University of Virginia | Curry School and Batten School
EdPolicyWorks

Working Paper:

Within-Year Teacher Turnover in Head Start and Children's School Readiness

Anna J. Markowitz[1]

Teachers in early childhood education (ECE) settings are central to providing children with high-quality experiences that promote both school readiness and long-term wellbeing; unfortunately, rates of teacher turnover are high in ECE settings. There are strong theoretical reasons to assume turnover has negative implications for development, but very few empirical studies test this hypothesis. Using two waves of data from the nationally-representative Head Start Family and Child Experiences Survey, this study provides the first national estimate of the relationship between within-year ECE teacher turnover and children's development in Head Start. We find an annual within-year turnover rate of 10%, more than twice that of K-12, and that turnover is meaningfully and negatively associated with children's language, literacy, and socioemotional development.

[1]University of Virginia

Updated May 2019

EdPolicyWorks
University of Virginia
PO Box 400879
Charlottesville, VA 22904

Acknowledgements: This research was supported by a grant from the American Educational Research Association which receives funds for its AERA Grants Program from the National Science Foundation under NSF award NSF-DRL #1749275. Opinions reflect those of the author and do not necessarily reflect those of AERA or NSF.

Within-Year Teacher Turnover in Head Start and Children's School Readiness

Children's early experiences can have a meaningful influence on their life trajectories. Research shows that early home and school environments have substantial impacts on early skills (Phillips et al., 2017; Shonkoff & Phillips, 2000), and that these skills, collectively known as school readiness skills, have long-run implications for educational, health, labor market, and wellbeing outcomes (Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002; Deming, 2009; Moffitt et al., 2011). For the majority of young children in America, these early experiences now include some form of regular, non-parental care. Enrollment in center-based early childhood education (ECE) programs has been rising for 40 years, and 70% of American children now attend some kind of ECE program during the preschool year (Magnuson & Waldfogel, 2016).

Developmental and educational research suggests that the adults who care for children in these settings play a pivotal role in shaping children's developmental trajectories. Children thrive in warm, secure settings staffed by skilled teachers able to respond to children's learning needs. There is strong empirical evidence for the importance of both small, everyday and specific instructional interactions between teachers and children (Hamre, Hatfield, Pianta, & Jamil, 2014; Hamre, 2014; Hatfield, Hestenes, Kintner-Duffy, & O'Brien, 2013; Weiland & Yoshikawa, 2013) and for the primacy of the child-caregiver bond (Elicker & Fortner-Wood, 1995; Howes, 1999), suggesting that for children to extract the developmental benefits of ECE settings, much rides on the quality and consistency of the adult in the room. Historically, however, the ECE workforce in the United States has been highly unstable; some estimates suggest that about 25% of ECE teachers leave their jobs each year, a turnover rate that is four times higher than that of elementary school teachers (Bassok, Fitzpatrick, Loeb, & Paglayan, 2013; Whitebook, Phillips, & Howes, 2014).

Consequently, there has been a rising public interest in developing a more well-trained, well-paid, and stable ECE workforce (Institute of Medicine, 2015), and large investments have been made in Head Start, public preschools, state Quality Rating and Improvement Systems (QRIS), and some localities to professionalize the ECE workforce and reduce turnover (Barnett & Friedman-Krauss, 2016; Chandler,

2017; Friedman-Krauss et al., 2018; QRIS Compendium, 2017). Given the estimated size of the turnover problem and the scope of these investments, there is a surprising lack of either descriptive or rigorous evidence about the consequences of teacher turnover in ECE settings. Indeed, although it is generally believed that turnover is negatively associated with children's development, just one study (Tran & Winsler, 2011) provides empirical evidence in support of this assumption. Tran and Winsler (2011) find a negative relationship between turnover and development in a large sample from Miami, but are hindered by large amounts of missing data, very few control variables, and no way to account for the across-program differences that may drive both turnover and outcomes.

Using two waves of large, nationally representative Head Start data, the present study examines the association between *within-year* teacher turnover—that is, when a teacher leaves Head Start during the academic year—and children's language, literacy, mathematics, and behavioral regulation skills in rigorous models that account for children's selection into Head Start programs. In doing so, the present study will provide the first national estimate of the relationship between within-year teacher turnover and children's development in a low-income sample, and allow for an estimation of the developmental cost of turnover in Head Start. As the use of ECE expands, policymakers and practitioners alike are hoping to use it as a lever for improving both the school readiness and long-term outcomes of America's lowest-income and most vulnerable youth. Understanding the role of caregiver stability in promoting children's development will provide invaluable information as to whether and how much to invest in maintaining a stable ECE workforce.

The Importance of Stable Caregivers for Children's Development

Several developmental theories assert that stable, supportive relationships between children and the adults who care for them are fundamental to children's development (Shonkoff & Phillips, 2000; Phillipa, Austin, & Whitebook, 2016), and thereby suggest that teacher turnover should have implications for children's development. Ecological Systems Theory (EST, Bronfenbrenner & Morris, 2007) contends that small, everyday interactions between children and their caregivers are the main influences on children's development, and that distal factors influence children through their ability to alter or disrupt

those interactions. According to EST, young children who receive frequent, high-quality verbal and cognitive stimulation in the context of a warm, responsive relationship will develop stronger readiness skills than children who do not (Hamre, 2014; Howes et al., 2008; Phillips et al., 2017), because these interactions facilitate children's willingness to explore new problems and situations, encourage thinking and problem solving, and create opportunities to practice social and behavioral skills. In line with this theory, the link between teacher-child interactions and children's readiness skills has been repeatedly explored in the developmental literature, and a large body of empirical research, both observational and experimental, has corroborated this link in early educational settings (e.g. Araujo, Carneiro, Cruz-Aguayo, & Schady, 2016; Hamre, Hatfield, Pianta, & Jamil, 2014; Hamre, 2014; Hatfield, Hestenes, Kintner-Duffy, & O'Brien, 2013; Phillips et al., 2017; Pianta, Mashburn, Downer, Hamre, & Justice, 2008; Weiland, Ulvestad, Sachs, & Yoshikawa, 2013; Whittaker, Kinzie, Williford, & DeCoster, 2016). For example, observational studies in two high-quality public preschool systems—Boston and Tulsa—have demonstrated that the estimated impact of these preschools on children's readiness covaries with measures of the quality of teacher-child interactions (Johnson, Markowitz, Hill, & Phillips, 2015; Weiland et al., 2013). Moreover, research on professional development interventions in ECE classrooms has linked intervention-based improvements in the quality of teacher-child interactions to enhanced school readiness (Pianta et al., 2008; Powell, Diamond, Burchinal, & Koehler, 2010). Such findings highlight the importance of the caregiver-child relationship in promoting children's development, and suggest that mid-year changes to this caregiver are likely to disrupt children's developmental progress.

Attachment theory (e.g. Bowlby, 2008) likewise emphasizes the importance of the consistency and stability of the adult in the room in early educational settings (Verschueren & Koomen, 2012). Attachment theory asserts that children who have a secure bond to a teacher or caregiver are more likely to engage with the developmental opportunities around them and have positive peer interactions (Elicker & Fortner-Wood, 1995), and are therefore more likely to progress academically, socially, and emotionally (Elicker & Fortner-Wood, 1995; Moss, Parent, Gosselin, Rousseau, & St-Laurent, 1996; Verschueren & Koomen, 2012). The difficulty in forming positive bonds with peers may be particularly problematic in

the ECE context, in which a central goal is to teach children positive social skills and other behaviors that will support their success in kindergarten and beyond (e.g. Bierman et al., 2008). However, strong child-caregiver bonds take time to form (Howes, 1999; Raikes, 1993), and teacher turnover within the academic year is likely to disrupt a child's ability form strong attachments (Cryer, Hurwitz, & Wolery, 2000; Howes, 1999; IOM, 2015). Children whose caregivers change may feel a sense of loss or resentment, and may struggle to form attachments to new caregivers (Cryer et al., 2000; Howes & Hamilton, 1992). Children who experience turnover may engage less with the caregiver, the environment, and their peers than children with a stable caregiver, and may not benefit from the developmental opportunities the program affords.

Both theories suggest that within-year teacher turnover, that is, when a teacher leaves in the middle of the academic year, should be disruptive for young children and may impede their developmental progress. Previous research has linked turnover to the quality of ECE settings (C Hale-Jinks, Knopf, & Kemple, 2006; Helburn, 1995; Phillips, Mekos, Scarr, McCartney, & Abbott–Shim, 2000), and quality, in turn, has been linked to children's development (e.g. Araujo et al., 2016; Burchinal, Peisner-Feinberg, & Clifford, 2000; Early et al., 2007; Helburn, 1995; Markowitz, Bassok, & Hamre, 2018). However, there is currently almost no research directly linking within-year teacher turnover to children's developmental outcomes. A related literature on child care instability suggests that children exposed to multiple caregivers over time and children who are currently experiencing a large number of caregivers may experience more behavior problems, providing suggestive evidence that the increased volume of caregivers experienced by a child due to turnover may also be negatively linked to children's socio-emotional development (e.g., Pilarz & Hill, 2014). Conversely, some research in K-12 settings suggests that the teachers who turn over are less capable than those who remain or who replace them, boosting overall classroom quality and, subsequently, children's academic growth (Adnot, Dee, Katz, & Wyckoff, 2017). If turnover provides children with the opportunity to engage with a more highly-skilled caregiver, it may not be developmentally detrimental, and indeed may promote academic development. The present provides an empirical test of these hypotheses.

Teacher Turnover in Early Childhood Settings

Teacher turnover has long been a concern in early childhood settings (Whitebook, Howes, & Phillips, 1998). The early childhood workforce has historically been made up of teachers and caregivers with relatively low levels of education who work for very low wages (Bassok et al., 2013; Whitebook, Phillips, & Howes, 2014b), two conditions that commonly lead to turnover. Both anecdotal evidence and data from large scale surveys of adults have provided evidence that turnover is likely to be sizable in the early childhood sector. For example, using data from the 2009 Current Population Survey (CPS) Bassok and colleagues (2013) found that nearly a quarter of teachers providing care for young children in 2009 were not doing so by 2010. Similarly, using data from the 2012 National Survey of Early Care and Education, Whitebook and colleagues (2014) find director-reported departure rates ranging from eight to 27 percent, depending on program sector. Such data has led to rising political and popular concern over teacher turnover in early childhood (e.g. Institute of Medicine, 2015; Interlandi, 2018). This concern has translated into increased investments in ECE teacher training as an antidote to turnover, including ever-increasing requirements for teacher education in Head Start and sate preschool, and the incentivization of teacher education and retention in QRIS (Barnett & Friedman-Krauss, 2016; Friedman-Krauss et al., 2018; QRIS Compendium, 2017).

Despite the scope of these investments, we ultimately know little about turnover in early childhood. Most ECE turnover data, including those cited above, come from staffing surveys which provide a one-time snapshot of the ECE workforce, or data like the CPS which provides information on overall industry turnover. These data portray an unstable workforce and raise concerns for how instability in programs may influence both program quality and children's development. However they cannot identify the prevalence of the within-year turnover that is likely to be the most detrimental for children's development because of its direct impacts on both teacher-child interactions and teacher-child attachment. Nevertheless, these prior studies provide an important foundation upon which the current paper builds.

Turnover in Head Start. Head Start is the largest federal investment in early childhood education. As such, it is subject to substantial regulation and collects more program data than other early

childhood settings. Annual Head Start Program Information Report (PIR) data provide a window into turnover in Head Start settings, and suggest that annual turnover in Head Start from 2002-2015 ranged from 12% to 18%. However PIR data provide an annual, director-reported snapshot of Head Start turnover, and thus do not provide information about within-year turnover specifically. Similarly, a recent Head Start report using data from the nationally representative Family and Child Experiences Survey reports that average program-level turnover in Head Start was 21.2%, 14.4%, and 13.7% in 2006, 2009, and 2014, respectively (Aikens, Bush, Gleason, Malone, & Tarullo, 2016). However, these estimates are generated from director report of 12-month turnover, which may be subject to error, and again do not capture within-year turnover specifically.

A recent study of a small sample (N= 65) of Head Start teachers in the Midwest provides some information with respect to within-year turnover, documenting that 36% of the newly-hired teachers who began at the start of the program year had left by January (Wells, 2015). It is not clear, however, if this estimate should be higher or lower than overall within-year teacher turnover in Head Start. On the one hand, it accounts only for teachers who turnover by January, which may underestimate within-year turnover; on the other hand, the sample includes only new teachers, among whom turnover is likely to be higher. To date, there are no national estimates of within-year turnover in Head Start, nor are there estimates of the relationship between turnover and children's development.

Empirical Evidence Linking Turnover to Children's Development

Although there are strong theoretical reasons to assume that teacher turnover has negative implications for children, there is surprisingly little rigorous evidence documenting this relationship, either in the early childhood or K-12 education sectors. In the early childhood literature there have been very few studies documenting links (causal or correlational) between teacher turnover and child outcomes (Whitebook & Sakai, 2003). Tran and Winsler (2011) offer the best existing evidence on the association between within-year teacher turnover and child outcomes. Using a large sample of low-income four-year-olds in Miami, they find evidence that children who experienced a change in teacher or caregiver between the fall and spring had more negative cognitive, social-emotional, and linguistic outcomes. Specifically,

children who did not experience turnover ended the year with higher levels of fine motor skills, cognitive skills, language skills, teacher- and parent-reported self-control, and parent reported-initiative, and lower levels of teacher-reported behavior concerns. Many of these associations disappeared when children's skills at the beginning of the program year were accounted for, however; in these more rigorous models, children who kept the same caregiver across the year had greater growth in teacher-reported initiative and teacher-reported closeness/attachment than children who experienced turnover. It is also worth noting that this study had high levels of missing data (22% of children were missing outcome data) and a relatively limited set of control variables, most notably the absence of any family process, teacher, or program covariates, which raises concerns for causal interpretation. In particular, it may be that the children and programs that experience turnover systematically differ from those that do not, and that those differences, rather than the turnover itself, drive negative associations.

Similarly, nearly all of the existing K-12 research on this topic is correlational, showing that, for example, schools with higher teacher turnover rates have lower student achievement (e.g. Bryk & Schneider, 2002; Guin, 2004). As in the ECE literature, these associations may not be capturing the impact of turnover, but rather other characteristics (e.g. poverty rates) that may be related both to turnover and outcomes. Ronfeldt, Loeb, and Wyckoff (2013) provided the first causal evidence on this topic. Using a unique quasi-experimental design, they showed that turnover rates among teachers led to lower learning outcomes for elementary school students (-0.08 and -0.05 of a standard deviation in math and reading, respectively), particularly in schools serving more low-performing and Black students. However, this study uses across-year turnover rather than within-year, and may thus underestimate the size of the link between turnover and outcomes. Moreover, preschool aged children whose learning is more dependent on teacher-child interactions and their attachment to their caregiver may struggle more with the experience of teacher turnover than older children.

Present Study

In sum, although there are strong theoretical reasons to believe that within-year teacher turnover in early childhood settings should have deleterious associations with children's development, we

currently have no national estimates of the prevalence of this within-year turnover, and very limited evidence linking any type of turnover to children's outcomes. The present study addresses these issues in the context of Head Start, the United States' largest investment in the early educational experiences of low-income children. Specifically, we use a nationally representative Head Start dataset to estimate the prevalence of within-year turnover in Head Start and to quantify the association between within-year turnover and children's math, literacy, social, and behavioral outcomes. We estimate these associations in rigorous models that include a measure of children's skills at Head Start entry, a rich set of child, family, teacher, and director covariates, and controls for program features that are linked to both turnover and children's development. Based on theory and previous research, we expect that turnover will be negatively associated with children's developmental gains over the course of the Head Start year, particularly for children's social and behavioral outcomes.

The present study adds to the small literature linking turnover to children's outcomes in several ways. First, it provides the first national study examining the relationship between ECE teacher turnover and the outcomes of low-income children. Second, it focuses on within-year turnover, which, due to data limitations, is understudied but more likely than other types of turnover to disrupt children's development. Finally, it compares the outcomes of children within the same program who do and do not experience teacher turnover within a given year. These within-program comparisons provide more plausibly causal estimates than existing studies.

The present study builds our understanding of the extent to which teacher turnover harms young children by providing evidence as to the scope of the problem and by quantifying the developmental loss associated with turnover. It provides important information about within-year teacher turnover for policymakers and practitioners hoping to give children the early experiences they need to propel them to lifelong wellbeing.

Method

Data were drawn from the Head Start Family and Child Experiences Surveys (FACES). FACES is an ongoing, nationally representative study of Head Start conducted every three years from 1997

through 2009, and again in 2014.[1] It uses a multi-stage probability sampling design with stratification to ensure a nationally representative sample at the program, center, classroom, and child level (West, Tarullo, Aikens, Malone, & Carlson, 2011). The FACES study tracks characteristics of the population served by Head Start; Head Start program features, including staff qualifications and characteristics, classroom characteristics and quality measures; and child and family outcomes including a direct assessment of children's skills in both the fall and spring of children's first Head Start year. A unique strength of the FACES data relative to other widely-used national, early childhood datasets (e.g. the Early Childhood Longitudinal Study-Birth Cohort), is that in most of the waves of data teachers are surveyed in both the fall and spring of the Head Start year, making it possible to assess whether the lead teacher present in the fall led the same classroom in the spring.

This study used FACES data from the 2006 and 2009 survey years, the two most recent years of FACES that have identical sampling frames and data collection strategies. Across these two waves there were 5,797 children with valid teacher identification information in both the fall and spring. Of these, 5,037 (86.9%) had the same teacher and stayed in the same classroom the entire year; 112 (1.9%) switched classrooms (and thereby teachers) over the course of the program year; 23 (0.4%) children who got a new teacher during the year, but whose teacher did not leave the Head Start program; and 625 (10.8%) children whose teachers left Head Start entirely over the course of the program year. Within this 625 were 45 (0.8%) children whose teacher left Head Start but who also switched classrooms. For these 45 children, we could not identify whether classroom switching or teacher turnover came first; that is, we could not say whether these 45 children left a classroom with a difficult teacher potentially precipitating turnover, or if the teacher's turnover led to the dissolution of the classroom and therefore classroom switching. Because the purpose of this analysis was to understand the association between turnover and children's development, we kept in our main analysis only the children that allow us to make the cleanest

[1] FACES recently released a new wave of data collection from 2014. However, FACES changed its sampling frame and surveys considerably. As a result, we do not use the 2014 data as our main analysis. However, in sensitivity analyses we do add the 2014 data to the 2006 and 2009, and they a similar pattern (Table 5).

comparison between turnover and the absences of turnover. We retained in the main analytic sample the 5,037 who had the same teacher in the same classroom for the entire year, and the 580 children who experienced teacher turnover, but stayed in the same classroom, for a total sample of 5,617. All analyses were weighted using FACES-provided sampling weights to account for the complex sampling design and to make estimates nationally representative. We thus further restricted the sample to children who have a valid value on the weight used in this analysis, P21RA2WT. This weight required that children have some parent interview data and some child assessment data in the spring, *N*= 5,476. Final sample restrictions occurred based on missingness in the dependent variable; sample sizes for our main analyses ranged from 5,054-5,429.

Some of these cases were missing data on child, family, teacher, and program covariates. Following Von Hippel (2007) and Johnson and Young (2011), missing data were multiple imputed using chained equations in Stata 14, and all estimates were combined across the 15 imputed datasets. We did not use imputed data on teacher turnover or on dependent variables in our analytic models, however they were included in the imputation model.

Measures

Within-year teacher turnover. Within-year teacher turnover was coded using the unique teacher identification (ID) numbers provided to Head Start teachers in the fall. A teacher was coded as having "turned over" if their teacher ID number was present at the fall data collection but not in the spring, indicating that they were no longer present at Head Start. We then conducted additional checks to ensure that this assumption is correct; we checked that the missing ID numbers were not due to missing data by ensuring that the students who had the missing teachers in the fall were present in the spring data, and found no evidence suggesting we miscoded attrition as turnover. We found that 10% of children in Head Start had a teacher that turned over during the program year.

Developmental outcomes. FACES included a rich set of child assessments that allowed us to test the relationship between turnover and development across several domains, including language, literacy, mathematics, social skills, and behavioral regulation.

Peabody Picture Vocabulary Test. The Peabody Picture Vocabulary Test (PPVT, Dunn & Dunn, 2007) is a receptive language measure in which children point to a picture that corresponds with the word spoken aloud by the assessor. The PPVT is widely used and has demonstrated strong reliability and validity in previous research (e.g. Dunn & Dunn, 2013) ; in the present data Cronbach's alpha ranged from 0.91 to 0.97. The PPVT was administered in both the fall and spring to all children who had adequate English language skills. We used the raw score in both fall and spring, and standardized each within the final sample such that coefficients can be interpreted as effect sizes.

Woodcock Johnson Letter Word Identification. The Woodcock Johnson (WJ) Letter-Word Identification test (Woodcock, Mather, & McGrew, 2001) is a literacy assessment that measures children's ability to identify isolated letters and words. The Woodcock Johnson battery is commonly used, and the Letter-Word subtest has shown high internal reliability in several studies (McGrew & Woodcock, 2001; Schrank et al., 2005) ; in the present data Cronbach's alpha ranged from 0.85 to 0.93. The FACES administration of this test used a stopping rule of 3 consecutive items incorrect, so most children were tested with about 10 items on this subtest. As with the PPVT, we used the raw score in both fall and spring, and standardized each within the final sample such that coefficients can be interpreted as effect sizes.

Woodcock-Johnson Spelling. The WJ Spelling subtest (Wookdcock et al., 2001) is a subtest that measures fine motor coordination and pre-writing skills (e.g. drawing lines, copying letters) and skill in providing written responses when asked to write specific upper or lower-case letters, words, phrases, and punctuation marks. Like the other WJ Letter-Word test, this test has shown high internal reliability in several studies (alpha=0.94, McGrew & Woodcock, 2001; Schrank et al., 2005); in the present data Cronbach's alpha ranged from 0.79 to 0.83. We used the raw score in both the fall and spring, and standardized each within the final sample such that coefficients can be interpreted as effect sizes.

Woodcock-Johnson Applied Problems. The WJ Applied Problems subtest (Woodcock et al., 2001) is a mathematics assessment that measures children's ability to analyze and solve math problems. As with the WJ Letter-Word test, children were stopped after they answered three consecutive items

incorrectly, for most this children this included at least 29 items. Like the other WJ subtests, this test has shown high internal reliability in several studies ((McGrew & Woodcock, 2001; Schrank et al., 2005); in the present data Cronbach's alpha ranged from 0.87 to 0.90. We used the raw score in both the fall and spring, and standardized each within the final sample such that coefficients can be interpreted as effect sizes.

Parent-reported behavior. In both the fall and the spring, parents rated their child on a set of items related to their social skills and behavioral regulation. Parents rated their child's behaviors on a set of 21 items such as "make friends easily" or "has a very strong temper and loses it easily" on a one to three scale ranging from "not true" to "very true or often true." These items were drawn from several commonly used scales: the Personal Maturity Scale (Entwisle & Alexander, 1987; Zill, Moore, Smith, & Stief, 1991), the Social Skills Rating System (SSRS, Gresham & Elliott, 1990), and the Behavior Problems Index (BPI, Zill & Peterson, 1986). The reliability of the Personal Maturity Scale is adequate (Cronbach's alpha ranges from 0.74 to 0.85), and the reliability of the SSRS and the BPI have been historically high (Cronbach's alpha > 0.85). FACES categorized these items into two summary scores: problem behaviors and social skills/positive behaviors. These scales were made by summing the component items (exact items are copyrighted), and scales had sufficient internal reliability, ranging from 0.72-0.79 for behavior problems and 0.68-0.72 for positive social skills. We used both FACES-created scales, and standardized each within the final sample such that coefficients can be interpreted as effect sizes.

Teacher-reported behavior. In both the fall and spring teachers rated children on a set of items drawn from the Personal Maturity Scale, the SSRS, and the BPI that were framed to reflect classroom behaviors such as "follows teachers directions," "is very restless," and "hits/fights with others" on a scale from 1 ("never") to 3 ("very often"). Teacher items were divided into several subscales. We used a 14-item general measure of problem behaviors taken from the sum of BPI items (alpha ranged from 0.81 to 0.85); a 13-item social skills subscale (alpha ranged from 0.88 to 0.90) that measured a child's interest, participation, cooperation, compliance, and attention in the classroom; and two subscales of the BPI that

measured a child's aggressive behavior ("hits/fights with others," alpha ranges from 0.81 to 0.85) and their anxious and withdrawn behaviors ("is unhappy," alpha ranges from 0.73 to 0.77). The aggressive and withdrawn subscales were created by taking the average of a set of BPI items. We standardized all scales within sample such that coefficients can be interpreted as effect sizes.

Covariates. A strength of the FACES data is that it provides a rich set of child, family, teacher, and program variables. To reduce, but not eliminate, the possibility of omitted variable bias, we used a host of these variables as covariates in all models. At the child and family level, we included a 4-group measure of child race/ethnicity (White, Black, Hispanic, other race/ethnicity); gender; disability status; child's age in months at the time of the spring assessment; an indicator for English as a second language in the household; a 6-level measure of family income; a dichotomous indicator for public benefit receipt; a 4-level indicator of maternal education (less than a high school degree, a high school degree, some college, BA or more); a dichotomous indicator of maternal employment; a dichotomous indicator for a single parent household; a continuous measure of mother's age at child's birth; a continuous measure of maternal depressive symptoms (constructed by FACES, taken from the Center for Epidemiological Studies Depression Scale, Radloff, 1977); a dichotomous indicator for whether a parent reads to the child at least three times a week; a dichotomous indicator of immigrant status, coded such that 1 indicates at least one parent is an immigrant, 0 otherwise; an indicator of child cohort (i.e., 3 or 4 years old at Head Start entry), and an indicator for whether the child's data came from the 2006 or 2009 wave of FACES.

At the teacher level, we created a 4-group measure of teacher race/ethnicity (White, Black, Hispanic, other race/ethnicity); a 3-level measure of teacher education (less than an AA, AA, or BA or more); a continuous measure of years of teacher experience; a continuous measure of teacher age; a continuous measure of class size; an indicator for full-day (rather than half-day) Head Start programs; and an indicator for whether the teacher teaches two classes daily. Teacher characteristics were taken from the spring survey; that is, they were the characteristics of the new teacher, and therefore the teacher the child had at the time of the spring direct assessments.

We also included a 4-level measure of director education (less than an AA, AA, BA, or BA or

more), a 4-group measure of director race/ethnicity (White, Black, Hispanic, other race/ethnicity); an indicator for whether the director is an NAEYC member; and a continuous measure of director experience. All covariates are listed in Table 1.

Analytic Strategy

To estimate the relationship between within-year teacher turnover and children's developmental outcomes in Head Start, we estimated four models. First, we regressed each individual child outcome for child *i* in program *j* on within-year teacher turnover and child, family, teacher, and director covariates, and indicators for data wave (i.e., 2006 or 2009), and cohort (3- or 4-year-old).

$$Developmental\ Outcome_{ij} = \alpha + \beta_1(turnover_{ij}) + \beta_2(child\ and\ family\ covariates_{ij}) + \beta_3(teacher\ covariates_{ij}) + \beta_4(director\ covariates_{ij}) + \beta_5(data\ wave_{ij}) + \beta_6(cohort_{ij}) + \varepsilon_i \quad (1)$$

In equation 1, the coefficient β_1 represents the association between within-year teacher turnover and children's developmental outcomes net of family, child, teacher, and director characteristics. Recall that in these models turnover was coded such that "0" indicated the child had the same teacher and stayed in the same classroom throughout the entirety of the program year; "1" indicated that the child's teacher left Head Start during the program year and the child did *not* switch classrooms. These estimates did not include children who switched classrooms during the Head Start year.

This standard regression model can only provide a causal estimate if all relevant variables are accounted for. Given the number of unobserved factors that contribute to students' development over time, this assumption seems unlikely to be true, despite the rich set of covariates included in this model. Thus, in a second model (below), we included a measure of children's skills as measured in the fall of the Head Start year. The inclusion of this variable accounts, in part, for children's innate skills and their experiences prior to Head Start entry, and allows us to measure children's growth over the Head Start year, making it a key adjustment for omitted variable bias (National Institute of Child Health and Human Development Early Child Care Research Network & Duncan, 2003).

$$Developmental\ Outcome_{ij} = \alpha + \beta_1(turnover_{ij}) + \beta_2(Fall\ Developmental\ Outcome_{ij}) + \beta_3(child\ and\ family\ covariates_{ij}) + \beta_4(teacher\ covariates_{ij}) + \beta_5(director\ covariates_{ij}) + \beta_6(data\ wave_{ij}) + \beta_7(cohort_{ij}) + \varepsilon_i \quad (2)$$

In equation 2, the coefficient β_1 represents the association between within-year teacher turnover and children's developmental outcomes net of family, child, teacher, and director characteristics, as well as children's developmental skills in the fall. Though the inclusion of the lagged measure does much to address unobserved differences in children's experiences prior to Head Start entry, it does not account for unobserved, omitted features of Head Start settings that may co-vary with turnover and thus bias the estimate. Head Start programs are likely to vary in a many ways that are correlated with both within-year teacher turnover and children's developmental gains. For example, it may be that programs with high levels of teacher turnover are also programs that have lower quality professional development, which might reduce children's developmental gains. To address such potential omitted variables, we estimated versions of equations 1 and 2 that included program fixed effects, which account for unobserved fixed program traits by making comparisons within program.

$$Developmental\ Outcome_{ij} = \alpha + \beta_1(turnover_{ij}) + \beta_2(child\ and\ family\ covariates_{ij}) + \beta_3(teacher\ covariates_{ij}) + \beta_4(director\ covariates_{ij}) + \beta_5(data\ wave_{ij}) + \beta_6(cohort_{ij}) + \gamma_j + \varepsilon_i \quad (3)$$

$$Developmental\ Outcome_{ij} = \alpha + \beta_1(turnover_{ij}) + \beta_2(Fall\ Developmental\ Outcome_{ij}) + \beta_3(child\ and\ family\ covariates_{ij}) + \beta_4(teacher\ covariates_{ij}) + \beta_5(director\ covariates_{ij}) + \beta_6(data\ wave_{ij}) + \beta_7(cohort_{ij}) + \gamma_j + \varepsilon_i \quad (4)$$

In equations 3 and 4, γ_j represents program fixed effects. In equation 3, the coefficient β_1 represents the within-program association between within-year teacher turnover and children's developmental outcomes net of family, child, teacher, and director characteristics; that is, it represents the average difference in children's developmental outcomes for children who experienced turnover as

compared to children who did not but who attend the same program. In equation 4, the coefficient β_1 represents the within-program association between within-year teacher turnover and children's developmental outcomes net of family, child, teacher, and director characteristics, as well as children's developmental skills in the fall. These models estimated the differential *gains* in skills for children who attended the same program but who had different experiences with respect to turnover.

The program fixed effects provide a more rigorous estimate of the impact of teacher turnover on children's outcomes than the OLS models. The tradeoff is that fixed effects models require a high level of within-program variability in within-year teacher turnover to generate precise estimates.[2] Given this tradeoff, we present results of estimates with and without fixed effects in all tables.

As discussed above, all models were estimated in imputed data, and were weighted using FACES-provided sampling weights to ensure nationally representative estimates.

Results

Table 1 presents descriptive statistics for the full sample, as well as tests for the bivariate relationship between turnover and all dependent variables and all covariates both between and within programs. Children whose teacher turned over had on average lower WJ Letter Word and Spelling scores, as well as higher levels of problem behavior, aggression, and withdrawal as reported by teachers, and lower levels of teacher-reported social skills. Notably, however, children whose teachers turned over also had higher levels of aggression and lower levels of social skills at the start of the Head Start program year. These differences were statistically significant within programs.

Demographically, the sample was 35% Black, 36% Hispanic, 22% White, and 8% other race, and 26% of the sample spoke a language other than English at home. Consistent with Head Start's mission, family incomes in the sample were low; about 55% of the sample had an income under $20,000, and about 60% received some kind of public benefit. Mothers also had low levels of education on average,

[2] Within-center variability in teacher turnover was insufficient for the use of center fixed effects models; whereas 63% of programs had variability in turnover, just 25% of centers did. The average number of centers per program in this sample was 3.38; the average number of teachers per center is 3.36 and per program 11.15.

and nearly 70% of children lived in a single parent household. Although there were no differences in child-level demographics across turnover groups, mothers whose children experienced turnover had lower levels of education, both across and within programs.

Teacher characteristics in Table 1 were taken from the spring survey; thus for the no turnover group (Turnover= 0) these describe the teacher that had been with children throughout the whole year, whereas for the turnover group (Turnover= 1) these are the characteristics of the new teacher (rather than the turning over teacher, for characteristics of turning over teacher, see Appendix A). About half of teachers in Head Start had a BA degree or higher, consistent with Head Start policy. Teacher experience differed by turnover both across and within programs such that children who did not experience turnover had on average more experienced teachers. Class size did not differ by turnover status and was roughly equal to 17 children. Children who experienced turnover were also far more likely to have a teacher who taught twice daily (43% vs. 14%, significant both across and within programs).

Head Start directors had relatively high levels of education on average; roughly 80% had a BA or more. Directors were 42% white, 36% Black, and 17% Hispanic. They had on average 13 years of experience and 73% were NAEYC members. Director characteristics did not differ significantly by turnover status.

Multivariate Models

Table 2 presents associations between within-year teacher turnover and children's developmental outcomes across all four models described above: the OLS model; the OLS model with the addition of a lagged dependent variable; the program fixed effects model; and the program fixed effects model with the addition of a lagged dependent variable. Recall that in these models the coefficient represents the difference between children who stayed in the same classroom with the same teacher all year, and children whose teacher left Head Start, excluding children who switched classrooms. Using this contrast, there were no associations between turnover and children's PPVT scores, WJ Applied Problems scores, or any parent-reported behavior problems across any of the four models.

There were associations between within-year teacher turnover and WJ Letter Word and Spelling

tests, and teacher-reported behavioral outcomes. Across all four WJ Letter Word models, within-year teacher turnover was associated with about a one-tenth of a standard deviation (SD) reduction in children's scores. This association was not significant in the OLS model, but was significant across all lagged and fixed effects models ($p< 0.05$). The pattern was similar for WJ Spelling: although associations were somewhat smaller (0.07 SD) and not conventionally significant in the OLS models (model 1 and model 2), the association in both of the program fixed effects models was again roughly -0.10 SD ($p< 0.05$ in model 3, $p= 0.06$ in model 4).

Across all four of the teacher-reported behavioral variables, there were large, statistically significant ($p< 0.05$) associations between turnover and maladaptive behavioral outcomes in models that did not include a lag (models 1 and 3). Associations ranged from 0.15 SD (aggressive behaviors) to -0.24 SD (social skills) in model 3. However, the addition of the lag substantially diminished these associations. In the most restrictive model (model 4, program fixed effects and a lagged dependent variable), teacher turnover was associated with a 0.15 and 0.16 SD increase in overall behavior problems and withdrawn behaviors, respectively. The association for social skills was similar in magnitude (-0.14 SD), but not statistically significant ($p= 0.13$). The association between turnover and aggressive behavior attenuated substantially to 0.05 SD, and was not significant.

Sensitivity analyses. We ran three sensitivity analyses. First, we changed the inclusion criteria for the turnover variable. In the main analyses, we conducted the cleanest test of the impact of teacher turnover by looking at children who had the same teacher in the same classroom across the full program year as compared to children whose teacher left Head Start and had a new teacher enter their classroom. In these models, children who switched classrooms were dropped from the analysis. However, in the FACES data a small portion of children (2%) switched classrooms during the academic year. We wanted to ensure that findings were not sensitive to the inclusion of children who voluntarily switch teachers during the program year. Thus, Appendix B presents all models from Table 2, but add the children who switched classrooms into the sample, coded as a "0" for turnover. We also added an indicator for classroom switching into this model. As expected, the findings in Appendix B mimic those of Table 2

nearly perfectly.

Our second sensitivity analysis was conducted only for the teacher-reported problem behaviors. A main concern with these models is that for teacher-reported problem behaviors the indicator of teacher turnover would be identical to an indicator for having a different rater across the two time points; that is, while for the children who did not experience turnover the rater was the same in the fall and the spring, for those who did experience turnover the rater was different. Moreover, we may be concerned that the ratings of turning over teachers may be systematically different than teachers that choose to stay. For example, these teachers may see children's behaviors as more problematic, and this may contribute to their decision to leave Head Start. Consistent with this hypothesis, in the fall children whose teachers turned over were rated as having higher levels of aggression and lower levels of social skills (Table 1). However, we were nonetheless conceptually interested in the in-school behavior of children who experience turnover. The disruption that they experience in school may lead to behavior problems in the Head Start classroom that do not manifest at home, where the child-caregiver bond has not changed. Thus, we conducted a sensitivity analysis in which we used the parent-reported behavioral measures as the lagged score, rather than the teacher-reported. These scores are imperfect because they do not capture children's classroom behaviors, but have the benefit of being consistent across turnover groups. Teacher and parent report behavior problems and social skills scores were modestly correlated, about 0.18. Using these measures as the lagged behavioral scores we found large, significant ($p< 0.05$) associations across all models and all behavioral outcomes (see Appendix C). In the most restrictive model (model 4, program fixed effects, lagged dependent variable), teacher turnover was associated with a quarter SD increase in overall behavioral problems; a quarter SD decrease in social skills; a 0.15 SD increase in aggressive behaviors; and a one-fifth SD increase in withdrawn behaviors. These estimates were very similar to those in the models without lagged dependent variables, and should be interpreted with caution. However, they were consistent with patterns in the models using lagged teacher reports, and in the parent reported behavior problems models.

Finally, we also ran models in which we appended the FACES 2014 data to the present data

(2006 and 2009) and ran all models in this new, larger dataset. As noted above the 2014 data were not collected in the same way as the 2006 and 2009 waves. Most notably, FACES no longer only sampled children who were new to Head Start. Thus, these models also included a covariate indicating whether the child was new to Head Start. The FACES 2014 also no longer asked parents to rate children's behavior. Results from these models are presented in Appendix D. The pattern of results was strikingly similar. We found statistically significant, negative associations between turnover and both letter-word and spelling subtests, with an effect size of about 0.15 of a standard deviation, and statistically significant increases in problem behaviors and withdrawn behaviors, and decreases in social skills, with effect sizes again about 0.15 of a standard deviation.

Discussion

There has been rising concern over the quality and stability of the ECE workforce in recent years, as evidence by national reports (e.g. IOM, 2015), early educational policies, and popular concern (e.g. Interlandi, 2018; Mooney, 2016; Paquette, 2016). A popular consensus has emerged that turnover is a significant problem in ECE settings, and that policymakers need to make substantial investments in the ECE workforce to reduce turnover and improve child outcomes. However, much of what we currently know about turnover is driven by assumptions—albeit theoretically sound assumptions—and almost none of our information about turnover is based on rigorous empirical research, specific to the within-year turnover that developmental theory suggests is likely the most detrimental to child development.

The present study addressed those gaps by estimating the prevalence of within-year turnover and its association with children's developmental gains in a nationally representative sample of Head Start children. It adds to the existing literature turnover by providing national estimates of the impact of turnover in a low-income sample; focusing on within-year turnover; and employing a more rigorous modeling strategy than has previously been used. That is, by comparing the outcomes of children within the same program who did and did not experience turnover in models with lagged measures of children's skills and a rich set of covariates, this study provides more plausibly causal—though still descriptive—estimates than previous studies.

This study found that about 10% of the Head Start teaching workforce turned over during both of the years from which our data were collected, 2006 and 2009.[3] Head Start serves nearly a million children annually; this amounts to about 100,000 children who do not have a stable caregiver throughout the program year. Moreover, this rate is about twice as large as recent estimates of within-year teacher turnover in K-12 settings in North Carolina (Redding & Henry, 2018).

As expected, and consistent with both theory and the existing empirical exploration of turnover (Tran & Winsler, 2011), we found that teacher turnover was negatively associated with children's developmental gains over the course of the Head Start year. Teacher turnover was associated with smaller gains in spelling and vocabulary (about 0.10 of a standard deviation), higher levels of teacher-reported problem behaviors (about 0.15 of a standard deviation), and lower levels of teacher-reported positive behaviors (about 0.15 of a standard deviation). These associations were stronger and more consistent with the addition of the newest wave of FACES data (Appendix D).

We did not find associations between turnover and children's applied problems scores. This may be because on average Head Start classrooms spend much more time on language and literacy skills than on mathematics skills (Walter & Lippard, 2017) so the disjunction created by turnover was less salient. It may also be because children build language and literacy skills particularly in relational contexts—through conversation and shared storytelling, for example—meaning that a disrupted teacher-child relationship is more salient in the activities used to build these skills (e.g. Pianta, 2007). Previous research suggests that the quality of instructional interactions in the classroom are linked with the emotional support in the classroom and the quality of the teacher-child relationship (Curby, Rimm-Kaufman, & Abry, 2013; Hamre & Pianta, 2005; Rudasill & Rimm-Kaufman, 2009), suggesting that when teachers leave and children need to form a relationship with a new caregiver, learning may be at least temporarily interrupted.

We did find consistent associations between turnover and children's behavior problems and social

[3] Turnover was substantially higher in the 2014 wave of FACES, about 17%.

skills, as hypothesized by developmental theory and related literatures (e.g. child care instability, Pilarz & Hill, 2014). Notably, we found strong associations between turnover and teacher report of children's withdrawn behavior. Attachment theory in particular highlights the role of secure attachment to a caregiver with children's exploration of the classroom and their willingness to engage in positive peer interactions (e.g. Elicker & Fortner-Wood, 1995; Verschueren & Koomen, 2012) If children who experience turnover feel a sense of loss and confusion with the departure of their former teacher, they may struggle to form an attachment to the new caregiver (e.g. Howes & Hamilton, 1992), and may thus fail to explore their surroundings and their peers, displaying more withdrawn behaviors in the classroom. These behaviors will not only prevent a child from building social skills, but may also prevent a child from taking advantage of important learning opportunities, such as those embedded in imaginary play or those gained through small group activities and teacher-child dialogue.

In interpreting the findings linking teacher turnover to children's behavioral problems and social skills, it is important to recognize that we only found associations in models in which child behavior was teacher-reported. This is a concern for two reasons. First, it seems likely that teachers who turn over rate children's behavior differently than teachers who stay. We do not know the nature of this difference, that is, we do not know if turning over teachers rate behavior as systematically better, systematically worse, or with more error. Table 1 shows that turning over teachers rated their students as more aggressive than teachers who stayed, but also that the standard deviation was higher. It could be that teachers perceived their children's behavior as worse than a staying teacher would have, however it may also be the case that children in these classrooms were more aggressive, which contributed to teacher departure. Because we do not know how the fall measures of behavior differ for turning over versus staying teachers, but feel confident that they do differ, we must interpret the lagged models with caution. Second, and similarly, in our analyses teacher turnover is perfectly correlated with having a different rater in the fall and the spring. Again, it is unclear how this factor would affect our estimates, raising concerns for interpretation. We addressed this problem in part by conducting sensitivity tests where we used parent reported behavior as the lagged measure of children's skills, however it is important to note that parent and teacher ratings

were only modestly correlated (r= 0.18) and that teacher-report and parent-report of child behavior are fundamentally different constructs. For these reasons, findings from all teacher-reported models should be interpreted with caution. The models that do not include lags are likely under-controlled and inflate associations between turnover and outcomes; but the lagged dependent variables included in the model may themselves vary systematically by turnover status, biasing the estimates in unknown ways.

Even if the finding that children who experienced teacher turnover were rated as systematically worse behaved than those who did not experience turnover does not represent a causal link between teacher turnover and children's social and behavioral regulation skills, it may still have important implications. Teachers' perceptions of students are important for their engagement with and cognitive stimulation of children, and for the development of the teacher-child relationship. If teachers perceive children as having high levels of problem behaviors, even if this is not the case, it may impede relationship development and children's developmental progress. Moreover, rates of preschool expulsion are surprisingly high (Gilliam, 2005) and are likely based, in part, on teachers' perceptions of children as having behavior problems (Gilliam, Maupin, Reyes, Accavitti, & Shic, 2016; Martin, Bosk, & Bailey, 2018). If teacher turnover leads to perceptions of children as badly behaved by the new teacher, regardless of children's actual regulatory skills, it may still influence children's preschool experiences and broader developmental trajectory.

This highlights the importance of understanding *how* teacher turnover may influence children's development. Although it is beyond the scope of this study to identify the mechanisms by which turnover affects children, some candidates include the initial difference in classroom quality for children in classrooms where a teacher ultimately leaves; changes to the teacher-child bond upon teacher departure; the difficulty of quickly establishing a warm, positive climate for new teachers; the difficulty of managing a classroom full of children who have experienced teacher departure; and the loss of instructional time as the new teacher must organize his or her classroom and get to know each child in order to tailor experiences to his or her own developmental level. To get a better understanding of the potential role of the turning over teacher, we compared the sociodemographic characteristics of turning over teachers as

compared to staying teachers, and found no significant differences in terms of race, education, experience, or age (see Appendix A). This suggests that the associations between turnover and children's development is likely not only due to children's experiences with the turning over teacher, but also stem from the disruption of turnover itself and its aftermath. Understanding how turnover impacts children is an important next step for future research, and will be integral to efforts to reduce the potential effect of turnover on children.

Limitations

Although this study uses nationally representative Head Start data and a rigorous modeling strategy there are several factors that should be kept in mind when interpreting these findings. We describe four limitations here. First and foremost, these findings are not causal, but instead provide much-needed descriptive information. Though our models included a large set of child, family, teacher, and director covariates and account for both children's skills at Head Start entry and time-invariant program characteristics, it is likely that some omitted variables still exist. Specifically, it is possible that omitted center-level factors influence both children's learning and teacher turnover. We used program rather than center fixed effects because the center fixed effects models were underpowered, however we acknowledge that in doing so we did not account for time-invariant center characteristics, for example center climate, that may be important in this analysis.

Second, as noted above, this study measures behavioral challenges in the Head Start setting using teacher-reported behavior problems. This is a problem both because staying and turning over teachers are likely systematically different from each other, making the value of the lagged dependent variable questionable, and because teacher turnover is perfectly correlated with having a new rater. It is unclear how time in the classroom may influence teacher ratings. We argue that the present findings, while likely present an inflated effect size, nonetheless suggest that understanding the relationship between turnover and children's social and behavioral outcomes is important. Future research should consider getting external raters of children's skills to address this issue.

Third, the findings of this study are generalizable to Head Start only, not the breadth of formal,

center-based ECE. It seems likely that turnover may be lower in Head Start than in other ECE sectors because Head Start is a well-known program that may attract teachers with a specific desire to work in Head Start. Additionally, the effect of turnover may vary in Head Start as compared to other sectors. Recent policy shifts have required Head Start teachers to be more educated, which may result in higher-quality replacement teachers than other sectors, or teachers who are more able to cope with the challenges of entering a classroom mid-year. Future research should examine this question in other sectors, particularly the large child care sector which serves many children from historically-underserved groups.

Finally, although a strength of this study is that it estimated the association between within-year turnover and children's development, it is also likely the case that between-year turnover influences children, and that this study therefore somewhat understates the total damage of churn in the ECE workforce on young children's development. For example, in their study of turnover in K-12 settings, Ronfeldt and colleagues (2013) found that some of their estimated turnover effects were due to the disruptive nature of turnover on the teachers that remained at the school site; this is likely to be the case in ECE settings as well.

Implications and Conclusions

Head Start serves nearly about one million children annually, and remains the federal government's largest investment in the early educational experiences of our most vulnerable children. Findings from the present study suggests that a large number of Head Start children experience within-year teacher turnover, and that this turnover if meaningfully associated with their developmental outcomes. Estimated effect sizes in this paper were about 0.15 of a standard deviation, or about 14% of the income-based achievement gap in reading, 27% of the white-Hispanic achievement gap, and 47% of the white-Black achievement gap (Reardon & Portilla, 2016), suggesting that for policymakers hoping to leverage Head Start as a tool to support school readiness and close early gaps, policy addressing teacher turnover should be a priority.

Head Start policy over the past 20 years has been focused on improving the qualifications of the workforce. For example the 1998 Community Opportunities, Accountability, and Training and

Educational Services Act mandated that 50% of all Head Start teachers nationwide have at least an associate's degree (AA) by 2003, and the 2007 Improving Head Start for School Readiness Act required that 50% of Head Start teachers nationwide hold a bachelor's (BA) degree by 2013. These laws dramatically increased the education levels and specialization of Head Start teachers, two factors hypothesized to lead to teacher retention. While in 1999, only 38 percent of teachers had an AA degree or more; today 96% of Head Start teachers meet that criterion; and since 1999 the percentage of Head Start teachers with a BA has tripled to 73% (Barnett & Friedman-Krauss, 2016).

Despite this rather remarkable change in the composition of the Head Start workforce, as of the spring of 2010, roughly 100,000 Head Start children experienced a change in a caregiver over the course of the program year, and between-year turnover remains higher than 10%. Moreover, within-year turnover estimates were nearly double those of 2006 and 2009 in the 2014 wave of FACES (17%). It seems likely, then, that policymakers hoping to improve the stability of the ECE workforce will need to consider turnover reduction strategies outside of professionalization. More research is needed on what causes ECE teachers to turnover, but research from K-12 suggests that center climate and leadership may be candidate explanations (e.g. Grissom, 2011; Grissom, Viano, & Selin, 2016). Research in ECE specifically also highlights the role of compensation and benefits—particularly access to health insurance—in addition to these factors (Bullough, Hall-Kenyon, & MacKay, 2012; Claudia Hale-Jinks & Knopf, 2006; Russell, Williams, & Gleason-Gomez, 2010). Policy solutions that focus on building leadership capacity and center supports may be effective, as well as policies that consider equalizing pay for ECE teachers and K-12 teachers given similar training and preparation. Future research should work to uncover policy strategies that can reduce turnover while simultaneously supporting the quality of care children receive.

Additionally, given that some turnover will always exist in Head Start settings, it is important to consider how programs can facilitate a successful transition for incoming replacement teachers. Strategies such as transitioning qualified assistant teachers who are already familiar with the program and potentially the children may be beneficial. Providing an extra or more experienced assistant teacher to

mid-year replacements may also be promising. Providing new teachers with specific supports for the transition, such as curriculum or behavior management coaching may also be effective, as may holding additional opportunities for children and parents to get to know the replacement teacher (e.g. parent-teacher night, home visiting, etc.). Research that unpacks the mechanisms by which turnover impacts children will be important in designing program policy to address the effects of turnover on children. Moreover, there is currently no research on how to mitigate the effects of within-year turnover in the ECE setting; this study suggests that such research is warranted.

Conclusion. This paper provides new evidence on the extent to which within-year teacher turnover in early educational settings harms children enrolled in Head Start. We find a nontrivial portion of Head Start teachers leave during the program year, and that their departure is negatively and meaningfully associated with children's language, literacy, social, and behavioral self-regulation skills. This study suggests that additional investments in the Head Start workforce are warranted; and that research should help clarify how Head Start funds can be most efficiently targeted to reduce both turnover and its effects on young children.

References

Adnot, M., Dee, T., Katz, V., & Wyckoff, J. (2017). Teacher Turnover, Teacher Quality, and Student Achievement in DCPS. *Educational Evaluation and Policy Analysis*, *39*(1), 54–76. https://doi.org/10.3102/0162373716663646

Aikens, N., Bush, C., Gleason, P., Malone, L., & Tarullo, L. (2016). *Tracking Quality in Head Start Classrooms: FACES 2006 to FACES 2014 (Technical Report)*. Retrieved from Mathematica Policy Research, Inc website: https://www.mathematica-mpr.com/our-publications-and-findings/publications/tracking-quality-in-head-start-classrooms-faces-2006-to-faces-2014-technical-report

Araujo, M., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *The Quarterly Journal of Economics*, *131*(3), 1415–1453.

Barnett, W. S., & Friedman-Krauss, A. H. (2016). *State(s) of Head Start*. Rutgers Graduate School of Education: National Institute for Early Education Research.

Bassok, D., Fitzpatrick, M., Loeb, S., & Paglayan, A. (2013). The early childhood care and education workforce from 1990 through 2010. Changing dynamics and persistent concerns. *Education Finance and Policy*, *8*(4), 581–601.

Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., … Gill, S. (2008). Promoting Academic and Social-Emotional School Readiness: The Head Start REDI Program. *Child Development*, *79*(6), 1802–1817. https://doi.org/10.1111/j.1467-8624.2008.01227.x

Bowlby, J. (2008). *A Secure Base: Parent-Child Attachment and Healthy Human Development*. Basic Books.

Bronfenbrenner, U., & Morris, P. A. (2007). The Bioecological Model of Human Development. In *Handbook of Child Psychology*. https://doi.org/10.1002/9780470147658.chpsy0114

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

Bullough, R. V., Hall-Kenyon, K. M., & MacKay, K. L. (2012). Head Start Teacher Well-Being: Implications for Policy and Practice. *Early Childhood Education Journal*, *40*(6), 323–331. https://doi.org/10.1007/s10643-012-0535-8

Burchinal, M. R., Peisner-Feinberg, E., & Clifford, R. M. (2000). Children's social and cognitive development and child care quality: Testing for differential associations related to poverty, gender, or ethnicity. *Applied Developmental Science*, *4*(3), 149–165.

Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early Childhood Education: Young Adult Outcomes From the Abecedarian Project. *Applied Developmental Science*, *6*(1), 42–57. https://doi.org/10.1207/S1532480XADS0601_05

Chandler, M. A. (2017, March 31). D.C. among first in nation to require child-care workers to get college degrees. *Washington Post*. Retrieved from https://www.washingtonpost.com/local/social-issues/district-among-the-first-in-nation-to-require-child-care-workers-to-get-college-degrees/2017/03/30/d7d59e18-0fe9-11e7-9d5a-a83e627dc120_story.html

Cryer, D., Hurwitz, S., & Wolery, M. (2000). Continuity of caregiver for infants and toddlers in center-based child care: Report on a survey of center practices. *Early Childhood Research Quarterly*, *15*(4), 497–514. https://doi.org/10.1016/S0885-2006(01)00069-2

Curby, T. W., Rimm-Kaufman, S. E., & Abry, T. (2013). Do emotional support and classroom organization earlier in the year set the stage for higher quality instruction? *Journal of School Psychology*, *51*(5), 557–569. https://doi.org/10.1016/j.jsp.2013.06.001

Deming, D. (2009). Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start. *American Economic Journal: Applied Economics*, *1*(3), 111–134.

Dunn, L., & Dunn, D. (2007). *PPVT-4: Peabody picture vocabulary test*.

Dunn, L. M., & Dunn, D. M. (2013). *Peabody picture vocabulary test, technical report*. Retrieved from htt p://images. pearsonassessments. com/images/assets/ppvt-4/2013-PPVT-Tech-RPT. pdf.

Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., … Zill, N. (2007). Teachers' Education, Classroom Quality, and Young Children's Academic Skills: Results From

Seven Studies of Preschool Programs. *Child Development*, *78*(2), 558–580. https://doi.org/10.1111/j.1467-8624.2007.01014.x

Elicker, J., & Fortner-Wood, C. (1995). Adult-child relationships in early childhood programs. *Young Children*, *51*, 69–78.

Entwisle, D., & Alexander, K. (1987). Long-term effects of Cesarean delivery on parents' beliefs and children's schooling. *Developmental Psychology*, *23*(5), 676–682. https://doi.org/10.1037/0012-1649.23.5.676

Friedman-Krauss, A., Barnett, W. S., Weisenfeld, G., Kasmin, R., DiCrecchio, N., & Horowitz, M. (2018). *The State of Preschool 2017*. New Brunswick, NJ: National Institute for Early Education Research.

Gilliam, W. S. (2005). *Prekindergarteners left behind: expulsion rates in state prekindergarten systems*. Retrieved from http://www.fcd-us.org/PDFs/NationalPreKExpulsionPaper03.02_new.pdf

Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., & Shic, F. (2016). *Do early educators' implicit biases regarding sex and race relate to behavior expectations and recommendations of preschool expulsions and suspensions.* [Research Study Brief]. New Haven, CT: Yale Child Study Center.

Gresham, F., & Elliott, S. (1990). *Social skills rating system: Manual.* American Guidance Services.

Grissom, J. A. (2011). Can good principals keep teachers in disadvantaged schools? Linking principal effectiveness to teacher satisfaction and turnover in hard-to-staff environments. *Teachers College Record*, *113*(11), 2552–2585.

Grissom, J. A., Viano, S. L., & Selin, J. L. (2016). Understanding Employee Turnover in the Public Sector: Insights from Research on Teacher Mobility. *Public Administration Review*, *76*(2), 241–251. https://doi.org/10.1111/puar.12435

Guin, K. (2004). Chronic Teacher Turnover in Urban Elementary Schools. *Education Policy Analysis Archives*, *12*(0), 42. https://doi.org/10.14507/epaa.v12n42.2004

Hale-Jinks, C, Knopf, H., & Kemple, K. (2006). Tackling teacher turnover in child care: Understanding causes and consequences, identifying solutions. *Childhood Education*, *82*(4), 219–226.

Hale-Jinks, Claudia, & Knopf, H. (2006). Tackling Teacher Turnover in Child Care: Understanding Causes and Consequences, Identifying Solutions. *Childhood Education*, *82*(4), 219–226. https://doi.org/10.1080/00094056.2006.10522826

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. https://doi.org/10.1111/cdev.12184

Hamre, B. K. (2014). Teachers' Daily Interactions With Children: An Essential Ingredient in Effective Early Childhood Programs. *Child Development Perspectives*, *8*(4), 223–230. https://doi.org/10.1111/cdep.12090

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*, *76*(5), 949–967. https://doi.org/10.1111/j.1467-8624.2005.00889.x

Hatfield, B. E., Hestenes, L. L., Kintner-Duffy, V. L., & O'Brien, M. (2013). Classroom Emotional Support predicts differences in preschool children's cortisol and alpha-amylase levels. *Early Childhood Research Quarterly*, *28*(2), 347–356. https://doi.org/10.1016/j.ecresq.2012.08.001

Helburn, S. W. (1995). *Cost, Quality and Child Outcomes in Child Care Centers. Technical Report, Public Report, and Executive Summary*. Retrieved from https://eric.ed.gov/?id=ed386297

Howes, C, Burchinal, M., Pianta, R. C., Bryant, D., Early, D. M., & Clifford, R. (2008). Ready to learn? Children's preacademic achievement in pre-kindergarten programs. *Early Childhood Research Quarterly*, *23*, 27–50.

Howes, Carollee. (1999). Attachment relationships in the context of multiple caregivers. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (pp. 671–687). New York, NY: Guilford Press.

Howes, Carollee, & Hamilton, C. E. (1992). Children's Relationships with Child Care Teachers: Stability and Concordance with Parental Attachments. *Child Development*, *63*(4), 867–878. https://doi.org/10.1111/j.1467-8624.1992.tb01667.x

Institute of Medicine. (2015). *Transforming the workforce for children, birth through age 8: A unifying foundation.* Washington DC: National Academies Press.

Interlandi, J. (2018, January 9). Why Are Our Most Important Teachers Paid the Least? *The New York Times*. Retrieved from https://www.nytimes.com/2018/01/09/magazine/why-are-our-most-important-teachers-paid-the-least.html

Johnson, A. D., Markowitz, A. J., Hill, C., J., & Phillips, D. A. (2016). Variation in impacts of Tulsa pre-k on cognitive development in kindergarten: The role of Instructional Support. *Developmental Psychology*, *52*(12), 2145.

Johnson, D. R., & Young, R. (2011). Toward Best Practices in Analyzing Datasets with Missing Data: Comparisons and Recommendations. *Journal of Marriage and Family*, *73*(5), 926–945. https://doi.org/10.1111/j.1741-3737.2011.00861.x

Magnuson, K., & Waldfogel, J. (2016). Trends in Income-Related Gaps in Enrollment in Early Childhood Education. *AERA Open*, *2*(2), 2332858416648933. https://doi.org/10.1177/2332858416648933

Markowitz, A. J., Bassok, D., & Hamre, B. (2018). Leveraging Developmental Insights to Improve Early Childhood Education. *Child Development Perspectives*, *12*(2), 87–92. https://doi.org/10.1111/cdep.12266

Martin, K. A., Bosk, E., & Bailey, D. (2018). Teachers' Perceptions of Childcare and Preschool Expulsion. *Children & Society*, *32*(2), 87–97. https://doi.org/10.1111/chso.12228

McGrew, K., & Woodcock, R. (2001). *Woodcock-Johnson Tehcnical Manual.* Itasca, IL: Riverside Publishing.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., … Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety.

Proceedings of the National Academy of Sciences, *108*(7), 2693–2698. https://doi.org/10.1073/pnas.1010076108

Mooney, N. (2016, October 4). Is this as good as childcare gets? *The Atlantic*. Retrieved from https://www.theatlantic.com/education/archive/2016/10/is-this-as-good-as-childcare-gets/502754/

Moss, E., Parent, S., Gosselin, C., Rousseau, D., & St-Laurent, D. (1996). Attachment and teacher-reported behavior problems during the preschool and early school-age period. *Development and Psychopathology*, *8*(3), 511–525. https://doi.org/10.1017/S0954579400007240

National Institute of Child Health and Human Development Early Child Care Research Network, & Duncan, G. J. (2003). Modeling the Impacts of Child Care Quality on Children's Preschool Cognitive Development. *Child Development*, *74*(5), 1454–1475. https://doi.org/10.1111/1467-8624.00617

Paquette, D. (2016, July). Half of America's childcare workers need food stamps, welfare payments, or Medicaid. *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/wonk/wp/2016/07/11/the-people-taking-care-of-our-kids-live-in-poverty/

Phillips, D., Lipsey, M., Dodge, K., Haskins, R., Bassok, D., Burchinal, M., … Weiland, C. (2017). *Puzzling it out: The current state of scientific knowledge of pre-kindergarten effects. A consensus statement.* Washington DC: Brookings Institution.

Phillips, D., Mekos, D., Scarr, S., McCartney, K., & Abbott–Shim, M. (2000). Within and beyond the classroom door: assessing quality in child care centers. *Early Childhood Research Quarterly*, *15*(4), 475–496. https://doi.org/10.1016/S0885-2006(01)00077-1

Pianta, R. C. (2007). Teacher-child relationships and early literacy. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of Early Literacy Research, Volume 2.* (pp. 149–162). New York, NY: Guilford Press.

Pianta, Robert C., Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (2008). Effects of web-mediated professional development resources on teacher–child interactions in pre-kindergarten

classrooms. *Early Childhood Research Quarterly*, *23*(4), 431–451. https://doi.org/10.1016/j.ecresq.2008.02.001

Pilarz, A. R., & Hill, H. D. (2014). Unstable and multiple child care arrangements and young children's behavior. *Early Childhood Research Quarterly*, *29*(4), 471–483. https://doi.org/10.1016/j.ecresq.2014.05.007

Powell, D., Diamond, K., Burchinal, M., & Koehler, M. (2010). Effects of an early literacy professional development intervention on head start teachers and children. *Journal of Educational Psychology*, *102*(2), 299–312. https://doi.org/10.1037/a0017763

QRIS Compendium. (2017). *What are the most common types of indicators included in QRIS?* Retrieved from qriscompendium.org/top-ten/question-4

Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement*, *1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Raikes, H. (1993). Relationship duration in infant care: Time with a high-ability teacher and infant-teacher attachment. *Early Childhood Research Quarterly*, *8*(3), 309–325. https://doi.org/10.1016/S0885-2006(05)80070-5

Reardon, S. F., & Portilla, X. A. (2016). Recent Trends in Income, Racial, and Ethnic School Readiness Gaps at Kindergarten Entry. *AERA Open*, *2*(3), 2332858416657343. https://doi.org/10.1177/2332858416657343

Redding, C., & Henry, G. T. (2018). New Evidence on the Frequency of Teacher Turnover: Accounting for Within-Year Turnover. *Educational Researcher*, *47*(9), 577–593. https://doi.org/10.3102/0013189X18814450

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, *50*(1), 4–36.

Rudasill, K. M., & Rimm-Kaufman, S. E. (2009). Teacher–child relationship quality: The roles of child temperament and teacher–child interactions. *Early Childhood Research Quarterly*, *24*(2), 107–120. https://doi.org/10.1016/j.ecresq.2008.12.003

Russell, E. M., Williams, S. W., & Gleason-Gomez, C. (2010). Teachers' Perceptions of Administrative Support and Antecedents of Turnover. *Journal of Research in Childhood Education*, *24*(3), 195–208. https://doi.org/10.1080/02568543.2010.487397

Schrank, F., McGrew, K., Ruef, M., Alvarado, C., Munoz-Sandoval, A., & Woodcock, R. (2005). *Overview and Technical Supplement Woodcock Johnson*. Itasca, IL: Riverside Publishing.

Shonkoff, J., & Phillips, D. (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. National Academies Press.

Tran, H., & Winsler, A. (2011). Teacher and center stability and school readiness among low-income, ethnically diverse children in subsidized, center-based child care. *Children and Youth Services Review*, *33*(11), 2241–2252.

Verschueren, K., & Koomen, H. M. Y. (2012). Teacher–child relationships from an attachment perspective. *Attachment & Human Development*, *14*(3), 205–211. https://doi.org/10.1080/14616734.2012.672260

Von Hippel, P. T. (2007). Regression with Missing Ys: An Improved Strategy for Analyzing Multiply Imputed Data. *Sociological Methodology*, *37*(1), 83–117. https://doi.org/10.1111/j.1467-9531.2007.00180.x

Walter, M. C., & Lippard, C. N. (2017). Head Start Teachers Across a Decade: Beliefs, Characteristics, and Time Spent on Academics. *Early Childhood Education Journal*, *45*(5), 693–702. https://doi.org/10.1007/s10643-016-0804-z

Weiland, C., Ulvestad, K., Sachs, J., & Yoshikawa, H. (2013). 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. https://doi.org/10.1016/j.ecresq.2012.12.002

Weiland, C., & Yoshikawa, H. (2013). Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills. *Child Development*, *84*, 2112–2130. https://doi.org/10.1111/cdev.12099

Wells, M. B. (2015). Predicting preschool teacher retention and turnover in newly hired Head Start teachers across the first half of the school year. *Early Childhood Research Quarterly*, *30*, 152–159. https://doi.org/10.1016/j.ecresq.2014.10.003

West, J., Tarullo, L., Aikens, N., Malone, L., & Carlson, B. (2011). *FACES 2009 study design*. Retrieved from Administration for Children and Families website: http://www.researchconnections.org/childcare/resources/21553

Whitebook, M., Howes, C., & Phillips, D. (1998). *Worthy Work, Unlivable Wages: The National Child Care Staffing Study, 1988-1997*. Retrieved from https://eric.ed.gov/?id=ED419614

Whitebook, M., Phillips, D., & Howes, C. (2014a). *Worthy work, STILL unlivable wages: The early childhood workforce 25 years after the National Child Care Staffing Study*. Berkeley, CA: Center for the Study of Child Care Employment, University of California, Berkeley.

Whitebook, M., Phillips, D., & Howes, C. (2014b). *Worthy Work, STILL Unlivable Wages: The Early Childhood Workforce 25 Years after the National Child Care Staffing Study* (pp. 1–124). Retrieved from Center for the Study of Child Care Employment, University of California, Berkeley website: http://www.irle.berkeley.edu/cscce/2014/report-worthy-work-still-unlivable-wages/

Whitebook, M., & Sakai, L. (2003). Turnover begets turnover: An examination of job and occupational instability among child care center staff. *Early Childhood Research Quarterly*, *18*(3), 273–293.

Whittaker, J., Kinzie, M. B., Williford, A., & DeCoster, J. (2016). Effects of MyTeachingPartner–Math/Science on Teacher–Child Interactions in Prekindergarten Classrooms. *Early Education and Development*, *27*(1), 110–127.

Woodcock, R., Mather, N., & McGrew, K. (2001). *Woodcock-Johnson III tests of cognitive abilities examiner's manual*. Riverside: Itasca.

Zill, N., Moore, K., Smith, E., & Stief, T. (1991). *Welfare mothers as potential employees: A statistical profile based on national survey data.* Washington, D.C.: Child Trends.

Zill, N., & Peterson, J. (1986). *Behavior Problems Index*. Washington, DC: Child Trends.

Table 1. Bivariate associations between within-year teacher turnover and child, family, teacher, and director characteristics.

					Bivariate Assn	
Developmental Outcomes	N	Turnover=0	N	Turnover=1	Btwn	Within
Spring PPVT	4708	51.34	501	49.91		
		(24.03)		(24.07)		
Fall PPVT	4948	37.76	528	36.56		
		(23.31)		(22.13)		
Spring WJ Letter Word	4431	6.39	479	5.85	*	**
		(4.48)		(4.52)		
Fall WJ Letter Word	4948	3.41	528	3.54		
		(3.31)		(3.31)		
Spring WJ Spelling	4436	6.45	481	6.41		**
		(3.17)		(3.23)		
Fall WJ Spelling	4948	4.42	528	4.65		
		(2.64)		(2.80)		
Spring WJ Applied Problems	4435	7.98	480	7.90		
		(4.65)		(4.59)		
Fall WJ Applied Problems	4948	4.91	528	4.77		
		(4.43)		(4.27)		
Spring Parent BPI	4488	5.29	487	5.64		
		(3.51)		(3.90)		
Fall Parent BPI	4948	5.53	528	5.74		
		(3.52)		(3.71)		
Spring Parent Positive Behavior	4497	12.32	489	12.37		
		(2.46)		(2.55)		
Fall Parent Positive Behavior	4948	11.97	528	12.01		
		(2.55)		(2.67)		
Spring Teacher BPI	4787	5.17	510	6.25	*	*
		(5.46)		(6.06)		
Fall Teacher BPI	4948	5.61	528	5.99		
		(5.37)		(5.53)		
Spring Teacher Social	4784	17.37	509	16.66		*
		(4.57)		(5.00)		
Fall Teacher Social	4948	15.40	528	14.96		+
		(4.74)		(5.16)		
Spring Teacher Aggress	4786	1.40	510	1.62	+	+
		(1.86)		(2.00)		
Fall Teacher Aggress	4948	1.45	528	1.70	*	*
		(1.84)		(2.07)		
Spring Teach Withdrawn	4786	1.39	510	1.78	+	+
		(1.88)		(2.08)		

Fall Teacher Withdrawn	4948	1.41	528	1.54		
		(1.86)		(1.97)		
Child Race						
White	4948	0.22	528	0.25		
Black	4948	0.35	528	0.26		
Hispanic	4948	0.36	528	0.37		
Other race	4948	0.08	528	0.13		
Female	4948	0.51	528	0.48		
Child has disability	4948	0.04	528	0.05		
Age at assessment	4948	53.28	528	53.75		
Child ESL	4948	0.26	528	0.34		
Family Income						
0-10k	4948	0.14	528	0.15		
10,001- 20k	4948	0.41	528	0.45		
20,001- 30k	4948	0.24	528	0.23		
30,001- 40k	4948	0.10	528	0.10		
40,001- 50k	4948	0.05	528	0.03	*	*
50,001 +	4948	0.06	528	0.05		
Receives public benefits	4948	0.60	528	0.62		
Mother's Education						
Less than HS	4948	0.35	528	0.46	*	+
HS	4948	0.33	528	0.29		
Some college	4948	0.25	528	0.18	**	*
BA or more	4948	0.06	528	0.07		
Mother unemployed	4948	0.18	528	0.17		
Single parent household	4948	0.69	528	0.68		
Mother's age at child's birth	4948	20.72	528	20.79		
Maternal depressive symptoms	4948	5.08	528	5.06		
Parent reads to child 3x/week	4948	0.74	528	0.79	+	
Parent is immigrant	4948	0.34	528	0.41		
Cohort	4948	0.42	528	0.48		
Year	4948	0.53	528	0.52		
Teacher Race						
White	4948	0.40	528	0.42		
Black	4948	0.33	528	0.22		
Hispanic	4948	0.21	528	0.17		
Other race	4948	0.06	528	0.19	+	
Teacher Education						
< AA	4948	0.15	528	0.22		
AA	4948	0.37	528	0.29		
BA or more	4948	0.48	528	0.49		
Teacher experience	4948	13.48	528	9.92	**	*
Teacher age	4948	40.84	528	39.70		

Class size	4948	17.57	528	17.02		
Full day class	4948	0.56	528	0.48		
Teaches two classes daily	4948	0.14	528	0.43	**	**
Director Education						
< AA	4948	0.05	528	0.04		
AA	4948	0.18	528	0.12		
BA	4948	0.35	528	0.47		
More than BA	4948	0.43	528	0.37		
Director Race						
White	4948	0.43	528	0.36		
Black	4948	0.37	528	0.29		
Hispanic	4948	0.17	528	0.22		
Other race	4948	0.03	528	0.13	+	+
Director is NAEYC member	4948	0.74	528	0.63		+
Director experience	4948	13.44	528	12.26		

Note. Data are drawn from the 2006 and 2009 waves of the Head Start Family and Child Experiences Survey (FACES). Means are calculated across 15 multiply imputed datasets and weighted using FACES provided sampling weights, imputed N= 5476. Standard deviations are in parentheses for continuous measures of children's developmental outcomes. Bivariate association columns indicate whether differences in means by within-year teacher turnover status are significant across programs and within programs. Teacher demographic characteristics are taken from the spring teacher; thus in the no turnover group, teacher characteristics are consistent across the year, for the turnover group teacher characteristics reflect the child's second teacher. ** $p< 0.01$, * $p< 0.05$, + $p< 0.10$

Table 2. Associations between within-year teacher turner and children's developmental outcomes.

		OLS		OLS, Lag		Program FE		Program FE, Lag	
	N	β	*p*	β	*p*	β	*p*	β	*p*
PPVT	5209	-0.02	0.64	-0.01	0.75	-0.03	0.53	-0.03	0.47
		(0.05)		(0.04)		(0.05)		(0.05)	
WJ Letter Word	4910	-0.08	0.18	-0.10	0.03	-0.10	0.05	-0.08	0.04
		(0.06)		(0.05)		(0.05)		(0.04)	
WJ Spelling	4917	-0.06	0.29	-0.07	0.20	-0.13	0.02	-0.10	0.06
		(0.06)		(0.05)		(0.05)		(0.05)	
WJ Applied Problems	4915	-0.04	0.47	-0.02	0.70	-0.06	0.28	-0.03	0.43
		(0.06)		(0.04)		(0.05)		(0.04)	
Parent Reported Behavior									
Behavior Problems (BPI)	4975	0.04	0.56	0.05	0.32	0.06	0.38	0.07	0.21
		(0.06)		(0.05)		(0.07)		(0.05)	
Positive Behaviors	4986	0.04	0.45	0.03	0.47	0.06	0.31	0.05	0.28
		(0.06)		(0.04)		(0.06)		(0.04)	
Teacher Reported Behavior									
Behavior Problems (BPI)	5297	0.22	0.01	0.15	0.04	0.23	0.00	0.15	0.04
		(0.08)		(0.07)		(0.08)		(0.07)	
Social Skills	5293	-0.22	0.02	-0.14	0.14	-0.24	0.01	-0.14	0.13
		(0.09)		(0.10)		(0.09)		(0.09)	
Aggressive Behavior	5296	0.17	0.00	0.07	0.24	0.15	0.02	0.05	0.51
		(0.06)		(0.06)		(0.06)		(0.07)	
Withdrawn	5296	0.22	0.02	0.17	0.02	0.21	0.02	0.16	0.02
		(0.09)		(0.07)		(0.09)		(0.07)	

Note. Data are drawn from the 2006 and 2009 waves of the Head Start Family and Child Experiences Survey (FACES). Estimates are calculated across 15 multiply imputed datasets and weighted using FACES provided sampling weights, total imputed N= 5476. Standard errors are in parentheses. The addition of "Lag" above indicates the addition of a measure of the dependent variable in the fall of preschool in the model. All models include a full set of child, family, teacher and director covariates. ** $p< 0.01$, * $p< 0.05$, + $p< 0.10$

Appendix A. Demographic characteristics of teachers who turned over versus the fall characteristics of teachers who did not turn over.

Teacher Race	Turned Over	Stayed
White	0.37	0.40
Black	0.21	0.33
Hispanic	0.29	0.21
Other Race	0.13	0.06
Teacher Education		
Less than AA	0.27	0.15
AA	0.39	0.37
BA or more	0.34	0.48
CDA	0.49	0.51
Experience	11.60	12.47
Age	39.29	40.80

Appendix B. Associations between within-year teacher turner and children's developmental outcomes, including children who switched classrooms.

		OLS		OLS, Lag		Program FE		Program FE, Lag	
	N	β	*p*	β	*p*	β	*p*	β	*p*
PPVT	5353	-0.02	0.63	-0.01	0.75	-0.03	0.53	-0.03	0.48
		(0.05)		(0.04)		(0.05)		(0.05)	
WJ Letter Word	5054	-0.08	0.17	-0.10	0.03	-0.10	0.06	-0.08	0.04
		(0.06)		(0.05)		(0.05)		(0.04)	
WJ Spelling	5061	-0.06	0.31	-0.07	0.22	-0.12	0.02	-0.10	0.07
		(0.06)		(0.05)		(0.05)		(0.05)	
WJ Applied Problems	5059	-0.04	0.46	-0.02	0.69	-0.05	0.30	-0.03	0.48
		(0.05)		(0.04)		(0.05)		(0.04)	
Parent Reported Behavior									
Behavior Problems (BPI)	5111	0.04	0.55	0.04	0.35	0.06	0.34	0.07	0.22
		(0.06)		(0.05)		(0.07)		(0.05)	
Positive Behaviors	5123	0.04	0.45	0.03	0.45	0.06	0.31	0.05	0.24
		(0.05)		(0.04)		(0.06)		(0.04)	
Teacher Reported Behavior									
Behavior Problems (BPI)	5429	0.22	0.01	0.15	0.03	0.23	0.00	0.15	0.04
		(0.08)		(0.07)		(0.07)		(0.07)	
Social Skills	5424	-0.22	0.02	-0.14	0.15	-0.24	0.01	-0.14	0.13
		(0.09)		(0.10)		(0.09)		(0.09)	
Aggressive Behavior	5428	0.17	0.00	0.07	0.23	0.15	0.02	0.05	0.48
		(0.06)		(0.06)		(0.06)		(0.07)	
Withdrawn	5428	0.22	0.02	0.17	0.02	0.21	0.03	0.16	0.02
		(0.09)		(0.07)		(0.09)		(0.07)	

Note. Data are drawn from the 2006 and 2009 waves of the Head Start Family and Child Experiences Survey (FACES). Estimates are calculated across 15 multiply imputed datasets and weighted using FACES provided sampling weights, total imputed N= 5476. Standard errors are in parentheses. The addition of "Lag" above indicates the addition of a measure of the dependent variable in the fall of preschool in the model. All models include a full set of child, family, teacher and director covariates. ** $p< 0.01$, * $p< 0.05$, + $p< 0.10$

Appendix C. Associations between within-year teacher turner and teacher-reported problem behavior, using parent-reported behavior in the fall as a covariate.

		OLS, BPI Lag		OLS, Pos Beh Lag		Prog FE, BPI Lag		Prog FE, Pos Beh Lag	
Teacher Reported Behavior	N	β	*p*	β	*p*	β	*p*	β	*p*
Behavior Problems (BPI)	5297	0.22	0.01	0.22	0.01	0.23	0.00	0.24	0.00
		(0.08)		(0.08)		(0.08)		(0.08)	
Social Skills	5293	-0.22	0.02	-0.22	0.02	-0.24	0.01	-0.25	0.01
		(0.09)		(0.10)		(0.09)		(0.09)	
Aggressive Behavior	5296	0.17	0.00	0.17	0.00	0.15	0.02	0.15	0.02
		(0.06)		(0.06)		(0.07)		(0.06)	
Withdrawn	5296	0.22	0.02	0.22	0.02	0.21	0.03	0.21	0.02
		(0.09)		(0.09)		(0.09)		(0.09)	

Note. Data are drawn from the 2006 and 2009 waves of the Head Start Family and Child Experiences Survey (FACES). Estimates are calculated across 15 multiply imputed datasets and weighted using FACES provided sampling weights, total imputed N= 5476. Standard errors are in parentheses. "BPI Lag" indicates the addition of the fall measure of parent-reported behavior problems as a covariate; "Pos Beh Lag" indicates the inclusion of the fall measure of parent-reported positive behaviors as a covariate. All models include a full set of child, family, teacher and director covariates.

** $p< 0.01$, * $p< 0.05$, + $p< 0.10$

Appendix D. Associations between within-year teacher turner and children's developmental outcomes, 2006-2014.

		OLS		OLS, Lag		Program FE		Program FE, Lag	
	N	β	*p*	β	*p*	β	*p*	β	*p*
PPVT	7064	-0.04	0.37	-0.01	0.79	-0.06	0.15	-0.05	0.20
		(0.05)		(0.04)		(0.04)		(0.04)	
WJ Letter Word	6677	-0.13	0.04	-0.10	0.02	-0.21	0.00	-0.15	0.00
		(0.06)		(0.04)		(0.07)		(0.04)	
WJ Spelling	6688	-0.07	0.25	-0.06	0.28	-0.16	0.01	-0.14	0.01
		(0.06)		(0.05)		(0.06)		(0.05)	
WJ Applied Problems	6690	-0.01	0.87	0.03	0.50	-0.03	0.51	-0.01	0.84
		(0.05)		(0.04)		(0.04)		(0.04)	
Teacher Reported Behavior									
Behavior Problems (BPI)	7208	0.22	0.00	0.13	0.05	0.27	0.00	0.17	0.02
		(0.07)		(0.07)		(0.07)		(0.07)	
Social Skills	7201	-0.19	0.03	-0.13	0.11	-0.25	0.00	-0.16	0.05
		(0.09)		(0.08)		(0.08)		(0.08)	
Aggressive Behavior	7207	0.19	0.00	0.08	0.12	0.19	0.00	0.10	0.13
		(0.05)		(0.05)		(0.07)		(0.06)	
Withdrawn	7207	0.22	0.01	0.18	0.01	0.25	0.01	0.20	0.01
		(0.08)		(0.07)		(0.09)		(0.07)	

Note. Data are drawn from the 2006, 2009, and 2014 waves of the Head Start Family and Child Experiences Survey (FACES). Estimates are calculated across 15 multiply imputed datasets and weighted using FACES provided sampling weights. Standard errors are in parentheses. The addition of "Lag" above indicates the addition of a measure of the dependent variable in the fall of preschool in the model. All models include a full set of child, family, teacher and director covariates, including a child level indicator for whether it was the child's first year in Head Start.
** $p< 0.01$, * $p< 0.05$, + $p< 0.10$