First-year maternal employment and child outcomes: Differences across racial and ethnic groups[*]

Lawrence Berger[a,*], **Jeanne Brooks-Gunn**[b], **Christina Paxson**[c], and **Jane Waldfogel**[d]

[a]University of Wisconsin, USA

[b]Columbia University, USA

[c]Princeton University, USA

[d]Columbia University, USA

Abstract

We use data from the Fragile Families and Child Wellbeing Study to examine associations between first-year maternal employment and child outcomes for 3-year-old White, Black, and Hispanic children (*N*=1483). Results from OLS regressions and propensity score matching models indicate that first-year maternal employment is associated with lower vocabulary scores for White, but not Black or Hispanic, children and with elevated levels of behavior problems for Hispanic, but not White or Black, children. Factors such as type of child care, maternal depressive symptoms and stress, and parenting behaviors (including measures of discipline, nurturance, and provision of cognitively stimulating materials) do not mediate these associations between first-year maternal employment and children's outcomes or explain the differential associations across racial and ethnic groups, suggesting the need to look at other explanations for these associations, as well as the need for better measurement of parenting, especially mother–child interaction.

Keywords

Maternal employment; Child outcomes

1. Introduction

Over the past few decades, a growing body of research has examined the associations between early maternal employment and child outcomes. Several studies have found that children whose mothers work in the first year of their lives have lower cognitive scores and more behavior problems than children whose mothers do not (see, most recently, Hill, Waldfogel, Brooks-Gunn, & Han, 2005). However, these findings have varied considerably by family characteristics, raising the question of whether associations between first-year maternal employment and child outcomes are the same across different types of families. In particular, negative associations between first-year maternal employment and child outcomes have been found more often in non-Hispanic White families than in non-Hispanic Black and Hispanic

*Corresponding author. The National Board of Health and Welfare, SE 106 30 Stockholm, Sweden. Tel.: +46 8 555 534 35; fax: +46 8 555 532 24. pia.kyhle.westermark@sos.se (L. Berger).

families. However, a common challenge in prior research is that datasets have tended to contain small samples of racial and ethnic minority families, making it difficult to estimate these associations with confidence for sub-groups of interest.

In this study, we examine associations between first-year maternal employment and two child outcomes – receptive vocabulary size and number of behavior problems – in a sample that includes large numbers of White, Black, and Hispanic families. Our sample is drawn from the Fragile Families and Child Wellbeing Study, a birth cohort study of children born to predominantly low income and single mothers (although some higher-income and married couple families are included) in 20 medium to large U.S. cities. The sample is racially and ethnically diverse: just over half of the sample is (non-Hispanic) Black, one fifth is Hispanic, and the other fifth is (non-Hispanic) White. This sample allows us to focus our analyses on differences in the associations between first-year maternal employment and child outcomes across these three racial and ethnic groups.

We begin by estimating associations of maternal employment in the first year of children's lives with measures of vocabulary and behavior problems when children are approximately three years old. We then test whether these associations differ across racial/ethnic groups by estimating models that allow first-year maternal employment to be interacted with race/ethnicity, thereby providing separate estimates of associations of first-year maternal employment with vocabulary and behavior problems for (non-Hispanic) White, (non-Hispanic) Black, and Hispanic children. These analyses extend prior work by examining a sample that is considerably more racially and ethnically diverse than has been used in the past. A second way in which the present study extends prior work is that we use propensity score matching models to better control for differences that may exist between families where the mother works by 12 months and families where she does not. Although propensity score matching models are an increasingly common approach used to adjust for selection factors in empirical analyses, they have (with the exception of one prior study) not been applied to this topic. Third, we extend prior work by examining several factors that might mediate the associations between first-year maternal employment and later child outcomes and that might explain the differential associations between first-year employment and child outcomes across racial/ethnic groups.

2. Background, prior research, and conceptual framework

Many studies have examined the associations between first-year maternal employment and children's later cognitive and behavioral outcomes (see reviews in Shonkoff & Phillips, 2000; Smolensky & Gootman, 2003). Such studies have typically found that children score more poorly on cognitive tests and/or measures of behavior problems if their mothers worked in the first year of their life (Bates et al., 1994; Baum, 2003; Baydar & Brooks-Gunn, 1991; Belsky & Eggebeen, 1991; Blau & Grossberg, 1992; Brooks-Gunn, Han, & Waldfogel, 2002; Desai, Chase-Lansdale, & Michael, 1989; Han, Waldfogel, & Brooks-Gunn, 2001; Haskins, 1985; Hill et al., 2005; James-Burdumy, 2005; Neidell, 2000; Ruhm, 2004; Waldfogel, Han, & Brooks-Gunn, 2002).

Most of these prior studies have focused on White and Black children, excluding Hispanic children due to small sample sizes. A common finding is that negative associations between first-year maternal employment and cognitive outcomes are observed for White children, but not for Black children. For instance, an early study using data on non-Hispanic children from the National Longitudinal Survey of Youth-Child Supplement (NLSY-CS) found that maternal employment in the first year of a child's life was negatively and significantly associated with cognitive outcomes at age three or four for White children but not Black children (Baydar & Brooks-Gunn, 1991). A follow-up study that also excluded Hispanic children analyzed outcomes for children from the NLSY-CS as late as age seven or eight and found that maternal

employment in the first year of a child's life continued to be significantly and negatively associated with cognitive outcomes for White children but not Black children (Han et al., 2001). In addition, the latter study also reported some adverse associations between maternal employment in the first year and behavioral problems as late as age seven or eight but, again, these associations were found only for White children and not Black children (Han et al., 2001).

Only a few studies to date have examined the associations between first-year maternal employment and children's later cognitive or behavioral outcomes for Hispanic children as well as non-Hispanic White and Black children. The first used data from the NLSY-CS and found persistent negative associations between first-year maternal employment and cognitive outcomes for White children but not for African American or Hispanic children (Waldfogel et al., 2002). A second study also used data from the NLSY-CS and estimated associations between maternal employment and children's cognitive and behavioral development for white, Black, and Hispanic children; however, this study concluded that the small samples of Black and Hispanic children made it impossible to determine whether the associations between first-year maternal employment and later child outcomes differed between Black and Hispanic children, or between these two groups of children and White children (Hill et al., 2005). Small sample sizes have also limited analyses using the National Institute of Child Health and Human Development Study of Early Child Care (NICHD SECC). The one study to date on this topic that used data from the NICHD SECC found significant negative associations between first-year maternal employment and later cognitive outcomes for non-Hispanic White children (N=900) but not for the small sample of non-Hispanic Black or Hispanic children (N=174 combined) (Brooks-Gunn et al., 2002).

Thus, prior research has not established conclusively whether the associations between first-year maternal employment and children's later outcomes differ for White, Black, and Hispanic children. We provide new evidence on this question by estimating regression models that include interaction terms that test for whether the association between first-year maternal employment and child outcomes differs for children from different racial and ethnic groups.

Prior research has also not established *why* first-year maternal employment might be differentially associated with child outcomes across White, Black, and Hispanic families. One possible explanation for variation in these associations across racial and ethnic groups is differential selection into employment. If, for example, White women who do not work in the first year after a birth tend to be primarily affluent married women, then those who do work by 12 months post-birth may be *negatively* selected. This might contrast with the situation for Black families, where those who do not work may disproportionately be low-income single mothers, in which case those who worked by 12 months post-birth would be *positively* selected. To address this possibility, we include extensive controls for work history and maternal characteristics, including a measure of the mother's cognitive development, in all of our regression models. Including these controls does not address the issue of selection on unobservables, but it can provide some information as to the direction of any remaining bias (Altonji, Elder, & Taber, 2005).

As an additional step to address possible selection bias, we also estimate propensity score matching models. As described in the methods section below, these models match our treatment group – women who worked by 12 months – with women who did not work by 12 months but are most like them in terms of the characteristics for which we have data. Again, these models cannot address selection on unobservables, but they are increasingly recognized as a good method to control for selection on observables in drawing causal inferences (see, e.g., Shonkoff & Phillips, 2000). With the exception of a study by Hill et al. (2005), propensity score matching

models have not previously been utilized in studies of the associations between early maternal employment and child outcomes.

A second possibility is that racial and ethnic differences in associations between first-year maternal employment and child outcomes are related to differences in the child care experiences or in-home experiences of children of different races or ethnicities. Specifically, we might hypothesize that such factors mediate the association between first-year maternal employment and later child outcomes, and that the role of such mediating influences differs by racial or ethnic group. For example, if maternal employment tends to be more stressful (on average) for women of a particular racial or ethnic group than it is for women of another group, then children whose mothers are in the former group and are employed in the first year may experience mothers who are more stressed or who exhibit harsher parenting. This, in turn, could lead to poorer outcomes for this group of children. Likewise, if maternal employment is more likely to alter child care arrangements in some racial or ethnic groups than in others, then differences in child care type or quality may partially or fully explain differential associations between first-year maternal employment and child outcomes across racial or ethnic groups. We test this hypothesis by estimating associations between first-year maternal employment and child outcomes, including interaction terms that test whether these associations differ by racial or ethnic group, while holding measures of non-parental child care, maternal depressive symptoms and stress, and selected aspects of parenting constant. We estimate these associations using both standard regressions and propensity score matching models.

3. Empirical strategy

Our research has two main objectives. The first is to estimate associations between first-year maternal employment and child outcomes, testing for whether these associations differ for different racial and ethnic groups. The second is to examine whether racial and ethnic differences in the associations between first-year maternal employment and child outcomes, where found, are explained by differences in child care experiences, maternal depressive symptoms and stress, or aspects of parenting.

To address our first objective – to determine whether associations between first-year maternal employment and later child outcomes differ by race/ethnic group – we estimate a series of ordinary least squares (OLS) regressions that focus on associations between maternal employment in the first year of a child's life and two child outcomes at age three: the child's receptive vocabulary, as measured by the Peabody Picture Vocabulary Test-Revised (PPVT-R), and the child's total number of behavior problems, as measured by a subset of items from the Child Behavior Checklist (CBCL). The PPVT-R and CBCL, which are described in detail below, have been widely used in prior work in this area, including several studies of children age three or four, using data from the NLSY-CS (see, e.g., Hill et al., 2005; Waldfogel et al., 2002) as well as the NICHD SECC (see, e.g., Brooks-Gunn et al., 2002). We focus our analyses on both measures because the prior literature has pointed to effects of early maternal employment on both cognitive and behavioral outcomes.

Consistent with prior research on this topic, we consider a mother as having worked in the first year after the focal child's birth if she worked any hours at all in the labor market that year (In supplemental models, we also distinguish between mothers who worked full-time vs. those who worked part-time, as some studies in this literature do). Roughly three-quarters of all women in this sample (78%) had worked by 12 months after the child's birth (Table 1). Descriptive statistics by racial and ethnic group, shown in Appendix Table 1, reveal that Black mothers were most likely to work by 12 months (81%), followed by White mothers (75%) and Hispanic mothers (72%) (the difference between the first-year employment rate of Black and Hispanic mothers is statistically significant; $p<.05$).

All of our models include controls for an extensive set of covariates (described below) that have been identified in theory and prior research as affecting child cognitive or behavioral development (see, e.g., Brooks-Gunn et al., 2002; Han et al., 2001; Waldfogel et al., 2002). Including these controls is important in order to address possible selection bias due to pre-existing differences between mothers who did or did not work by 12 months post-birth. Descriptive statistics shown in Table 1 indicate that mothers who worked by 12 months differ significantly from those who did not work by 12 months in several ways. For instance, the group of mothers that worked in the year following the focal child's birth was comprised of a higher proportion of Black mothers than the group that did not work in the year following the birth, while the opposite was true vis-à-vis Hispanic mothers. Those mothers who worked in the year following the birth are also more likely to have been married at the time of the birth, to have higher levels of education, and to have incomes above the poverty line than those who did not.

To further address the issue of possible selection bias, we estimate propensity score matching models. Here, we first use a probit model to estimate *propensity scores* representing the conditional probability that each mother worked in the 12 months following the birth as a function of all of the observed covariates that we use in our regression models (results not shown but available on request). Once these scores have been computed, each mother who worked by 12 months is matched to the comparison mother (who did not work by 12 months) with the closest propensity score. We perform this task using a matching with replacement strategy, such that comparison group mothers can be used as matches for multiple treatment group mothers. We then compare covariate distributions across treatment group (working by 12 months) and comparison group (not working by 12 months) mothers. This comparison (see Appendix Table 1) suggests that, in the matched sample in which all unmatched (*N*=155) mothers have been discarded, the treatment and comparison groups are similar vis-à-vis their background characteristics. Although the means of two of our covariates (mother and father cohabiting at the child's birth and family income above 200% of poverty at the child's birth) differ significantly in the two groups, results from a two-group Hotelling's *T*-squared generalized means test (*p*=.703) reveal that the set of means on the covariates, on the whole, are equal in the two groups. Thus, we re-estimate our regression models using the matched sample. This method is useful for estimating causal effects in the presence of selection bias (Rosenbaum and Rubin, 1983, 1984, 1985) in that it restricts inferences to the sample in which the treatment and comparison group covariate distributions overlap. However, like linear regression, it is only able to adjust for selection on observables and our estimates continue to be subject to possible omitted variable bias.

Our second objective is to assess whether the associations between first-year maternal employment and child outcomes, where found, are partially or fully explained (i.e., mediated) by differences between racial and ethnic groups in child care patterns, maternal depressive symptoms and stress, and aspects of parenting. To test for mediation, we re-estimate our matched and unmatched regressions including measures (detailed below) of the type of child care used, maternal depressive symptoms and stress, and aspects of parenting, assessed when the children were 12 and 36 months old. We compare the coefficients from models with and without the mediators to assess how their inclusion affects our estimates of associations between first-year maternal employment and the child outcomes both for the full sample and differentially for White, Black, and Hispanic mothers (by assessing the influence of controlling for the mediators on the coefficients of the interaction terms). Because our models include a considerable number of controls and mediators, which are likely to be correlated, we have computed variance inflation factors for each model to test for multi-collinearity. In all of our models, the mean variance inflation factor is below 2.2, suggesting that multi-collinearity is not a problem.

4. Data and measures

4.1. Data

Our data are drawn from the Fragile Families and Child Wellbeing Study (FFCW) and its associated in-home module. FFCW is a longitudinal birth cohort study of approximately 4900 children born between 1998 and 2000 in 20 medium to large U.S. cities. The study includes a substantial over-sample of children born to unmarried parents, resulting in a sample of children who are more likely to live in low income or poor families, to have absent fathers, and to be Black or Hispanic than would be the case in a nationally representative sample (see Reichman, Teitler, Garfinkel, & McLanahan, 2001, for a complete description of the sample and study design). The advantage of using this dataset is that it provides us with large samples of Black, Hispanic, and non-Hispanic white children. We note, however, that our results will be most applicable to low-income and primarily single-parent families, and may not generalize to more affluent and two-parent families.

FFCW interviewed families in person shortly after the focal child was born and conducted follow-up interviews by telephone when the focal child was approximately 12- and 36-months of age. Subsequent to the 36-month telephone interview, parents were asked to participate in an in-home module designed to assess multiple domains of parenting, the home environment, mother–child interactions, and child cognitive and emotional/behavioral development through both a questionnaire and a set of interviewer-observed items. Mothers who refused an in-home visit were asked to complete the questionnaire portion of the in-home module by telephone. Our analysis sample is limited to White, Black, and Hispanic families in which the mother was interviewed in all three waves of FFCW and in which the in-home assessment was completed in person (because our cognitive outcome measure cannot be completed by telephone). We further limit our sample to children for whom we have non-missing data on key variables included in our analyses.

In generating our analysis sample from the full FFCW sample of 4898 observations, we exclude cases for 6 reasons. First, we exclude 567 (11.6% of the full FFCW sample) observations from the two FFCW "pilot" cities (Austin and Oakland) because data on the timing of maternal work in the year after the focal child's birth was not collected in those cities. Second, we drop 667 (13.6%) cases because the mother did not complete interviews in all three waves. Third, we exclude 1375 (28.1%) cases in which families did not participate in the in-home assessment. Fourth, we drop 86 (1.7%) observations because the family's race or ethnicity was other than White, Black, or Hispanic. Fifth, we exclude an additional 75 (1.5%) observations with missing data on our outcome measures (described below). Finally, we exclude 645 (13.2%) observations with missing data on the maternal work in the year following the birth, mediator, or control variables. This yields a total analysis sample of 1483 children.

Despite the exclusion of these cases, our analysis sample is quite similar to the larger FFCW sample from which it is drawn. A majority of our sample is low income (36% had incomes below 100% of poverty at birth, while another 27% had incomes between 100 and 200% of the poverty line), and 79% of the children were born to unmarried parents (42% had single mothers and another 37% had mothers who were cohabiting with, but not married to, their fathers). A majority of the children (*N*=885) are Black (60%). Hispanic children (*N*=298) and White children (*N*= 300) each make up 20% of the sample.

4.2. Outcome variables

Cognitive development is measured by the child's score on the Peabody Picture Vocabulary Test-Revised (PPVT-R) when children were approximately 36-months old. The PPVT-R is a receptive vocabulary test that has been widely used to measure children's language and

cognitive ability (Dunn & Dunn, 1981). The PPVT-R must be administered in person. Fourteen of the children included in our analysis sample (4.7% of the Hispanic children in our sample), who spoke mainly Spanish, were administered the TVIP, a Spanish version of the test. We include these children in our analyses.

Child behavior problems are measured using a subset of items from the Child Behavior Checklist (CBCL) (Achenbach, 1992; Achenbach & Rescorla, 2000). The CBCL is a commonly used measure of children's behavioral problems and has been used extensively in prior studies of associations between maternal employment and child outcomes. It consists of a series of 99 items scored on a 0–2 point scale, resulting in total scores ranging from 0 to 188. For each item, individuals are assigned a score of 0 if they respond that the statement is "not true", 1 if they respond that the statement is "sometimes or somewhat true," or 2 if they respond that the statement is "very true or often true" of the focal child. Scores on each item are then summed to create a total score. The items in the CBCL are answered by the adult respondent to the survey, typically the child's mother. Items on the CBCL can be grouped into two broad subscales, representing internalizing behavior problems and externalizing problems respectively. The broad internalizing behavior problems subscale is comprised of four narrower subscales representing emotionally reactive, anxious/depressed, and withdrawn behavior problems, as well as somatic complaints; the broad externalizing behavior problems subscale is comprised of three narrower subscales representing sleep problems, attention problems, and aggressive behaviors.

FFCW includes a 35 item subset of the 99 item CBCL for children ages one and a half to five (Achenbach & Rescorla, 2000). These 35 items comprise the anxious/depressed, withdrawn, and aggressive subscales. Alphas for each of these subscales in FFCW are .64, .66 and .88. In our analyses, we utilize a combined measure of the three subscales (0 to 70 points) to approximate total behavior problems. In supplemental analyses, we also estimate (separately) associations of first-year maternal employment with both the aggressive subscale, to approximate externalizing behavior problems, and the combined anxious/depressed and withdrawn subscales, to approximate internalizing behavior problems.

To assist in the interpretation of our estimates, we standardized the child outcomes to have means of 0 and standard deviations of 1. Descriptive statistics by racial and ethnic group, shown in Appendix Table 2, suggest that White children in this sample score nearly a half of a standard deviation above the sample mean on the PPVT-R, whereas Black children score about one fifth of a standard deviation below the mean and Hispanic children score about eight-hundredths of a standard deviation below the mean. In terms of behavior problems, White children score at about the mean while Hispanic and Black children score slightly above the mean (0.17 and 0.10 standard deviations respectively).

4.3. Control variables

All models include the following controls (measured at birth unless otherwise specified): an indicator for whether the mother worked in the year before the birth; indicators for the child's gender (male=1) and whether the child had any older siblings; indicators for whether the child's parents were married or cohabiting at the time of the child's birth (with not married or cohabiting the excluded category); indicators for whether income was one to two times the poverty line, or more than twice the poverty line (with below poverty the excluded category); indicators for whether the mother had less than a high school education or more than a high school education (with just a high school education the excluded category); a control for the mother's age at the time of the child's birth; and an indicator for whether the mother was born in the United States. Because children were of varying ages at the time of the in-home assessment, we also include a control for the child's age (in months) at the time of the assessment in all of our models (In supplemental models, we stratify our sample by child gender

to see whether the results differ for boys and girls, as some prior studies, though not all, have found gender differences in how first-year maternal employment relates to later child outcomes).

All models also include the mother's PPVT-R score (measured at 36 months). The inclusion of the mother's PPVT-R score, which is a proxy for cognitive ability, is particularly important since differences in maternal cognitive ability are likely to be correlated with both maternal employment and children's vocabulary development. The 70 cases (less than 5% of our sample) with missing data on the mother's PPVT-R score have been assigned the sample mean of the PPVT-R; we include an indicator that the mother's PPVT-R is missing in our regressions. Our results are similar if we exclude observations for which the mother's PPVT-R is missing (although we achieve better balance in our matched sample when these cases are included). We have also estimated our models without controlling for the mother's PPVT-R and achieve results that are consistent with those presented here.

In models predicting children's behavior problems, we also include a control for maternal depressive symptoms, as measured at 36 months using an 8-point scale drawn from the Composite International Diagnostic Interview-Short Form (CIDI-SF) (Nelson, Kessler, & Mroczek, 1998). The CIDI-SF depression measure has been widely used in prior research and can be coded either as a dichotomous measure of major depression "caseness" or as an index of depressive symptoms. We use the latter approach, assigning a respondent one point for each affirmative response to items assessing whether the respondent was sad, blue, or depressed for two or more weeks during the last 12 months and whether, during that time period, she lost interest in things, felt more tired than usual, experienced a weight change of 10 or more pounds without trying, had more trouble sleeping than usual, had more trouble concentrating than usual, felt worthless, and thought a lot about death. Controlling for maternal depressive symptoms at 36 months is important because the measure of children's behavior problems is reported by the mother and thus could potentially be affected by her level of depressive symptoms at the time of the survey (see, e.g., Klebanov, Brooks-Gunn, McCarton, & McCormick, 1998; Klebanov, Brooks-Gunn, & McCormick, 2001; NICHD Early Child Care Research Network, 1999). We have also estimated our models controlling for major depression caseness in lieu of the depressive symptoms index. Results are consistent with those presented here.

Means for the control variables for the full sample are shown in Table 1, and means by racial and ethnic group are shown in Appendix Table 2. Of particular interest, we see in Appendix Table 2 that Black mothers are less likely than White or Hispanic mothers to be married and more likely to be poor at the time of the child's birth; they also have lower PPVT-R scores (at 36 months) than White mothers. Hispanic mothers are more likely than both White and Black mothers to have less than a high school degree at the time of the birth.

4.4. Mediating variables

As discussed above, maternal employment in the first year may affect the care children receive from non-parental caregivers or the care provided by parents, and these associations may differ by race and ethnic group. If so, controlling for these potential mediating factors might help explain differential associations, where found, between maternal employment and child outcomes across racial and ethnic groups. We examine this hypothesis using several different measures of non-parental child care, maternal depressive symptoms and stress, and aspects of parenting, gathered in the 12-month phone interview, the 36-month phone interview, and the 36-month home visit.

We measure children's child care experience using information reported by the mother about the type of child care in which the child was placed at 12 months and at 36 months. We use

this information to distinguish four categories of care—center-based child care, relative child care, non-relative child care, or other child care. Children who received less than 10 h per week of non-parental child care or no non-parental care serve as the reference category. Descriptive statistics (Table 1) indicate that the use of center-based care at 36 months is more common among families where the mother worked by 12 months (as is the use of relative care and non-relative care). Prior research has found that children who attend center-based care tend to have better cognitive outcomes, but may also have more behavior problems (prior research on relative child care and non-relative child care has produced less consistent results) (NICHD Early Child Care Research Network, 2007; see, also, review in Smolensky & Gootman, 2003). Thus, type of child care is a potentially important mediator to control for in our analyses.

We also control for maternal depressive symptoms, an important aspect of mental health which prior research has found may be affected by first-year maternal employment and which has also been shown to have important effects on children's development and in particular children's behavior problems (see, e.g., Klebanov et al., 1998, 2001; NICHD Early Child Care Research Network, 1999; see also review in Shonkoff & Phillips, 2000). We measure maternal depressive symptoms, at both 12 months and 36 months, using an 8-point maternal depressive symptoms scale drawn from the CIDI-SF (Nelson et al., 1998), as described above. In our sample, we do not see any differences in mean levels of reported depressive symptoms at 12 or 36 months between mothers who did or did not work by 12 months (Table 1).

Our models also control for parenting stress, which may be affected by employment and may, in turn, affect both cognitive and behavioral development (Crnic & Acevedo, 1995; Crnic, Gaze, & Hoffman, 2005; Crnic & Greenberg, 1990). Our measure of parenting stress, which is collected at both 12 months and 36 months, is a 16-point scale based on 4 items from the Panel Study of Income Dynamics—Child Development Supplement's Aggravation in Parenting Scale (Mainieri & Grodsky, 2006). Items are measured on a 4-point scale ascertaining the extent to which the mother agrees that being a parent is harder than she expected, she feels trapped by her responsibilities as a parent, she finds taking care of her children much more work than pleasure, and she often feels tired, worn out, or exhausted from raising a family. As shown in Table 1, in our data parenting stress levels do not differ between mothers who did or did not work by 12 months as reported at the 12-month interview. But, at the 36-month interview, mothers who worked by 12 months reported significantly lower levels of parenting stress than those who did not work by 12 months.

We also include controls for several other aspects of parenting, again focusing on those that might be affected by maternal employment in the first 12 months and that might, in turn, affect children's cognitive or behavior development. We include three controls for discipline, an important aspect of parenting (Baumrind, 1996; Deater-Deckard & Dodge, 1997; Deater-Deckard, Dodge, & Sorbring, 2005; Smith & Brooks-Gunn, 1997). We control for the mother's report of whether she spanked the child in the month prior to the interview. Prior research has suggested that there may be some associations between maternal employment and the use of spanking (see e.g. Berger, 2007). We see such an association in our sample, with mothers who worked by 12 months significantly more likely to report spanking their child in the past month at both 12 months and 36 months (Table 1). Although the links between spanking and child cognitive and behavioral outcomes are not clear cut and may vary by racial or ethnic group (see e.g. Deater-Deckard et al., 2005; Deater-Deckard & Dodge, 1997), a considerable body of research has found associations between spanking and adverse developmental outcomes for children (see, e.g., Gershoff, 2002). Thus, it is of interest to include spanking as a possible mediator in our analyses.

We also include measures of mothers' psychological aggression and physical discipline, drawn from the Conflict Tactics Scales (CTS) (Straus et al., 1998; Straus & Gelles, 1990). The CTS

is the most widely used measure of parental disciplinary practices and has been extensively used to measure child abuse. The measure of psychological aggression is designed to assess the frequency (during the last year) with which the mother shouted, yelled, or screamed at the focal child, threatened to spank or hit him or her but didn't actually do it, swore or cursed at, called him or her dumb or lazy (or something similar), and said she would send him or her away or would kick him or her out of the house. The measure of physical discipline assesses the frequency with which the mother spanked the focal child on the bottom with her bare hand, hit him or her on the bottom with something like a belt, hairbrush, stick or some other hard object, slapped him or her on the hand, arm, or leg, pinched him or her, and shook him or her. As shown in Table 1, mothers in our sample who worked by 12 months score significantly higher on both psychological aggression and physical discipline at 36 months than those who did not work by 12 months.

In addition to the discipline measures, our models include two measures of nurturance, another key aspect of parenting (Bornstein, 2002; Collins, Maccoby, Steinberg, Hetherington, & Bornstein, 2000; Ispa et al., 2004; Maccoby & Martin, 1983). Specifically, we control for harsh parenting and lack of maternal responsivity, drawing on subscales of the Home Observation for Measurement of the Environment (HOME), a very widely used measure of the quality of the home environment for learning and cognition (Bradley, 1993; Caldwell & Bradley, 1984). The particular subscales we use have been developed to measure specific constructs that are related to nurturance and that are thought to be particularly consequential for children's cognitive or behavioral development (Fuligni, Han, & Brooks-Gunn, 2004; Leventhal, Martin, & Brooks-Gunn, 2004). Harsh parenting consists of the sum of 5 dichotomous interviewer-observed items, including whether the mother shouted at the focal child, expressed annoyance with or hostility toward the child, slapped or spanked the child, scolded or criticized the child, and interfered with or restricted the child more than 3 times (with the exception of protecting the child from harm) during the in-home interview. We would expect harsh parenting to be related to child behavior problems. Our lack of maternal responsivity measure is comprised of 6 dichotomous interviewer-observed items, including whether the mother did not spontaneously vocalize to child twice or more during the in-home visit, did not respond verbally to child's vocalizations, did not tell the child the name of an object or person during visit, did not spontaneously praise the child at least twice, did not caress or kiss the child at least once, and whether her voice did not convey positive feelings toward the child. Lack of maternal responsivity, which is the closest measure we have to the constructs of parental sensitivity and responsiveness, might be related to both vocabulary and behavior. In our sample, as shown in Table 1, there is no difference at 36 months in the mean level of harsh mothering between mothers who did or did not work by 12 months, but mothers who worked by 12 months score significantly lower in terms of lacking maternal responsivity (that is, they are more responsive) (Because the 12-month FFCW interview was conducted by telephone, the HOME, which must be administered in person, could not be administered at that time point. Thus we do not have measures of these two constructs at 12 months.).

Finally, we include a measure of the lack of cognitively stimulating materials available to the child, also drawn from a subscale of the HOME . This subscale is comprised of 11 dichotomous items assessing specific aspects of parenting that are particularly consequential for children's cognitive and language development, including parents' language, teaching, and provision of materials (Brooks-Gunn & Markman, 2005; Hart & Risley, 1995, 1999; Rebello-Britto & Brooks-Gunn, 2001; Rebello-Britto, Fuligni, & Brooks-Gunn, 2001; Snow, 1991). The individual items focus on the numbers and types of toys and reading materials available in the home. In our sample, as shown in Table 1, mothers who worked by 12 months score significantly lower in terms of lacking cognitively stimulating materials at 36 months (that is, they provide more cognitively stimulating materials) (As noted above, the HOME was not administered in FFCW at 12 months, so we do not have a measure of cognitively stimulating

materials at that time point.) (An appendix with further details on our measures is available on request.).

5. Results

5.1. Do the associations between first-year maternal employment and child receptive vocabulary vary by racial and ethnic group?

Table 2 shows results from OLS regressions of children's PPVT-R scores on indicators of first-year maternal employment, holding constant the control variables discussed earlier. Full results for the controls are shown in the table but we do not discuss them here, as our focus is on the first-year maternal employment coefficients.

Column1 shows results for our full sample of 1483 Black, White, and Hispanic children. Here we see that the coefficient on working by 12 months is not significantly different from zero. This indicates that, for our overall sample, there is no significant association between whether a mother worked in the 12 months following a child's birth and the child's PPVT-R score. However, the results in Column 2 indicate that this conclusion does not extend to all racial and ethnic groups.

Column 2 shows results from a model for the same sample, and with the same controls, but adding interaction terms that test for whether the association between working by 12 months and child PPVT-R score differs by racial or ethnic group. Specifically, we include interaction terms for White*working by 12 months and Hispanic*working by 12 months. The main effect of working by 12 months in the model provides the estimated association for Black children, who are the reference group. Consistent with prior research, we find that there is a significant negative association between first-year maternal employment and the child's PPVT-R score for White, but not Black or Hispanic children. To calculate the total effect of working by 12 months for White children, we must add the coefficient on the main effect for working by 12 months (.062, n.s.) and the coefficient on the White*working by 12 months interaction term (−.323, $p<.05$). Doing so indicates that maternal employment within 12 months of a focal child's birth is associated with a 0.24 standard deviation lower PPVT-R score at age 3 for White children.

Although our models include an extensive set of controls for variables that are associated with selection into employment, we might still worry that we have not adequately controlled for possible selection bias. For this reason, as detailed above, we next estimated a propensity score matching model. This model is also estimated using OLS but includes only those in the matched sample (N=1328). The results for the matched sample are shown in Column 3. Restricting our analysis to the matched sample does not alter the conclusion that there is a significant negative association between first-year maternal employment and child PPVT-R for White children but not Black or Hispanic children. However, it does increase the magnitude of that association. The results in column 3 (again adding the main effect of working by 12 months to the coefficient for White*working by 12 months) indicate that working by 12 months is associated with about a .50 standard deviation lower PPVT-R score at age 3 for White children. This change in results – as we move from the full sample to the propensity score matched sample – suggests that, if anything, selection bias may lead to our under-estimating the negative association between first-year maternal employment and PPVT-R scores for White children.

Our reference group in these models includes all women who did not work by 12 months, some of whom went to work by 36 months and some of whom did not. It is possible that our results might be biased by not including a control for whether the mother worked at all by 36 months. Accordingly, we re-estimated the models adding a control for whether the mother had worked by 36 months and found that our results (not shown but available on request) were nearly

identical to those reported in Table 2. We also estimated supplemental models in which we allowed the effect of first-year maternal employment to vary by whether that employment was full-time or part-time. In results not shown but available on request, we found that both part-time and full-time employment in the first year were associated with significantly lower PPVT-R scores for White children. In addition, we estimated supplemental models stratifying the sample by child gender. In these models (results not shown but available on request), significant negative associations between first-year maternal employment and lower PPVT-R scores for White children were found for girls, but not for boys.

5.2. Do the associations between first-year maternal employment and child behavior problems vary by racial or ethnic group?

Table 3 presents results for the analyses of the associations between first-year maternal employment and child behavior problems. In column 1, we see that there is no significant association between working by 12 months and child behavior problems at 36 months in our overall sample (*N*=1483). However, results shown in column 2 indicate that, when we test for interactions between working by 12 months and racial or ethnic group (again in the full sample), we find that the association between working by 12 months and later child behavior problems does differ by race and ethnicity. Specifically, Hispanic children whose mothers worked by 12 months have significantly more behavior problems at 36 months than comparable children whose mothers did not work by 12 months. To estimate the full association for Hispanic children, we must add the main effect of working by 12 months (−.104, n.s.) and the interaction effect for Hispanic*working by 12 months (.298, *p*<.05). Doing so indicates that working by 12 months is associated with 0.19 standard deviations higher behavior problem scores for Hispanic children.

Again, we also estimated propensity score matching models to address possible selection bias associated with differences between mothers who do and do not work by 12 months. The results for the propensity score matched sample (*N*=1328) shown in column 3 still point to a significant elevation of behavior problems for Hispanic children whose mothers work by 12 months and, as with the PPVT-R effect for White children, the magnitude of the effect is larger than in the unmatched sample. Adding the main effect and the interaction term from the matched model in column 3 indicates that working by 12 months is associated with 0.449 standard deviations higher behavior problem scores for Hispanic children. Another point of difference in the propensity score matched results in column 3, as compared to the full sample (unmatched) results in column 2, is that in column 3 the main effect of working by 12 months is nearly twice as large and is now statistically significant. This main effect represents the association between working by 12 months and child behavior problems for Black children, who are the reference category in the regression. Thus, the results in column 3 indicate that working by 12 months is associated with significantly lower levels of behavior problems for Black children; the reduction is nearly .20 of a standard deviation.

Again, these results are robust to adding a control for whether the mother was employed by 36 months (results not shown but available on request). In addition, we estimated supplementary models where we analyzed the associations of working by 12 months and our proxies for children's scores on the internalizing and externalizing subscales of the CBCL (rather than the total behavior problem score). In these models (results not shown but available on request) we found that the above pattern of results held up in the propensity score matched models for both the internalizing and externalizing subscales.

In supplemental models (results not shown but available on request), we found significant associations between first-year maternal employment and later behavior problems for Hispanic children regardless of whether that employment was full-time or part-time. We also found that these significant associations were present for both boys and girls.

5.3. Do child care experiences, maternal depressive symptoms or stress, or aspects of parenting mediate the associations between first-year maternal employment and child receptive language?

The results reported in Table 2 indicate that there is a negative association between working by 12 months and child PPVT-R scores for White, but not Black or Hispanic, children and that this association is present in both full sample (unmatched) regressions and propensity score matched regressions. Our next set of analyses – again using both OLS models and propensity score models – address the question of whether child care experiences, maternal depressive symptoms and stress, or aspects of parenting mediate that association. We show two sets of results in Table 4. The results in the first two columns are from full sample (unmatched) and propensity score matched models that add controls for mediators assessed at 12 months, while those in the last two columns are from full sample (unmatched) and propensity score matched models that add controls for mediators assessed at both 12 and 36 months.

Regardless of whether just the 12 month variables (columns 1 and 2), or both the 12 and 36 month variables (columns 3 and 4), are entered, the striking result is that the coefficients on working by 12 months and the interaction of White*working by 12 months are basically unchanged from their values in Table 2. In the full sample model, the combined effect of working by 12 months and White*working by 12 months predicts a .27 standard deviation reduction in a White child's PPVT-R score; in the propensity score matched model, the predicted reduction is twice as large, ranging from −.53 (column 2) to −.57 (column 4). Thus, the potential mediators for which we control do not explain the association between working by 12 months and lower PPVT-R scores at 36 months for White children. This may be because the potential mediators themselves have fairly weak associations with the outcome. The 12-month mediators in particular have few significant associations with children's PPVT-R scores. Among the 36-month mediators, the most robust associations are that children whose mothers are rated as more lacking in maternal responsivity have lower PPVT-R scores, while children whose mothers are rated as more psychologically aggressive have higher scores. This latter result is a bit puzzling but might suggest that mothers who use more psychological techniques for discipline may be more verbal, in general, with their children and thus may be promoting language development.

5.4. Do child care experiences, maternal depressive symptoms or stress, or aspects of parenting mediate the associations between first-year maternal employment and child behavior problems?

Our final set of analyses, reported in Table 5, analyzes the role of child care type, maternal depressive symptoms and stress, and aspects of parenting in mediating the associations between first-year maternal employment and child behavior problems. We found, in models shown in Table 3, that there is a significant association between working by 12 months and higher behavior problem scores for Hispanic, but not Black or White, children. This association was present in both full sample (unmatched) regressions and propensity score matched regressions. In addition, when we moved to the propensity score matched regressions, we found a significant negative main effect of working by 12 months, indicating that working by 12 months is associated with lower behavior problem scores for Black children, the reference group in our models.

Our next set of analyses addresses the role of our potential mediators in explaining those associations. The results in the first two columns of Table 5 are from full sample (unmatched) and propensity score matched models that add controls for mediators assessed at 12 months, while those in the last two columns are from full sample (unmatched) and propensity score matched models that add controls for mediators assessed at both 12 and 36 months.

Adding the 12 month or 12 month and 36 month variables to the full sample (unmatched) model (columns 1 and 3) results in the Hispanic*working by 12 months association becoming larger; this association was .298 in Table 3 and is .372 (column 1) or .380 (column 3) in Table 5. This result suggests that the variables we are adding play a suppressor role, such that excluding them from the model was suppressing part of the effect of working by 12 months among Hispanics. This indicates that Hispanic families tend to have better characteristics on these variables and that, when we do not control for them, the relative disadvantage faced by Hispanic children whose mothers work in the first year is understated. We do not see evidence of this type of suppression in the propensity score matched results, perhaps because these models do a better job of controlling for selection bias. In the propensity score matched results in Table 3, the coefficient on Hispanic*working by 12 months was larger (.646) than in the full sample (matched) model, and we see coefficients of roughly the same magnitude here (columns 2 and 4). The significant main effect for working by 12 months in the propensity score matched model (that we saw in Table 3) is also relatively unaffected by the addition of the potential mediators here in Table 5. Thus, adding the potential mediators affects the working by 12 months coefficients in the full sample (unmatched) model, but not in the propensity score matched model.

The child care controls have generally weak associations with child behavior problem scores. We find more significant effects of the parent-related variables. In particular, parenting stress at 12 months and 36 months is associated with higher levels of behavior problems, as are psychological aggression at 36 months and harsh mothering at 36 months.

Again, in results not shown (but available on request), we repeated these analyses for our proxies for children's internalizing and externalizing problems (rather than total behavior problems) and found that the same overall pattern of results held.

6. Discussion

Although we caution that, like previous studies in this area, our results may not represent causal estimates, they nevertheless provide the first evidence on the extent to which associations between first-year maternal employment and child outcomes vary by race and ethnicity in a sample that contains large numbers of Black, White, and Hispanic children. Our results indicate that, as suggested by prior research, these associations are not uniform across racial and ethnic groups. We find that maternal employment in the first year after birth is associated with lower PPVT-R scores for White children, but not Black or Hispanic children. We also find significant associations between first-year maternal employment and elevated levels of child behavior problems for Hispanic children, but not White or Black children.

These differential associations by racial and ethnic group are not easily explained. To address the possibility that they reflect differential selection into employment by White, Black, and Hispanic women, we estimated models that include a rich set of controls to account for selection into employment. We also estimated propensity score matching models, whereby we limit the analysis to only those sample members who are most like our treatment group — families where the mothers worked by 12 months. The propensity score matched results point to, if anything, larger associations between first-year maternal employment and child outcomes than those from the full sample models, reinforcing the conclusion that working by 12 months is associated with lower PPVT-R scores for White children, and more behavior problems for Hispanic children. They also suggest that these associations are not likely to be due to differential selection into employment. In addition, the propensity score matching models point to significant benefits of first-year maternal employment for Black children, in the form of lower levels of behavior problems.

To explore the reasons for these differential associations across racial and ethnic groups, we estimated mediating models where we add controls for several potential mediators, measured at 12 months and 36 months post-birth. The mediators we control for include type of child care, maternal depressive symptoms and stress, and several aspects of parenting (including measures of discipline, nurturance, and the provision of cognitively stimulating materials), all factors that are likely to be affected by maternal employment and that also might be related to child outcomes. These variables had generally weak associations with the child's PPVT-R score and adding them to the model did not alter the estimated associations between working by 12 months and the PPVT-R.

Adding the potential mediators to the behavior problem model, in contrast, did affect the estimated associations with working by 12 months. Specifically, in the (unmatched) full sample models, adding the child care, maternal depressive symptoms and stress, and aspects of parenting variables resulted in a larger estimated association for Hispanic*working by 12 months, suggesting that omitting these variables suppresses some of the effect of that interaction. However, adding these variables did not affect the associations between working by 12 months and child behavior problems in the propensity score matched models, perhaps because those models better account for selection bias in the first place.

Although the results presented answer the questions we set out to address – whether associations of first-year maternal employment with child outcomes vary by race and ethnicity and whether these associations, were found, can be explained by differences in child care experiences, maternal depressive symptoms and stress, or parenting behaviors – they leave other questions unanswered. In particular, the question remains as to what accounts for the differential patterns of associations between working by 12 months and child outcomes by race and ethnicity. Why would working by 12 months be associated with lower vocabulary scores for White children but not Black or Hispanic children? And why would working by 12 months lead to more behavior problems for Hispanic children, but not White or Black children? Having examined a plausible set of mediators related to the type of child care, maternal depressive symptoms and stress, and several aspects of parenting, and found that they do not explain these associations, we can only speculate as to what other factors might play a role.

Turning first to the PPVT-R results for White children, we note that in our sample, White mothers have significantly higher PPVT-R scores themselves and also are significantly less likely to be rated as lacking in the provision of cognitively stimulating materials than Black mothers or Hispanic mothers (see Appendix Table 2). If these mean differences imply that the average White mother is more engaged or more productive in promoting her child's vocabulary acquisition than the average Black or Hispanic mother, then that might, at least in part, explain why the association between working by 12 months and lower vocabulary scores is uniquely found for White mothers and not the other two groups. It is also possible that parents from different racial or ethnic groups engage in different activities with their young children, or emphasize different activities in early childhood. Children's language development has been linked to the amount and quality of verbal interactions, such as conversation and shared book reading, with their mothers, and prior research suggests that these interactions vary considerably by race and ethnicity (as well as SES), with White children (and higher SES children) tending to be read to and engaged in conversation more often (Brooks-Gunn & Markman, 2005). If White mothers are more likely to engage in activities such as book reading with their infants (see e.g. Raikes et al., 2006), this might help explain why working by 12 months is associated with lower levels of vocabulary for White children but not the other two groups. We are unable to fully test these possibilities given that FFCW, like most studies of children's development, lacks direct measures of maternal language output. This will be an important direction for future research.

Turning to the behavior problem results for Hispanic children, there are several possible explanations for why working by 12 months would be associated with more behavior problems for that group, but not for White or Black children. The Hispanic women in our sample differ from the other two groups in that they are on average less educated and less likely to be born in the United States (see Appendix Table 2). We do not have information about the type of occupation or job that the mother holds, but it may be that the Hispanic mothers who are working by 12 months hold jobs that are more stressful or less rewarding than those held by the other mothers, perhaps because they are less educated and have less experience in the US labor market. If so, the stress in the workplace may spill over to the children, resulting in more behavior problems (see, e.g., Parke et al., 2004). It may also be that the meaning of work, and the mother's satisfaction with work, varies by racial or ethnic group. In our sample, Hispanic mothers were least likely to have been working before the child's birth (Appendix Table 2). Evidence also suggests that values and behaviors regarding work outside the home, family roles, and child rearing may differ for Hispanics as compared to Blacks or Whites, as well among sub-groups within the Hispanic population (see, e.g., Oropesa, 1996; Stier & Tienda, 1992).

Another possibility is that the differential associations reflect differences in how Hispanic mothers engage with their children. In our sample, Hispanic mothers have the lowest mean scores for psychological aggression, physical discipline, and harsh mothering (Appendix Table 2). Prior research has suggested that Hispanic mothers tend to be more physically involved with their infants, providing more hands-on support with tasks such as feeding and drinking (Carlson & Harwood, 2003; Ispa et al., 2004). It may be that, due to these types of differences, the disruption to a child's social and emotional development caused when a Hispanic mother works by 12 months is greater than when a comparable White or Black mother works. This, too, is an important direction for future research.

Another result to be explained is that, across both outcomes examined, we find no evidence of adverse effects of first-year maternal employment for Black children. Indeed, we find some evidence of beneficial effects in terms of reduced behavior problems (in the propensity score matching models). Black women, of course, have historically had higher levels of employment than White or Hispanic women, and it may be that working by 12 months is seen as more normative in Black families than it is in White or Hispanic families. It is also possible that maternal employment leads to fewer changes for the typical Black child than it does for the typical White or Hispanic child. Researchers have long noted the greater involvement of kin in the rearing of Black children (see, e.g., Stack, 1974). If friends and family members are already caring for children part of the time, alongside the mother, the impact on the child of the mother going to work in the first year may be less consequential than in a situation where the mother was providing sole or primary care with less involvement on the part of kin. Further research, in particular ethnographic studies that explore what parents do with their infants and young children across different racial and ethnic groups, would be very useful here.

In interpreting our results, it is important to note that our sample consists primarily of low income, urban families. Future research with more representative samples will be an important next step. It is also important to consider that there is tremendous heterogeneity within racial and ethnic groups, much of which has likely been obscured in our analyses. Our sample has allowed us to estimate average effects for large samples of Black, White, and Hispanic children. But we have not been able to look within racial and ethnic groups where surely there is also considerable variation to be examined. This within group variation is particularly important for Hispanic families, who come from a range of different countries and cultures, but is also important for Black and White families who should not be seen as homogeneous groups. We also note that there is variation within mothers who work in the first year of life, as to when

they begin work and how intensively they work. We have examined some of this variation here, but these aspects of maternal employment should be examined further in future research.

Further research should also examine a wider set of possible mediators. In spite of the richness of the FFCW data, the set of mediators included in this study is nevertheless limited, and it is possible that different results would have been obtained with a richer set of measures. We do not, for example, have measures of the quality of non-parental child care to which children are exposed, and this could be an important factor for both cognitive and behavioral outcomes. For instance, if Hispanic children are attending poorer quality child care than Black or White children, this might help explain the association between maternal employment in the first year and later child behavior problems given that low quality child care has been linked to increased behavior problems (see, e.g., Loeb, Bridges, Bassok, Fuller, & Rumberger, 2005; NICHD Early Child Care Research Network, 2003).

We also do not have complete controls for the type of child care that children have experienced to date, but rather snapshot measures of the type of care children are enrolled in at 12 months and 36 months. The lack of specificity of our child care measures may help explain why we find generally weak associations between child care type and the child outcome variables, in contrast to other research which has found associations between, in particular, center-based care and both cognitive and behavior outcomes (NICHD Early Child Care Research Network and Duncan, 2003; NICHD Early Child Care Research Network, 2007).

Additionally, of particular relevance for children's cognitive outcomes, we do not have a measure of cognitive and language stimulation or maternal–child interaction in the first year of life. If, as discussed earlier, White mothers (on average) provide more such stimulation in the first year of life, controlling for this would be important. In this regard, it would be particularly useful to examine the role of maternal–child book reading, which has been found to affect PPVT scores in other studies (Rebello-Britto & Brooks-Gunn, 2001; Snow, 1991), but which is not measured in FFCW.

Another significant omission is that we do not include measures of the fathers' parenting, or the quality of interactions between fathers and mothers. Such measures could be particularly important for the results for child behavior among Hispanic children. Prior research (with White non-Hispanic families) has found that fathers are less satisfied with their marriage when their wife works in the first year after a birth (Hyde, Essex, & Horton, 1993). If the same dynamics hold for Hispanic families, assessing the father's parenting and the quality of his relationship with the mother could be important. These factors, and other unobserved differences, could possibly account at least in part for the associations we find, as well as their differences across racial and ethnic groups.

Finally, it is important to understand the long-run implications of first-year maternal employment for children's development. Research that tracks children over time will allow us to learn whether the associations we have found for children at age three, and their differences across racial and ethnic groups, persist over time, and also whether associations that are not apparent at age three emerge as children age.

7. Conclusion

Despite the limitations noted above, this study provides new evidence of racial and ethnic differences in associations of first-year maternal employment and cognitive and behavioral outcomes for children at age three among primarily low income, urban families. We find that first-year maternal employment is associated with lower vocabulary scores for White, but not Black or Hispanic, children and with elevated levels of behavior problems for Hispanic, but not White or Black, children. These associations are not explained by differences in the types

of child care, levels of maternal depressive symptoms and stress, or aspects of parenting (including discipline, nurturance, and provision of cognitively stimulating materials) to which children in various racial and ethnic groups are exposed. Future research should further examine the mechanisms through which these associations may operate, as well as whether these mechanisms differ by racial and ethnic groups, in order to illuminate the processes driving differential associations of first-year maternal employment with child outcomes for White, Black, and Hispanic children.

References

Achenbach, T. Manual for the child behavior checklist/2–3 and 1992 profile. Burlington, VT: University of Vermont Department of Psychiatry; 1992.

Achenbach, T.; Rescorla, L. Manual for the ASEBA preschool forms and profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families; 2000.

Altonji J, Elder T, Taber C. Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. Journal of Political Economy 2005;113:151–184.

Bates JE, Marvinney D, Kelly T, Dodge KA, Bennett DS, Pettit GS. Child-care history and kindergarten adjustment. Developmental Psychology 1994;30:690–700.

Baum CL. Does early maternal employment harm child development? An analysis of the potential benefits of leave-taking. Journal of Labor Economics 2003;21(2):409–448.

Baumrind D. The discipline controversy revisited. Family Relations 1996;45:405–414.

Baydar N, Brooks-Gunn J. Effects of maternal employment and child care arrangements in infancy on preschoolers' cognitive and behavioral outcomes: Evidence from the children of the NLSY. Developmental Psychology 1991;27:918–931.

Belsky J, Eggebeen D. Early and extensive maternal employment/child care and 4–6 year olds socioemotional development: Children of the National Longitudinal Survey of Youth. Journal of Marriage and the Family 1991;53:1083–1099.

Berger, LM. Socioeconomic Factors and Substandard Parenting. Madison: University of Wisconsin Manuscript; 2007.

Blau FD, Grossberg AJ. Maternal labor supply and children's cognitive development. Review of Economics and Statistics 1992;74(3):474–481.

Bornstein, MH., editor. Handbook of Parenting: Children and Parenting. Mahway NJ: Lawrence Erlbaum Associates; 2002.

Bradley RH. Children's home environments, health, behavior, and intervention efforts: A review using the HOME inventory as a marker measure. Genetic, Social, and General Psychology Monographs 1993;119:437–490.

Brooks-Gunn J, Han W, Waldfogel J. Maternal employment and child cognitive outcomes in the first three years of life: The NICHD Study of Early Child Care. Child Development 2002;73:1052–1072. [PubMed: 12146733]

Brooks-Gunn J, Markman LB. The contribution of parenting to ethnic and racial gaps in school readiness. The Future of Children 2005;15:139–168. [PubMed: 16130545]

Caldwell, BM.; Bradley, RH. Administration manual (revised edition): Home observation for measurement of the home environment. Little Rock, AR: University of Arkansas; 1984.

Carlson V, Harwood R. Attachment, culture, and the caregiving system: The cultural patterning of everyday experiences among Anglo and Puerto Rican mother–infant pairs. Infant Mental Health Journal 2003;24:53–73.

Collins WA, Maccoby EE, Steinberg L, Hetherington EM, Bornstein MH. Contemporary research on parenting: The case for nature and nurture. American Psychologist 2000;55(2):218–232. [PubMed: 10717969]

Crnic, K.; Acevedo, M. Everyday stresses and parenting. In: Bornstein, MH., editor. Handbook of parenting. Mahwah, NJ: Lawrence Erlbaum; 1995. p. 277-297.

Crnic KA, Gaze C, Hoffman C. Cumulative parenting stress across the preschool period: Relations to maternal parenting and child behaviour at age 5. Infant and Child Development 2005;14:117–132.

Crnic KA, Greenberg MT. Minor parenting stresses with young children. Child Development 1990;61:1628–1637. [PubMed: 2245752]

Deater-Deckard K, Dodge KA. Externalizing behavior problems and discipline revisited: Nonlinear effects and variation by culture, context, and gender. Psychological Inquiry 1997;8:161–175.

Deater-Deckard, K.; Dodge, KA.; Sorbring, E. Cultural difference in the effects of physical punishment. In: Rutter, M.; Tienda, M., editors. Ethnicity and causal mechanisms. New York: Cambridge University Press; 2005.

Desai S, Chase-Lansdale L, Michael R. Mother or market? Effects of maternal employment on cognitive development of four year old children. Demography 1989;26:545–561. [PubMed: 2583316]

Dunn, LM.; Dunn, LM. Peabody Picture Vocabulary Test-Revised. Circle Pines, MN: American Guidance Service; 1981.

Fuligni AS, Han WJ, Brooks-Gunn J. The infant–toddler HOME in the second and third years of life. Parenting: Science and Practice 2004;4(23):139–159.

Gershoff E. Corporal punishment by parents and associated child behaviors and experiences: A meta-analytic and theoretical review. Psychological Bulletin 2002;128:530–579.

Han W, Waldfogel J, Brooks-Gunn J. The effects of early maternal employment on later cognitive and behavioral outcomes. Journal of Marriage and the Family 2001;63:336–354.

Hart, B.; Risley, T. Meaningful differences in the everyday experience of young American children. Baltimore: Paul Brookes Publishing; 1995.

Hart, B.; Risley, T. The social world of children learning to talk. Baltimore: Paul Brookes Publishing; 1999.

Haskins R. Public school aggression among children with varying day care experience. Child Development 1985;56:689–703. [PubMed: 4006574]

Hill J, Waldfogel J, Brooks-Gunn J, Han W. Maternal employment and child development: A fresh look using newer methods. Developmental Psychology 2005;41:833–850. [PubMed: 16351331]

Hyde JS, Essex MJ, Horton F. Fathers and parental leave: Attitudes and experiences. Journal of Family Issues 1993;14:616–638.

Ispa JM, Fine M, Halgunseth LC, Harper S, Robinson J, Boyce L, et al. Maternal intrusiveness, maternal warmth, and mother–toddler relationship outcomes: Variations across low-income ethnic and language groups. Child Development 2004;75:1613–1631. [PubMed: 15566369]

James-Burdumy S. The effect of maternal labor force participation on child development. Journal of labor Economics 2005;25(1):177–211.

Klebanov PK, Brooks-Gunn J, McCarton C, McCormick MC. The contribution of neighborhood and family income to developmental test scores over the first three years of life. Child Development 1998;69:1420–1436. [PubMed: 9839425]

Klebanov PK, Brooks-Gunn J, McCormick MC. Maternal coping strategies and emotional distress: Results of an early intervention program for low birth weight young children. Developmental Psychology 2001;37:654–667. [PubMed: 11552761]

Loeb, S.; Bridges, M.; Bassok, D.; Fuller, B.; Rumberger, R. How much is too much? The influence of preschool centers on children's social and cognitive development. National Bureau of Economic Research Working Paper; 2005. p. 11812

Leventhal T, Martin A, Brooks-Gunn J. The EC-HOME across five national datasets in the third to fifth year of life. Parenting: Science and Practice 2004;4(23):161–188.

Maccoby, E.; Martin, JA. Socialization in the context of the family: parent–child interaction. In: Hetherington, EM., editor. Handbook of Child Psychology, Vol. 4, Socialization, Personality, and Social Development. 4th Edition. New York: Wiley; 1983.

Mainieri, T.; Grodsky, M. The Panel Study of Income Dynamics Child Development Supplement: User Guide Supplement for CDS-I. Ann Arbor, MI: Institute for Social Research, University of Michigan; 2006.

National Institute of Child Health and Human Development Early Child Care Research Network. Chronicity of maternal depressive symptoms, maternal sensitivity, and child functioning at 36 months. Child Development 1999;35:1297–1310.

National Institute of Child Health and Human Development Early Child Care Research Network. Does amount of time spent in child care predict socioemotional adjustment during the transition to kindergarten? Child Development 2003;4:976–1005.

National Institute of Child Health and Human Development Early Child Care Research Network. Are there long-term effects of early child care? Child Development 2007;78:681–701. [PubMed: 17381797]

National Institute of Child Health and Human Development Early Child Care Research Network, & Duncan, Greg. Modeling the impacts of child care quality on preschool children's cognitive development. Child Development 2003;74:1454–1475. [PubMed: 14552408]

Neidell M. Early parental time investments in children's human capital development: Effects of time in the first year on cognitive and non-cognitive outcomes. UCLA mimeo. 2000

Nelson, CB.; Kessler, RC.; Mroczek, D. Scoring the World Health Organization's Composite International Diagnostic Interview Short Form (CIDI-SF; v 1.0 Nov 98). Geneva, Switzerland: Unpublished manuscript, Epidemiology, Classification and Assessment Group, World Health Organization; 1998.

Oropesa RS. Normative beliefs about marriage and cohabitation: A comparison of non-Latino whites, Mexican Americans, and Puerto Ricans. Journal of Marriage and the Family 1996;58(1):49–62.

Parke R, Coltrane S, Duffy S, Buriel R, Dennis J, Powers J, et al. Economic stress, parenting, and child adjustment in Mexican American and Euro-American families. Child Development 2004;75:1632–1656. [PubMed: 15566370]

Raikes H, Luze G, Brooks-Gunn J, Raikes HA, Pan BA, Tamis-LeMonda CS, et al. Mother–child bookreading in low-income families: Correlates and outcomes during the first three years of life. Child Development 2006;77:924–953. [PubMed: 16942498]

Rebello-Britto P, Brooks-Gunn J. Beyond shared book reading: dimensions of the home process and low-income African American preschoolers' skills. New Directions for Child Development 2001;92:73–90.

Rebello-Britto, P.; Fuligni, A.; Brooks-Gunn, J. An open book? Effects of home-based approaches on children's literacy development. In: Dickinson, D.; Neuman, S., editors. Handbook of Early literacy Research. Vol. Vol. 2. New York: Guilford Press; 2001.

Reichman NE, Teitler JO, Garfinkel I, McLanahan SS. Fragile Families: Sample and design. Children and Youth Services Review 2001;23(4–5):303–326.

Rosenbaum P, Rubin D. The central role of the propensity score in observational studies for causal effects. Biometrica 1983;70(1):41–55.

Rosenbaum P, Rubin D. Reducing bias in observational studies: Using subclassifications on the propensity score. Journal of the American Statistical Association 1984;79:516–524.

Rosenbaum P, Rubin D. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. The American Statistician 1985;39:33–38.

Ruhm C. Parental employment and child cognitive development. Journal of Human Resources 2004;39 (1):155–192.

Shonkoff, JP.; Phillips, DA., editors. From neurons to neighborhoods: The science of early childhood development. Washington, DC: National Academy Press; 2000.

Smith JR, Brooks-Gunn J. Correlates and consequences of harsh discipline for young children. Archives of Pediatrics and Adolescent Medicine 1997;151:777–786. [PubMed: 9265878]

Smolensky, E.; Gootman, J., editors. Working families and growing kids: Caring for children and adolescents. Washington, DC: National Academy Press; 2003.

Snow C. The theoretical basis for relationships between language and literacy development. Journal of Research in Childhood Education 1991;6(1):5–10.

Stack, C. All our kin. New York: Basic Books; 1974.

Stier H, Tienda M. Family, work and women: The labor supply of Hispanic immigrant wives. International Migration Review 1992;26:1291–1313.

Straus, M.; Gelles, RJ., editors. Physical violence in American families: Risk factors and adaptations to violence in 8145 families. New Brunswick, NJ: Transaction Publishers; 1990.

Straus M, Hamby SL, Finkelhor D, Moore DW, Runyan D. Identification of child maltreatment with the Parent–Child Conflict Tactics Scales: Development and psychometric data for a national sample of American parents. Child Abuse and Neglect 1998;22(4):249–270. [PubMed: 9589178]

Waldfogel J, Han W, Brooks-Gunn J. The effects of early maternal employment on child cognitive development. Demography 2002;39(2):369–392. [PubMed: 12048957]

Table 1

Means and standard deviations for full sample and by maternal work status by 12-months

	All families	Mother working by 12-months	Mother not working by 12-months
Outcomes			
PPVT-R (standardized)	−0.034	−0.026	−0.063
	(0.964)	(0.951)	(1.008)
Total behavior problems (standardized)	0.090	0.071	0.158
	(1.009)	(0.997)	(1.049)
Predictors			
Working by 12 months	0.777		
White	0.202	0.194	0.230
Hispanic	0.201	0.186[*]	0.252
Black	0.597	0.619[**]	0.518
Mother worked in year before birth	0.822	0.883[***]	0.609
Child male	0.514	0.524	0.482
Older siblings	0.622	0.610	0.664
Mother and father married at child's birth	0.206	0.193[*]	0.252
Mother and father cohabiting at child's birth	0.374	0.380	0.352
Income above poverty to 200% of poverty at child's birth	0.269	0.284[*]	0.218
Income above 200% of poverty at child's birth	0.372	0.389[**]	0.309
Mother's PPVT-R (standardized)	−0.027	0.030[***]	−0.226
	(1.175)	(1.134)	(1.292)
Mother had less than high school education at child's birth	0.313	0.275[***]	0.445
Mother had more than high school education at child's birth	0.350	0.369[**]	0.285
Mother's age at child's birth	24.609	24.585	24.691
	(5.806)	(5.755)	(5.986)
Mother US born	0.944	0.957[***]	0.900
Child age (months) at in-home assessment	37.748	37.825[*]	37.478
	(2.598)	(2.634)	(2.453)
Maternal depressive symptoms at 36-months (standardized)	0.052	0.056	0.037
	(1.052)	(1.062)	(1.019)
Mother's PPVT-R missing	0.047	0.043	0.061
12-month mediators			
Center-based child care	0.163	0.195[***]	0.048
Relative child care	0.211	0.252[***]	0.070
Non-relative child care (excludes center based)	0.061	0.074[***]	0.018
Other child care arrangement	0.030	0.036[*]	0.009

	All families	Mother working by 12-months	Mother not working by 12-months
Mother spanked child in last month	0.301	0.326***	0.212
Parenting stress (standardized)	−0.009	−0.031	0.069
	(1.004)	(0.996)	(1.028)
Maternal depressive symptoms (standardized)	0.000	−0.013	0.046
	(1.001)	(0.984)	(1.062)
36-month mediators			
Center-based child care	0.322	0.359***	0.191
Relative child care	0.219	0.234**	0.167
Non-relative child care (excludes center based)	0.071	0.077	0.048
Other child care arrangement	0.005	0.004	0.006
Mother spanked child in last month	0.570	0.598***	0.473
Parenting stress (standardized)	−0.008	−0.038*	0.098
	(0.998)	(0.990)	(1.022)
Psychological aggression (standardized)	0.129	0.165*	0.003
	(1.029)	(1.026)	(1.032)
Physical discipline (standardized)	0.110	0.171***	−0.104
	(1.051)	(1.067)	(0.967)
Harsh mothering (standardized)	0.003	−0.019	0.080
	(1.001)	(0.947)	(1.168)
36-month mediators			
Lack of maternal responsivity (standardized)	−0.009	−0.038*	0.093
	(0.992)	(0.969)	(1.065)
Lack of cognitively stimulating materials (standardized)	0.004	−0.034**	0.135
	(0.989)	(0.980)	(1.010)

Note: 1483 observations. Standard deviations in parentheses.

*
 p<0.05;

**
 p<0.01;

***
 p<0.001 for *t*-statistics testing mean differences for mothers working and not working by 12-months.

Table 2

First-year maternal employment and receptive vocabulary (PPVT-R), basic model for unmatched and matched samples

	Without interactions	With interactions	Matched
Working by 12 months	−0.040	0.062	0.037
	(0.058)	(0.076)	(0.059)
White	0.288**	0.538**	0.749**
	(0.067)	(0.126)	(0.093)
Hispanic	0.065	0.167	−0.096
	(0.062)	(0.122)	(0.098)
White*working by 12 months		−0.323*	−0.533**
		(0.138)	(0.120)
Hispanic*working by 12 months		−0.127	0.124
		(0.137)	(0.129)
Mother worked in year before birth	0.055	0.059	−0.055
	(0.064)	(0.065)	(0.079)
Child male	−0.136**	−0.138**	−0.160**
	(0.045)	(0.045)	(0.047)
Older siblings	−0.083	−0.087	−0.083
	(0.053)	(0.053)	(0.054)
Mother and father married at child's birth	−0.034	−0.041	−0.071
	(0.076)	(0.076)	(0.082)
Mother and father cohabiting at child's birth	−0.023	−0.024	0.105
	(0.053)	(0.053)	(0.054)
Income above poverty to 200% of poverty at child's birth	−0.019	−0.023	0.061
	(0.060)	(0.060)	(0.061)
Income above 200% of poverty at child's birth	0.104	0.101	0.026
	(0.065)	(0.065)	(0.067)
Mother's PPVT-R (standardized)	0.229**	0.229**	0.268**
	(0.023)	(0.023)	(0.024)
Mother had less than high school education at child's birth	0.133*	0.141*	0.076
	(0.060)	(0.060)	(0.061)
Mother had more than high school education at child's birth	0.196**	0.195**	0.163**
	(0.061)	(0.061)	(0.062)
Mother's age at child's birth	0.012**	0.013**	0.020**
	(0.005)	(0.005)	(0.005)
Mother US born	0.104	0.107	0.067
	(0.107)	(0.108)	(0.124)
Child age (months) at in-home assessment	0.018*	0.018*	0.025**
	(0.009)	(0.009)	(0.010)

	Without interactions	With interactions	Matched
Mother's PPVT-R missing	−0.397**	−0.401**	−0.189
	(0.107)	(0.107)	(0.126)
Intercept	−1.194**	−1.264**	−1.576**
	(0.381)	(0.385)	(0.427)
Observations	1483	1483	1328
R-squared	0.20	0.20	0.27

Note: Coefficients (and standard errors) from OLS regressions presented. Outcomes are standardized to have means of 0 and standard deviations of 1.

*
$p<0.05$;

**
$p<0.01$;

***
$p<0.001$.

Table 3

First-year maternal employment and behavior problems, basic model for unmatched and matched samples

	Without interactions	With interactions	Matched
Working by 12 months	−0.028	−0.104	−0.197**
	(0.064)	(0.084)	(0.065)
White	0.218**	0.181	0.095
	(0.074)	(0.139)	(0.102)
Hispanic	0.138*	−0.090	−0.406**
	(0.069)	(0.135)	(0.107)
White*working by 12 months		0.045	0.223
		(0.152)	(0.132)
Hispanic*working by 12 months		0.298*	0.646**
		(0.151)	(0.142)
Mother worked in year before birth	0.048	0.036	−0.017
	(0.071)	(0.071)	(0.087)
Child male	0.170**	0.169**	0.094
	(0.050)	(0.050)	(0.052)
Older siblings	0.121*	0.119*	0.288**
	(0.058)	(0.058)	(0.059)
Mother and father married at child's birth	−0.209*	−0.210*	−0.387**
	(0.084)	(0.084)	(0.090)
Mother and father cohabiting at child's birth	−0.032	−0.030	−0.092
	(0.058)	(0.058)	(0.059)
Income above poverty to 200% of poverty at child's birth	0.009	0.007	0.161*
	(0.066)	(0.066)	(0.067)
Income above 200% of poverty at child's birth	−0.076	−0.073	0.083
	(0.072)	(0.072)	(0.073)
Mother's PPVT-R (standardized)	−0.091**	−0.091**	−0.120**
	(0.026)	(0.026)	(0.027)
Mother had less than high school education at child's birth	0.094	0.093	0.070
	(0.067)	(0.067)	(0.067)
Mother had more than high school education at child's birth	−0.164*	−0.170*	−0.101
	(0.068)	(0.068)	(0.068)
Mother's age at child's birth	−0.014*	−0.014**	−0.017**
	(0.005)	(0.005)	(0.005)
Mother US born	−0.076	−0.109	−0.409**
	(0.118)	(0.119)	(0.136)
Child age (months) at in-home assessment	−0.023*	−0.023*	−0.040**
	(0.010)	(0.010)	(0.010)

	Without interactions	With interactions	Matched
Mother's PPVT-R missing	0.021	0.020	0.223
	(0.119)	(0.119)	(0.138)
Maternal depressive symptoms at 36-months (standardized)	0.123[**]	0.122[**]	0.126[**]
	(0.024)	(0.024)	(0.025)
Intercept	1.216[**]	1.339[**]	2.323[**]
	(0.422)	(0.426)	(0.469)
Observations	1483	1483	1328
R-squared	0.10	0.11	0.14

Note: Coefficients (and standard errors) from OLS regressions presented. Outcomes are standardized to have means of 0 and standard deviations of 1.

[*]
 $p<0.05$;

[**]
 $p<0.01$.

Table 4

First-year maternal employment and receptive vocabulary (PPVT-R), net of mediators for matched and unmatched samples

	12-month mediators		36-month mediators	
	Unmatched	**Matched**	**Unmatched**	**Matched**
Employment and race variables				
Working by 12 months	0.040	0.033	0.033	0.035
	(0.078)	(0.062)	(0.077)	(0.062)
White	0.549**	0.786**	0.483**	0.730**
	(0.127)	(0.095)	(0.126)	(0.094)
Hispanic	0.186	−0.064	0.161	−0.099
	(0.122)	(0.099)	(0.121)	(0.098)
White*working by 12 months	−0.309*	−0.565**	−0.309*	−0.603**
	(0.139)	(0.122)	(0.137)	(0.120)
Hispanic*working by 12 months	−0.148	0.087	−0.190	0.053
	(0.137)	(0.130)	(0.135)	(0.127)
12-month mediators				
Center-based child care	0.032	0.027	−0.018	−0.035
	(0.067)	(0.075)	(0.069)	(0.076)
Relative child care	0.095	0.110	0.084	0.098
	(0.062)	(0.067)	(0.062)	(0.068)
Non-relative child care (excludes center based)	−0.111	−0.113	−0.153	−0.183
	(0.100)	(0.125)	(0.100)	(0.123)
Other child care arrangement	0.051	0.336*	0.029	0.210
	(0.136)	(0.146)	(0.136)	(0.146)
Mother spanked child in last month	0.077	−0.000	0.114*	0.031
	(0.052)	(0.054)	(0.053)	(0.056)
Parenting stress (standardized)	−0.023	−0.018	−0.030	−0.004
	(0.023)	(0.025)	(0.027)	(0.028)
Maternal depressive symptoms (standardized)	−0.007	0.039	−0.003	0.049*
	(0.023)	(0.024)	(0.023)	(0.024)
36-month mediators				
Center-based child care			0.096	0.108
			(0.059)	(0.061)
Relative child care			0.036	0.048
			(0.062)	(0.065)
Non-relative child care (excludes center based)			0.133	0.275**
			(0.097)	(0.098)
Other child care arrangement			0.300	0.593
			(0.328)	(0.339)

	12-month mediators		36-month mediators	
	Unmatched	**Matched**	**Unmatched**	**Matched**
Mother spanked child in last month			−0.103[*]	−0.097
			(0.051)	(0.053)
Parenting stress (standardized)			0.017	0.002
			(0.027)	(0.028)
Psychological aggression			0.077[**]	0.143[**]
			(0.028)	(0.028)
Physical discipline			−0.035	−0.110[**]
			(0.029)	(0.028)
Harsh mothering			−0.072[**]	0.000
			(0.024)	(0.025)
Lack of maternal responsivity			−0.120[**]	−0.128[**]
			(0.025)	(0.025)
Lack of cognitively stimulating materials			−0.047	−0.074[**]
			(0.026)	(0.028)
Observations	1483	1328	1483	1328
R-squared	0.20	0.27	0.24	0.32

Note: Coefficients (and standard errors) from OLS regressions presented. Outcomes are standardized to have means of 0 and standard deviations of 1. These models also control for all of the covariates listed in Table 1.

[*] $p < 0.05$;

[**] $p < 0.01$;

[***] $p < 0.001$.

Table 5

First-year maternal employment and behavior problems, net of mediators for matched and unmatched samples

	12-month mediators		36-month mediators	
	Unmatched	**Matched**	**Unmatched**	**Matched**
Employment and race variables				
Working by 12 months	−0.088	−0.170[*]	−0.090	−0.147[*]
	(0.085)	(0.067)	(0.080)	(0.065)
White	0.174	0.125	0.215	0.157
	(0.138)	(0.103)	(0.131)	(0.099)
Hispanic	−0.115	−0.327[**]	−0.016	−0.226[*]
	(0.133)	(0.107)	(0.127)	(0.103)
White[*]working by 12 months	0.053	0.208	0.069	0.204
	(0.151)	(0.132)	(0.143)	(0.126)
Hispanic[*]working by 12 months	0.372[*]	0.653[**]	0.380[**]	0.630[**]
	(0.149)	(0.141)	(0.141)	(0.133)
12-month mediators				
Center-based child care	−0.096	0.034	−0.057	0.074
	(0.073)	(0.081)	(0.072)	(0.080)
Relative child care	−0.121	−0.151[*]	−0.147[*]	−0.169[*]
	(0.067)	(0.072)	(0.065)	(0.072)
Non-relative child care (excludes center based)	−0.043	−0.067	−0.057	−0.061
	(0.108)	(0.135)	(0.104)	(0.129)
Other child care arrangement	−0.120	−0.042	−0.112	0.004
	(0.148)	(0.158)	(0.142)	(0.154)
Mother spanked child in last month	0.103	0.195[**]	0.009	0.016
	(0.056)	(0.058)	(0.056)	(0.059)
Parenting stress (standardized)	0.163[**]	0.130[**]	0.070[*]	0.058[*]
	(0.025)	(0.027)	(0.028)	(0.029)
Maternal depressive symptoms (standardized)	0.045	0.062[*]	0.045	0.037
	(0.027)	(0.028)	(0.026)	(0.027)
36-month mediators				
Center-based child care			−0.107	−0.263[**]
			(0.061)	(0.064)
Relative child care			0.104	−0.009
			(0.065)	(0.069)
Non-relative child care (excludes center based)			−0.048	−0.204[*]
			(0.101)	(0.103)
Other child care arrangement			0.560	0.616
			(0.343)	(0.356)

	12-month mediators		36-month mediators	
	Unmatched	**Matched**	**Unmatched**	**Matched**
Mother spanked child in last month			0.076	0.127[*]
			(0.053)	(0.056)
Parenting stress (standardized)			0.085[**]	0.078[**]
			(0.028)	(0.030)
Psychological aggression			0.218[**]	0.187[**]
			(0.030)	(0.029)
Physical discipline			0.040	0.064[*]
			(0.030)	(0.030)
Harsh mothering			0.104[**]	0.101[**]
			(0.025)	(0.026)
Lack of maternal responsivity			−0.007	0.005
			(0.026)	(0.026)
Lack of cognitively stimulating materials			0.076[**]	0.018
			(0.027)	(0.029)
Observations	1483	1328	1483	1328
R-squared	0.14	0.18	0.24	0.27

Note: Coefficients (and standard errors) from OLS regressions presented. Outcomes are standardized to have means of 0 and standard deviations of 1. These models also control for all of the covariates listed in Table 2.

[*] $p<0.05$;

[**] $p<0.01$;

[***] $p<0.001$.

Berger et al.

Page 32

Appendix A Appendix Table 1

Mean statistics for matched and unmatched samples

	Means	Mother not working by 12-months	Mother working by 12-months	t-statistic	(p-value)
White	Unmatched	0.194	0.230	−1.44	(0.151)
	Matched	0.199	0.201	−0.11	(0.915)
Hispanic	Unmatched	0.186	0.252	−2.60	(0.009)
	Matched	0.187	0.162	1.52	(0.129)
Mother worked in year before birth	Unmatched	0.883	0.609	12.00	(0.000)
	Matched	0.905	0.894	0.85	(0.394)
Child male	Unmatched	0.524	0.482	1.35	(0.178)
	Matched	0.523	0.520	0.13	(0.898)
Older siblings	Unmatched	0.610	0.664	−1.78	(0.075)
	Matched	0.608	0.633	−1.23	(0.218)
Mother and father married at child's birth	Unmatched	0.193	0.252	−2.34	(0.019)
	Matched	0.191	0.178	0.77	(0.441)
Mother and father cohabiting at child's birth	Unmatched	0.380	0.352	0.94	(0.348)
	Matched	0.383	0.432	−2.31	(0.021)
Income above poverty to 200% of poverty at child's birth	Unmatched	0.284	0.218	2.37	(0.018)
	Matched	0.282	0.320	−1.96	(0.051)
Income above 200% of poverty at child's birth	Unmatched	0.389	0.309	2.67	(0.008)
	Matched	0.393	0.369	1.19	(0.235)
Mother's PPVT-R (standardized)	Unmatched	0.030	−0.226	3.50	(0.000)
	Matched	0.038	0.015	0.46	(0.644)
Mother had less than high school education at child's birth	Unmatched	0.275	0.445	−5.96	(0.000)
	Matched	0.269	0.262	0.39	(0.699)
Mother had more than high school education at child's birth	Unmatched	0.369	0.285	2.82	(0.005)
	Matched	0.370	0.339	1.52	(0.129)
Mother's age at child's birth	Unmatched	24.585	24.691	−0.29	(0.771)

NIH-PA Author Manuscript   NIH-PA Author Manuscript   NIH-PA Author Manuscript

	Means				
		Mother not working by 12-months	Mother working by 12-months	*t*-statistic	(*p*-value)
	Matched	24.647	24.879	−0.95	(0.341)
Mother US born	Unmatched	0.957	0.900	3.96	(0.000)
	Matched	0.961	0.954	0.74	(0.458)
Child age (months) at in-home assessment	Unmatched	37.825	37.478	2.14	(0.032)
	Matched	37.748	37.795	−0.44	(0.658)
Mother's PPVT-R missing	Unmatched	0.043	0.061	−1.30	(0.193)
	Matched	0.041	0.031	1.26	(0.208)
2-group Hotelling's *T*-squared		12.722			
F-statistic		0.786			
(*p*-value)		(0.703)			

Note: Means for the matched (obs.=1328) and unmatched (obs.=1483) samples, by maternal work by 12 months following the focal child's birth, are presented.

Appendix Table 2

Means and standard deviations by race and ethnicity

	White(obs=300)	Black(obs=885)	Hispanic(obs=298)
Outcomes			
PPVT-R (standardized)	0.476_a	-0.191_b	-0.081_b
	(1.047)	(0.902)	(0.881)
Total behavior problems (standardized)	-0.005_a	0.096_a	0.167_a
	(0.986)	(1.010)	(1.026)
Predictors			
Working by 12 months	0.747_{ab}	0.807_b	0.721_a
Mother worked in year before birth	0.870_a	0.829_a	0.752_b
Child male	0.503_a	0.516_a	0.520_a
Older siblings	0.560_a	0.652_b	0.594_{ab}
Mother and father married at child's birth	0.473_a	0.116_b	0.201_c
Mother and father cohabiting at child's birth	0.340_a	0.348_a	0.483_b
Income above poverty to 200% of poverty at child's birth	0.233_a	0.277_a	0.282_a
Income above 200% of poverty at child's birth	0.637_a	0.292_b	0.342_b
Mother's PPVT-R (standardized)	0.870_a	-0.289_b	-0.151_b
	(1.151)	(1.009)	(1.221)
Mother had less than high school education at child's birth	0.170_a	0.330_b	0.406_c
Mother had more than high school education at child's birth	0.573_a	0.293_b	0.295_b
Mother's age at child's birth	26.673_a	24.089_b	24.074_b
	(6.086)	(5.623)	(5.603)
Mother US born	0.980_a	0.977_a	0.809_b
Child age (months) at in-home assessment	37.640_a	37.561_a	38.414_b
	(2.373)	(2.526)	(2.907)
Maternal depressive symptoms at 36-months (standardized)	0.058_a	0.093_a	-0.075_a
	(1.065)	(1.072)	(0.972)
Mother's PPVT-R missing	0.033_a	0.052_a	0.047_a
12-month mediators			
Center-based child care	0.147_a	0.176_a	0.138_a
Relative child care	0.163_a	0.209_{ab}	0.265_b
Non-relative child care (excludes center based)	0.093_a	0.056_{ab}	0.044_b
Other child care arrangement	0.043_a	0.032_a	0.013_a
Mother spanked child in last month	0.193_a	0.373_b	0.195_a
Parenting stress (standardized)	-0.094_a	0.027_a	-0.028_a
	(0.858)	(1.055)	(0.980)
Maternal depressive symptoms (standardized)	-0.032_a	0.029_a	-0.053_a
	(1.005)	(1.016)	(0.953)

	White(obs=300)	Black(obs=885)	Hispanic(obs=298)
36-month mediators			
Center-based child care	0.253$_a$	0.369$_b$	0.248$_a$
Relative child care	0.190$_a$	0.212$_a$	0.268$_a$
Non-relative child care (excludes center based)	0.133$_a$	0.055$_b$	0.054$_b$
Other child care arrangement	0.017$_a$	0.002$_b$	0.000$_b$
Mother spanked child in last month	0.540$_{ab}$	0.605$_a$	0.500$_b$
Parenting stress (standardized)	−0.035$_a$	0.000$_a$	−0.004$_a$
	(0.875)	(1.037)	(0.999)
Psychological aggression (standardized)	0.086$_a$	0.231$_a$	−0.132$_b$
	(1.078)	(1.028)	(0.930)
Physical discipline (standardized)	0.007$_a$	0.264$_b$	−0.245$_c$
	(0.995)	(1.103)	(0.829)
Harsh mothering (standardized)	−0.184$_a$	0.128$_b$	−0.182$_a$
	(0.803)	(1.092)	(0.827)
Lack of maternal responsivity (standardized)	−0.387$_a$	0.170$_b$	−0.160$_c$
	(0.592)	(0.592)	(1.093)
Lack of cognitively stimulating materials (standardized)	−0.501$_a$	0.201$_b$	−0.072$_c$
	(0.656)	(1.035)	(0.941)

Note: 1483 observations. Standard deviations in parentheses. Means in the same row that do not share subscripts differ at $p<0.05$ in ANOVA analyses with Scheffe post-hoc tests.