty ratio: P1POVRTO
- Household size: P1HHSIZE
- Home language: P1RHHLNG
- Mother ed. Level: P1RMOMED
- Mother's age: P1RMAGE
- Mother's race: MRACE
- Mother's employment status: P1MOMEMP
- Father's employment status: P1DADEMP
- Whether parents were born in the U.S: P1PBRNUS
- Whether child was born in the U.S: P1CBRNUS
- Parent Depression symptoms: P1DEPCAT

1.b. Family Activities.

- Parents read to child 3+ times a week P1READS
- Number of parent-child activities in past week P1PWKAC2
- Time spent watching TV P1TIMETV
- Hours of sleep at night P1SLPTM
- Any child care before of after HS P1SLPTM

2. Child-Level Covariates

- Child race CRACE
- Child gender CHGENDER
- Child age in months P1RCAGE

- ECLS–B mathematics T- score with par+ weight A1ECP1WT
- ECLS–B Letter- Sound Knowledge t-score A1ELP1WT

*T-scores in FACES illustrated a child's performance relative to the population of first-time Head Start children as a whole, with a mean of 50 and a standard deviation of 10.

Note: I will also include the variable "child's age at time of assessment" (AnCAGE) as a covariate in the analysis however NOT as in the PSM because it is not a pre-existing characteristic.

3. Program-Level covariates

- Type of Curriculum TyCurr
- Teacher depressive symptoms (categories) T1DEPCAT

Descriptive Statistics of Covariates

Family-Level.

Poverty ratio **P1POVRTO:**

P1: Ratio of Income to Poverty	Freq.	Percent	Cum.
1= Less than 50% of Poverty Threshold	684	21.92	21.92
2= Between 50% and 100% of Poverty Thre	1,261	40.42	62.34
3= Between 101% and 130% of Poverty Thr	490	15.71	78.04
4= Between 131% and 185% of Poverty Thr	407	13.04	91.09
5= Between 186% and 200% of poverty thr	49	1.57	92.66
6= Above 200% of the Poverty Threshold	229	7.34	100.00
Total	3,120	100.00	

P1HHSIZE : Household size

Variable	Obs	Mean	Std. Dev.	Min	Max
P1HHSIZE	3,120	4.616346	1.645748	2	14

P1RHHLNG: Home language

P1: Primary Language Spoken to Child	Freq.	Percent	Cum.
0=English	2,272	70.41	70.41
1=Non-English	955	29.59	100.00
Total	3,227	100.00	

P1RMOMED: Mother educational level

P1: Mother Highest Education	Freq.	Percent	Cum.
-9/.M=Not Ascertained	225	7.00	7.00
1=Less than HS Diploma	1,088	33.84	40.84
2=HS Diploma or GED	1,023	31.82	72.66
3=Voc/Tech-Assoc-Some College Degree	705	21.93	94.59
4=Bachelor Degree or Higher	174	5.41	100.00
Total	3,215	100.00	

P1RMAGE: Mother's age

Variable	Obs	Mean	Std. Dev.	Min	Max
P1RMAGE	3,219	27.97981	7.991988	-9	50

MRACE: Mother's race

P1: Mother's Race/Ethnicity	Freq.	Percent	Cum.
-9/.M=Not Ascertained	2	0.06	0.06
1=White, Non-Hispanic	791	24.52	24.58
2=African American, Non-Hispanic	1,019	31.59	56.17
3=Hispanic/Latino	1,209	37.48	93.65
4=American Indian or Alaska Native	24	0.74	94.39
5=Asian or Pacific Islander	62	1.92	96.31
6=Multi-Racial/Bi-Racial,Non-Hispanic	94	2.91	99.23
7=Other Race	25	0.77	100.00
Total	3,226	100.00	

P1MOMEMP: Mother's employment status

P1:Mother Employment Status	Freq.	Percent	Cum.
-9/.M=Not Ascertained	220	7.05	7.05
1=Working Full Time	752	24.10	31.15
2=Working Part Time	609	19.52	50.67
3=Looking for Work	630	20.19	70.87
4=Not in Labor Force	909	29.13	100.00
Total	3,120	100.00	

P1DADEMP: Dad's employment status

P1:Father Employment Status	Freq.	Percent	Cum.

	Freq.	Percent	Cum.
-9/.M=Not Ascertained	1,719	55.11	55.11
1=Working Full Time	811	26.00	81.12
2=Working Part Time	211	6.76	87.88
3=Looking for Work	225	7.21	95.09
4=Not in Labor Force	153	4.91	100.00
Total	3,119	100.00	

P1PBRNUS: Whether parents were born in the U.S

P1:Whether both parents were botn in U.S.	Freq.	Percent	Cum.
-9/.M=Not Ascertained	78	2.50	2.50
1= Both Parents Born in US	1,887	60.50	63.00
2= One Parent Born Outside US	272	8.72	71.72
3= Both parents Born Outside US	882	28.28	100.00
Total	3,119	100.00	

P1CBRNUS: Whether child was born in the U.S

P1:Whether child was born in the U.S.	Freq.	Percent	Cum.
-9/.M=Missing	8	0.26	0.26
0=No	63	2.02	2.28
1=Yes	3,049	97.72	100.00
Total	3,120	100.00	

P1DEPCAT: Parent depression

P1:Parent Depress Score-CES-D Short Form Categories	Freq.	Percent	Cum.
1=not depressed	1,917	62.16	62.16
2=mildly depressed	644	20.88	83.04
3=moderately depressed	308	9.99	93.03
4=severely depressed	215	6.97	100.00
Total	3,084	100.00	

P1READS: Parents read to child 3 or more times a week

P1:Read to child 3+ times in past week	Freq.	Percent	Cum.
-9/.M=Not Ascertained	1	0.03	0.03
0=No	767	24.58	24.62
1=Yes	2,352	75.38	100.00
Total	3,120	100.00	

P1PWKAC2: Number of parent-child activities in past week

Variable	Obs	Mean	Std. Dev.	Min	Max
P1PWKAC2	3,120	11.2888	2.06532	2	14.3

P1TIMETV: Time spent watching TV

P1:Time watching TV	Freq.	Percent	Cum.
-9/.M=Missing	20	0.64	0.64
0	271	8.69	9.33
.5	761	24.39	33.72
1.5	1,482	47.50	81.22
2.5	586	18.78	100.00
Total	3,120	100.00	

P1SLPTM: Hours of sleep at night

Variable	Obs	Mean	Std. Dev.	Min	Max
P1SLPTM	3,120	10.37489	1.388182	-9	15

P1ANYCCR: Any child care before of after HS

P1:Any child care before or after Head Start	Freq.	Percent	Cum.
0=No	1,933	62.19	62.19
1=Yes	1,175	37.81	100.00

Total	3,108	100.00

Child-Level.

CRACE: Child race

Child Race/Ethnicity	Freq.	Percent	Cum.
-9/.M=Not Ascertained	3	0.09	0.09
1=White, Non-Hispanic	664	20.60	20.69
2=African American, Non-Hispanic	1,025	31.79	52.48
3=Hispanic/Latino	1,275	39.55	92.03
4=American Indian or Alaska Native	20	0.62	92.65
5=Asian or Pacific Islander	56	1.74	94.39
6=Multi-Racial/Bi-Racial,Non-Hispanic	175	5.43	99.81
7=Other Race	6	0.19	100.00
Total	3,224	100.00	

CHGENDER: Child gender

Child Gender	Freq.	Percent	Cum.
0=Female	1,608	49.83	49.83
1=Male	1,619	50.17	100.00
Total	3,227	100.00	

P1RCAGE: Child age in months

Variable	Obs	Mean	Std. Dev.	Min	Max
P1RCAGE	3,226	45.77154	6.537207	32	60

Mean child age by number of years in HS

numyearsinHS	mean
1	52.18205
2	41.27273
Total	46.84054

A1ECP1WT: ECLS–B mathematics T- score with par+ weight

Variable	Obs	Mean	Std. Dev.	Min	Max
A1ECP1WT	3,055	49.75736	9.820329	27.42	90.19

A1ELP1WT: ECLS–B Letter- Sound Knowledge t-score

Variable	Obs	Mean	Std. Dev.	Min	Max
A1ELP1WT	707	49.81373	10.05294	36.94	100

Program/Center-Level.

TyCurr: Type of curriculum

TyCurr	Freq.	Percent	Cum.

Creative Curr.	1,597	55.80	55.80
High Scope	477	16.67	72.47
High Reach	103	3.60	76.07
Montessori	66	2.31	78.37
Scholastic Curr.	183	6.39	84.77
Locally designed curr	436	15.23	100.00
Total	2,862	100.00	

T1DEPCAT: Teacher depressive symptoms, categories.

T1: Teacher Depress Score CES-D Shrt Form-Categories	Freq.	Percent	Cum.
1=not depressed	1,842	63.47	63.47
2=mildly depressed	766	26.40	89.87
3=moderately depressed	213	7.34	97.21
4=severely depressed	81	2.79	100.00
Total	2,902	100.00	

Part 2:

Challenges and How I will Address Them

As addressed in the first part of this memo, one of the biggest challenges in PSM is the selection of the “right” variables, meaning the variables that will help maximize the reduction of selection bias. To deal with this challenge I have examine all available covariates and using the Causal Structural Model method (Pearl, 2009) and added the covariates that work in general (Steiner, Cook, Shadish, & Clark, 2010) to the list. Using these strategies I was able to reduce the list of covariates from 69 to 21.

I have selected variables that fit into three categories: Family-level, child-level and program-level. The Family-level variables are important in accounting for endogenous family differences between those families who choose to start sending their children to HS at age three, versus those who choose to start sending their child at age four. The families in these two groups can be different in ways that relate to the outcome, severely affecting the ability to make inferences about the relation between the amount of preschool and self-regulation. By selecting covariates that account for observable differences and that relate to unobservable differences we can significantly reduce the selection bias and compare the different groups.

The child-level variables are important because these account for pre-existing differences in the children's characteristic and abilities, which are related to the outcome of study (self-regulation). Lastly, for program-level characteristics I have chosen to include the type of curriculum used and teacher's depression levels. Even though teacher depression levels is not necessarily a program-level characteristic it is a variable that can be grouped in the bucket of "other school factors". Prior evidence indicates that teacher's depression and burn out can be detrimental to children's learning.

Conclusion

I will use a list of 21 covariates in the analyses. I will include these covariates in the OLS, for constructing the propensity scores, and in the estimations using PSM. The covariates have been selected to reflect the theoretical causal structure between the context and the outcome, and to include the most commonly used covariates in these types of studies.

References

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys, 3*, 96-146.

Steiner, P.M., Cook, T.D., & Li, W. (2012, November 10). Covariate selection for PS designs in the absence of substantive theories. Paper presented at the Annual Fall Research Conference of the Association for Public Policy Analysis and Management, Baltimore, Maryland.

Steiner, P.M., Cook, T.D., Shadish, W.R., & Clark, M.H. (2010). The importance of covariate selection in controlling for selection bias in observational studies. *Psychological Methods, 15*(3), 250-267.

Cook, T.D., Steiner, P.M., & Pohl, S. (2009). How bias reduction is affected by covariate choice, unreliability, and mode of data analysis: Results from two types of within-study comparisons. *Multivariate Behavioral Research, 44*, 828-847.

Steiner, P.M., Cook, T.D., & Li, W. (2012, November 10). Covariate selection for PS designs in the absence of substantive theories. Paper presented at the Annual Fall Research Conference of the Association for Public Policy Analysis and Management, Baltimore, Maryland.

Memo 5

Reviewed Research Questions and Analytic Plan

Question 1.a: Do children that attend Head Start for two years (starting at three years of age) have significant better self-regulation in kindergarten than those who attended one year (starting at four years of age)? Does attendance moderate the relation between number of years in HS and self-regulation?

Question 1.b.: Does the number of hours a week in HS significantly relate to self-regulation in kindergarten? Does attendance moderate de relation between number of hours a week in HS and self-regulation?

Question 2.: Does the quality of teacher-child interaction in Head Start predict self-regulation skills in Kindergarten? Do specific domains of Teacher-Child Interactions differentially predict gains in self-regulation?

Question 3: Does quality of teacher-child interaction moderate dosage relation to self-regulation?

I will start the analyses by examining if a bifactor approach to examine the CLASS fits the data. A Bifactor approach to CLASS, which has been validated previously (Hamre et al., 2014), presents the advantage of generating orthogonal latent factors or teacher-child interaction domains: one general domain and some domain-specific factors. These orthogonal factors can then be examined concurrently in the predictive models without the problem of multicollinearity. I will do this using Mplus 7.4. and following the steps and models detailed by Hamre et al. (2014).

First I will examine one general factor and three domain-specific factors, which map onto the traditional three domains of the class, then I will examine a model with one general factor and two domain-specific factors: positive management and routines, and cognitive facilitation. I will examine the model of fit through fit statistics (Chi-squared, RMSEA, SRMR, and CFI) and by comparing the fit between the bifactor and the traditional 3-domain solution.

If the bifactor model fits the data I will use the latent factors to predict children's self-regulation, preferably using the same model as Hamre et al., (2014). If the model does not fit, I will proceed using the traditional three domains.

For answering all the questions I will run two sets of analyses: Ordinary-least-squares (OLS), and OLS with Propensity score matching to reduce selection bias in the comparison between number of years in HS.

Ordinary-least-squares (OLS).

Ordinary-least-squares (OLS) regression analyses will be conducted, using Stata 14, to examine the relation between three forms of dosage (number of years, number of hours a week, and absenteeism), quality of HS experiences and children's self-regulation in kindergarten. I will include all the 23 selected covariates in the model and I will examine each of the three outcomes (as determined in Memo 1) separately, that is: teacher-reported self-regulation, assessor-reported self-regulation and executive functions. The model is as shown below:

Self-Regulation= Constant + Covariates + CLASS bifactor factors + N of years in HS + N of hours a week + Absenteeism

Next, I will include interaction terms between number of years and absenteeism and between numbers of hours a week and absenteeism, to examine how absenteeism moderates the relation between the main dosage variables and self-regulation. Additionally I'll include an interaction term between the two main dosage variables (number of years, number of hours a week) and quality, to examine how quality moderates the relation between dosage and self-regulation. The model is as shown below:

Self-Regulation= Constant + Covariates + CLASS factors + N of years in HS + N of hours a week + Absenteeism + N of years in HS*Absenteeism + N of hours a week*Absenteeism + N of years in HS*CLASS factors + N of hours a week* CLASS factors

Propensity Score Matching.

To control for potential selection bias, given that the sample of children was not randomized to attend one or two years of preschool, and thus there might be important family characteristics that may be driving this decisions, I will use *Propensity Score Matching* (PSM) (Rosenbaum & Rubin, 1984). PSM reduces selection bias by adjusting for variables that relate to children and contextual characteristics that could be correlated to the decision of some families to enroll their children in preschool at 3 versus 4 years old. This quasi-experimental technique intend to imitate random assignment by matching cases (in the one condition) with possible counterfactual (in the other condition) as defined by their propensity to select into treatment, in this case attending 2 years of HS. This probability is defined by a set of covariates that are used in the estimation of the *propensity score* (*p)*. The propensity score in this case will indicate a child's propensity to attend two years of HS as a function of the included covariates.

There are four steps in utilizing PSM: estimate the propensity score, matching, check for balance, estimate the treatment effect (Lanza, Moore, & Butera, 2013). For the first step I will use logistic regression for predicting the propensity of selecting into treatment (T_i) as a function of the chosen set of covariates (X) and generating a p score for each individual (p_i), where $p_i = Pr(T_i = 1|\mathbf{x})$. I will use all the set of child and family level covariates described above, as well as child school readiness measures at the beginning of Head Start.

For matching the sample of children in each condition (those who attended one versus those who attended two years of Head Start) I will use the nearest-neighbor matching technique, with replacement, using a caliper of 0.1, which allows enough overlap between the two groups, reducing this way the selection bias in over 98% (Stuart, 2010). 1:1 Nearest neighbor matching selects, for each individual i in the treated group, the individual in the control group with the smallest distance from individual i . This technique has been described as one of the most simple and common methods for matching and when using a small caliper optimally reducing selection bias. Furthermore using nearest-neighbor matching with replacement can often reduce bias because individuals in the control group that look similar to those in the treatment group can be used multiple times (Stuart, 2010).

The next step in PSM is checking balance of the matched samples. I will do this by using two of the methods recommended by Harder (2010), checking standardized mean differences for all the covariates across each group, and comparing the ratio of the variances of the propensity score in the treated and control groups.

Once the samples are matched and adequate balance is achieved, then I will add the outcomes variables back to the data set and estimate the effects of two versus one year of HS and quality as a moderator. The model will be specified as below:

Self-regulation= constant + N of years in HS + CLASS factors + N of years in HS*CLASS factors

Missingnes.

To address issues of missing data I will use multiple imputation with chained equations (White, Royston, & Wood, 2011), as described on Memo 1.

Accounting for clustering in classrooms.

Once intra class correlations (ICC) are estimated, I will most likely account for nesting of children in classrooms by using robust standard errors clustered at the classroom level, like other studies using this sample and data set have used (see Ansari, Purtell, & Gershoff, 2016). Prior studies using this data have shown that there are small ICCs and that clustering the standard error is sufficient. Like multilevel modeling, clustered robust standard errors correct for the shared variance of observations due to children being in the same classrooms.